Title
Target site recognition in Tn7 transposition

Author
Kubo, Kenneth M.

Publication Date
1992

Peer reviewed|Thesis/dissertation

Target Site Recognition in Tn7 Transposition

by

Kenneth M. Kubo

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

Acknowledgements

I am deeply grateful to many who have helped me through my years at UCSF: to Nancy Craig, who has been a wonderful mentor, teaching me much about science and life; to the members of the Craig laboratory and the Hooper Foundation, who have made "the lab with the great view" such a special place to work and play; to Bruce Alberts and Rick Myers, for their valuable advice and encouragement; to my parents, who will always believe I am young enough to be cared for and fed and in need of money; and to my wife Cindy, for her continual love and patience through the ups and downs of my graduate school education -- I couldn't have come this far without her.

Chapter 2 is adapted from Kubo and Craig, 1990 (*J. Bacteriol.* 172:2774-2778).

Target Site Recognition in Tn7 Transposition

by

Kenneth M. Kubo

ABSTRACT

The bacterial transposon Tn7 is unusual in its ability to transpose at high frequency into a specific site in the Escherichia coli chromosome, called attTn7. I have examined Tn7 insertion sites in the E. coli chromosome and identified two distinct classes of target sites. Target site choice was determined by two overlapping ensembles of Tn7-encoded genes. tnsABC + tnsE display little target site selectivity, directing insertion into many different sites unrelated in sequence. By contrast, tnsABC + tnsD display high target sequence selectivity, directing insertion into attTn7 and a limited number of other sites related in sequence to attTn7.

To examine the role of tnsD in Tn7 insertion into attTn7, I have purified the TnsD protein and show that TnsD directly binds to attTn7, recognizing sequences located at a distance from the specific point of Tn7 insertion. I also show that TnsD directly participates in and is essential for transposition to attTn7 in a cell-free recombination system. I provide evidence that TnsC, an ATP-dependent non-sequence-

specific DNA binding protein, also associates with the attTn7 target site when TnsD is present. These studies indicate that TnsD is a critical determinant in the selection of attTn7 as a specific Tn7 insertion site and suggest that TnsD communicates with the rest of the transposition machinery through its interaction with TnsC.

Table of Contents

Chapter One:

Introduction.....1

Chapter Two:

Bacterial transposon Tn7 utilizes two different classes of target sites 44

Chapter Three:

Purification and characterization of TnsD, a Tn7 transposition protein that binds specifically to the attTn7 target site 62

Chapter Four:

Tn7 transposition: Target DNA recognition by multiple Tn7-encoded proteins 99

Chapter Five:

Summary and Perspectives 152

References 166

List of Tables

CHAPTER TWO

Table 1:

Strains and plasmids used in this study 52

Table 2:

Transposition frequency to the *E. coli* chromosome 53

Table 3:

Physical analysis of miniTn7Km insertions in the blocked *attTn7* chromosome 54

CHAPTER THREE

Table 1:

TnsD Purification 82

Table 2:

Evaluation of *attTn7* target activity 83

List of Figures

CHAPTER ONE

Figure 1:

Steps involved in transposition and retroviral integration 38

Figure 2:

Cleaved donor and strand transfer complexes formed during transposition 41

CHAPTER TWO

Figure 1:

Tn7, attTn7, and miniTn7Km 56

Figure 2:

Distribution of chromosomal insertion sites 58

Figure 3:

Sequence analysis of Tn7 insertion sites 60

CHAPTER THREE

Figure 1:

Purification of TnsD 84

Figure 2:

Mobility shift analysis of TnsD binding to attTn7 86

Figure 3:

Cosedimentation of TnsD polypeptide and TnsD-dependent binding activity to attTn7 88

Figure 4:

Footprint analysis of TnsD and TnsD+Host complexes with attTn7 90

Figure 5:

Summary of TnsD protection of attTn7 nucleotide sequences against nuclease attack 94

Figure 6:

TnsD is required for in vitro Tn7 transposition to attTn7 96

CHAPTER FOUR

Figure 1:

Tn7 transposition substrates, intermediates, and products 132

Figure 2:

Purified Tns proteins reconstitute Tn7 transposition *in vitro* 134

Figure 3:

Tn7 transposition *in vitro* is stimulated by a preincubation step that lacks MgAc 137

Figure 4:

Formation of TnsC + TnsD complexes with *attTn7* 140

Figure 5:

DNase I footprint analysis of TnsD and TnsC + TnsD complexes with *attTn7* 143

Figure 6:

Summary of DNase I footprint analysis 145

Figure 7:

Tn7 transposition in the presence of AMP-PNP 148

Figure 8:

Tn7 transposition immunity *in vitro* is altered by the presence of AMP-PNP 151

Chapter 1

Introduction

Transposable elements are specialized DNA segments which can move from one site to another site, called a target site (reviewed in Berg and Howe 1989). Barbara McClintock first identified these elements, which she named controlling elements, from her cytological and genetic studies of maize (McClintock 1951, 1956). However, these elements were not widely recognized until it was found that many spontaneous mutations in the lac and gal operons of Escherichia coli resulted from transposable element insertions into these loci (Malamy 1966; Jordan et al. 1967; Adhya and Shapiro 1969; Shapiro and Adhya 1969). Since then, examples of these elements have been characterized in all prokaryotic and eukaryotic organisms in which they have been sought. It is now clear that these elements can comprise a significant portion (from 1-10%) of the genome of an organism (Finnegan and Fawcett 1986; Fanning and Singer 1987; Ajioka and Hartl 1989).

Functional transposable elements encode one or more proteins important for transposition and contain special cis-acting DNA sites, usually located at the ends of the element, required for movement of the element (Grindley and Reed 1985; Craig and Kleckner 1987). Some elements also encode other accessory genes, such as antibiotic resistance or biosynthetic functions, that can provide some selective advantage to the organism (Kopecko 1980). Many transposable elements, however, contain only the determinants needed for their movement and can proliferate within a genome, which suggests that these elements are "selfish DNAs" whose major function is to maintain themselves within the host genome (Doolittle and Carpienza 1980; Orgel and Crick 1980). Despite their

capacity to increase in copy number, many elements tightly regulate their transposition to avoid an extremely high level of proliferation that would prove detrimental to the organism (Doolittle et al. 1984; Kleckner 1990).

Target site selection is one important aspect of transposable elements. In general, these elements insert into many different genomic sites, which can have dramatic consequences for the host organism (for reviews, see Syvanen 1984; Finnegan 1989; Galas and Chandler 1989). These insertions may disrupt host gene function or affect gene expression near the insertion site, for example, by providing regulatory sequences to adjacent host genes. In addition, these elements can produce gross genetic changes, including deletions, inversions, and replicon fusions. These chromosomal rearrangements can occur by homologous recombination between multiple copies of the element in the genome. Because of their potential to promote genetic change, transposable elements have been postulated to play an important role in genome evolution (Chao et al. 1983; Syvanen 1984). The ability of transposable elements to disrupt gene function has also been exploited as a tool for genetic analysis, since the element can serve as a convenient insertional mutagen that marks the site of mutation (Berg et al. 1989; Soriano et al. 1989; Kleckner et al. 1991).

The overall goal of this thesis is to understand the mechanism of target site selection in transposition. The focus of this work is the bacterial transposon Tn7, which displays an unusual target site selectivity. To provide a conceptual framework for examining Tn7

transpostion, this introduction will review two major aspects of target site selection in transposition:

1) The determinants of target specificity displayed by transposable elements. Generally, transposable elements can insert into many different sites, but many do show at least some bias in their selection of target sites. Both the nucleotide sequence and general structure of the target DNA can affect the choice of insertion sites. This target site specificity reflects the particular interactions between the transposition apparatus and the target DNA.

2) The mechanism of target site recognition and utilization during transposition. The recent development of *in vitro* transposition systems for several elements has provided a mechanistic understanding of the target site transactions involved in transposition. Surprisingly, these studies have revealed that these transposable elements, although derived from a wide spectrum of organisms, share many features in their transposition mechanisms (Sherratt 1991), which may indicate that transposition is a process conserved through evolution.

Finally, an introduction to Tn7 is presented to address the significance of target site selection in Tn7 transposition.

Types of transposable elements

Transposable elements can be categorized on the basis of their genetic organization, DNA sequence homologies, and mechanism of transposition. For prokaryotic elements, the following major groups have emerged:

1. IS elements and composite transposons. Insertion sequences (IS elements) are small DNA segments, from about 800 to 1500 bp in length, that usually encode a single protein, a transposase, which is required for transposition. These elements also contain short terminal inverted repeats that are recognized by the element-encoded transposase (reviewed in Galas and Chandler, 1989). IS elements can transpose as individual entities, but frequently act in pairs to mobilize an unrelated DNA segment lying between the elements. A composite transposon contains the two IS elements and the central DNA segment, which often encodes a determinant of selective advantage to the organism, such as antibiotic resistance or sugar utilization.

In some cases the IS elements encode inhibitors to regulate their transposition. IS1, for example, produces a fusion protein InsAB, which is the active transposase and results from translational frameshifting between insA and insB (reviewed in Plasterk 1991). Moreover, IS1 produces in much greater abundance the InsA polypeptide, which acts as an inhibitor for transposition, perhaps by competing with InsAB for binding to the IS1 ends (Sekine and Ohtsubo 1989; Escoubas et al. 1991). Another element, IS50, which forms the composite transposon Tn5, also contains a transposition inhibitor (p2) that is a shortened form of the transposase (p1) (Yin and Reznikoff 1988).

2. Noncomposite transposons. These elements can mobilize antibiotic resistance determinants, but otherwise are distinct from IS-flanked composite transposons. In particular, an individual end of a noncomposite transposons cannot by itself translocate, as found with the IS modules composite elements; instead, the two ends must act in

combination to promote transposition. One predominant group of such elements includes Tn3 and its relatives (Arthur and Sherratt 1979; Grindley 1983; Grindley and Reed 1985). These elements encode a transposase which carries out transposition replicatively, generating a cointegrate that contains two copies of the element. The cointegrate results from the transposase-mediated fusion of the original donor element-carrying molecule and the target molecule, which can then be resolved by another element-encoded protein, a resolvase, that promotes recombination between a specific site in each element. Another group of noncomposite elements includes Tn7 and closely related elements (Craig 1989). In contrast to Tn3, Tn7 translocates via a nonreplicative mechanism (Bainton et al. 1991).

3. Bacteriophage Mu. A third group of elements comprises the bacteriophage Mu and related elements, which utilizes transposition during the lytic and lysogenic phases of its life cycle (for reviews, see Mizuuchi and Craigie 1986; Pato 1989). Mu transposes replicatively during the lytic phase, repeatedly forming cointegrates as a means to replicate its genome. Alternatively, during lysogeny Mu inserts into the bacterial chromosome by nonreplicative transposition. For both transposition events, Mu encodes two transposition proteins: MuA, the transposase protein; and MuB, which plays an important role in target DNA interactions.

Studies of prokaryotic elements from these three classes has provided a mechanistic understanding of transposition, which has been helped to a large extent by the development of cell-free transposition systems for a few of these elements, e.g. the bacteriophage Mu (Mizuuchi

1983), the composite transposon Tn10 (Morisato and Kleckner 1987), and the noncomposite transposon Tn7 (Bainton et al. 1991).

Eukaryotic transposable elements can be divided into two major classes based on their transposition mechanism. Class I elements require reverse transcription of an RNA intermediate (a DNA-RNA-DNA mechanism) in order to transpose, whereas class II elements, like prokaryotic elements, utilize a direct DNA to DNA transposition mechanism (Finnegan 1989a).

1. Class I elements. Class I transposable elements transpose through an RNA intermediate transcribed from the donor element. A DNA copy is produced by reverse transcription, which then integrates into a target site. The most extensively studied examples of these elements are retroviruses and the structurally related retrotransposons, such as the yeast Ty elements (reviewed in Boeke 1989) and the Drosophila copia-like elements (reviewed in Bingham and Zachar 1989). These elements contain at each end a long terminal repeat (LTR) that flanks an internal domain encoding long open reading frames similar in function and sequence between retroviruses and retrotransposons. Gene products from this internal region include reverse transcriptase (RT) and integrase (IN), which carries out the actual transposition process.

Another group of class I elements are more distantly related to retroviruses. These elements, which include the mammalian LINE elements (reviewed in Hutchison et al. 1989) and the Drosophila I factors (reviewed in Finnegan 1989b), lack long terminal repeats and contain an A-rich tract at the 3′ end of one strand. They do contain long open

reading frames, one of which encodes a putative reverse transcriptase. In one case, the mammalian LINE-1 has been shown to encode an active reverse transcriptase (Mathias et al. 1992) and a version of this element which is likely to be functional for transposition has also been isolated (Dombroski et al. 1992).

2. Class II elements. Like bacterial transposable elements, class II elements transpose directly through a DNA to DNA mechanism. Examples of these elements include P elements in Drosophila (reviewed in Engels 1989), Ac and En/Spm elements in maize (reviewed in Federoff 1989), and Tam elements in Antirrhinum majus (reviewed in Coen et al. 1989). These elements contain short inverted repeats at their termini and encode one or more genes required for transposition, since elements containing internal deletions in these genes mobilize only when functional elements are present elsewhere in the genome.

It is important to distinguish transposable elements from another class of mobile elements that undergoes conservative site-specific recombination (CSSR), a DNA rearrangement that can also mediate the translocation of a discrete DNA segment (reviewed in Craig 1988). CSSR essentially involves a reciprocal exchange between two specific DNA sites that must share a short region of homology. In contrast, transposition does not require any homology between the target site and the transposable element. Like transposition, CSSR requires specific DNA sequences and specialized proteins to direct the precise breakage-reunion of DNA strands. However, these two recombination reactions differ in the reciprocal nature of DNA exchange and the role of DNA

synthesis. CSSR is a reciprocal event whose breakage and joining reaction requires no DNA synthesis, whereas transposition does not involve any reciprocal exchange between participating DNAs and requires at least some DNA synthesis at the newly formed element/target DNA junctions. Unlike the products of CSSR, the element/target DNA junctions formed during transposition contain single-strand gaps resulting from a staggered break in the target DNA. These gaps are subsequently repaired, producing a characteristic duplication of the target sequence, with a copy of the duplicated sequence flanking each end of the element (Grindley and Sherratt 1978; Kleckner 1981). In some cases, DNA synthesis continues so that the entire element is replicated.

Target site specificity in transposition

Although transposable elements generally insert into many different target sites, closer examination has revealed that these elements display at least some degree of insertion bias. In one study, the distribution of several IS elements (IS1, IS2, IS3, IS4, IS5, IS30, and IS150) in the E. coli chromosome was compiled, and it was found that these elements, taken together, had inserted preferentially into particular regions of the E. coli chromosome, whereas a striking paucity of insertions was observed in another chromosomal region (Birkenbihl and Vielmetter 1989). For individual transposable elements the insertion specificity varies widely. In a few cases the insertion specificity is extremely restricted, such as the bacterial element IS4, which has been observed to insert into only four unique sites in the E. coli chromosome (Habermann et al. 1979; Klaer and Starlinger 1980; Klaer et al. 1981,

Mayaux et al. 1984). In contrast, bacteriophage Mu can insert into many different sites in the E. coli chromosome during its lytic phase, generating approximately 100 copies of the phage genome (Pato and Waggoner 1987). Nonetheless, Mu can preferentially insert into particular sites within a defined target DNA, and closer inspection has indicated a limited similarity in sequence at the insertion site (Kamp and Kahmann 1980; Mizuuchi and Craigie 1986; Castilho and Casadaban 1991).

Eukaryotic transposable elements also display different levels of insertion specificity. For example, the Drosophila P elements preferentially insert into particular genomic loci (Simmons et al. 1984; Berg and Spradling 1991). These elements disrupt the singed locus at an unusually high frequency, whereas other loci, such as alcohol dehydrogenase, are poor targets for insertion (Engels 1989). Within the singed gene, nearly 50% of the P element insertions analyzed were clustered within a 100 bp region; moreover, repeated insertions were found at a single site (Roiha et al. 1988). At a genomic level, the Rous sarcoma virus (RSV) can display a remarkable insertion specificity. Analysis of genomic insertions from turkey embryo fibroblasts revealed a small number of sites that were utilized at a 10^6-fold higher frequency than expected for random insertion, and independent insertions occurred repeatedly at the same nucleotide position (Shih et al. 1988). Thus, RSV shows a high preference for certain genomic sites even in the absence of any target gene selection.

In general, two patterns of insertion specificity have been characterized for transposable elements. In some cases, repeated

insertions occur into particular target site, which is often attributed to the presence of a consensus nucleotide sequence at or near the target DNA duplicated upon insertion. Such preferred target sequences for insertion apparently reflect the recognition of the target site by the transposition machinery. In other cases, transposable elements display regional specificity, where insertions are clustered within a small region of the target DNA. This insertion specificity can result from a general feature of the target DNA region, such as DNA structure, gene activity, and DNA binding by host proteins.

A local consensus target sequence as a determinant for insertion specificity

For several prokaryotic and eukaryotic elements, the target DNA information responsible for insertion site specificity occurs locally at the point of insertion. This specificity appears to reflect the preferred recognition and cleavage of particular target sites by the transposition machinery.

A notable example of this type of insertion specificity is the bacterial transposon Tn10. Although Tn10 inserts into many different sites in the bacterial genome, Tn10 preferentially inserts into certain target sites, or hotspots (Kleckner et al. 1979). Sequence analysis of these hotspots has revealed a consensus sequence within the 9 bp duplicated target DNA, where the strand breakage and rejoining events occur. The 6 bp consensus sequence 5'-NGCTNAGCN'-3' is symmetrically located within the duplicated target DNA (Halling and Kleckner 1982). In addition, the palindromic structure suggests that this site may be

recognized and cleaved by symmetrically positioned subunits of a single protein, akin to Type II restriction enzymes, which bind and cleave a short symmetrical sequence (Rosenberg et al. 1987). The Tn10 transposase plays an important role in target site specificity, since specific transposase mutations can alter the target site specificity, which may indicate a perturbation in the transposase-target DNA interaction (Bender and Kleckner 1992).

Analysis of many insertion sites demonstrated a correlation between the preference for a given site and its similarity to the hot spot consensus sequence (Halling and Kleckner 1982). In addition, one insertion hotspot, HisD4, has been directly examined in vivo to explore the importance of the target sequence for insertion specificity (Lee et al., 1987). The HisD4 target sequence 5'-NGCCAGGCN-3' deviates from the consensus hotspot (5'-NGCTNAGCN-3')at the third positions of its half-sites, where non-consensus cytosines are present. However, these cytosines are methylated by the DNA cytosine methylase (dcm), which recognizes the 5'-CCAGG-3' sequence embedded within the HisD4 site. The resulting 5-methylcytosines restores the 5-methyl group normally found at these positions in the consensus thymines and contributes to its use as a hot spot. When the HisD4 site was examined in dcm$^-$ host, the efficiency of insertion into this site was dramatically reduced. Thus, the 5-methyl groups present in the thymines or 5-methylcytosines of the consensus sequence are important for Tn10 insertion specificity.

Although the target consensus sequence is a critical determinant for Tn10 insertion, other features of the target DNA can also contribute to the insertion specificity (Halling and Kleckner 1982). For example,

context effects from sequences flanking the target duplication can alter the insertion efficiency, although the basis of these effects still remains to be determined (cited in Bender and Kleckner 1992).

Many transposable elements insert into target sites with a more relaxed sequence specificity. In a well-studied example, the bacterial element Tn5 inserts preferentially into target sites with G-C or C-G base pairs at each end of the target duplication, i.e. 5'-(G/C)NNNNNNN(G/C)-3' (Berg et al., 1983). The importance of the G/C pairs in target specificity was directly examined by mutating the terminal guanines for a preferred target site (5'-GCTTTAATG-3') in pBR322, called hotspot I, and then evaluating transposition to the altered target site (Lodge et al. 1988). Changing the G to A or T at either end of the duplicated target sequence caused a dramatic reduction in insertion frequency into the target site. Thus, the G/C pairs can contribute to target site usage, which may reflect the preferential recognition and/or cleavage at the target site by the transposition machinery. It should be noted that many other sets of G/C pairs spaced nine bp apart in pBR322 are not used as Tn5 insertion hotspots; hence, other features of the target DNA also contribute to target site selection (see below).

Structural features of the target DNA that influence regional insertion specificity

Many transposable elements can insert into many different sites within a target DNA region; however, detailed mapping has revealed clusters of insertions within preferred stretches of DNA. This regional

specificity may reflect some structural feature of the target DNA. For instance, insertion of the bacterial element IS1 into the lacZ region of E. coli did reveal a limited sequence specificity for insertion adjacent to G/C base pairs (Galas et al. 1980). However, IS1 showed a more remarkable preference for particular regions in lacZ which are relatively rich in A-T base pairs (Miller et al. 1980). Moreover, addition of a poly (dA-dT) segment into a plasmid creates a synthetic hotspot for insertion, resulting in a cluster of IS1 insertions in the neighborhood of the A+T rich segment (Zerbib et al. 1985). The tendency for insertion into A+T rich regions may indicate a preference for easily denaturable sequences that facilitate IS1 insertion.

The presence of special DNA sites may also contribute to regional specificity. Analysis of IS1 transposition into pBR322 revealed a striking insertion hotspot in a 200 bp region that contains multiple binding sites for the E. coli protein IHF (Gamas et al. 1987). Approximately 80% of all insertions into pBR322 occur in this region, and the plasmid contains IHF binding sites only in this region. IHF may thus influence the distribution of IS1 insertions in a target molecule. One model for IHF function in target site specificity is that IHF may alter DNA conformation so as to favor insertion. Consistent with this view, IHF induces DNA bending both in the pBR322 insertion hotspot and the IS1 ends (Prentki et al. 1987). Alternatively, since IHF also binds to the IS1 ends, IHF may stabilize the nucleoprotein complex formed with IS1 and the target site through protein-DNA and protein-protein interactions (Gamas et al. 1987).

Despite the correlation between IHF binding sites and insertion specificity, IHF$^-$ strains still display wild type transposition frequency, indicating that IHF is not essential for IS1 transposition, although it may play an accessory role (Gamas et al. 1985). Comparison of the distribution of insertions in pBR322 using IHF$^+$ and IHF$^-$ backgrounds will be needed to support the proposed role of IHF in determining regional specificity.

Tn5 insertion is also influenced by general features of the target DNA. Examination of Tn5 insertions into pBR322 revealed a cluster of insertion hotspots within the first 300 bp of the tet gene (Berg et al. 1983). In particular, the major insertion hotspot, called hotspot I, is dramatically affected by DNA topology (Lodge and Berg 1990). Notably, hotspot I lies within a region of high negative supercoiling, which results from the specific transcription and translation of the tet gene (Lodge et al. 1989). Mutations in tet which reduced negative supercoiling in pBR322 also blocked insertion into hotspot I (Lodge and Berg 1990). This result is consistent with earlier studies which have demonstrated that Tn5 insertion into bacteriophage λ DNA is reduced in cells blocked for the production of supercoiled DNA, either by examining mutant cells deficient in DNA gyrase or wild-type cells treated with the gyrase inhibitor coumermycin (Isberg and Syvanen 1982). It should be noted that another pBR322 insertion hot spot (hotspot II), which is also located in the tet gene near hot spot I, is not affected by high negative supercoiling (Lodge and Berg 1990). Thus, target site specificity is a complex process, where the determinants that contribute to target site usage can differ from site to site.

Preferential transposition to nearby target sites: Maize Ac

The maize transposable element Activator (Ac) also displays regional specificity, preferentially inserting into target sites genetically linked to the donor site (Greenblatt and Brink 1962). Examination of transposition events from the unstable P-vv allele, which contains the donor Ac element, revealed that Ac often transposes from a replicated donor site into non-replicated target DNA near the donor site (Greenblatt and Brink 1962; Greenblatt 1984). This observation may indicate that replication of the Ac element activates the translocation of the element (Chen et al. 1992).

Notably, this Ac translocation generates dissimilar daughter chromosomes. Although both daughter chromosomes contain the newly transposed Ac insertion, one chromosome contains an empty P-vv donor site resulting from Ac excision, whereas the other retains the Ac element at the donor site. The two mitotic daughter lineages from this event produce detectable twin sectors, whose DNAs have been analyzed to confirm the disposition of the Ac elements (Chen et al. 1987, 1992). Insertion by this means allows Ac to proliferate in the host genome by a non-replicative mode of transposition.

Introduction of Ac into a heterologous system, such as tobacco, showed that Ac also preferentially transposes to genetically linked sites (Jones et al. 1990). This tendency to insert into nearby sites reflects the cis-acting preference of the Ac transposition machinery and may also result from the chromatin structure that limits the availability of potential insertion sites. Like Ac, Spm elements in

maize (Nowick and Peterson 1981) and P elements in Drosophila (Raymond and Simmons 1981; Hawley et al. 1988) also preferentially insert into nearby target sites.

DNA accessibility: retrotransposons and retroviruses

The chromatin structure of the target DNA, which affects DNA accessibility, can also influence regional specificity. Detailed mapping of insertions by the yeast retrotransposon Ty1 into particular genomic loci (CAN1, URA3, and LYS2) revealed a preference for the 5′ region of targeted genes, as compared to the rest of the coding sequence (Natsoulis et al. 1989; Wilke et al. 1989). In some cases, repeated insertions occurred at the same nucleotide position, although alignment of these preferred target sites revealed very little sequence similarity. Insertion into the 5′ region of these genes may indicate preference for the more accessible chromatin structure in this region, as suggested by the presence of nuclease hypersensitive sites generally found in the 5′ region of genes (reviewed in Elgin, 1988). The uncondensed chromatin structure would then allow the target DNA to interact with the integration machinery. Interestingly, avian and murine retroviruses apparently display a similar target specificity, as both of these elements insert into regions associated with DNase I hypersensitive sites and transcriptionally active regions (Vijaya et al. 1986; Rohdewohld et al. 1987; Scherdin et al. 1990; Mooslehner et al. 1990).

Insertion preference for accessible DNA regions can operate in a more localized manner. Analyses of retroviral integration in vitro into

naked DNA targets did not reveal any obvious target site specificity (Brown et al. 1987). However, in vitro integration into minichromosome target plasmids produced insertions in both nucleosome-free and nucleosome-containing regions; intriguingly, the insertion sites tended to be spaced approximately 10 bp from each other in a periodic fashion (Pryciak et al. 1992). By contrast, insertions into naked DNA did not follow this pattern. A reasonable model is that the DNA wrapping around a nucleosome core produces a periodic spacing of exposed positions that are preferentially accessible to the integration machinery.

Positional specificity: Yeast retrotransposon Ty3

The yeast retrotransposon Ty3 displays a remarkable insertion specificity in that it integrates preferentially near tRNA genes distributed throughout the genome (Chalker and Sandmeyer 1990; reviewed in Sandmeyer et al. 1990). Notably, these insertions were located exclusively within several nucleotides upstream of the transcriptional start site. Detailed examination of Ty3 integration into a plasmid target carrying the SUP2 tRNA gene also showed this positional specificity (Chalker and Sandmeyer 1992). An attractive view for this target site preference is that the Ty3 integration machinery recognizes the RNA polymerase III transcription complex assembled at tRNA genes. Consistent with this view, a mutation within the SUP2 gene that disrupts the assembly of the transcription apparatus also dramatically reduces Ty3 target activity. Moreover, no specific nucleotide sequence at the transcription initiation region is required for Ty3 insertion, as Ty3 still inserts near the SUP2 transcription start site even when replaced

with heterologous sequences. In addition, Ty3 integration displays a similar positional specificity into target plasmids containing the 5S or U6 RNA genes, which are also transcribed by RNA pol III (Chalker and Sandmeyer 1992). Further studies may determine whether the Ty3 integration machinery directly interacts with the RNA pol III transcription apparatus.

Mechanism of target recognition and utilization during transposition

Several manipulations of the target site occur during transposition. After a target site is identified by the transposition machinery, the target sequence is cleaved at staggered positions and concurrently joined to the incoming transposable element. The resulting strand transfer product then undergoes at least some DNA synthesis to repair the single-stranded gaps formed at the element/target junctions; in some cases, replication proceeds through the entire element. As discussed below, in vitro characterization of several transposable elements has elucidated many of the mechanistic details underlying these target site interactions.

General classes of transposition mechanisms

Two general transposition mechanisms, replicative and non-replicative, have been characterized among the transposable elements studied (for reviews, see Grindley and Reed 1985; Derbyshire and Grindley 1986; Craig and Kleckner 1987).

In replicative transposition, the entire element is replicated as an essential part of the transposition mechanism. Insertion of an

element from a donor molecule into a target molecule generates a cointegrate: this structure consists of a fusion between the donor and target DNA molecules, where a copy of the element as a direct repeat is located at each of the two donor/target junctions (Shapiro 1979; Arthur and Sherratt 1979). In some instances, the cointegrate is then resolved by recombination between the two element copies to form 1) a molecule identical to the original donor molecule and 2) a target molecule that contains a copy of the element with a short target duplication at each side of the element. A few elements (e.g. Tn3) contain a specialized protein, a resolvase, that performs this resolution event at specific DNA sites (Arthur and Sherratt 1979; Grindley 1983); in other cases, the host recombination system resolves the cointegrate (Grindley and Sherratt 1978; Shapiro 1979).

In non-replicative transposition, the element is excised from the donor DNA and then inserted into the target DNA without any extensive replication. Some DNA synthesis does occur at the newly formed element/target junctions to repair the single-strand gaps resulting from the staggered break in the target DNA. In one example, the bacterial transposon Tn10 translocates via a non-replicative "cut-and-paste" process in which the transposon is released from the donor DNA by double-strand cuts at each end and inserted into the target site (Morisato and Kleckner 1984, 1987; Benjamin and Kleckner 1989). Genetic studies using a heteroduplex Tn10 element also showed that any DNA synthesis at the transposon-target junctions did not extend significantly into the transposon (less than 70 bp) (Bender and Kleckner 1986).

The broken donor DNA, resulting from excision of the element, can now undergo several fates. As demonstrated in Tn10 (Bender et al. 1991) and Drosophila P element (Engels et al. 1990) transposition, the broken donor molecule may be restored to its original element-containing state by gap repair with a sister chromosome. Particularly with plant transposable elements, the broken donor DNA can also be rejoined, albeit imprecisely, presumably by host-encoded functions (Sadler and Nevers 1985); this process rarely if ever occurs within transposable elements in bacteria (Bender et al. 1991). In other cases, the broken donor molecule is rapidly degraded and lost (Berg 1977; Weinert et al. 1984; Morisato and Kleckner 1984).

Some transposable elements employ either a replicative or non-replicative mechanism, whereas other elements, notably bacteriophage Mu, use both. During Mu transposition, a recombination intermediate, called the strand transfer intermediate, is generated which can proceed through either mode of transposition (Shapiro 1979; Ohtsubo et al. 1981; Craigie and Mizuuchi 1985; Figure 1A). In this intermediate the donor and target molecules are joined: the two 3′ ends of the element are ligated to the 5′ ends of the target site, whereas the other strand at each end of the element remains uncleaved. Formation of the intermolecular strand transfer intermediate is dependent on the Mu transposition proteins MuA and MuB.

After the Mu strand transfer intermediate is generated, it can then be resolved into a cointegrate structure, by DNA replication through the Mu element from the free 3′ ends of the target DNA, or into either a simple insertion product, by cleavage of the donor DNA flanking

the element and DNA repair of the donor/target junctions (reviewed in Mizuuchi and Craigie 1986; Figure 1A). Host proteins are sufficient to resolve the Mu transposition intermediate, since incubation of the deproteinized DNA intermediate with an E. coli extract, lacking any Mu proteins, can generate both types of transposition products (Craigie and Mizuuchi 1985). It is tempting to speculate, however, that the Mu-encoded transposition proteins, Mu A and Mu B, may regulate whether the transposition intermediate is resolved replicatively, which occurs during lytic growth, or non-replicatively, which occurs during lysogeny. Consistent with this view, a truncated Mu B protein, containing a short carboxy terminal deletion, can no longer support Mu transposition replicatively in vivo, but can still promote Mu lysogeny via non-replicative simple insertion (Chaconas et al 1985b). An attractive hypothesis is that the carboxy terminus of Mu B recruits the host replication machinery to the transposition intermediate in order to promote replicative transposition.

Proteins that participate in transposition

Many transposable elements encode only one protein which is involved in transposition. Tn10, for example, encodes a single transposase protein which mediates its translocation. In vitro studies have also demonstrated that Tn10 transposition also requires either of two E. coli histone-like proteins, HU or IHF, which may contribute to the DNA conformational changes that facilitate transposition (Morisato and Kleckner 1987). Biochemical studies have demonstrated that the transposase mediates Tn10 excision from the flanking donor DNA by

precise double strand cleavages at the Tn10 ends (Morisato and Kleckner 1984; Haniford et al. 1989, 1991; Figure 1B). Moreover, the transposase also directs the strand transfer between the Tn10 ends and the target site, which involves the ligation of the 3′ strand at each Tn10 end to the 5′ ends of the target DNA (Benjamin and Kleckner 1989; Figure 1B). Genetic analysis has also provided evidence for an interaction by transposase with the target DNA. Several transposase mutants have been isolated which generate high levels of an excised transposon fragment (ETF), resulting from double strand cuts at the transposon ends (Haniford et al. 1989). Although the ETF is normally an intermediate in Tn10 strand transfer (Haniford et al. 1991), the transposase mutants cannot promote strand transfer between the Tn10 ends and the target site. These properties suggest a defect in the transposase in some aspect of target DNA interaction whose overall consequence is a block to target site insertion.

Retroviral elements also encode a single protein, the IN protein, which is required for integration. The IN protein from Moloney murine leukemia virus (MoMLV) or human immunodificiency virus (HIV) possesses a site-specific endonuclease activity that removes a dinucleotide from the 3′ ends of the retroviral DNA produced by reverse transcriptase (Craigie et al. 1990; Bushman and Craigie 1991; Figure 1C). The IN protein from MoMLV and HIV also promotes the joining of the recessed 3′ ends to the target DNA (Figure 1C). Although the IN protein alone can generate the concerted integration of a pair of retroviral end DNA substrates, single-end integrations, where only one retroviral end has inserted into the target DNA, are another significant reaction product. The relatively

high number of these non-concerted integration events found *in vitro* may reflect the use of purified IN protein and synthetic end substrates, instead of the elaborate nucleoprotein complex where the retroviral DNA and several proteins, including IN, are located *in vivo* (Bowerman et al. 1989). The nucleoprotein complex may play a role in ensuring the coordinated integration of both ends of the retroviral element into the target DNA. In addition, the organization of the viral DNA within the complex may inhibit intramolecular "autointegration" events that would disrupt the integrity of the element (Craigie et al. 1990).

Several transposable elements encode more than one protein to mediate their translocation. Bacteriophage Mu contains two transposition proteins, MuA and MuB, that play important roles in transposition. Two host DNA binding proteins, HU (Craigie et al. 1985) and IHF (Surette and Chaconas 1989), also assist in Mu transposition, but operate at an early step, probably by influencing the conformation of the Mu donor DNA. The MuA protein recognizes specific DNA sites at the Mu ends (Craigie et al. 1984) and a separate site, called the internal activation sequence, located within Mu (Surette et al. 1989; Mizuuchi and Mizuuchi 1989; Surette and Chaconas 1989). Importantly, MuA also functions as the transposase, mediating the cleavage and joining events between the target DNA and the Mu ends (Maxwell et al. 1987). MuA therefore performs the essential target site manipulations required for strand transfer. The MuB protein also interacts directly with target DNA and affects target site selection. MuB directs MuA to promote insertions into intermolecular target sites; in the absence of MuB, transposition *in vitro* by MuA occurs predominantly into intramolecular sites, but at a

significantly reduced overall efficiency (Mizuuchi 1983; Maxwell et al. 1987). Biochemical studies have shown that MuB binds nonspecifically to DNA in the presence of ATP (Maxwell et al. 1987), and MuB also forms oligomers on DNA which may be the active form of MuB that attracts MuA to promote strand transfer (Adzuma and Mizuuchi 1991).

Mu inserts into a preferred target consensus sequence: 5'-N(Pyr)(G/C)(Pur)N-3' (Kamp and Kahmann 1981; cited in Mizuuchi and Craigie 1986). Although MuB recruits MuA to the target DNA, transposition promoted by MuA and HU, in the absence of MuB, still displays the same sequence specificity at the target site (Mizuuchi and Craigie 1986). Since HU operates at an early step in Mu transposition involving only the Mu donor DNA (Lavoie and Chaconas 1990; see below), MuA is likely responsible for this insertion specificity. Thus, MuB directs Mu to particular target DNAs, whereas MuA plays a critical role in selecting preferred sequences for insertion within the target DNA.

In addition to affecting target site choice, MuB can also modulate the activity of MuA by stimulating cleavage of the Mu ends or intramolecular strand transfer (Surette et al. 1991; Baker et al. 1991; Surette and Chaconas 1991). MuA activation by MuB appears to be distinct from the DNA binding properties of MuB, since chemical modification of MuB disrupts its DNA binding activity, but does not affect its ability to stimulate MuA. The MuB enhancement of MuA activity also implicates direct MuB-MuA interactions. In agreement with this view, a carboxy terminal deletion of MuA still promotes strand transfer, but fails to undergo activation by MuB, implying that the MuA deletion mutant cannot interact with MuB (Baker et al. 1991). In addition, another closely

related MuA deletion mutant can still direct insertions into intramolecular sites, but is defective in utilizing intermolecular target sites identified by MuB (Leung and Harshey 1991). In vivo, these MuA deletion mutants promote transposition at a reduced frequency in the presence or absence of MuB, consistent with the inability to interact with MuB (Harshey and Cuneo 1986).

Communication between the target site and the ends of the element

During transposition three separate DNA segments, the two ends of the element and the target site, must be brought together in order for strand transfer between the ends and the target site to proceed accurately. An emerging view is that specific protein-DNA complexes containing the element ends and the target site mediate the transactions between these DNA sites. The involvement of protein-DNA complexes suggests that an elaborate set of protein-DNA and protein-protein interactions operate during transposition.

For Mu transposition two distinct protein-DNA complexes have been identified: the type I, or cleaved donor, complex; and the type II, or strand transfer, complex (Surette et al. 1987; Craigie and Mizuuchi 1987; Figure 2). The cleaved donor complex consists of a Mu donor molecule in which the two protein-bound Mu ends are held stably together and contains a single-stranded nick at each 3′ end, separating Mu from the flanking donor sequences. The bound Mu ends partition the donor molecule into two topological domains: the Mu DNA remains negatively supercoiled while the flanking donor DNA is relaxed. To complete the strand transfer reaction, the cleaved donor complex is converted to the

strand transfer complex, consisting of the strand transfer product in which the two free 3′ ends of Mu are ligated to the 5′ ends of the cleaved target DNA (Figure 2). In this complex the resultant transposon/target junctions are also non-covalently bound together by protein, with the Mu DNA supercoiled and the flanking donor and target DNAs relaxed.

Mu DNA strand transfer in vitro requires the phage-encoded proteins, MuA and MuB, and the host protein HU (Craigie et al. 1985). Formation of the cleaved donor complex requires MuA and HU. MuA specifically binds to the Mu ends and maintains the stable synapsis of the ends in the complex (Mizuuchi et al. 1991), whereas HU facilitates the assembly of this higher-order complex, perhaps by its ability to wrap or coil DNA (Lavoie and Chaconas 1990). Another host protein, IHF, although not absolutely required for strand transfer, can also assist in the formation of the cleaved donor complex through its interaction with an enhancer-like sequence, called the internal activation sequence (IAS), which lies about 1 kb from the Mu left end (Surette et al. 1989; Mizuuchi and Mizuuchi 1989; Surette and Chaconas 1989). IHF relieves the high negative supercoiling requirement for IAS stimulation of cleaved donor complex formation. Conversion of the cleaved donor complex into a strand transfer complex requires MuB, which binds to target DNA and promotes intermolecular target capture by the cleaved donor complex (Surette et al. 1987; Craigie and Mizuuchi 1987). The strand transfer complex contains MuA tightly bound noncovalently to the target/element junctions, whereas HU and MuB are more loosely associated with the complex (Lavoie and Chaconas 1990). HU is not required for this step, as

the cleaved donor complex containing MuA, but selectively depleted of HU by high salt, can efficiently generate strand transfer product (Lavoie and Chaconas 1990). In addition, a precut Mu element containing two right ends can complete strand transfer in the absence of HU (Craigie and Mizuuchi 1987).

Interestingly, Tn10 transposition also proceeds through two stable protein-DNA complexes which are transposition intermediates analogous to the cleaved donor and strand transfer complexes found in Mu transposition (Haniford et al. 1991; Figure 2). One Tn10 complex consists of an excised transposon fragment (ETF), in which Tn10 has undergone transposase-mediated double-strand breaks at both ends to excise the transposon precisely from the flanking donor DNA. The ETF protein-DNA complex closely resembles the Mu cleaved donor complex in that the transposon ends are held tightly together by the non-covalently bound transposase (Figure 2). Similarly, the Tn10 sequences in this complex remain supercoiled; however, the flanking donor DNA is no longer present, released from the Tn10 sequences by double-strand breaks at the transposon/donor DNA junctions. The ETF is a precursor to the strand transfer intermediate, which exists predominantly as a transposon circle generated by Tn10's utilization of a target site within the transposon. This intramolecular, transposase-dependent rearrangement is a deletion/inversion, since the flanking donor DNA outside the Tn10 ends is deleted, and the transposon sequences between one Tn10 end and the target site are inverted (Benjamin and Kleckner 1989; Figure 2). The protein-DNA complex containing the Tn10 deletion/inversion circle is analogous to the Mu strand transfer complex. In both cases, the

transposon ends have been joined to the cleaved target site, and a stable complex results from protein tightly bound at the target/transposon junctions. As found with the target molecule in the Mu strand transfer complex, the Tn10 circle is relaxed, since the transposon itself serves as an intramolecular target.

Target site cleavage and DNA strand transfer

Insertion of a transposable element into a target site generally produces a short duplication of the target sequence, and each element generates a target site duplication of a specific length (Grindley and Sherratt 1978; Kleckner 1981). This target DNA repeat, several base pairs long, results from a staggered break on each strand of the target sequence mediated by one or more element-encoded proteins. Cleavage of the target DNA is coupled to the joining of the transposable element to the target site, and the single-strand gaps at the element/target junctions are subsequently restored by host-dependent DNA synthesis, resulting in the characteristic target duplication flanking each end of the element.

Analysis of Tn10 transposition, which generates a 9 bp duplication upon transposon insertion, has provided physical evidence for the staggered cleavage at the target site. Intramolecular strand transfer products were generated in vivo with a Tn10 substrate containing an internal HisG1 hotspot (Bender and Kleckner 1989). Examination of the strand transfer products revealed that the transposon-HisG1 target site junctions contained a 9 nt. single strand gap extending from the 5′ transposon terminus to the 3′ cleaved target terminus. On the other

strand opposite the gap, the 5′ end of the target strand is joined to the 3′ transposon end. The structure of the transposition intermediate is consistent with the occurrence of 9 bp staggered nicks in the target site.

The development of highly purified systems for transposition *in vitro* has also demonstrated that element-encoded proteins mediate the target site cleavage and joining to the ends of the element. For instance, integration intermediates were generated by purified MoMLV IN protein and single-strand gaps at the viral end/target junctions were repaired by passage of the plasmid containing the junctions through *E. coli* (Craigie et al. 1990). Analysis of the viral end/target junctions indicated that IN protein presumably generates the staggered cleavage of the target DNA that results in the 4 bp duplication observed for integration events *in vivo*.

Intriguingly, an emerging view is that breakage of the target site is coupled to strand transfer between the ends of the element and the target site. Support for this model has come from stereochemical studies of DNA strand transfer. During bacteriophage Mu transposition, DNA strand transfer proceeds chemically through the joining of the 3′OH ends of the Mu element to the 5′ phosphoryl ends of the target DNA (Mizuuchi 1984). No high energy cofactor is required for the strand transfer reaction, indicating that the phosphodiester bond energy for target DNA cleavage is probably conserved throughout this step (Maxwell et al. 1987). Strand transfer reactions with target DNAs containing chiral phosphothiorates were performed to examine the stereochemistry of the phosphate group at the target site junction that was joined to the Mu

end (Mizuuchi and Adzuma 1991). During the course of strand transfer, the chirality of the phosphothiorate linkage between the donor and target DNA was inverted. This result suggests a one-step transesterification mechanism, where the exposed 3'OH of the cleaved Mu end acts as a nucleophile that attacks the phosphodiester linkage in the target DNA. In this view, the nucleophilic attack simultaneously breaks the target DNA strand and joins the 3' end of the element ot the 5' end of the target DNA. Similar studies of HIV retroviral strand transfer in vitro have demonstrated that this reaction also proceeds through a one step transesterification mechanism (Engelman et al. 1991), which may indicate its applicability to a wide range of transposable elements.

The concerted cleavage-ligation reaction, where the 3'OH from the donor DNA is used directly as the attacking nucleophile, bypasses the requirement for a covalent protein-DNA intermediate in strand transfer. If a protein-DNA covalent linkage were involved, as argued for site-specific recombinases and DNA topoisomerases, then strand transfer would proceed via a two-step transesterification mechanism (reviewed in Craig 1988). In fact, stereochemical analysis of bacteriophage lambda integration indicated a two step mechanism, which is consistent with the view that lambda integration proceeds through a covalent Int-DNA intermediate (Mizuuchi and Adzuma 1991).

The coupling of target DNA cleavage to strand transfer is consistent with the observation that Mu and retroviruses do not require an exogenous energy source for this step, indicating that the phosphodiester bond energy from target DNA cleavage is likely conserved (Maxwell et al. 1987; Brown et al. 1987; Ellison et al. 1990). Indeed,

this cleavage-ligation step may generally prove to be reversible, because of its isoenergetic nature. In one study, HIV IN protein could accurately convert a synthetic integration intermediate into a restored target DNA segment and viral end (Chow et al. 1992). This reaction, referred to as "disintegration", likely proceeds through a one-step mechanism analogous to the forward cleavage-ligation step.

Mechanism of target immunity

For a few transposable elements, the presence of an element or even an element end sequence in a DNA molecule may render the potential target molecule immune, so that additional insertions of the same element into the target DNA are inhibited (Robinson et al 1978; Lee et al. 1983; Hauer and Shapiro 1984; Darzins et al. 1988). The mechanism of target immunity should provide insight into how the transposition apparatus mediates communication between the transposable element and the target DNA. This phenomenon has been extensively characterized in bacteriophage Mu, where immunity has been studied <u>in vitro</u> using purified components. During Mu strand transfer, when two target DNAs are provided in which one target molecule carries a Mu end sequence, the Mu element inserts into only the non-immune target DNA lacking the Mu end (Adzuma and Mizuuchi 1988, 1989).

How is such target discrimination by Mu achieved? Biochemical studies (Adzuma and Mizuuchi 1988, 1989) have demonstrated that DNA bound by MuB protein is preferentially used as a target by Mu. However, MuA protein causes MuB, when ATP is present, to dissociate from immune target DNA molecules which contain Mu end sequences. MuB selectively

dissociates from immune DNAs because of its proximity to MuA, bound specifically to the Mu end. Immunity also involves ATP hydrolysis by MuB, as Mu inserts into both immune and non-immune DNAs in the presence of the non-hydrolyzable analogue, ATP-γS. When ATP-γS is present, MuB is stably associated with both immune and non-immune target DNAs, which promotes Mu insertion into both target molecules.

Studies on the Tn3-like transposon γδ indicate that target immunity is dependent on γδ transposase binding to the target DNA (Wiater and Grindley, 1990a,b). A γδ end carrying a deletion at the terminus is deficient in promoting recombination when coupled with a normal γδ end; however, the mutant end still retains the ability to confer target immunity. The γδ mutant end can also bind transposase, consistent with the view that transposase binding to the end sequences is important for conferring immunity. Moreover, IHF stimulates the ability of a γδ end to confer immunity, as shown by assaying immunity in IHF-deficient strains or utilizing mutant ends which are deleted for the IHF binding site (Wiater and Grindley 1990a). Footprint studies of IHF and transposase interaction with the γδ ends have demonstrated that IHF and transposase bind cooperatively to adjacent sites (Wiater and Grindley 1988, 1991). Thus, the IHF stimulation of immunity may be explained by its ability to increase transposase binding to γδ end sequences on an immune target. The transposase bound to the immune target perhaps confers immunity by interacting with the incoming transposon complex and aborting the transposition reaction (Lee et al. 1983; Wiater and Grindley 1990b).

Bacterial Transposon Tn7

The bacterial transposon Tn7, which confers resistance to trimethoprim, spectinomycin, and streptomycin, is particularly interesting because of its unusual target site selectivity. Originally identified in E. coli as a determinant for trimethoprim resistance (Hedges et al. 1972), Tn7 was subsequently recognized for its ability to insert at unusually high frequency into a single chromosomal site (Barth et al. 1976). Tn7 insertions into a preferred chromosomal site have later been found in a wide range of bacteria (reviewed in Craig 1989). In addition to this site-specific mode of transposition, Tn7 also inserts at low frequency into many different target sites in plasmids (Barth and Grinter 1977; Barth et al. 1978) .

In E. coli, the specific Tn7 insertion site, called attTn7, is located at minute 84 between the two host genes glmS, which is involved in cell wall biosynthesis, and phoS, which participates in phosphate transport (Barth et al. 1976; Lichtenstein and Brenner 1981; Walker et al. 1984). As observed for other transposable elements, Tn7 produces a short target DNA duplication (5 bp) upon insertion into attTn7 (Lichtenstein and Brenner 1982; McKown et al. 1988). An interesting feature is that Tn7 transposes to attTn7 in a unique orientation, with the right end of Tn7 adjacent to glmS and the left end adjacent to phoS. Examination of Tn7 transposition to attTn7 has shown that plasmids carrying a chromosomal attTn7 DNA segment can confer the orientation- and site-specific insertion observed for Tn7 transposition (McKown et al. 1988). Thus, the target DNA information necessary for Tn7 insertion is contained within the attTn7 sequence. Interestingly, the sequence

information required for attTn7 target activity lies entirely to one side of the specific point of Tn7 insertion, and the target DNA sequences immediately surrounding the Tn7 insertion point can be changed without obviously affecting target activity (Gringauz et al. 1988). Sequence comparisons between the ends of Tn7 and the attTn7 target site also showed no significant DNA homology, another feature indicative of transposition (Gay et al. 1986; McKown et al. 1988).

Tn7 also displays target site selectivity by evaluating potential target DNA molecules for the presence of another Tn7 element; i.e. Tn7 exhibits target immunity, where the presence of Tn7 in a target molecule greatly inhibits a second insertion of the element (Hauer and Shapiro 1984; Arciszewska et al. 1989). Examination of various Tn7 end segments revealed that the right end of Tn7 is sufficient to confer target immunity, and the cis-acting Tn7 right end sequences required for transposition are also necessary for immunity (Arciszewska et al. 1989). Tn7 insertion into attTn7 also inactivates any subsequent insertion into the occupied attTn7 sequence, thus avoiding tandem element insertions (Arciszewska et al. 1989).

The relatively large size of Tn7 (14 kb) suggested that it carries a complex array of DNA sequence information required for transposition, and Tn7 derivatives containing internal deletions were first examined to define these determinants (Smith and Jones 1984; Hauer and Shapiro 1984; Ouartsi et al. 1985). These studies indicated that the functional ends of Tn7 extended into the element beyond the terminal inverted repeats. In a later study, the size and complexity of the Tn7 ends required for

transposition was established by analysis of miniTn7 elements containing various Tn7 end segments (Arciszewska et al. 1989).

The initial studies with Tn7 deletions also indicated that transposition required several Tn7-encoded genes (Hauer and Shapiro 1984). Subsequent studies, using complementation with various Tn7 DNA segments (Rogers et al. 1986) and using insertional mutations in Tn7 (Waddell and Craig 1988), determined that Tn7 encoded five transposition genes, designated tnsA, tnsB, tnsC, tnsD, and tnsE. Moreover, analysis of transposition to plasmids showed that tnsABC + tnsD direct high-frequency, site-specific insertion into attTn7, whereas tnsABC + tnsE promote low-frequency insertions into other plasmid target sites. Thus, Tn7 transposition occurs along two overlapping, but distinct pathways.

To understand the roles of the Tns proteins in recombination, a biochemical approach has been undertaken. Purification and characterization of individual Tn7 recombination proteins has revealed that TnsB directly binds to the ends of Tn7 (Arciszewska et al. 1991; Tang et al. 1991; Arciszewska and Craig 1991) and TnsC is an ATP-dependent, non-sequence-specific DNA binding protein (Gamas and Craig, in press). In addition, a cell-free system that reconstitutes Tn7 transposition to attTn7 has been developed using protein fractions from cells containing the individual tns genes (Bainton et al. 1991). Analysis of Tn7 transposition in vitro will prove invaluable in dissecting the mechanism of Tn7 transposition.

Objectives

The major objective of this thesis is to determine the mechanism of target site selection in Tn7 transposition. To this end, I have addressed the following questions:

1. What are the features of Tn7 insertion sites that influence target site selection? I examined Tn7 insertions into the *E. coli* chromosome and show that Tn7 directs insertion into two distinct classes of target sites that differ dramatically in nucleotide sequence. These studies are presented in Chapter 2, which was published as: K.M. Kubo and N.L. Craig, "Bacterial transposon Tn7 utilizes two different classes of target sites." *J. Bacteriol.* (1990) 172:2774-2778.

2. How does Tn7 direct site-specific insertion into attTn7? I purified the TnsD protein and show that TnsD directly recognizes attTn7. Moreover, TnsD participates directly in and is required for transposition to attTn7. These studies argue that TnsD recognition of attTn7 is an important determinant for Tn7 insertion. These results are described in Chapter 3.

3. How do the protein-target DNA interactions with attTn7 promote Tn7 insertion? These studies, which are a collaboration between Roland Bainton and myself, are presented in Chapter 4 with a preface that summarizes the details of our collaboration. We show that TnsD recruits another transposition protein, TnsC, to the attTn7 target site. We provide evidence that TnsC can associate with the target DNA and likely plays a role in target selection.

The concluding chapter summarizes our current understanding of target site selection in Tn7 transposition and its relevance to other specialized protein-DNA transactions.

Figure 1. Steps involved in transposition and retroviral integration.

Insertion of Mu (A), Tn10 (B), and the retrovirus MoMLV (C) into a target site are shown. The transposable element and target DNAs are shown by clear and hatched boxes, respectively. Solid lines indicate the flanking donor DNA. Arrowheads indicate cleavage sites on each DNA strand of the transposable element, and asterisks indicate positions of strand transfer in the target DNA. DNA synthesis after strand transfer is shown by a dashed line.

Figure 1

A. Mu

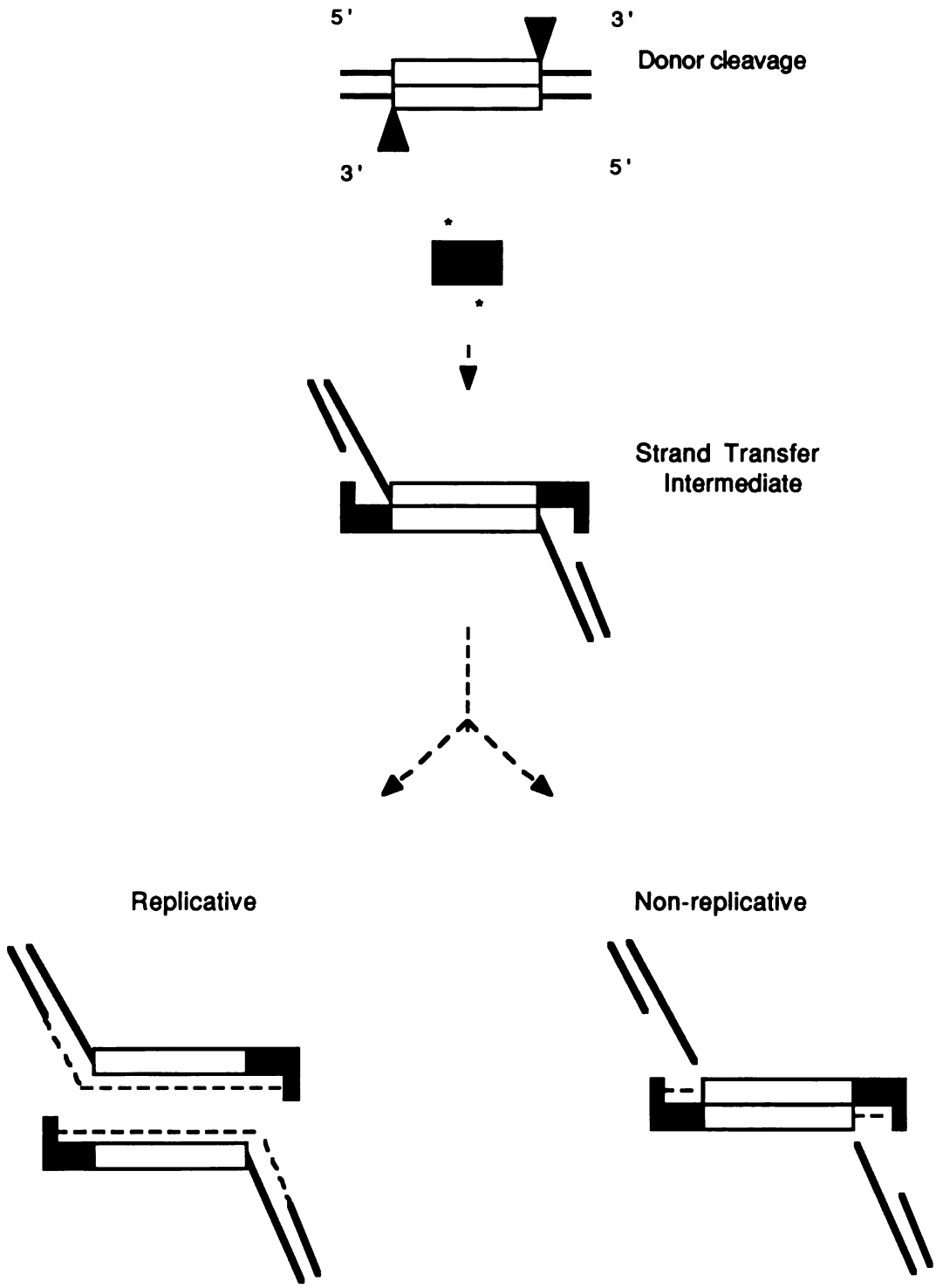

5'
3'
Donor cleavage
3'
5'
Strand Transfer
Intermediate
Replicative
Non-replicative

Figure 1

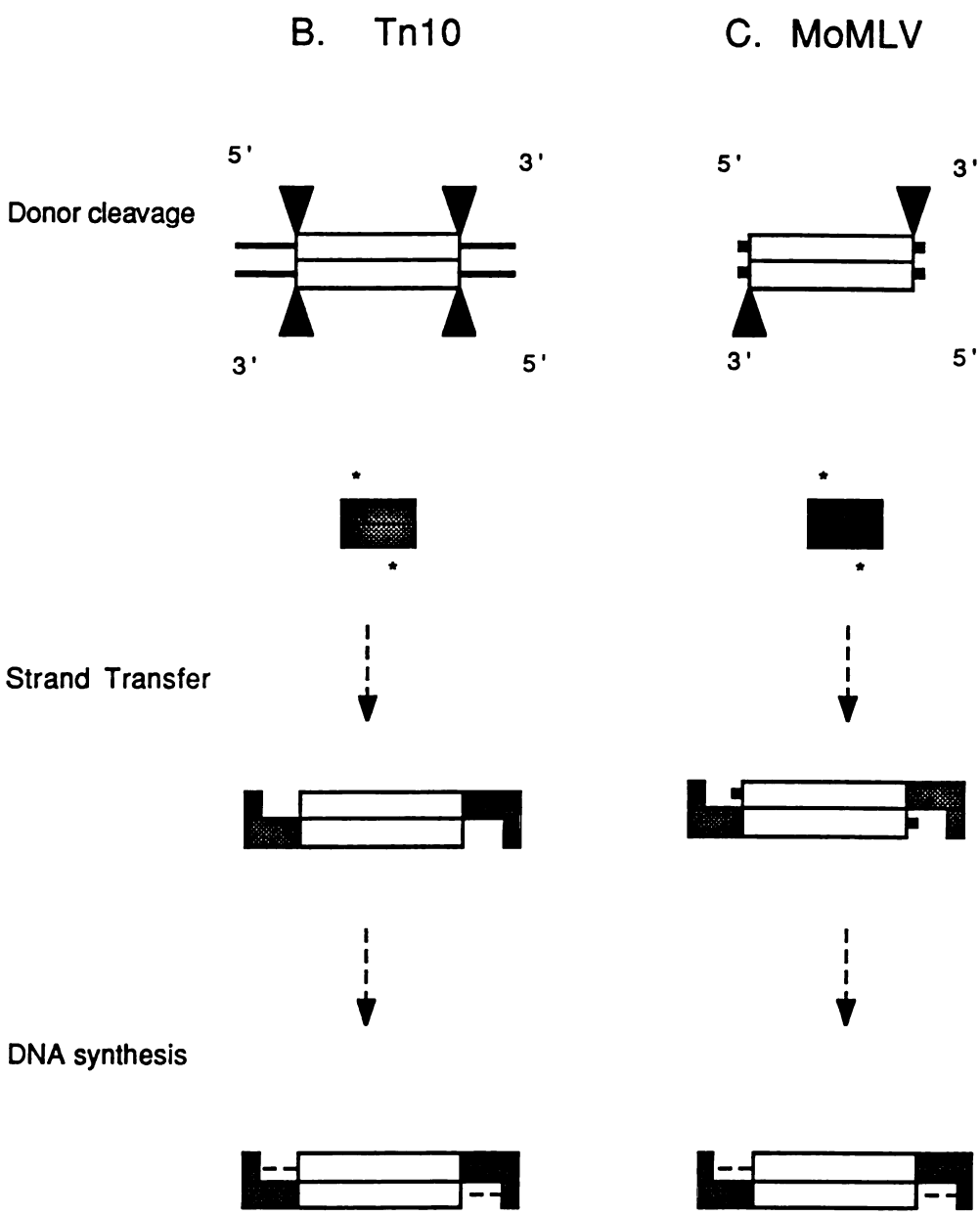

B. Tn10
C. MoMLV
Donor cleavage
5'
3'
3'
5'
5'
3'
3'
5'
Strand Transfer
DNA synthesis

Figure 2. Cleaved donor and strand transfer complexes formed during transposition.

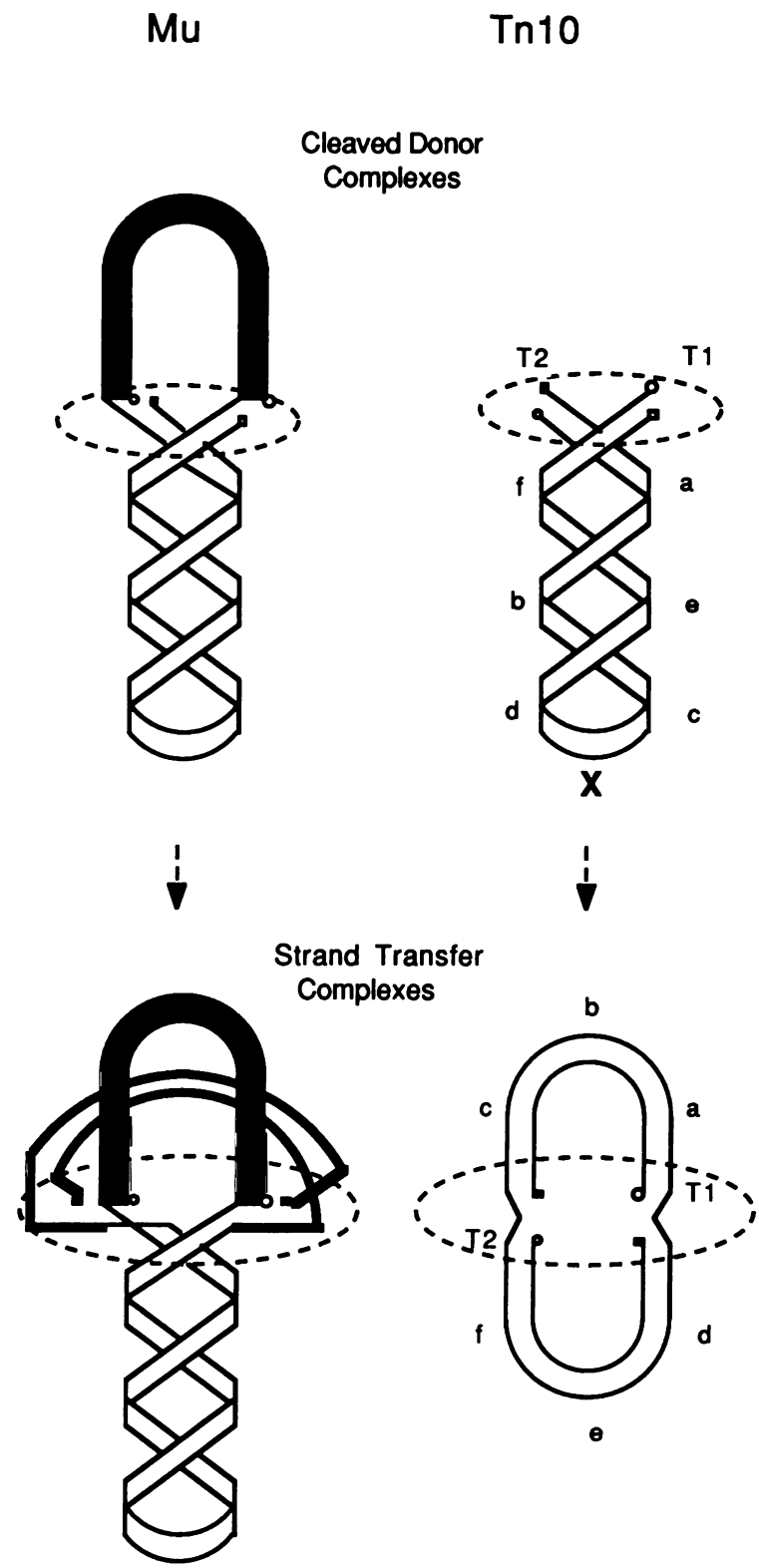
The transposon (Mu or Tn10) and flanking donor DNAs are clear and solid black, respectively. The synaptic structures formed at the transposon ends and the strand transfer junctions are indicated by the dashed ovals, and the free 5'PO4 (circles) and 3'OH (squares) groups at the ends of the DNA strands are also shown. Figure was adapted from Haniford et al. (1991).

Mu: The cleaved donor complex contains two single-strand breaks at the 3' ends of the transposon. Transposon sequences are constrained as negative supercoils by the synaptic structure; the donor DNA is relaxed. In the strand transfer complex, the 3'OH ends of Mu are joined to the 5'PO4 ends of the intermolecular target DNA (shown as boldfaced strands). The transposon sequences are constrained as negative supercoils by the synaptic structure formed at the Mu/target junctions, but the flanking donor and target DNAs are relaxed.

Tn10: The Tn10 termini (T1,T2) and sequences within the transposon (a,b,c,d,e,f) are indicated. The cleaved donor complex consists of an excised transposon or ETF (excised transposon fragment) generated by transposase-mediated double-strand breaks at each transposon end. Tn10 is excised from the flanking donor DNA, and the transposon sequences are constrained as negative supercoils by the synaptic structure. The Tn10 strand transfer reaction shown here is intramolecular. DNA strand transfer occurs within the transposon (marked by "X"), resulting in inversion of the Tn10 sequences and the formation of 9 bp single-strand

gaps at the transposon/target junctions. Intramolecular strand transfer also releases the negative supercoiling in the transposon DNA.

Figure 2


Mu
Tn10
Cleaved Donor Complexes
T2
T1
f
a
b
e
d
c
X
Strand Transfer Complexes
b
c
a
T1
T2
f
d
e

Chapter 2

Bacterial transposon Tn7 utilizes two different classes of target sites

ABSTRACT

Sites of transposon Tn7 insertion in the Escherichia coli chromosome were examined, and two distinct classes of target sites differing in nucleotide sequence were identified. The target site choice was determined by Tn7-encoded transposition genes.

Most transposable elements move at low frequency and insert into many different target sites (Berg and Howe 1989). The bacterial transposon Tn7 (Barth et al. 1976; Craig 1989; Figure 1A) is unusual in that it can transpose at high frequency to a specific site in the Escherichia coli chromosome called attTn7, located near the bacterial glmS gene (Barth et al. 1979; Lichtenstein and Brenner 1981; Lichtenstein and Brenner 1982; Gay et al. 1986; McKown et al. 1988; Figure 1B). Tn7 can also transpose at low frequency to many different sites in plasmids (Barth et al. 1976; Barth et al. 1978; Moore and Krishnapillai 1982). Tn7 insertion is accompanied by a 5-bp duplication of target sequences (Lichenstein and Brenner 1982; Gay et al. 1986; Krishnapillai et al. 1987; Nash and Krishnapillai 1987; McKown et al. 1988). Tn7 encodes five genes - tnsABCDE - that mediate two distinct but overlapping transposition pathways (Rogers et al. 1986; Waddell and Craig 1988; Figure 1A). Previous studies have shown that tnsABC+tnsD promote transposition to attTn7, whereas tnsABC+tnsE promote transposition to other sites.

We have analyzed Tn7 insertions in the E. coli chromosome; its large size provides a wider variety of target sites than do the plasmids used in previous studies. We report here that the target site sequences used by the tnsD-dependent (tnsABC+tnsD) and tnsE-dependent (tnsABC+tnsE) pathways differ dramatically. We show that the tnsD-dependent pathway has high target site selectivity, promoting transposition to attTn7 and a limited number of other sites similar in sequence to attTn7. By contrast, the tnsE-dependent pathway displays

little target site selectivity, promoting transposition to many different sites not related in sequence to the tnsD-dependent sites.

We evaluated the frequency of Tn7 transposition to the E. coli chromosome (Table 2) by measuring the translocation of miniTn7Km, a Tn7 derivative containing a kanamycin resistance marker but lacking the tns genes (McKown et al. 1988; Arciszewska et al. 1989; Figure 1C), from the replication- and integration-defective bacteriophage λ derivative KK1 into bacteria containing the tns genes and selecting for kanamycin-resistant colonies (McKown et al. 1988). Strains and plasmids used are shown in Table 1. We also physically analyzed many of these miniTn7Km chromosomal insertions (Figure 2 and Table 3). Genomic DNA was isolated (Raleigh and Kleckner 1984), digested with restriction enzymes, resolved by agarose gel electrophoresis, transferred to Nytran (Schleicher and Schuell), and filters hybridized with nick-translated kanamycin fragment or nick-translated Tn7R fragment.

tnsABC+tnsD promote relatively high-frequency transposition to the chromosome (Table 2, line 4), and all such insertions are located in attTn7 (Figure 2A). To determine whether tnsABC+tnsD can also promote transposition to sites other than attTn7, we used a slightly different target chromosome, one already containing a Tn7 derivative in attTn7. We used this "blocked attTn7" chromosome because multiple Tn7 insertions in attTn7 have not been observed (Lichtenstein and Brenner 1981; Hauer and Shapiro 1984; Arciszewska et al. 1989). We detected tnsD-dependent transposition to the "blocked attTn7" chromosome at a frequency about 10^4 fold lower than to vacant attTn7 (Table 2, line 4). Analysis of these tnsD-dependent insertions revealed that they occurred into a

limited array of sites (Table 3) with a marked preference for the DC-1, DC-2 and DC-3 sites (DC = tnsABC+tnsD Chromosomal).

Comparison of the sequences of several DC sites to each other and to attTn7 (Figure 3) reveals considerable sequence similarity among all these tnsD-dependent target sites. This similarity is most pronounced to one side (rightwards) of the Tn7 insertion points; little similarity is apparent at the actual insertion points. Strikingly, this region of similarity between the DC sites and attTn7 includes a region which has been shown to be required for high-frequency, site-specific insertion into attTn7 (Gringauz et al. 1988; McKown et al. 1988; Qadri et al. 1989). Because of their similarity to attTn7, we call the DC sites "pseudo-attTn7" sites. Thus tnsABC+tnsD direct transposition to target sites of related nucleotide sequence, promoting high-frequency insertion in attTn7 and low-frequency insertion in pseudo-attTn7 sites.

tnsABC+tnsE promote low frequency transposition to the chromosome (Table 2, line 5). Physical analysis showed that the tnsABC+tnsE pathway ignores chromosomal attTn7 (compare Figure 2A and 2B). A more detailed analysis of tnsE-dependent insertions in a "blocked attTn7" chromosome demonstrated that EC (tnsABC+tnsE Chromosomal) sites were located in many different positions (Table 3). Notably, no two tnsE-dependent insertions occurred at the same target site, nor were any tnsE-dependent insertions located in the identified pseudo-attTn7 sites.

To examine the Tn7 insertions at higher resolution, we determined the nucleotide sequence for several DC and EC sites (Figure 3). Chromosomal DNA segments containing miniTn7Km insertions were isolated by cloning into pUC18 with selection for kanamycin resistance (Maniatis et

al. 1982). The sequences flanking the insertions were determined by the chain terminator method using Sequenase (USB) following alkali denaturation (Sanger et al. 1977). Oligonucleotides specific to the left and right Tn7 ends were used as primers.

Comparison of the DC sites with the EC sites (Figure 3) does not reveal any striking sequence similarities between these two classes of target sites, although both result in a 5-bp target sequence duplication. A modest similarity among our EC sites is detectable about 15 bps leftwards of the Tn7 insertion point; however, its functional significance is unclear. We note that previously characterized plasmid insertions of Tn7 (Lichtenstein and Brenner 1982; Krishnapillai et al. 1987; Nash and Krishnapillai 1987), which likely resulted from tnsABC+tnsE transposition, do not all share this sequence similarity in this region. Analysis of many more tnsE-dependent insertions will be required to determine if tnsE-dependent sites actually do share a common determinant. Nevertheless, the emergent picture is that tnsABC+tnsE promotes low-frequency transposition to many different target sites with little obvious target sequence selectivity.

When tnsD and tnsE are present individually with tnsABC, they direct Tn7 transposition to two different classes of target sites. What target site class is used when both tnsD and tnsE are present? When attTn7 is available, transposition to this site - a tnsABC+tnsD reaction - is exclusively observed even in the presence of both tnsD and tnsE (Lichtenstein and Brenner 1981; Hauer and Shapiro 1984; Rogers et al. 1986; Waddell and Craig 1988), likely reflecting the high frequency of this reaction. When attTn7 is unavailable, i.e. in a "blocked attTn7"

chromosome, we have observed similar frequencies of tnsD- and tnsE-dependent transposition (Table 2, lines 3-6). The distribution of insertions in a "blocked attTn7" chromosome obtained in the presence of both tnsD and tnsE also suggests that both pathways are active under these conditions (Table 3). Several insertions (D/EC sites 1-3) apparently lie in tnsD-dependent sites DC 1-3, and it is not unreasonable to suggest that the other D/EC sites represent both tnsD- and tnsE-dependent chromosomal sites. In previous studies the observed frequency of tnsD-dependent transposition to plasmids is much less than that of tnsE-dependent transposition (Rogers et al. 1986; Waddell and Craig 1988), perhaps reflecting a paucity of attTn7-like sequences in these relatively small DNA molecules.

We have shown that Tn7 can transpose to two distinct classes of target sites. tnsABC+tnsD transposition utilizes target sites of related sequence, directing insertion into attTn7 and pseudo-attTn7 sites. By contrast, tnsABC+tnsE transposition utilizes many different target sites not markedly related to each other or to the tnsD-dependent target sites. The emergent picture (this work; Rogers et al. 1986; Waddell and Craig 1988) is that tnsABC provide functions common to all Tn7 transposition events whereas tnsD and tnsE determine the target site. An attractive hypothesis is that TnsD and TnsE are target DNA binding proteins, TnsD being a specific DNA binding protein recognizing attTn7 and pseudo-attTn7 sites and TnsE being a non-specific DNA binding protein. In support of this view, a tnsD-dependent binding activity that specifically recognizes attTn7 has been identified (Waddell and Craig 1989); the TnsD polypeptide copurifies with this attTn7-binding activity

(Chapter 3). Moreover, the tnsD-dependent recognition region in attTn7 (from about +30 to +50) corresponds to the region of sequence simlarity among the tnsD target sites, i.e. attTn7 and the pseudo-attTn7 sites.

Is Tn7's ability to use two distinct classes of target sites advantageous? Target site specific insertion into attTn7 provides a union between the host and Tn7 in which no bacterial gene is inactivated although sequences in a bacterial gene (glmS) are required for Tn7 insertion (Gay et al. 1986). Tn7's ability to transpose with little site selectivity is perhaps most useful in its transfer to and among plasmids.

ACKNOWLEDGMENTS

We thank members of the laboratory for many helpful discussions and Candy Waddell, Pascal Gamas, Roland Bainton, and Rick Myers for comments on the manuscript.

K.M.K. was supported by a Public Health Service genetics training grant from the National Institute of General Medical Sciences. This work was supported by a Public Health Sevice grant from the National Institute of Allergy and Infectious Diseases to N.L.C.

Table 1. Strains and plasmids used in this study

Strain	Genotype and Reference
NLC51	F⁻ *araD*139A Δ(*argF*-*lac*)U169 *rpsL*150 *relA*1 *flbB*5301 *deoC*1 *ptsF*25 *rbsR* *val*^{R} *recA*56 (Waddell and Craig 1988)
LA214	NLC51 *attTn7*::miniTn7Cm[a] (Arciszewska et al. 1989)

Plasmid	Description and Reference
pCW4	pACYC184 derivative containing *tnsABCDE* (Waddell and Craig 1988)
pCW4miniMuΩ76	same as pCW4 except for insertion in *tnsE* inactivating that gene (Waddell and Craig 1988)
pCW21	pACYC184 derivative containing *tnsABC* (Waddell and Craig 1988)
pCW23	pUC18 derivative containing *tnsD* (Waddell and Craig 1988)
pCW30	pUC18 derivative containing *tnsE* (Waddell and Craig 1988)

[a]miniTn7Cm is a Tn7 derivative lacking the *tns* genes and containing a chloramphenicol resistance marker (Arciszewska et al. 1989; McKown et al. 1988).

Table 2. Transposition frequency to the E. coli chromosome[a]

		Transposition Frequency (KmR colonies/PFU) <Mean ± SEM (n)>	
tns genes	Plasmids	Vacant attTn7	attTn7::miniTn7Cm
none	none	$< 4.3 \times 10^{-8}$ (3)	$< 6.3 \times 10^{-9}$ (3)
tnsABC	pCW21	$< 4.3 \times 10^{-8}$ (3)	$< 6.3 \times 10^{-9}$ (3)
tnsABC::miniMu	pCW4miniMuΩ76	n.d.	$(1.0 \pm 0.3) \times 10^{-7}$ (6)
tnsABC+tnsD	pCW21 + pCW23	$(3.9 \pm 2.8) \times 10^{-4}$ (6)	$(3.0 \pm 1.1) \times 10^{-8}$ (3)
tnsABC+tnsE	pCW21 + pCW30	$(3.0 \pm 0.0) \times 10^{-7}$ (3)	$(9.1 \pm 2.4) \times 10^{-8}$ (3)
tnsABCDE	pCW4	$(8.0 \pm 1.0) \times 10^{-5}$ (3)	$(1.1 \pm 0.3) \times 10^{-7}$ (6)

[a]E. coli used were derivatives of NLC51 or LA214 containing the indicated plasmids (Table 1).

Table 3. Physical analysis of miniTn7Km insertions in the blocked attTn7 chromosome[a]

Target Site Class[b]	Number of Independent Products	Restriction Fragment Size (kb)[c]			
		EcoRI	BamHI	HindIII a	HindIII b
tnsD					
DC-1	12	13.5-20	7.3-7.8	11-11.5	10.5
DC-2	10	11.5-15.5	15-17.5	18.5-20	7.1-8.0
DC-3	7	10.5-13	12-13.5	11-12	n.d.
DC-4	2	14	16.5-17	17-17.5	n.d.
DC-5	2	19	17	n.d.	n.d.
DC-6	1	19	6.5	9.0-9.8	3.4
DC-7	1	17	13.5	n.d.	n.d.
DC-8	1	12.5	12	10	n.d.
DC-9	1	11.5	20	18	n.d.
tnsE					
EC-1	1	7.3	17.5	3.0-3.1	2.1
EC-2	1	13	15	16	1.8-1.9
EC-3	1	4.4	17.5	16.5-19	8.1
EC-4	1	19	19	10.5-12.5	2.9
EC-5	1	9.3	15	16.5	4.2-4.6
EC-6	1	9.3	15	8.0	n.d.
EC-7	1	n.d.	16.5	9.0	n.d.
EC-8	1	11	15.5	17	n.d.
EC-9	1	7.7	17.5	18.5	8.0-8.7
EC-10	1	7.0	6.0	7.2	1.9-2.0
EC-11	1	13	15.5	5.4	2.4-2.6
EC-12	1	12	11.5	3.3	n.d.
EC-13	1	n.d.	15.5	6.1	1.8
EC-14	1	n.d.	15.5	5.8	2.3
EC-15	1	7.5	7.0	2.2-2.3	2.1
EC-16	1	8.0	17	3.3-3.5	2.9
tnsD + tnsE					
D/EC-1 (= DC-1)	3	16.5	7.5	11-11.5	10.5
D/EC-2 (= DC-2)	4	14	17.5	20	n.d.
D/EC-3 (= DC-3)	1	11.5	13.5	11.5	n.d.
D/EC-4	1	4.9	14.5	10	n.d.
D/EC-5	1	11.5	11	2.2	n.d.
D/EC-6	1	9.8	11	8.8	n.d.
D/EC-7	1	9.8	3.8	19.5	n.d.
D/EC-8	1	10.5	6.7	n.d.	n.d.
D/EC-9	1	15	10.5	3.4	n.d.
D/EC-10	1	8.6	11	5.1	n.d.
D/EC-11	1	7.3	16	4.8	n.d.
D/EC-12	1	18	14.5	2.3	n.d.

[a]Genomic DNA from various miniTn7Km insertions in a "blocked attTn7" chromosome obtained as described in Table 2 (tnsABC+tnsD as in line 4; tnsABC+tnsE as in line 5; tnsABCDE as in line 6) was physically analyzed as described in text.

[b]Independent insertions are classified as in the same target site if they share similar restriction fragment sizes; size ranges are from determinations repeated on separate blots. Insertions producing restriction fragments of related size were examined on the same blot, and additional restriction fragments were analyzed to confirm target site classes (data not shown). DC = tnsABC+tnsD Chromosomal, EC = tnsABC+tnsE Chromosomal, and D/EC = tnsABCDE Chromosomal.

[c]The lengths of miniTn7Km-containing restriction fragments are shown, as detected by using a nick-translated kanamycin fragment or a nick-translated Tn7R fragment. Both probes identify EcoRI and BamHI fragments containing intact miniTn7Km and flanking DNA. For HindIII, the kanamycin probe identifies two fragments (a,b for HindIII) and the Tn7R probe identifies one fragment (a for HindIII) containing portions of miniTn7Km and flanking DNA.

Figure 1. Tn7, attTn7 and miniTn7Km.

(A) Tn7.

The left (Tn7L) and right (Tn7R) ends of Tn7 are indicated. Tn7 contains two resistance determinants, dhfr encoding a dihydrofolate reductase providing trimethoprim resistance, and aadA encoding 3''(9)-O-nucleotidyl-transferase providing spectinomycin and streptomycin resistance (Fling and Richards 1983; Simonsen et al. 1983; Fling et al. 1985). The five Tn7 transposition genes tnsABCDE (Rogers et al. 1986; Waddell and Craig 1988) are shown.

(B) Chromosomal attTn7.

The point of Tn7 insertion is indicated (vertical arrow) and lies between the phoS and glmS genes as previously described (Lichtenstein and Brenner 1981; Lichtenstein and Brenner 1982; Gay et al. 1986; McKown et al. 1988). Tn7 insertion into attTn7 is orientation-specific such that Tn7L is proximal to phoS and Tn7R is proximal to glmS (Lichtenstein and Brenner 1981).

(C) MiniTn7Km.

The ends of Tn7 are shown as open boxes and the segment between them encodes a gene that confers resistance to kanamycin (McKown et al. 1988; Arciszewska et al. 1989). Restriction sites: H, HindIII.

Figure 1

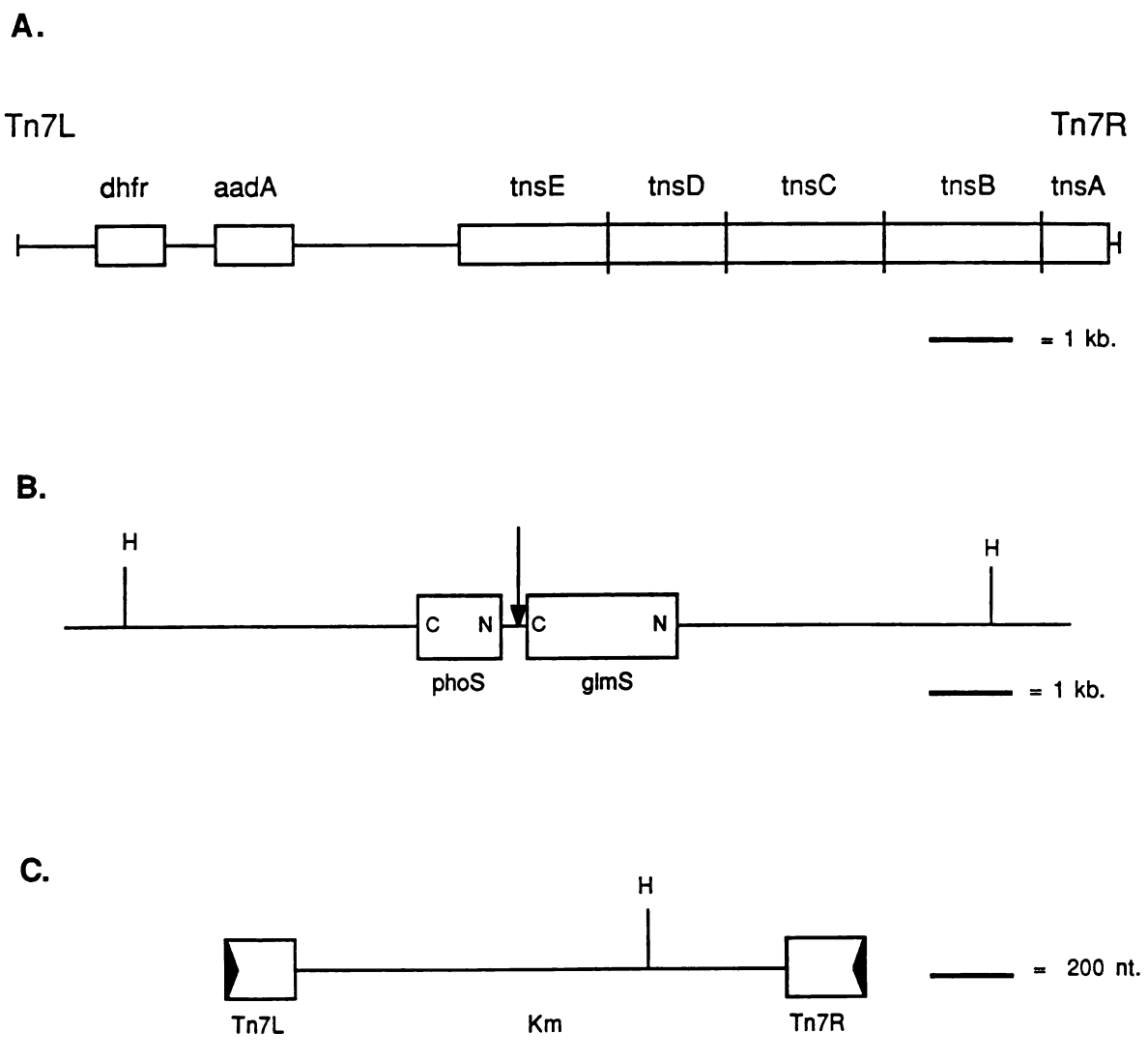

A.
Tn7L
Tn7R
dhfr
aadA
tnsE
tnsD
tnsC
tnsB
tnsA
= 1 kb.
B.
H
H
C
N
C
N
phoS
glmS
= 1 kb.
C.
H
Tn7L
Km
Tn7R
= 200 nt.

Figure 2. Distribution of chromosomal insertion sites.

Each numbered lane contains genomic DNA from a single transposition product which was physically analyzed as described in text. Genomic DNA was digested with HindIII and hybridized with nick-translated kanamycin fragment. The lanes marked "*" are from an attTn7::miniTn7Km chromosome; digestion with HindIII generates two 6.5 kb fragments (Figure 1) not resolved here which hybridize to the kanamycin probe. The lane marked "M" has molecular weight markers. Panel A: tnsABC+tnsD insertions generated in vacant attTn7 chromosome as in Table 2, line 4. Panel B: tnsABC+tnsE insertions in a vacant attTn7 chromosome generated as in Table 2, line 5.

Figure 2

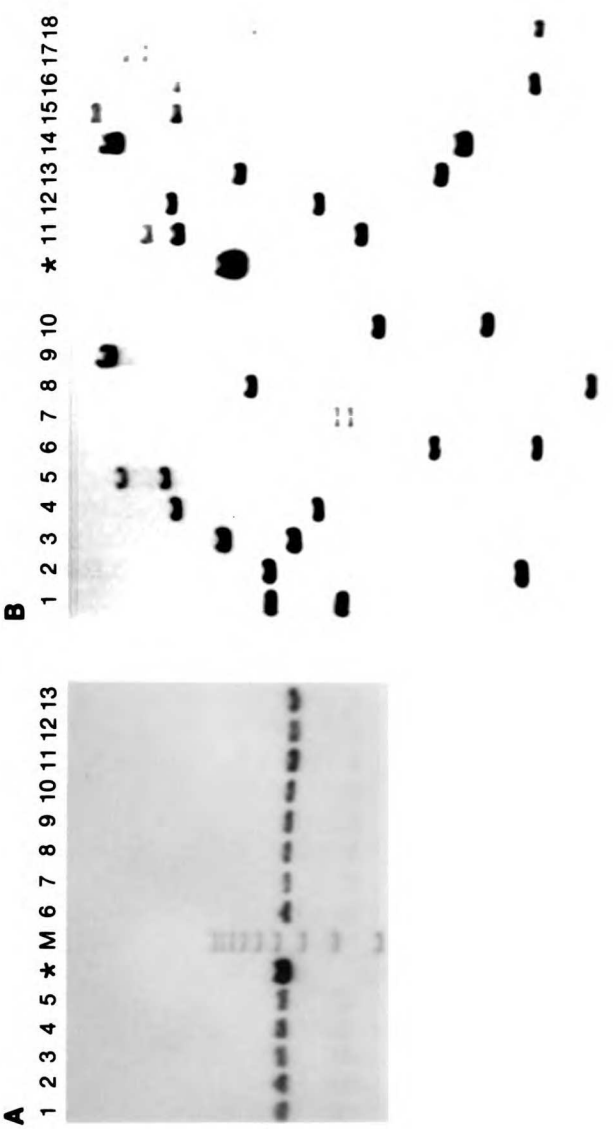


Figure 3. Sequence analysis of Tn7 insertion sites.

The boxed nucleotides indicate the 5-bp chromosomal sequence duplicated upon Tn7 insertion (Lichtenstein and Brenner 1982; Gay et al. 1986; McKown et al. 1988). The upper panel also shows the sequence of attTn7; the central bp of the duplicated sequence is designated "0", sequences leftwards (towards phoS) are given "-" values and sequences rightwards (towards glmS) are given "+" values (see McKown et al. 1988); the COOH terminus of the glmS gene product is also indicated (Gay et al. 1986). The bar underlines the region containing the nucleotides required for attTn7 target activity (McKown et al. 1988; Gringauz et al. 1988; Qadri et al. 1989). Among the tnsD-dependent (DC) sites, sequence identities with attTn7 are indicated in boldface type and positions with 3 sequence identities among attTn7 and the pseudo-attTn7 sites are indicated with "*". The sequence of two independent DC-1 insertions was determined; a difference of 1 bp in the positions of the duplicated nucleotides was observed (other duplication = GGTTG). A small degree of "wobble" has also been observed upon insertion in attTn7 (McKown et al. 1988). The lower panel shows the sequences of tnsE-dependent (EC) sites. Positions of sequence identify among at least 3 of the EC sites are indicated by boldface type and "*". The chromosomal locations of the DC and EC sites are not known, although sequence comparisons suggest that the DC-3 insertion may lie within the bacterial aceK gene (Cortay et al. 1988; Klumpp et al. 1988). Nucleotide sequences of the DC and EC sites have been deposited in the GenBank data base (accession nos. M31529-M31536).

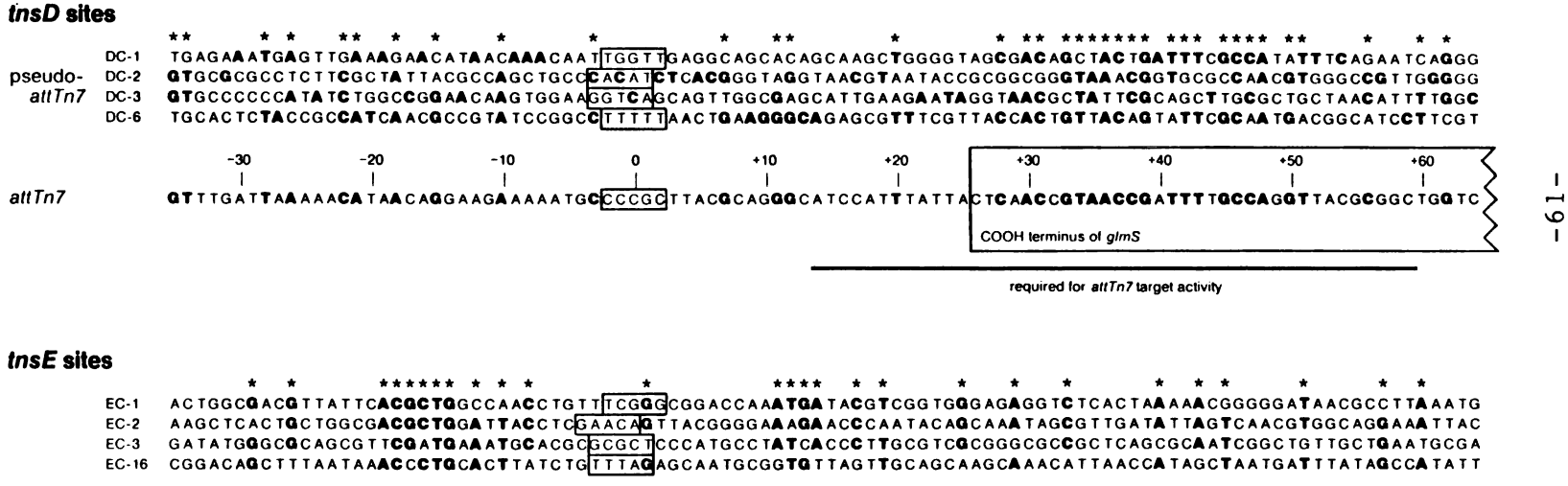

tnsD sites
pseudo-attTn7
DC-1
DC-2
DC-3
DC-6
-30
-20
-10
0
+10
+20
+30
+40
+50
+60
attTn7
COOH terminus of glmS
required for attTn7 target activity
tnsE sites
EC-1
EC-2
EC-3
EC-16

Figure 3

Chapter 3

Purification and characterization of TnsD, a Tn7 transposition protein that binds specifically to the attTn7 target site

ABSTRACT

The bacterial transposon Tn7 encodes four transposition genes, tnsABCD, that are required for high frequency insertion into a specific site, called attTn7. We have purified TnsD protein and show that TnsD directly binds to attTn7, recognizing sequences located at a distance from the specific point of Tn7 insertion. We also demonstrate that TnsD directly participates in and is essential for transposition to attTn7 in a cell-free recombination system. These studies indicate that TnsD is a critical determinant for the selection of attTn7 as a specific site of Tn7 insertion.

INTRODUCTION

Most transposable elements insert into many different target sites at low frequency (for review, see Berg and Howe 1989). The bacterial transposon Tn7 is unusual in that it can transpose at high frequency to a specific site in a variety of bacterial genomes (Barth et al. 1976; for review, see Craig 1989). In Escherichia coli this site, called attTn7, is located at minute 84 near glmS, a gene involved in cell wall biosynthesis (Lichtenstein and Brenner 1981; Gay et al. 1986; Gringauz et al. 1988). We are interested in understanding how Tn7 recognizes and inserts specifically into attTn7.

Tn7 encodes five genes, tnsABCDE, which mediate transposition to two distinct classes of target sites (Rogers et al. 1986; Waddell and Craig 1988; Kubo and Craig 1990). tnsABC + tnsD direct insertion at high frequency into attTn7 and at low frequency into pseudo-attTn7 sites, which are related in sequence to attTn7. By contrast, tnsABC + tnsE promote low frequency insertion into many apparently random sites which are unrelated to attTn7. Thus, tnsABC provide functions common to all Tn7 insertions, whereas tnsD and tnsE mediate target site selection.

We have proposed that TnsD protein is a specific target DNA binding protein which recognizes attTn7 (Waddell and Craig 1989; Kubo and Craig 1990). Analysis of crude cell lysates has identified a tnsD-dependent DNA binding activity that specifically recognizes attTn7 (Waddell and Craig 1989). We report here the purification of TnsD protein and show that it directly interacts with attTn7. We also show that the tnsD-dependent attTn7 binding activity observed in crude cell

lysates reflects the interaction of both TnsD and a host protein with attTn7. Using a cell-free system (Bainton et al. 1991) we also demonstrate that TnsD directly participates in and is essential for Tn7 insertion into attTn7. Thus, TnsD promotes Tn7 insertion by directly interacting with attTn7.

We have also determined the region of attTn7 that contains the specific nucleotide sequences required for TnsD binding and for attTn7 transposition activity. In our nomenclature for attTn7, the central base pair of the 5 bp target sequence duplicated by Tn7 insertion (Lichtenstein and Brenner 1982; Gay et al. 1986; McKown et al. 1988) is designated "0", with sequences toward glmS given positive values, and sequences away from glmS given negative values. Other studies have shown that sequences required for attTn7 target activity are arranged in an interesting way: the sequence information lies at a distance, within the region from +14 to +58, so that no special sequences at the specific point of Tn7 insertion are required (Gringauz et al. 1988; McKown et al. 1988; Waddell and Craig 1989; Qadri et al. 1989). We show here that a small region of attTn7 from +23 to +58, which contains the sequences for TnsD interaction and lies within glmS, can efficiently promote Tn7 transposition. We suggest that DNA recognition by TnsD is the critical determinant for attTn7 target activity.

RESULTS

Purification of TnsD

As a source of TnsD protein, we used an *E. coli* strain carrying an expression plasmid in which the the proposed *tnsD* initiation codon (Flores et al. 1990; Orle and Craig 1991) is fused just downstream of a heterologous ribosome binding site and a highly expressed, IPTG-inducible peptide. *tnsD* overexpression in this construction apparently results from translational coupling to the upstream peptide open reading frame (Ghrayeb et al. 1984). Using this construction, the TnsD polypeptide is approximately 0.05% of total protein in a cleared cell lysate; this source produced about ten-fold more *tnsD*-dependent attTn7 binding activity than other *tnsD* expression plasmids and at least several hundred fold more binding activity than Tn7 in the chromosomal *attTn7* site. (Waddell and Craig 1989; Orle and Craig 1991; K.M.K., C.S. Waddell, and N.L.C., unpublished data).

The TnsD purification procedure involves a sequential polyethylenimine and ammonium sulfate fractionation of a cleared, sonicated cell lysate, which is then followed by three chromatography steps: Affigel heparin, green-1, and hydroxylapatite. A typical purification is shown in Figure 1A and Table 1. We purified TnsD protein by following the fractionation of the *tnsD*-dependent *attTn7* binding activity using a previously developed mobility shift assay in which TnsD fractions are supplemented with crude extract from *E. coli* cells lacking Tn7 (Waddell and Craig 1989). We also monitored the TnsD polypeptide (Orle and Craig 1991) by immunoblotting with affinity purified TnsD

antisera (Figure 1B). The apparent molecular weight of TnsD as determined by SDS-PAGE is 54 kD, which is in close agreement with the value of 59 kD from the proposed amino acid sequence of TnsD (Flores et al. 1990). In all fractionation steps, the TnsD polypeptide copurified with the attTn7 binding activity. From this procedure 57 g cell paste yielded about 140 μg TnsD protein of nearly 85% purity, and the specific activity of TnsD increased over 2000-fold (Table 1). Purified TnsD has no detectable nuclease activity with both supercoiled and linear substrates that contain or lack attTn7 sequences (R. Bainton, K.M.K., and N.L.C., unpublished data).

TnsD directly recognizes attTn7

Using a mobility shift assay (Fried and Crothers 1981; Garner and Revzin 1981), we examined whether purified TnsD protein binds to attTn7 (Figure 2A). Using an end-labeled attTn7 (-25 to +64) segment, we found that purified TnsD with added crude extract from cells lacking Tn7 produces the tnsD-dependent complex observed in crude TnsD lysates (Figure 2A, lanes 3 and 4). This TnsD+Host complex is not present in crude host extract alone (lane 2). Purified TnsD alone also produces a discrete retarded species (lane 5); however, this TnsD complex migrates faster than the TnsD+Host complex. We evaluated the attTn7 specificity of the TnsD and TnsD+Host species by substituting an end-labeled DNA fragment of similar length but lacking attTn7 sequences (lane 8); in this case, no retarded complexes were observed. Furthermore, the attTn7-binding activity also cosedimented with the TnsD polypeptide on a

glycerol gradient, indicating that TnsD participates directly in binding to attTn7 (Figure 3).

In addition to changing the mobility of the TnsD complex, the inclusion of host extract with purified TnsD stimulates by as much as 10-fold the amount of total complex formed as compared to the TnsD complex alone for a given amount of TnsD protein (compare Figure 2A, lanes 4 and 5; K.M.K. and N.L.C., data not shown). The host activity is heat-stable, trypsin-sensitive, phenol-sensitive, and nuclease-resistant, suggesting that it is protein (K.M.K., O. Hughes, and N.L.C., unpublished data).

Formation of the TnsD and TnsD+Host complexes requires relatively simple buffer and salt conditions; ATP does not obviously affect these species. Although these complexes are stable to relatively high salt concentrations (> 0.3 M KCl), they show a particular sensitivity to $MgCl_2$ (> 5 mM) (data not shown). We note that these complexes can stably form when incubated with other divalent cations, such as in $MnCl_2$ and $CaCl_2$. We thus performed nuclease footprint analyses of the complexes (as described below and in Experimental Procedures) utilizing low $MgCl_2$ concentrations or other divalent cations.

Analysis of TnsD interaction with attTn7 by nuclease protection

We examined the interaction between TnsD and attTn7 by footprint analysis of TnsD-attTn7 complexes. We incubated purified TnsD (in the absence of crude host extract) with end-labeled attTn7 and then treated the mixture briefly with either micrococcal nuclease or DNase I. After resolving the treated TnsD-attTn7 complexes by polyacrylamide gel

electrophoresis, we analyzed the recovered DNA on a sequencing gel. The results of this analysis are shown in Figure 4 and summarized in Figure 5.

Micrococcal nuclease analysis revealed a region of protection in attTn7 located entirely rightward of the specific Tn7 insertion point (Figure 4A). The extent of protection spans from about +24 to +53; some positions within this region do remain accessible to cleavage. Notably, the region encompassing the specific point of Tn7 insertion remains accessible to attack, indicating that TnsD does not directly contact these sequences. We obtained a similar result from footprint analysis with DNase I, where the region of TnsD protection extends from about +30 to +50 (Figure 4B). We note that several positions within this protected region also remain sensitive to attack.

We have also examined the TnsD+Host interaction with attTn7 by DNase I protection. We observed only a modest protection on the top strand that was similar to the TnsD protection (data not shown); however, analysis of the bottom strand revealed that the TnsD+Host footprint, which extends from about +18 to +59, is significantly longer and more strongly protected than that of TnsD alone (Figure 4C). Thus, a host activity can contribute to the TnsD interaction with attTn7. Whether the extended protection represents additional contacts by a host protein(s) or TnsD remains to be established.

Nucleotide sequences required for TnsD-attTn7 binding and attTn7 target activity

Previous studies have shown that the sequence information required for attTn7 target activity is displaced entirely rightward of the Tn7 insertion towards glmS. A plasmid containing attTn7 sequences (+7 to +64) efficiently promoted Tn7 transposition in vivo (Gringauz et al., 1988). We found that a plasmid containing a short attTn7 segment (+23 to +58) has target activity comparable to a larger attTn7 segment (-52 to +64) which provides wild-type activity (Table 2). As previously shown, a smaller attTn7 segment (+28 to +55) is substantially diminished in target activity, although Tn7 can insert specifically into this attTn7 target (Waddell and Craig 1989; Table 2). Thus, the sequences essential for attTn7 target activity correspond to those protected by TnsD (see above).

The sequences that provide attTn7 target activity are also very similar to those specifically bound by TnsD as shown by mobility shift assay (Figure 2B). We found that TnsD efficiently promotes complex formation with a DNA fragment containing attTn7 (+23 to +58) (Figure 2B, lane 3), which also served as an effective target. A DNA segment with less attTn7 sequences (+28 to +55), which is reduced in target activity, is also defective in TnsD binding (lane 6), although both the attTn7 (+28 to +55) and attTn7 (+23 to +58) segments are bound efficiently by TnsD+Host (lanes 2 and 5). These results indicate that the TnsD binding region closely approximates the sequences required for attTn7 target activity.

TnsD participates in transposition *in vitro* to *attTn7*

We asked whether purified TnsD is active in Tn7 transposition. We assayed transposition with a cell-free system (Bainton et al. 1991) in which a miniTn7 element transposes from a donor plasmid into an *attTn7* target plasmid in the presence of protein fractions from cells carrying *tnsA*, *tnsB*, *tnsC*, and *tnsD*, i.e. the genes required for *in vivo* insertion into *attTn7* (Rogers et al. 1986; Waddell and Craig 1988). Purified TnsD provides the *tnsD* function required for Tn7 transposition *in vitro* to *attTn7*; in the absence of TnsD, no insertions into *attTn7* were generated (Figure 6A). Transposition activity was comparable in crude TnsD lysates or with purified TnsD at equivalent concentrations of TnsD, suggesting that the TnsD activity is not significantly altered during purification (data not shown). Moreover, transposition activity cosedimented on a glycerol gradient with the TnsD-*attTn7* binding activity (Figure 6B). Thus, TnsD participates directly in both Tn7 transposition and specific binding to *attTn7*.

DISCUSSION

We have previously proposed that TnsD is a target DNA binding protein which specifically recognizes *attTn7* (Waddell and Craig 1989; Kubo and Craig 1990). In this work, we have purified TnsD and show that TnsD directly recognizes *attTn7*. We have also determined that purified TnsD also participates directly in and is essential for transposition *in vitro* to *attTn7*. TnsD therefore has a direct role in Tn7 transposition which explains the requirement for the *tnsD* gene in Tn7 insertion into

attTn7 in vivo (Rogers et al. 1986, Waddell and Craig 1988). Biochemical activities have also been established for other Tns proteins that participate directly in Tn7 transposition: TnsB is a sequence-specific DNA binding protein that recognizes the ends of Tn7 (McKown et al. 1987; Arciszewska et al. 1991; Tang et al. 1991), and TnsC is an ATP-dependent nonspecific DNA binding protein (Gamas and Craig, in press).

A particularly interesting feature of the attTn7 target activity is the requirement for sequences within the glmS gene which are displaced from the specific point of Tn7 insertion. We have found that a small attTn7 segment (+23 to +58) can confer high frequency target activity. In addition, attTn7 target activity closely correlates with the sequences required for TnsD interaction: the TnsD footprint extends from approximately +24 to +53, and TnsD binds efficiently to the small attTn7 segment (+23 to +58). It is tempting to speculate that TnsD binding is the only sequence-specific determinant for attTn7 target activity. Notably, the TnsD recognition site lies within the glmS coding sequence, but Tn7 insertion occurs downstream of glmS, thereby avoiding inactivation of this essential gene. The likely presence of glmS-like genes in a number of different bacteria may also explain the ability of Tn7 to insert in a site-specific manner in a wide host range (reviewed in Craig 1989; Qadri et al. 1989).

In addition to directing transposition at high frequency into attTn7, tnsABC + tnsD also promote insertions at low frequency into pseudo-attTn7 sites, which are related in sequence to attTn7 (Kubo and Craig, 1990). Notably, the pseudo-attTn7 and attTn7 sites are considerably similar in sequence from approximately +30 to +50. This

region of homology coincides with the TnsD recognition sequence in attTn7, suggesting that TnsD directs insertion into pseudo-attTn7 sites by directly binding to these sequences. The reduced target activity of the pseudo-attTn7 sites presumably reflects the differences in nucleotide sequence between these sites and attTn7. It will be interesting to determine whether these sites are reduced in TnsD binding or reflect some other feature of the target DNA.

Our studies have also revealed that the attTn7 binding activity in crude TnsD lysates actually results from the interaction of TnsD and a host protein(s) with attTn7. The identity of the host protein(s) and its activity in Tn7 recombination remains to be established. We have shown that neither IHF nor HU affects TnsD-attTn7 binding (data not shown). Tn7 transposition to attTn7 can also occur in a reconstituted in vitro system with purified TnsA, TnsB, TnsC, and TnsD (Chapter 4). Thus, host proteins do not play an obviously essential role in Tn7 recombination, although they may act as accessory proteins (Drlica and Rouviere-Yaniv, 1987; Friedman, 1988; Glasgow et al., 1989; Landy, 1989). In addition, E. coli mutations which affect Tn7 transposition in vivo have been identified (O. Hughes and N.L.C., unpublished data).

We have shown that TnsD directly recognizes attTn7 and may be the unique specificity determinant for target activity. TnsD apparently lacks catalytic activity, i.e. transposon end cleavage and strand transfer, since TnsABC with AMP-PNP can promote non-target site selective insertions (Chapter 4). It is tempting to speculate that TnsD activates the transposition machinery in a manner analogous to the MuB protein in bacteriophage Mu transposition, where MuB can both direct

target selection and stimulate Mu end cleavage and strand transfer (Baker et al. 1991; Surette et al. 1991; Surette and Chaconas 1991). TnsD may also direct assembly of the transposition apparatus onto the attTn7 target site. In this view, TnsD plays a role similar to replication or transcription initiation proteins which bind to DNA in a sequence-specific manner to facilitate assembly of the catalytic machinery at the proper initiation site (Bramhill and Kornberg 1988; Echols 1990; Greenblatt 1991). Elsewhere, we have demonstrated that TnsD recruits TnsC, an ATP-dependent non-specific DNA binding protein, to the attTn7 target sequence (Chapter 4). TnsC may thus mediate the communication between TnsD, bound specifically to attTn7, and the rest of the transposition machinery. Further biochemical studies of Tn7 transposition will provide more insight into the assembly and activity of multiprotein-DNA complexes.

EXPERIMENTAL PROCEDURES

Plasmids and bacterial strains

The TnsD expression plasmid pKK18 was constructed by inserting a NcoI-HindIII *tnsD* fragment from pKA041 (Orle and Craig 1991), after partial digestion with HindIII, between the NcoI and HindIII sites of pGD108 (Ghrayeb et al. 1984). In pKA041 the proposed *tnsD* initiation ATG was modified to produce a NcoI site, which produces a change in the second amino acid from the N-terminus of TnsD (Gly to Ser).

Plasmids containing *attTn7* sequences were also used; *attTn7* sequences are numbered as shown in Figure 5. pKK20 contains the BamHI-EcoRI *attTn7* (-25 to +64) fragment inserted between BamHI and EcoRI sites of pUC18. pKK24 contains the BamHI-EcoRI *attTn7* (-25 to +64) inserted between BamHI and EcoRI sites of pUC19. pKK25 contains the BamHI-EcoRI *attTn7* (-52 to +64) fragment from pEG31 inserted between the BamHI and EcoRI sites of pUC19. pKK28, which contains *attTn7* (+23 to +58), an EcoRI linker, and HindIII and SalI staggered ends, was obtained by annealing two 46 nt. synthetic oligonucleotides:

```
5'-AGCTTTACTCAACCGTAACCGATTTTGCCAGGTTACGCGGAATTCG
      AATGAGTTGGCATTGGCTAAAACGGTCCAATGCGCCTTAAGCAGCT-5'
```

followed by insertion between the HindIII and SalI sites of Bluescript-SK (Stratagene).

TnsD protein was prepared from NCM533 = *E. coli* K-12 λ+ *lacZ*::Tn5 *lacI*^{q1} (Shand et al. 1991) carrying the pKK18 *tnsD* plasmid. Host lysate for analysis of TnsD binding and transposition activity was prepared from MC4100 = F- *ara*D139A Δ(*arg*F-*lac*)U169 *rps*L150 *rel*A1 *flb*B5301 *deo*C1 *pts*F25 *rbs*R. For λ hop assays, LA3 (= MC4100 *val*^{R} *rec*A56) was used.

Preparation and labeling of DNA fragments

Plasmid DNA was digested with appropriate restriction enzymes and end-labeled at 3' ends with (α-32P) dATP and the Klenow fragment of DNA polymerase I. End-labeled fragments were purified by separation in a nondenaturing polyacrylamide gel and visualized by autoradiography. DNA fragments were recovered by electroelution followed by isopropanol precipitation.

For mobility shift assays, fragments containing the following sequences were used: pUC18, the 122 bp PvuI-HindIII pUC18 fragment; attTn7 (-25 to +64), the 115 bp EcoR1-BamHI pKK20 fragment; attTn7 (+23 to +58), the 92 bp BamHI-ApaI pKK28 fragment. For nuclease protection studies, the pKK20 EcoR1-HindIII fragment (151 bp), labeled at either the HindIII or EcoR1 end for analysis of the top and bottom strand (as shown in Figure 5), respectively, was used. In addition, the EcoRI-PvuI fragment (260 bp), labeled at the EcoRI end, from pKK25, containing attTn7 (-52 to +64), was used to analyze the bottom strand for TnsD+Host footprint analysis.

Mobility shift assays

Reaction mixtures (20 μl) contained 50 mM Tris-HCl (pH 8.0), 140 mM KCl, 1 mM EDTA, 1.8 mM DTT, 9.8% glycerol (v/v), 340 μg/ml BSA, 17 μg/ml sheared salmon sperm DNA, and approximately 0.01 pmol end-labeled DNA fragment. After incubation for 15 minutes at room temperature, the reaction mixtures were electrophoresed through 5% polyacrylamide gels (29:1 acrylamide/N,N'-methylene-bisacrylamide) in Tris borate/EDTA

buffer at 11.5 V/cm for 1.75 hr. Gels were vacuum dried and exposed to X-ray film. Quantification of protein-DNA complexes was done by exposure of gels to a Molecular Dynamics phosphoimager.

Purification of TnsD

Unless otherwise indicated, all steps were done at 4°C.

Cell growth

Cells used for purification were grown at 37°C in a 200 L fermentor in LB broth (Miller, 1972) supplemented with 100 μg/ml carbenicillin. At OD 600 = 1.0, IPTG was added to 100 μM. Growth was continued for an additional 2 hours and cells were harvested by centrifugation. Cell paste was frozen and stored at -80°C.

Cleared cell lysate preparation

2 ml/g cells of buffer B (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 2 mM DTT, 10% (v/v) glycerol) with 1 M KCl was added to 57 g cells. Cells were subjected to sonication, cell debris was removed by centrifugation for 30 min at 16,800 x g, and supernatant was collected (Fraction I).

Polyethylenimine and ammonium sulfate fractionation

Polyethylenimine (pH 7.5) was added to 0.9% (v/v), the mixture incubated for 10 min, centrifuged for 30 min at 16,800 x g, and supernatant collected. Solid ammonium sulfate was added (423 mg/ml) with stirring over 40 min, the mixture was incubated an additional 20 min, and the resulting pellet was collected by centrifugation for 30 min at 16,800 x g. The pellet was resuspended in buffer B with 0.3 M KCl using volume equal to the 2 ml/g cells used for purification, followed by dialysis against buffer B with 0.3 M KCl with two 10 hr changes. After

centrifugation for 30 min at 16,800 x g, the supernatant was collected and diluted to 0.18 M KCl by the addition of buffer B with 0.1 M KCl. The diluted sample was centrifuged for 30 min at 16,800 x g, and supernatant collected (Fraction II).

Affigel heparin chromatography

Fraction II was applied to an Affigel heparin (Bio-Rad Laboratories) column (2.5 x 20 cm) equilibrated in buffer B with 0.18 M KCl. The column was washed with the same buffer, and eluted with a 500 ml linear gradient of 0.18-0.3 M KCl in buffer B, followed by a 250 ml elution with 0.3 M KCl in buffer B. TnsD binding activity eluted beginning at 0.27 M KCl, and active fractions were pooled (Fraction III).

Green-1 chromatography

Fraction III was applied to a Protrans Green-1 (ICN Biomedicals) dye column (1.0 x 7.0 cm) equilibrated in buffer B in 0.4 M KCl. The column was washed with same buffer and eluted with a 120 ml linear gradient of 0.4-1.0 M KCl in buffer B. TnsD binding activity eluted in a broad peak beginning at 0.65 M, and active fractions were pooled (Fraction IV).

Hydroxylapatite chromatography

Fraction IV was applied to a Bio-Gel HT hydroxylapatite (Bio-Rad Laboratories) column (0.8 x 2 cm) equilibrated in buffer B with 1.0 M KCl. The column was successively eluted with 0.5 M KCl, 0.5 M KCl + 25 mM KPO_4 (pH 8.0), and 0.5 M KCl + 50 mM KPO_4 (pH 8.0) in buffer B. TnsD binding activity eluted at 0.5 M KCl + 50 mM KPO_4 (pH 8.0); active fractions were dialyzed against 0.5 M KCl in buffer B with 25% (v/v)

glycerol instead of 10% glycerol. Dialyzed fractions (Fraction V) were frozen in liquid nitrogen and stored at -80°C. TnsD fractions were stable for at least several months.

Glycerol gradient sedimentation

5 ml of 20-40% (v/v) glycerol gradients were prepared in 50 mM Tris-HCl (pH 8.0), 2 mM DTT, 1 mM EDTA, and 0.5 M KCl. Approximately 3 μg TnsD (Fraction V) with 2 mg BSA and 2 mg ovalbumin were layered onto the gradients and centrifuged for 44 hr at 2°C at 40,000 rpm in a SW50.1 rotor. Fractions were collected from bottom of gradient.

Nuclease protection experiments

DNase I protection assays were performed using the method of Andrews et al. (1987), and micrococcal nuclease protection assays were done according to Zhang and Gralla (1989). The reactions (100 μl) contained 45 mM Tris-HCl pH 8.0, 90 mM KCl, 0.1 mM EDTA, 1.8 mM DTT, 320 μg/ml BSA, 20 μg/ml sheared salmon sperm DNA, 2 μg/ml TnsD (Fraction V),and approximately 0.15 pmol 3′ end-labeled DNA fragments. After incubation for 20 min at 30°C, reactions were subjected to either DNase I or micrococcal nuclease attack. For DNase I treatment, 5 μl of 20 mM $MgCl_2$ + 2 mM $CaCl_2$ solution and 5 μl of 4 μg/ml DNase I (Boehringer Mannheim) were added. (In some cases, $MnCl_2$ was substituted for $MgCl_2$, as described below). Incubation was continued for 1 min, and reaction was stopped by addition of EDTA to 10 mM. For micrococcal nuclease treatment, 5 μl of 20 mM $CaCl_2$ and 5 μl of 1.4 μg/ml micrococcal nuclease (Boehringer Mannheim) were added. Incubation was continued for

2 min, and reaction was stopped by addition of EDTA to 10 mM. TnsD-DNA complexes were resolved by electrophoresis, and gel slices containing DNA of interest were isolated. DNA was electroeluted, recovered by isopropanol precipitation, and then analyzed on a sequencing gel. DNA sequencing was performed using the method of Maxam and Gilbert (1980).

TnsD+Host complexes were analyzed by DNase I attack as described above, with the following modifications: $MnCl_2$ was substituted for $MgCl_2$, and 0.4 μg/ml TnsD (Fraction V) was used with 2 μg MC4100 host extract, which was prepared by boiling MC4100 crude cell lysate for 2 min and collecting the supernatant after centrifugation for 2 min at room temperature (Bainton et al. 1991).

Tn7 transposition in vitro

Reactions were performed according to Bainton et al. (1991). Reaction mixtures (100 μl) contained 25 mM HEPES pH 8.0, 3.5 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, 27 mM KCl, 5 mM NaCl, 2.1 mM DTT, 15 mM magnesium acetate (MgAc), 1.4% (v/v) glycerol, and 5% polyvinyl alcohol MW 8000 in the presence of 100 ng of donor plasmid containing a miniTn7 element and 2 μg of an attTn7 target plasmid. The reactions also contained 5 μg of crude tnsA extract (Bainton et al. 1991), 50 ng of fraction IV TnsB (Arciszewska et al. 1991), 50 ng of fraction III TnsC (Gamas and Craig, in preparation), and 2 μg host extract (Bainton et al. 1991). The indicated reaction in Figure 6A also contained 15 ng TnsD (fraction V).

After incubation in the absence of MgAc for 7 min at 30°C, 4 μl 375 mM MgAc was added and incubations continued for 30 min at 30°C. As described in Bainton et al. (1991), reactions were stopped and the DNA

recovered by spermine precipitation. An aliquot of the recovered DNA was digested with EcoRI (Figure 6A) or Bgl II (Figure 6B), electrophoresed through 0.7% agarose, transferred to Nytran (Schleicher and Schuell), and analyzed by Southern hybridization using a miniTn7-specific DNA probe. Reaction products were quantified by exposing blots on a Phosphorimager (Molecular Dynamics).

λ hop assays

attTn7 target activity was evaluated by a Tn7 λ hop assay (McKown et al. 1988) in which the attTn7-containing plasmids to be assayed were introduced into LA3, a strain containing a chromosomal insertion of Tn7 to provide transposition proteins. The replication- and integration-defective λKK1 (Chapter 2) contains miniTn7Km, which translocates into the attTn7 target plasmid from λKK1 upon infection.

Protein analysis

Protein concentration was determined using Bio-Rad Protein Assay with BSA as standard. SDS-PAGE using 10% gels was carried out by method of Laemmli. Prior to loading, TnsD fractions were concentrated by 10% TCA precipitation with insulin (Sigma) as carrier. Anti-TnsD antibodies were prepared as described elsewhere (Orle and Craig, 1991), except that TnsD antisera was affinity purified against a TnsD protein fusion to glutathione-S-transferase. Immunoblot analysis was carried out as described (Orle and Craig, 1991), except that 35S-labeled Protein A was used after incubation with anti-TnsD antibodies. Quantification of TnsD was performed by exposure to a Phosphorimager (Molecular Dynamics).

Table 1: TnsD Purification

	Fraction	Volume	Total protein	Total activity	Specific activity	Purification	Yield
		ml	mg	units	units/mg	-fold	%
I	Cleared lysate	130	6240	240,000	38	(1)	(100)
II	Polyethylenimine/ Ammonium Sulfate	360	3600	180,000	50	1.3	75
III	Affi-gel Heparin	270	25.7	78,000	3,000	79	32
IV	Green-1	52	1.3	42,000	32,000	840	18
V	Hydroxylapatite	5.4	0.16	16,000	100,000	2630	6.7

Specific activities of the pooled fractions were compared using the mobility shift assay to determine TnsD binding to attTn7. A unit of activity was defined as the amount of protein required to shift about 10% of attTn7 in the presence of an E. coli lysate.

Table 2. Evaluation of attTn7 target activity.

Plasmid	attTn7 sequence	Transposition/PFU <mean ± SEM (n)>
pSK	none	$(2.5 \pm 2.5) \times 10^{-7}$ (2)
pKK26	-52 to +64	$(2.5 \pm 2.1) \times 10^{-3}$ (3)
pKK28	+23 to +58	$(3.7 \pm 2.7) \times 10^{-3}$ (3)
pKK27	+28 to +55	$(2.9 \pm 1.6) \times 10^{-6}$ (2)

λ hop assays were done in LA3 containing the indicated plasmids. PFU, plaque-forming units.

LA3 cells lacking any plasmids show a similar transposition frequency (Kubo and Craig, 1990).

Figure 1. Purification of TnsD.

A) SDS-PAGE analysis of TnsD fractions. Lane 1: Fraction I, 95 μg. Lane 2: Fraction II, 95 μg. Lane 3: Fraction III, 10 μg. Lane 4: Fraction IV, 5 μg. Lane 5: Fraction V, 1.5 μg. The numbers at the left indicate the positions of the molecular weight markers, and the arrow indicates the position of the TnsD polypeptide.

B) Immunoblot analysis of TnsD. Lane 1, NCM533 crude lysate lacking TnsD, 50 μg; lanes 2 and 3, NCM533 crude lysate containing TnsD, 100 μg; lane 4, TnsD protein (Fraction V), 100 ng. The numbers at the left and right indicate the positions of the molecular weight markers, and arrows indicate position of TnsD polypeptide.

Figure 1

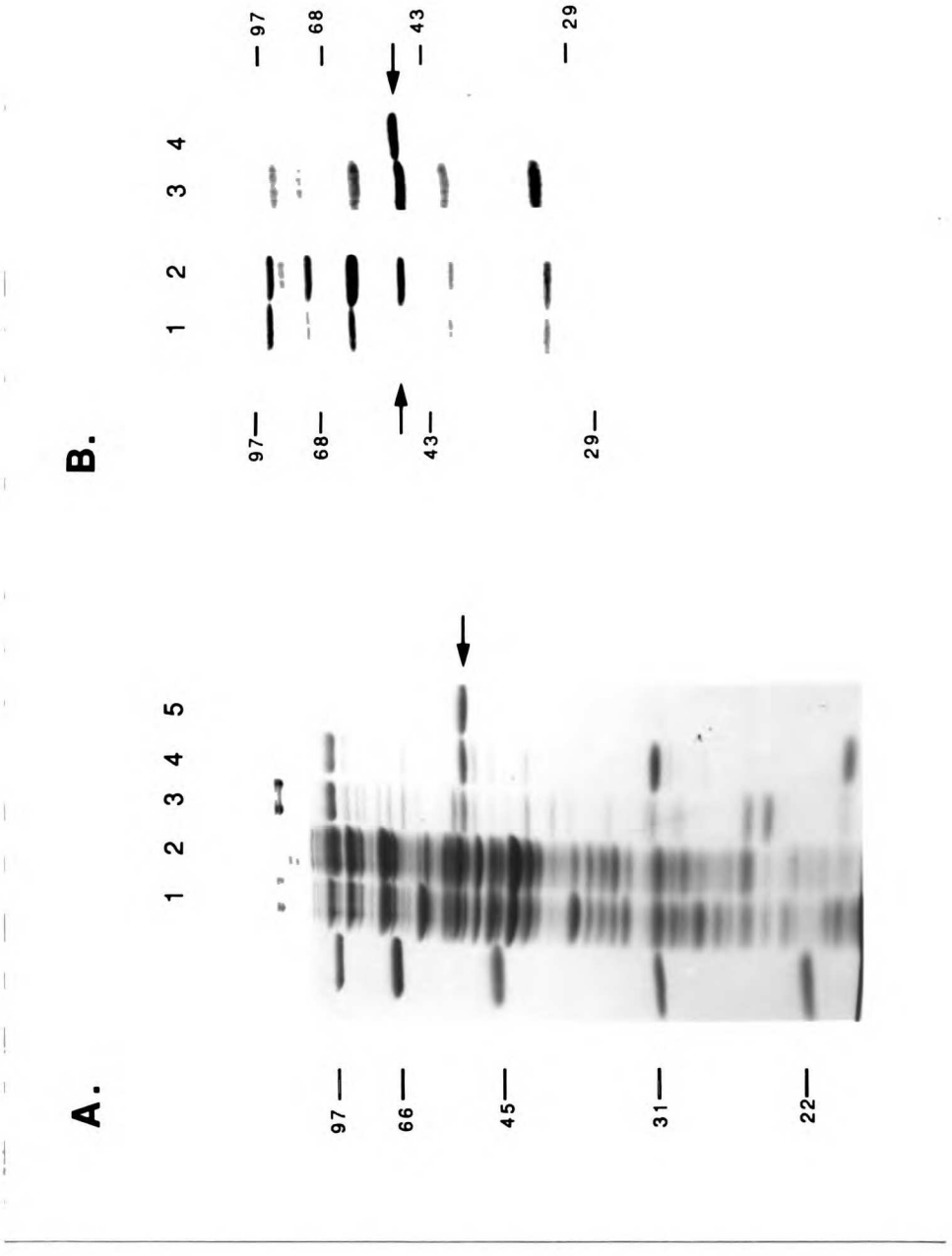


Figure 2. Mobility shift analysis of TnsD binding to attTn7.

A) Lanes 1-5 contain attTn7 (-25 to +64); lanes 6-8 contain pUC18 (122 bp). Lanes 1 and 6, no host lysate or TnsD. Lane 2, crude host lysate (5 μg). Lane 3, crude TnsD lysate (5 μg). Lanes 4 and 7, crude host lysate (5 μg) and TnsD (15 ng fraction V). Lanes 5 and 8, TnsD (15 ng fraction V).

B) Lanes 1-3 contain attTn7 (+23 to +58); Lanes 4-6 contain attTn7 (+28 to +55). Lanes 1 and 4, no TnsD or host lysate. Lanes 2 and 5, TnsD (5 ng fraction V) + host lysate (5 μg). Lanes 3 and 6, TnsD (15 ng fraction V).

Figure 2

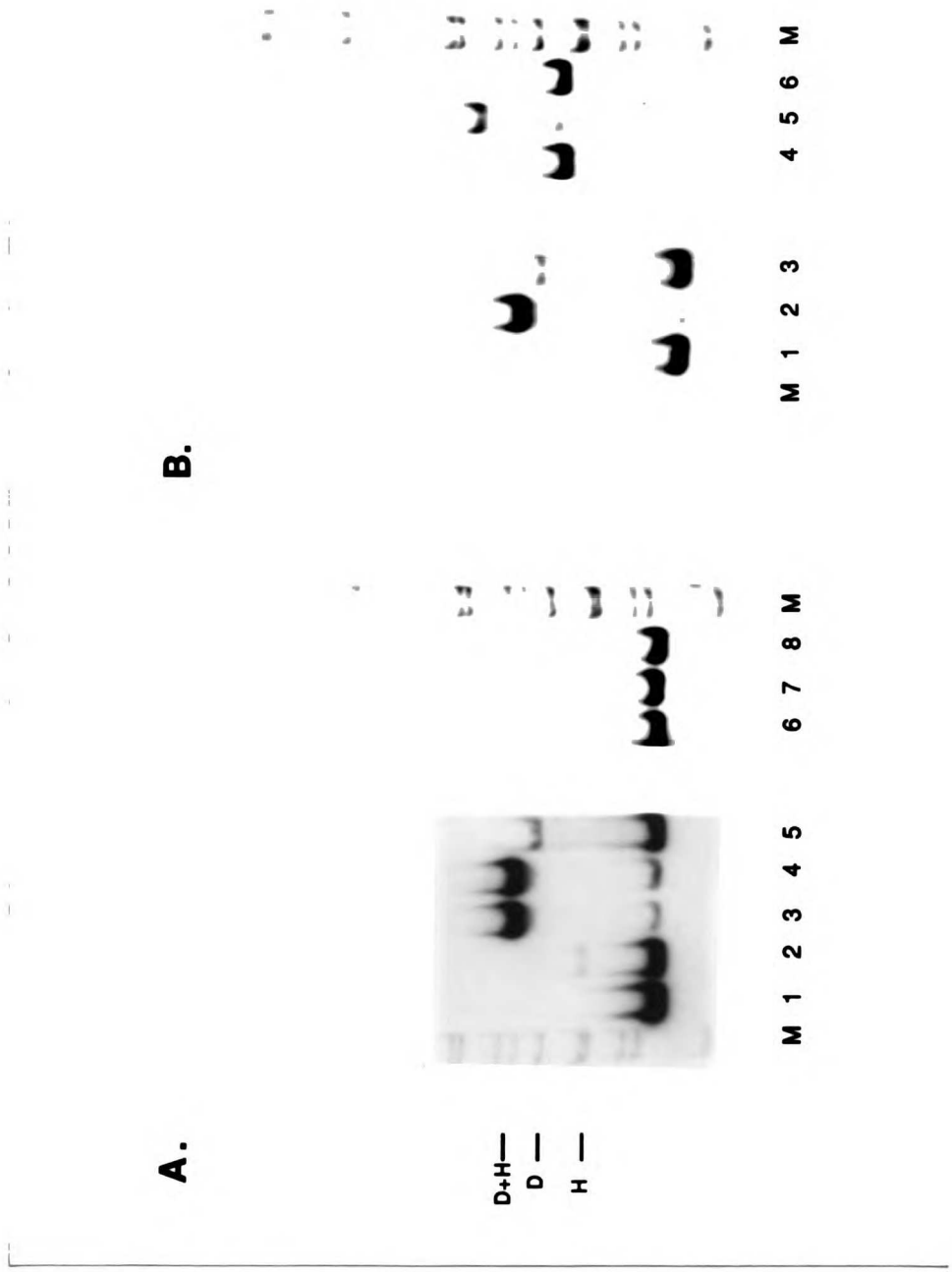


Figure 3. Cosedimentation of TnsD polypeptide and TnsD-dependent binding activity to attTn7.

TnsD (fraction V) was fractionated on a glycerol gradient, and each fraction was evaluated for TnsD binding to attTn7 using a mobility shift assay. The amount of TnsD in each fraction was quantified by immunoblot analysis followed by exposure of the resulting blot to a Phosphorimager. The amount of TnsD polypeptide in gradient fraction 16 was arbitrarily designated 1.0; the actual TnsD concentration was about 2.5 μg/ml. attTn7 binding activity is shown as the ratio of bound complex to total attTn7 DNA.

Figure 3

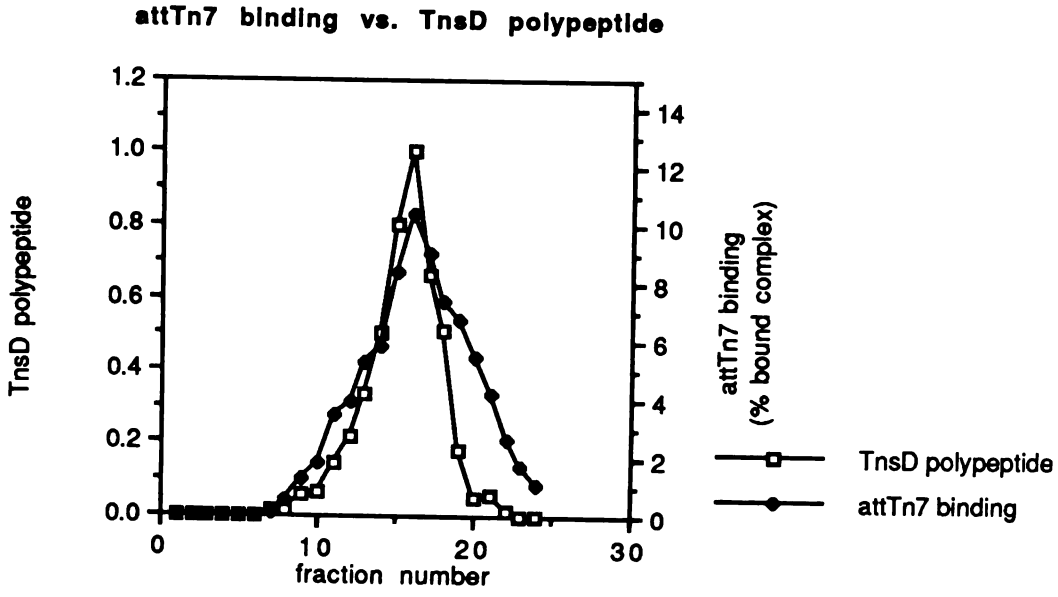

attTn7 binding vs. TnsD polypeptide
TnsD polypeptide
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
fraction number
attTn7 binding
(% bound complex)
14
12
10
8
6
4
2
0
TnsD polypeptide
attTn7 binding

Figure 4. Footprint analysis of TnsD and TnsD + Host complexes with attTn7.

TnsD and TnsD+Host complexes with attTn7 were formed, treated with micrococcal nuclease (TnsD complexes) or DNase I (TnsD and TnsD+Host complexes), and the DNA recovered from the isolated complexes. Top and bottom strands of attTn7, and the numbering of the nucleotide positions are as shown in Figure 5. The vertical lines indicate the extent of protection, and asterisks indicate individual protected positions observed outside the contiguous protection region. Lanes marked "AG" contain Maxam-Gilbert A+G reaction of same fragment.

A) Micrococcal nuclease footprint of TnsD-attTn7 complex. Left and right panels show top and bottom strands, respectively. Lane 1 (both panels), TnsD at 2 μg/ml; lane 2, no TnsD added.

B) DNase I footprint of TnsD-attTn7 complexes. Left and right panels show top and bottom strands, respectively. Lanes 1 and 2, no TnsD added; lanes 2 and 3, TnsD at 2 μg/ml.

C) DNase I footprint of of TnsD and TnsD+Host complexes with attTn7. Bottom strand is shown in both panels. Lane 1, TnsD at 2 μg/ml; lane 2, no TnsD added; lane 3, TnsD at 2 μg/ml and host lysate at 0.4 μg/ml; lane 4, no TnsD or host lysate added.

Figure 4A

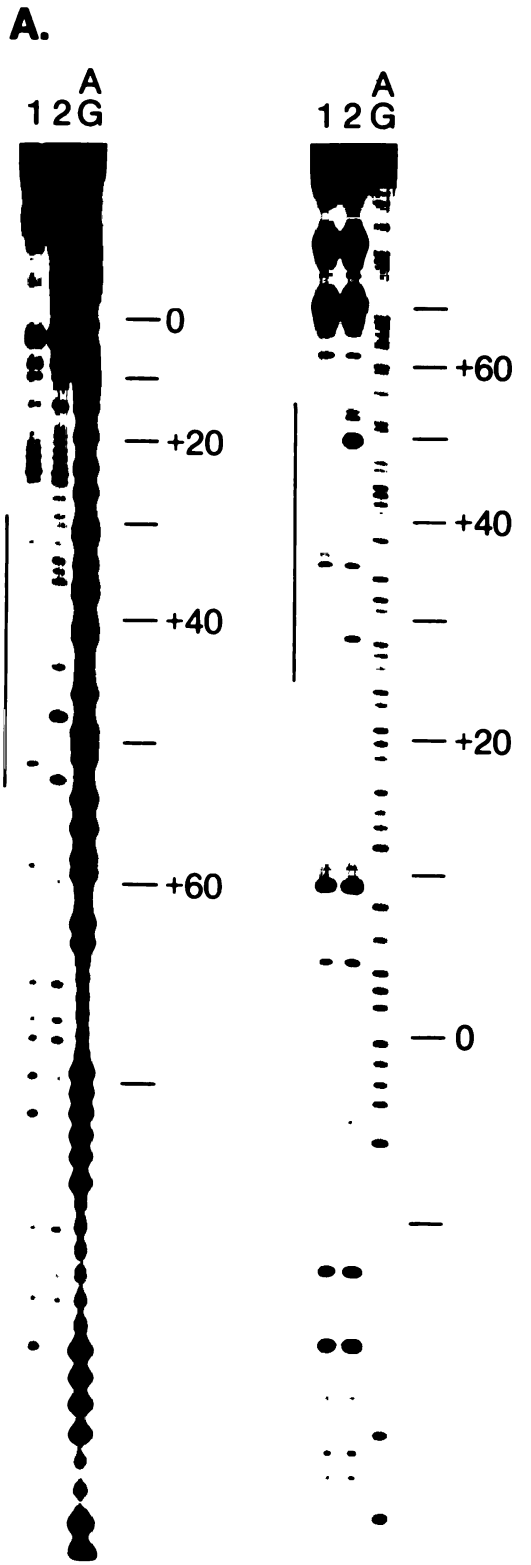


Figure 4B

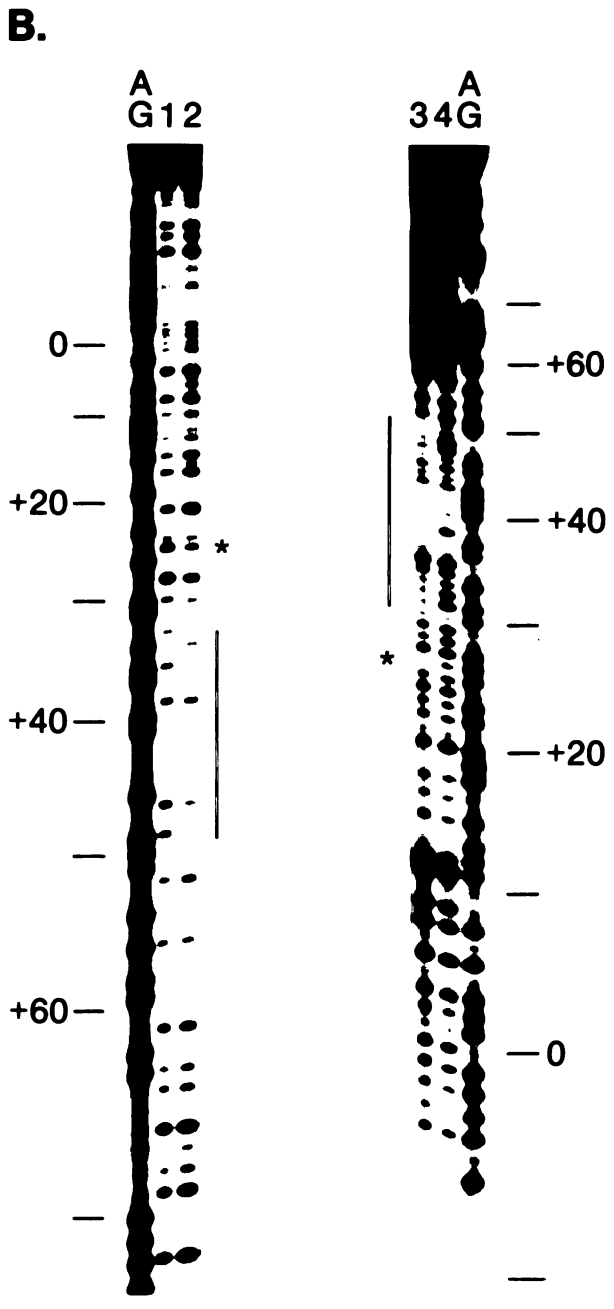


Figure 4C

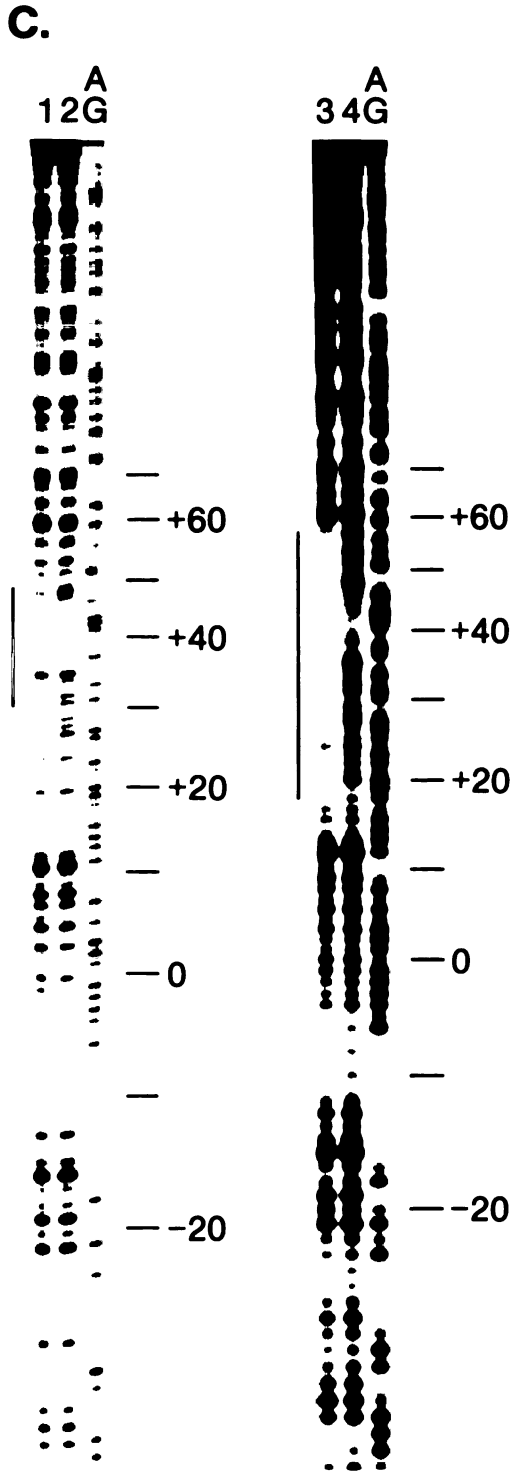


Figure 5. Summary of TnsD protection of attTn7 nucleotide sequences against nuclease attack.

The attTn7 sequence is shown, with boxed nucleotides indicating the 5 bp sequence duplicated upon Tn7 insertion (Lichtenstein and Brenner 1982; Gay et al. 1986; McKown et al. 1988). The central base pair of the target duplication is designated "0"; sequences extending leftward are given negative values, whereas sequences extending rightward are given positive values. Solid bars above and below the sequence show the extent of protection on the top and bottom strands, respectively, in the presence of the indicated Tns proteins. The bars extend to the outermost positions that are inaccessible to DNase I, and arrows designate accessible positions within protected regions (arrows marked with X indicate enhanced cleavage positions). Open circles mark the positions nearest the protected regions that are accessible.

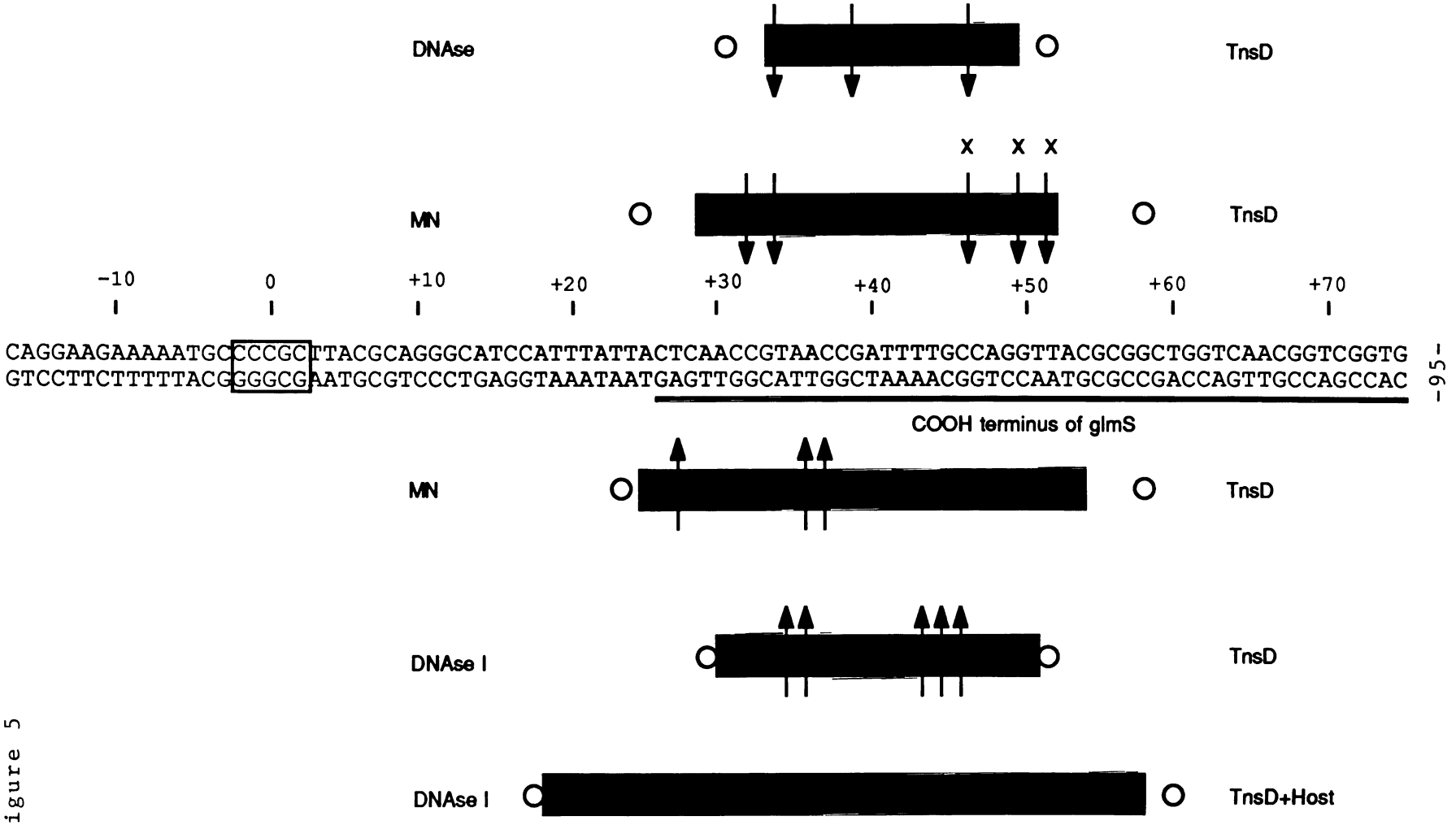

DNAse
TnsD
X X X
MN
TnsD
-10
0
+10
+20
+30
+40
+50
+60
+70
CAGGAAGAAAAATGCCCCGCTTACGCAGGGCATCCATTTATTACTCAACCGTAACCGATTTTGCCAGGTTACGCGGCTGGTCAACGGTCGGTG
GTCCTTCTTTTTACGGGGCGAATGCGTCCCTGAGGTAAATAATGAGTTGGCATTGGCTAAAACGGTCCAATGCGCCGACCAGTTGCCAGCCAC
COOH terminus of glmS
MN
TnsD
DNAse I
TnsD
DNAse I
TnsD+Host

Figure 5

Figure 6. TnsD is required for *in vitro* Tn7 transposition to *attTn7*.

A) Purified TnsD is active for *in vitro* transposition to *attTn7*. *In vitro* Tn7 transposition to *attTn7* was done in the presence (+) or absence (-) of TnsD protein. DNAs were detected by Southern hybridization with a miniTn7-specific probe, and transposition was evaluated by translocation of the miniTn7 element from a donor plasmid (indicated by "D") into an *attTn7* target plasmid to produce a simple insertion product (indicated by "P").

B) Cosedimentation of *tnsD*-dependent transposition and *TnsD*-*attTn7* binding. TnsD (fraction V) was fractionated on a glycerol gradient, and each fraction was evaluated for *tnsD*-dependent transposition activity and TnsD binding to *attTn7*. Transposition activity is shown as the ratio of insertion product to the sum of the insertion product and remaining donor plasmid. *attTn7* binding is shown as the ratio of bound complex to total *attTn7* DNA.

Figure 6A

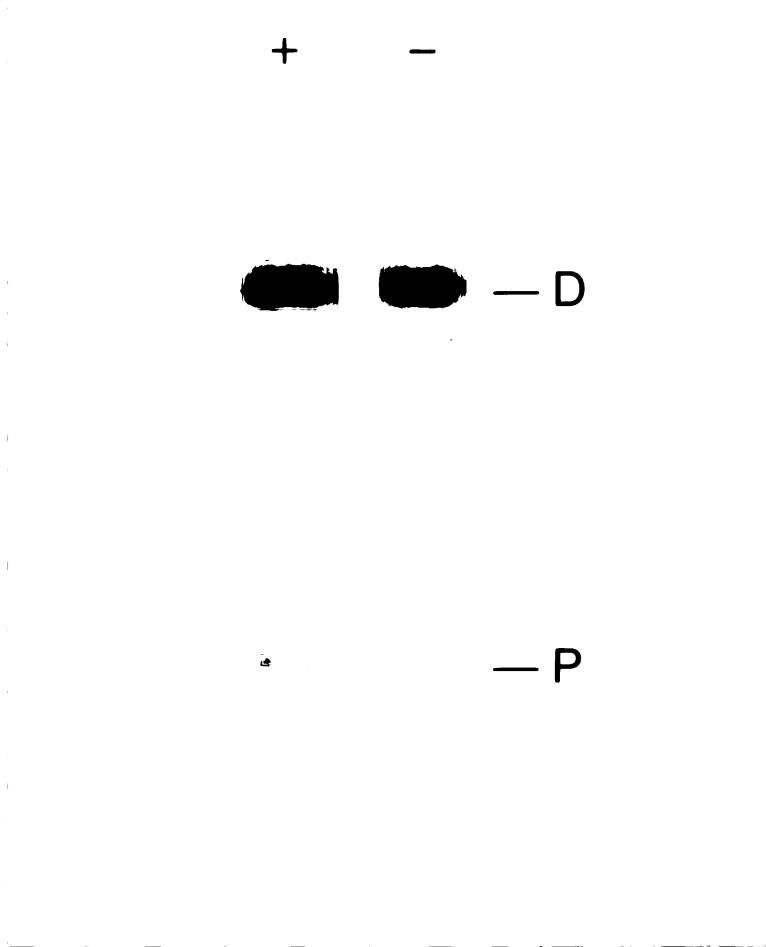
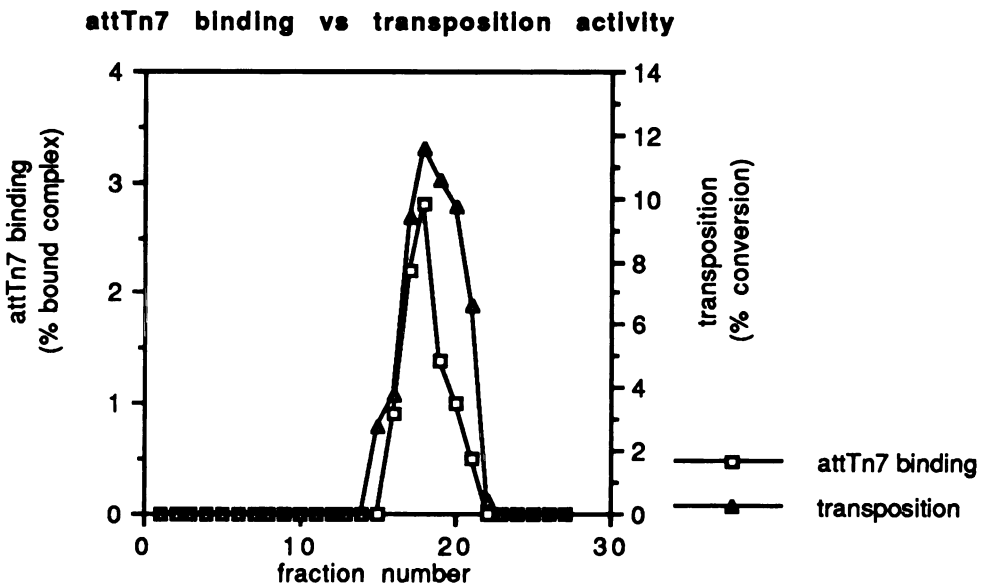


Figure 6B


attTn7 binding vs transposition activity
attTn7 binding
(% bound complex)
transposition
(% conversion)
fraction number
attTn7 binding
transposition
0
1
2
3
4
0
2
4
6
8
10
12
14
0
10
20
30

Chapter 4

Tn7 Transposition: Target DNA recognition by multiple Tn7-encoded proteins

This work was done in collaboration with Roland Bainton. The mobility shift and footprinting studies were done by me, and we both performed the *in vitro* transposition reactions using a linear attTn7 target (shown in Figure 4B). The remaining *in vitro* transposition experiments and the purification of TnsA were done by Roland Bainton.

ABSTRACT

We have reconstituted the transposition of the bacterial transposon Tn7 into its specific insertion site attTn7 with purified proteins. attTn7 insertion in vitro requires four Tn7-encoded transposition proteins, TnsA, TnsB, TnsC, and TnsD, and ATP. Insertion without target site selectivity is observed with TnsA + TnsB + TnsC and AMP-PNP; thus, these proteins mediate the donor cleavage and strand transfer stages of recombination. We provide evidence that TnsC, an ATP-dependent non-sequence-specific DNA binding protein, can associate with the target DNA and likely plays an important role in target selection. During transposition to attTn7, TnsC associates with the target DNA in the presence of TnsD, an attTn7-specific DNA binding protein. TnsC also appears to be required for Tn7 to insert selectively into target DNAs which do not already contain Tn7, i.e are not immune. We suggest that TnsC plays a central role in communication between the target DNA and the other components of the transposition machinery.

INTRODUCTION

Transposition is a DNA rearrangement that involves three distinct DNA segments: two transposon ends and an insertion site in the target DNA (reviewed in Berg and Howe 1989). Critical steps in transposition include the choice of an insertion site and the DNA strand breakage and joining reactions at the transposon ends which separate the transposon from the donor DNA and join it to the target DNA.

We are interested in understanding the mechanism of transposition of the bacterial transposon Tn7 which displays unusual target site selectivity (Barth et al. 1976; Craig 1989). Tn7 is distinguished by its ability to insert at high frequency into a specific site called attTn7 in the chromosomes of Escherichia coli and many other bacteria (Barth et al. 1976; Lichtenstein and Brenner 1982; Craig 1989). As is true in all other transposition reactions, no DNA sequence homology between the Tn7 ends and attTn7 is involved in the choice of attTn7 as an insertion site (Gay et al. 1986; McKown et al. 1988). When attTn7 is unavailable, Tn7 resembles most other transposable elements, inserting at low frequency into many different sites with little obvious sequence selectivity (Rogers et al. 1986; Waddell and Craig 1988; Kubo and Craig 1990). The frequency of Tn7 insertion into a particular target DNA molecule is also highly influenced by whether the target replicon already contains a copy of Tn7. If the target already contains Tn7, the frequency of insertion of a second Tn7 into that DNA is greatly reduced (Hauer and Shapiro 1987; Arciszewska et al. 1989). Such "target immunity" dictated by the presence of another copy of the transposon in the target DNA is also

displayed by the bacterial transposons bacteriophage Mu and the Tn3 class of mobile elements (Lee et al. 1983; Darzins et al. 1983; Adzuma and Mizuuchi 1988).

Tn7 transposition to its different classes of target sites is mediated by two distinct but overlapping sets of Tn7-encoded transposition genes (Rogers et al. 1986; Waddell and Craig 1988; Kubo and Craig 1990). tnsABC + tnsD promote "target site-specific" insertions into sites that are related in nucleotide sequence, i.e. into attTn7 at high frequency and into pseudo-attTn7 sites at low frequency. In Tn7's "random target site" pathway, tnsABC + tnsE promote low frequency insertion into many different sites that are distinct from attTn7 and not apparently related to each other. Thus, tnsABC provide functions common to all types of Tn7 transposition, and tnsD and tnsE are alternative target site-determining functions. Tn7 transposition immunity is active in both the tnsD and tnsE-dependent pathways (Arciszewska et al. 1989).

Dissection of the mechanism and control of Tn7 transposition requires a biochemical approach. We have extended our earlier biochemical analysis of Tn7 insertion into attTn7 via the tnsABC + tnsD pathway (Bainton et al. 1991) by reconstituting this reaction in vitro with purified TnsA, TnsB, TnsC, and TnsD, which indicates that these Tns proteins participate directly in recombination. What roles do the Tns proteins play in recombination? Analysis of recombination reactions using the alternative cofactor AMP-PNP instead of the usual cofactor ATP reveals that TnsA + TnsB + TnsC can perform the DNA strand breakage and joining reactions in the absence of TnsD. Other experiments demonstrate

that TnsD binds specifically to attTn7 and thus plays a critical role in the identification of attTn7 as a specific and efficient site for Tn7 insertion (Chapter 3). We show here that TnsC, an ATP-dependent non-sequence-specific DNA binding protein (Gamas and Craig, in press), also interacts with the target DNA, likely through binding to TnsD. We also suggest that TnsC plays a key role in evaluating whether a target DNA is immune, i.e. already contains a copy of Tn7. Our results suggest that TnsC plays an important role in target selection and mediates communication between the transposon ends and the target site.

RESULTS

Reconstitution of Tn7 transposition in vitro with purified Tns proteins and ATP

We have previously reported (Bainton et al. 1991) the development of a cell-free Tn7 transposition system using crude extracts that reproduces many features of Tn7 transposition in vivo to attTn7. We established that a miniTn7 element can transpose from a donor plasmid into attTn7 in a target plasmid upon incubation of these substrates in crude extracts derived from cells containing individual tns genes (tnsA, tnsB, tnsC and tnsD) and ATP. We demonstrated that transposition proceeds via a non-replicative mechanism in which Tn7 is excised from the donor and specifically inserted into attTn7 (Figure 1). We have now reconstituted Tn7 insertion into attTn7 using four highly purified protein fractions containing TnsA, TnsB, TnsC, or TnsD.

We report here the purification of TnsA, as described in detail in the Experimental Procedures. We constructed a fusion protein in which

TnsA is fused to the carboxy-terminus of glutathione-S-transferase (GST) (Figure 2A, lane 1), and purified the GST-TnsA fusion protein (54 kd) by binding it to glutathione agarose beads and releasing TnsA (29 kd) from the GST domain by cleavage with thrombin (lane 2). Removal of the GST domain by affinity chromatography results in highly purified TnsA protein whose only obvious contaminant is residual GST domain (lane 3). The transposition activity of purified TnsA protein when added to crude extract lacking TnsA is comparable to authentic crude TnsA protein (R. Bainton and N.L.C., unpublished data). We have previously described the purifications of TnsB, TnsC and TnsD (Arciszewska et al. 1991; Gamas and Craig, in preparation; Chapter 3).

Transposition in vitro is evaluated by extraction of DNA from the incubation mixtures, digestion with restriction enzymes, separation by gel electrophoresis and hybridization with specific probes. We find that Tn7 transposes efficiently in vitro when the substrate miniTn7 donor and attTn7 target plasmids (Figure 1) are incubated at 30°C in reaction mixtures with highly purified Tns protein fractions, ATP, and magnesium acetate (MgAc). Using a transposon-specific probe, we observe that more than 80% of the donor substrate can be converted to simple insertion product (Figure 2B); the other product of transposition, the donor backbone from which the Tn7 element is excised, is also present in these reactions, but is not detectable by this hybridization probe.

Also detectable with the reconstituted system are several DNA species indicated in the crude extract studies to be intermediates in Tn7 transposition: donor molecules broken by a single double-strand break between the backbone and either transposon end (DSB.L = double-

strand break.left end or DSB.R=double-strand break.right end) and excised linear transposons (ELT=excised linear transposon) cut away from the donor backbone by two double-strand breaks (see Figure 1). Time courses demonstrate that the DSBs in the reconstituted system display the kinetic properties of transposition intermediates (data not shown).

We have also observed in the reconstituted system a DNA species in which the exposed left transposon ends of DSB.Ls are apparently joined to a single DNA strand of attTn7 in the target molecule (see Figure 1); we call these species DSB-SEJ.L (=single end join of left end of DSB.L). The identity of this species has been determined by digestion with a variety of restriction enzymes and hybridization to specific probes (data not shown). Using different restriction enzymes, these experiments have shown that the DSB-SEJ.R species (=single end join of the right end of a DSB.R) is detectable in amounts equivalent to that of the DSB-SEJ.Ls. Also, ELTs which have joined one transposon end to one strand of attTn7 in the target DNA are observed (ELT-SEJ.L and ELT-SEJ.R), albeit at very low levels (see Figure 1). All of these SEJ species appear late in the reaction and accumulate in reaction time courses (data not shown), suggesting that they are not key transposition intermediates. Although these other species are present, we emphasize that the simple insertion molecules are by far the predominant transposition product.

We find that the presence of all four Tns proteins is required for the generation of the simple insertion transposition product (Figure 2B, lanes 2-5). It is notable that no transposition intermediates of any sort are detectable when TnsA, TnsB, or TnsC is omitted from the

incubation mixture, indicating that these proteins are critical to the initiation of recombination, i.e. the double strand breaks which separate the transposon from the donor backbone. However, very low levels of the transposition intermediates (DSBs and ELTs) are observed when TnsD is omitted from the reaction, i.e. with only TnsA, TnsB and TnsC present (Figure 2B, lane 5). (The activities of TnsA + TnsB + TnsC are considered below.) Also notable is that the presence of attTn7 is required for the efficient production of transposition intermediates or products: although donor cleavage apparently precedes joining of the transposon to the target DNA during transposition, no donor cleavage occurs in the absence of attTn7 (Figure 2B, lane 7).

As transposition does occur with highly purified Tns proteins, Tn7 transposition in vitro does not obviously require any E. coli proteins; however, it should be noted that the addition of crude extracts from cells lacking the tns genes can stimulate transposition, particularly under sub-optimal conditions (data not shown). Thus, although not apparently essential, host proteins may be able to act as accessory proteins in Tn7 transposition.

We find that ATP is an essential cofactor for Tn7 transposition in vitro in the reconstituted system with purified Tns proteins; no recombination intermediates or products are observed when ADP is substituted for ATP (Figure 2B, lane 6). (We cannot completely omit ATP from the reaction mixture as it is present in the TnsC fraction to increase TnsC's solubility.) We have established elsewhere that TnsC is an ATP binding protein that binds to DNA in an ATP-dependent fashion and thus imagine that TnsC is the likely site of ATP action during Tn7

transposition (Gamas and Craig, in preparation).

These experiments provide evidence that the four Tns proteins participate directly in transposition and also establish that ATP is an essential cofactor for this reaction.

TnsC and TnsD can interact with target DNA: Efficient Tn7 transposition *in vitro* requires preincubation of attTn7, TnsC, TnsD, and ATP

Our reaction mixtures include MgAc; no recombination is observed when MgAc is omitted (Figure 3A, lane 1). Recombination is most efficient, however, when all the reaction components except MgAc are mixed, incubated at 30°C, and then incubated further at 30°C after the addition of MgAc (Figure 3A, lane 3; also see Figure 3B). Little recombination occurs if MgAc is present for the entire incubation period (Figure 3A, lane 2).

Not all of the reaction components are required to be present in the "preincubation" step that lacks MgAc. The presence of attTn7, TnsC, TnsD and ATP is required in the preincubation step for high-level recombination; if any one of these components is omitted, little recombination is observed (Figure 3A, lanes 4-7). Efficient recombination is still observed if TnsA, TnsB and the donor DNA are added after the precincubation step, i.e. when MgAc is added (lane 8). The requirement for the simultaneous presence of TnsC, TnsD, attTn7, and ATP in the preincubation step suggests that the intial interaction of these components is stimulated by the low MgAc conditions.

Formation of protein-DNA complexes involving TnsC and TnsD on attTn7

The observation that recombination is stimulated by the preincubation of TnsD with attTn7 is not unanticipated; we have shown elsewhere that TnsD is a sequence-specific DNA binding protein that recognizes attTn7 (Chapter 3). Other studies have also revealed that TnsC is an ATP-dependent non-sequence-specific DNA binding protein. We have used gel mobility shift assays to establish that in the presence of ATP, TnsC and TnsD can form a distinctive novel complex on attTn7.

Specific binding of TnsD to a labeled attTn7 fragment is observed by the formation of a complex of retarded electrophoretic mobility (Figure 4A, lane 3). No binding of TnsC alone to the labeled attTn7 fragment is observed because of the presence of a large excess of unlabeled competitor DNA in these assays (lane 2). However, when TnsC and ATP are mixed with TnsD and the attTn7 fragment, a novel protein-DNA complex is observed (lane 4). Formation of this novel complex requires ATP, and no "supershifted" complex is observed in presence of ADP (lanes 6 and 7). (Reaction mixtures to which ATP is not added do not entirely lack ATP because of the presence of ATP in the TnsC fraction.) Formation of the ATP-dependent TnsC-TnsD-attTn7 complex is specific to TnsC; no such complex is observed when TnsA or TnsB are substituted for TnsC (data not shown). Formation of the TnsC-TnsD-attTn7 complex is also dependent on attTn7, as no protein-DNA complexes are observed with a DNA fragment lacking attTn7 sequences (lane 9).

The observation of an ATP-dependent TnsC-TnsD-attTn7 complex by mobility shift assay suggests that a similar complex may be formed during the stimulatory preincubation step of transposition that lacks

MgAc and also requires the presence of TnsC, TnsD, attTn7 and ATP. This view is supported by the finding that formation of the TnsC-TnsD-attTn7 complex is blocked by the presence of MgAc in the reaction mixture (Figure 4A, lane 5), a condition that also blocks transposition (Figure 3A, lane 2).

The ATP-dependent TnsC-TnsD-attTn7 complexes are formed under conditions somewhat different from those of the normal preincubation step of transposition; for example, in the standard transposition reaction attTn7 is part of a supercoiled plasmid whereas the attTn7 used in analyzing complex formation is a short linear DNA fragment; also, the ionic conditions of the incubations are different. We have therefore repeated the transposition reactions under conditions similar to those used for the gel mobility shift assays. We find that efficient insertion of Tn7 into attTn7 DNA fragments is observed when TnsA, TnsB, donor DNA substrate, and MgAc are mixed with ATP-dependent TnsC-TnsD complexes formed on the attTn7 fragments (Figure 4B, lane 1). No Tn7 insertions are observed when only TnsD-attTn7 complexes are formed (lane 2), and transposition into these small linear targets also requires a preincubation step without MgAc (lane 3). These observations support the hypothesis that a key feature of the stimulatory preincubation step in Tn7 transposition in vitro is the ATP-dependent formation of a TnsC-TnsD complex on attTn7.

We have explored the stability of the TnsC-TnsD complexes formed on attTn7 during the preincubation step by challenging the reactions mixtures with competitor attTn7 plasmid (at 10-fold molar excess compared to linear attTn7 target) prior to or after the precincubation

step. The protein complex formed on the target is apparently stable as no reduction in Tn7 insertion is observed by the addition of competitor attTn7 plasmid after the preincubation step, whereas competitor attTn7 DNA added from the beginning of the reaction successfully competes with linear attTn7 target DNA (Figure 4B, compare lanes 4 and 5). This finding is consistent with the view that the complexes formed during the preincubation step, which are detectable as TnsC-TnsD complexes by mobility shift assays, are the targets for Tn7 insertion.

Footprinting analysis of TnsC-TnsD complexes on attTn7

We have also characterized the TnsC-TnsD-attTn7 complex by footprinting analysis. We formed complexes with end-labeled attTn7 DNA, treated these mixtures briefly with DNase I, isolated complexes after polyacrylamide gel electrophoresis, and then analyzed the recovered DNA on a DNA sequencing gel. The results of this analysis are shown in Figure 5 and summarized in Figure 6.

In the TnsC-TnsD complex, we find that a region approximately 50 bp in length, extending rightwards from about 10 nt to the right of the specific point of Tn7 insertion (designated +10) to about 60 nt to the right of the Tn7 insertion point (+60), is protected from DNase I attack. This pattern is notably different from that observed with TnsD alone (Figure 5, also see Chapter 3), the TnsC + TnsD pattern being both more extensive and the degree of protection much greater. On the bottom strand of attTn7, protection by TnsD alone extends from about +25 to +50, whereas the TnsC + TnsD protection extends from about +10 to +60. On the top strand, although a striking protection is observed with TnsC

+ TnsD, only a modest and limited protection is seen with TnsD alone. It remains to be established which positions in attTn7 may be most directly contacted by TnsC and TnsD; it is not possible to discern the actual dispositions of the proteins from such protection experiments. A simple view however, is that TnsC occupies the region of attTn7 DNA that extends from the observed edge of the binding region for TnsD alone (about +30) to the border of the TnsC + TnsD protection that is closest to the point of insertion (about +10). An important feature of the TnsC + TnsD protection is that the specific point of Tn7 insertion is accessible to the attack reagent and thus does not appear to be directly contacted by either TnsC or TnsD.

TnsA + TnsB + TnsC can promote transposition in the presence of AMP-PNP

A notable feature of Tn7 transposition in vitro using either the crude system or the reconstituted system is the highly coupled nature of the reaction, i.e. generation of the simple insertion transposition product requires the presence of all four Tns proteins and ATP (Figure 7A, lane 2). Likely because of the high efficiency of the reconstituted system, we observe in these reactions low levels of the transposition intermediates, DSBs and ELTs, when TnsD or attTn7 are omitted from the reaction (Figure 7A, lanes 1 and 3, respectively). This observation reveals that the active site for strand cleavage of the donor molecule lies within TnsA + TnsB + TnsC. In contrast to the lack of detectable insertion product in TnsA + TnsB + TnsC reactions that use ATP as a cofactor, simple insertion transposition products and high levels of recombination intermediates are generated when AMP-PNP is substituted

for ATP, even in the absence of TnsD (Figure 7A, compare lanes 4 and 6). Determination of the DNA sequence for several of these AMP-PNP insertions after transformation reveals that they, like products obtained with ATP, generate 5 bp target sequence duplications (data not shown), as observed for Tn7 insertions in vivo (Lichtenstein and Brenner 1982; Chapter 2). We suspect that TnsC, an ATP-dependent DNA binding protein, is the site of action for the nucleotide cofactor; the location of the active site(s) for DNA strand cleavage and transfer is not known.

The simple insertions produced with TnsA + TnsB + TnsC in the presence of AMP-PNP display little target site-selectivity, since efficient insertion into target molecules that lack attTn7 is observed (Figure 7A, lane 5). As expected when TnsD is absent, there is no apparent preference for attTn7 insertion even when this site is present in a target molecule (Figure 7A, lane 4). Insertions generated by TnsA + TnsB + TnsC appear to be randomly distributed in the target molecule (Figure 7B). Another indication of the apparently random distribution of insertions is that the quantity of insertion product generated is proportional to the size of the target DNA segment examined (data not shown).

Tn7's ability to insert specifically and selectively into attTn7 is not entirely abolished when AMP-PNP is used as a cofactor instead of ATP. In TnsA + TnsB + TnsC + TnsD reactions performed in the presence of AMP-PNP, specific insertions into attTn7 and random insertions into the target backbone are both observed (Figure 7A, lane 6). AMP-PNP is also an effective cofactor in promoting the formation of TnsC-TnsD attTn7 complexes as evaluated in mobility shift assays (data not shown). These

results suggest that although AMP-PNP can dramatically change the insertion specificity, it does not abolish the ability of TnsD to direct TnsC to attTn7.

A prominent feature of the standard Tn7 transposition reactions performed with ATP is that recombination is most efficient when TnsC, TnsD, attTn7, and ATP are preincubated in the absence of MgAc and the other reaction components to facilitate the formation of the TnsC-TnsD complex on the target DNA. In TnsA + TnsB + TnsC reactions with AMP-PNP, recombination does proceed more efficiently when TnsC, the target DNA, and nucleotide are precincubated prior to the addition of MgAc and the other reaction components (data not shown). Although there is a much greater degree of stimulation by preincubation in the ATP reaction compared to the AMP-PNP reaction, the fact that recombination is increased in both cases argues that interaction of TnsC with the target is a critical step in recombination.

Lack of target immunity in Tn7 transposition in vitro with AMP-PNP.

Tn7 displays transposition immunity in vivo; that is, DNA molecules that already contain Tn7 are poor targets for transposition. The presence in the target DNA of the sequences from the right end of Tn7 that participate directly in recombination are sufficient to provide immunity (Arciszewska et al. 1989). We previously established that Tn7 immunity is active in vitro in the crude extract system (Bainton et al. 1991). We have found that Tn7 transposition immunity is active in the reconstituted TnsA + TnsB+ TnsC + TnsD system when ATP is used as a

cofactor (Figure 8). In these reactions with ATP, which include both a standard attTn7 target plasmid and an immune attTn7 target plasmid, i.e. one containing both attTn7 and a Tn7R end segment, no insertions into the immune target are observed although the standard attTn7 plasmid remains an effective target (lanes 1 and 2). We find, however, that transposition immunity is virtually abolished when AMP-PNP is used as a cofactor instead of ATP (lanes 3 and 4). In the presence of AMP-PNP, insertions into both the non-immune and immune target molecules are readily detected. A similar result is also seen in the absence of TnsD, although insertions into attTn7, as expected, are significantly reduced (lanes 7 and 8).

Thus, substitution of AMP-PNP for ATP dramatically changes the Tn7 transposition machinery's sensing of the target DNA: 1) a target DNA is no longer evaluated for the presence or absence of an existing copy of Tn7, i.e. target immunity is no longer operative, and 2) the target sequence selectivity is reduced, i.e. attTn7 is no longer the exclusive site of transposon insertion. As TnsC is an ATP binding protein which associates with DNA in an ATP-dependent manner (Gamas and Craig, in press), it seems reasonable to suggest that differences in the interaction of TnsC with these alternative adenine cofactors and the resultingly different interactions of TnsC with the target DNA underlie the observed differences in recombination. These observations further support the view that the interaction of TnsC with the target DNA is a key step in Tn7 transposition.

DISCUSSION

Tn7 transposition is executed by four Tn7-encoded proteins.

We have extended our biochemical dissection of Tn7 transposition to its preferred insertion site attTn7 through in vitro reconstitution of this reaction. We have established that Tn7 inserion into attTn7 in vitro can occur efficiently in the presence of and requires four purified Tn7-encoded transposition proteins, TnsA, TnsB, TnsC, and TnsD, and ATP. These findings have revealed that each of these Tns proteins are likely to participate directly in Tn7 transposition, accounting for the genetic observation that the tns genes encoding these proteins are essential to Tn7 transposition in vivo (Rogers et al. 1986; Craig and Waddell 1988). The observation that ATP is an essential cofactor in the reconstituted system reveals that this molecule also has a direct role in recombination.

The ability of the four purified Tns proteins to promote Tn7 transposition in vitro suggests that these proteins contain both the specificity determinants for transposition, i.e. the ability to specifically recognize the transposon ends and attTn7, and the active site(s) that mediates strand exchange, i.e. the cleavage reactions which separate the transposon from the donor DNA and the strand transfer reactions which join the transposon ends to the target DNA. The interactions of some individual Tns proteins with the substrate DNAs has suggested their activities in recombination: TnsB is a sequence-specific DNA binding protein that recognizes the ends of Tn7 (Arciszewska et al. 1991; Tang et al. 1991), and TnsD is a sequence-specific DNA binding

protein that recognizes attTn7 (Chapter 3). Because TnsC is an ATP-dependent non-sequence-specific DNA binding protein (Gamas and Craig, in press), a reasonable hypothesis is that it is the site of action of the essential ATP cofactor in Tn7 transposition. As discussed below, we have provided evidence that TnsC can interact with the target DNA and with TnsD, and propose that TnsC mediates communication between the target DNA and the transposon ends. The role of TnsA protein in Tn7 transposition has not yet been explicitly identified.

The roles of E. coli-encoded proteins in Tn7 transposition remain to be determined. Recombination in the reconstituted reaction occurs in the absence of any obvious host protein, suggesting that a host protein is not essential to this reaction. However, it is important to note that we cannot exclude the possibility that a contaminating host protein present at low levels in our Tns fractions is required for Tn7 recombination. Also, the possibility that a host protein(s) may participate in recombination in an accessory role should not be dismissed. Indeed, we observe that transposition can be stimulated under some conditions by the addition of crude host extract to the purified Tns proteins; also, our optimization of recombination in the reconstituted system may have bypassed an important host contribution. The observation that mutations in some E. coli genes do alter Tn7 transposition in vivo (O. Hughes and N.L.C., unpublished data) also supports the view that host proteins contribute to this reaction. Many reactions involving extrachromosomal DNA elements are mediated by collaborations between element and host-encoded proteins (Drlica and Rouviere-Yaniv 1987; Friedman 1988; Glasgow et al. 1989; Landy 1989).

TnsC can interact with the target DNA

Characterization of the reconstituted Tn7 transposition system has revealed that efficient recombination to apparently random target sites occurs with TnsA + TnsB + TnsC and AMP-PNP. Also, cleavage of the transposon from the donor backbone is observed, albeit at very low levels, in the presence of TnsA + TnsB + TnsC and ATP. These observations demonstrate that the active site(s) for donor cleavage and strand transfer are contained within TnsA + TnsB + TnsC. Because TnsC can bind ATP, AMP-PNP is likely influencing the activity of TnsC in its reaction with TnsA + TnsB + TnsC. Moreover, because TnsC is an ATP-dependent DNA binding protein whose apparent affinity for DNA is increased in the presence of non-hydrolyzable ATP analogues such as AMP-PNP (Gamas and Craig, in press), it is reasonable to suggest that AMP-PMP promotes recombination under conditions where ATP does not, i.e. in the absence of TnsD, by promoting a more stable interaction of TnsC with DNA. The finding that insertion into a target DNA is most efficient when that DNA is preincubated with TnsC and AMP-PNP in the absence of the other reaction components (the donor DNA, TnsA, and TnsB) suggests that TnsC interacts directly with the target DNA during recombination.

The hypothesis that the ATP-dependent interaction of TnsC with the target DNA plays a central role in Tn7 transposition is supported by the different effects of ATP and AMP-PNP on target immunity. The frequency of Tn7 insertion <u>in vivo</u> into a target DNA is much reduced by the presence of Tn7 or even just the right end of Tn7 in that DNA, i.e. the Tn7-containing target is immune (Arciszewska et al. 1989). In the

presence of ATP, Tn7 target immunity is active in the crude extract (Bainton et al. 1991) and reconstituted Tn7 transposition systems. By contrast, target immunity is abolished when AMP-PNP is used as the cofactor in transposition. Thus, the nature of the adenine cofactor has a profound impact on target immunity, i.e. the ability of the Tn7 transposition machinery to evaluate the target DNA for the presence of a Tn7 end. An attractive hypothesis is that TnsC mediates the recognition of immune and non-immune target DNAs through its ATP-dependent binding to DNA. We postulate that in the presence of AMP-PNP, TnsC binds stably to all target DNAs, including those containing Tn7 end sequences, and that when ATP is used as a cofactor, the binding of TnsC to immune targets is discouraged. In this view, immune DNA molecules, i.e. those that contain a Tn7 end, are poor substrates for tranposition because TnsC cannot effectively interact with these DNAs.

The picture that has emerged from these experiments is that TnsC can interact with the target DNA during transposition, Tn7 insertion being directed to DNAs on which TnsC is stably associated. An attractive view is that TnsC both contacts the target DNA and also communicates with the transposon ends, perhaps via protein-protein contacts with TnsB and/or TnsA, thereby interacting with the transposon ends.

Comparison of Tn7 TnsC protein and bacteriophage Mu transposition protein MuB

The view we have presented for the central role of the ATP-dependent interaction of TnsC with the target DNA is very reminiscent of the proposed role of MuB protein in bacteriophage Mu transposition. MuA protein, a Mu-end binding protein (Craigie et al. 1984), with ATP and MuB collaborate to promote Mu recombination (Surette et al. 1987; Craigie and Mizuuchi 1987). Like TnsC, MuB is a non-specific DNA binding protein (Chaconas et al. 1985) that requires ATP for stable association with DNA (Maxwell et al. 1987); upon ATP hydrolysis, MuB dissociates from DNA (Adzuma and Mizuuchi 1988, 1989). MuB can stimulate Mu transposition by interacting with both the target DNA and MuA bound to the transposon ends (Baker et al. 1991; Surette et al. 1991; Surette and Chaconas 1991).

MuB has been directly shown to play a key role in target immunity during Mu transposition (Adzuma and Mizuuchi 1988, 1989). MuB selectively dissociates from immune target DNAs through stimulation of its ATPase activity which is provoked by an interaction with MuA bound to the transposon ends on the immune target. Transposition immunity is abolished in the presence of the non-hydrolyzable ATP-γS which promotes stable association of MuB with both non-immune and immune DNAs, making them both effective targets for Mu insertion.

In spite of the apparently strong similarities between TnsC and MuB in their activities and roles in recombination, there is no obvious similarity in the amino acid sequences of these proteins beyond the

limited regions likely to be directly involved in ATP binding (Miller et al. 1984; Flores et al. 1990).

TnsC and TnsD can act on attTn7

Several observations provide evidence that during TnsD-dependent transposition, the interaction of TnsC with the attTn7 target DNA is promoted by TnsD. A reasonable hypothesis is that TnsC, an ATP-dependent non-sequence-specific DNA binding protein, is directed to attTn7 by protein-protein contact with TnsD, a protein which binds directly to attTn7.

We have observed that TnsC and TnsD can form specialized complexes with attTn7. We have detected these TnsC-TnsD attTn7 complexes in the presence of ATP by mobility shift assays. We have also found that recombination is stimulated by conditions that favor the formation of such a complex. Recombination in vitro is most efficient in the presence of high MgAc, but efficient recombination also requires the preincubation of TnsC, TnsD, attTn7, and ATP in the absence of MgAc. The binding of TnsC and TnsD to DNA is individually sensitive to MgAc, and the initial presence of MgAc blocks the formation of the TnsC-TnsD attTn7 complex (data not shown). We hypothesize that the preincubation step promotes the formation of a TnsC-TnsD attTn7 complex which is relatively resistant to MgAc. The subsequent presence of high MgAc perhaps stimulates some other step in recombination, for example, the DNA strand cleavage and strand transfer steps.

We propose that TnsC is directed to attTn7 by interaction with TnsD. Although TnsC alone can bind to DNA, no specific interaction of

TnsC with attTn7 is observed in the absence of TnsD. Furthermore, no specific sequences in attTn7 outside the region of attTn7 which can interact with TnsD alone appear to be required for attTn7 target activity (Chapter 3). Of course, we cannot exclude the possibility that TnsD induces a particular conformational change in the attTn7 DNA which is specifically recognized by TnsC or that there are additional specific TnsC contacts within this TnsD region. However, a reasonable view is that TnsD binds specifically to attTn7 and that TnsC also interacts with attTn7 through specific protein-protein contacts with TnsD and non-specific contacts with DNA. The ability of TnsC alone to bind DNA is influenced by ATP; it remains to be determined how this ATP-dependent interaction of TnsC with DNA is modulated by TnsD.

How might TnsC interact with the rest of the recombination machinery, i.e. the ends of Tn7, TnsA and TnsB? We have observed that in the ATP-dependent TnsC-TnsD attTn7 complexes, the region of the target DNA immediately surrounding actual point of Tn7 insertion is not contacted by either TnsC or TnsD; this region is accessible to attack by DNase I. The region of TnsC-TnsD protection is about 50 nt in length and extends rightwards from a position about 10 nt from the specific point of Tn7 insertion. Because the region of attTn7 that can be contacted by TnsD alone, which also contains the nucleotide sequence information required for attTn7 target activity (Gringauz et al. 1988; Qadri et al. 1990; Chapter 3), lies within the rightward portion of the region of TnsC-TnsD protection, we speculate that the region of the TnsC-TnsD protection that adjoins but does not directly contact the Tn7 point of insertion reflects the interaction of TnsC with the target DNA. We are

intrigued by the hypothesis that during transposition the region immediately surrounding the point of insertion is actually occupied by the transposon ends, TnsA, and TnsB. Given the spatial separation between the actual point of Tn7 insertion and apparent position of TnsC, we speculate that TnsC may act in transposition by protein-protein contact with TnsB and/or TnsA, thereby juxtaposing the transposon ends with the target DNA and promoting the DNA strand cleavages at the transposon ends and the strand transfer reactions which join the transposon to the target DNA.

Tn7 transposition occurs in an elaborate nucleoprotein complex

We have observed in both the crude extract cell-free Tn7 transposition system (Bainton et al. 1991) and the reconstituted system that Tn7 transposition is highly coupled: the presence of all the recombination proteins is required to produce significant amounts of recombination intermediates and products, and, furthermore, that the presence of attTn7 is required to provoke the first steps in transposition, i.e. cleavage of the donor DNA. Thus, we argue that Tn7 transposition actually occurs within an elaborate nucleoprotein complex that contains all the Tns proteins and the three DNA recombination substrates, i.e. the Tn7 ends and attTn7. Multiple protein-DNA and protein-protein interactions must underlie the assembly and activity of this elaborate nucleoprotein complex. It is not unreasonable to consider that the assembly of the final complex that actually executes strand exchange can occur via several alternative pathways. For example, our particular reaction conditions appear to favor the formation of a

distinct TnsC-TnsD complex on attTn7 as a critical step in recombination; under other conditions, for example, formation of a distinct complex containing the ends of Tn7 and TnsA + TnsB + TnsC may precede the interaction of TnsC with target DNA and/or TnsD.

We propose that in the recombination complex TnsC plays a key role in communication between the target DNA and the transposon ends. Several types of evidence have revealed that TnsC can interact with the target DNA: we have found that TnsC can form a special complex with TnsD on the attTn7 target, transposition appears to be directed towards DNA to which TnsC is bound, and also TnsC appears to play a central role in evaluating whether or not a target DNA is immune. We speculate that TnsC may also interact via protein-protein contacts with TnsA and/or TnsB to communicate with the transposon ends. We imagine that the initiation of recombination in this complex, i.e. the execution of the strand cleavages at the transposon termini which is followed by the strand transfer reactions which join the transposon ends to the target DNA, results from multiple coordinate interactions between the components of the complex.

A proposal for the mechanism of tnsE-dependent Tn7 transposition.

We have suggested that during TnsD-dependent transposition, the target site is specifically recognized by the binding of TnsD and that the interaction of TnsD with TnsC then stably positions TnsC on the target DNA, therby provoking transposition. What relationship might the mechanism of tnsE-dependent transposition, whose target sites are apparently random, have to the mechanism proposed for TnsD-dependent

transposition? An attractive hypothesis is that TnsE directs the basic TnsA + TnsB + TnsC machinery to promote Tn7 insertion into the target DNA; like TnsD-dependent transposition, tnsE-dependent transposition in vivo is a non-replicative reaction (K. Orle, M. Lopata, and N.L.C., unpublished data). It is not unreasonable to suggest that the overall mechanisms of TnsD- and TnsE- dependent transposition reactions are similar, being most distinguished in how target sites are recognized. In this view, we would expect that during TnsE-dependent transposition, TnsC would also play a key role in interacting with the target DNA. One possibility is that TnsE, like TnsD, is a DNA binding protein, but one that lacks the sequence preference of TnsD. TnsE might then, like TnsD, interact simultaneously with TnsC and DNA, thereby fixing TnsC to the target DNA at many different positions and provoking transposition. Another alternative is that TnsE does not interact with DNA but rather interacts only with TnsC as a sort of allosteric effector, perhaps modulating TnsC's interaction with ATP, to promote TnsC's stable interaction with DNA and thereby provoke transposition.

These models suggest that two alternative types of nucleoprotein complexes, one with TnsD and one with TnsE, may form and mediate Tn7 transposition. Both types of complexes would share the same basic protein machinery, TnsA + TnsB + TnsC and ATP, but a TnsD nucleoprotein complex would direct insertions into attTn7, whereas the TnsE-dependent nucleoprotein complex would promote insertions into random target sites. This combinatorial assembly of nucleoprotein complexes to provide regulated and differential activities is analogous to other processes in recombination, replication, transcription, and translation (Echols 1990;

McCarthy and Gualerzi 1990; Hershey 1991). The study of Tn7 transposition *in vitro* provides a defined and readily manipulable system that can be used to dissect the macromolecular interactions that underlie such elaborate protein-nucleic acid transactions.

EXPERIMENTAL PROCEDURES

DNA substrates

The donor plasmid pEM-1 contains an E. coli chromosomal insertion of a miniTn7 element flanked by sequences unrelated to attTn7 (Arciszewska et al. 1989). The donor plasmid pEMΔ contains the miniTn7 element and nearby flanking sequences (51 nt and 83 nt adjacent to the Tn7L and Tn7R ends, respectively) of pEM-1. The miniTn7 element and flanking donor DNAs were generated by PCR amplification using pEM-1 as a template with oligonucleotides complementary to the flanking donor sequences positioned at 51 nt to 30 nt from the Tn7L end and 83 nt to 64 nt from the Tn7R end. The amplified DNA was treated with Klenow enzyme and introduced into the ScaI site of pTRC99 (Pharmacia). The attTn7 target plasmid pKAO4-3 (McKown et al. 1988) contains a 150 bp segment that includes the sequences necessary for attTn7 target activity. pLA11 is an attTn7 plasmid which also contains the Tn7R199 segment (Bainton et al. 1991). pKS+ (Stratagene) is used as a target DNA lacking attTn7 sequences.

Linear DNA substrates from indicated plasmids were prepared and end-labeled with [α-32P] dATP and Klenow enzyme as described in Chapter 3. For mobility shift assays, the 115 bp EcoRI-BamHI pKK20 DNA fragment contained attTn7 (-25 to +64) sequences, whereas the 122 bp PvuI-HindIII pUC18 fragment contained sequences unrelated to attTn7. For DNase I protection studies, the pKK24 HindIII-PvuI fragment (235 bp), labeled at the HindIII end, was used to examine the bottom strand of attTn7. pKK24 contains the BamHI-EcoRI attTn7 (-25 to +64) fragment from pKOA4-3

inserted between the BamHI and EcoRI sites of pUC19. The EcoRI-PvuI pKK25 fragment (260 bp), labeled at the EcoRI end, contained attTn7 (-52 to +64) and was used to analyze the bottom strand. This DNA fragment was also used as linear attTn7 target for Tn7 transposition in vitro.

Preparation of Tns proteins

Protein preparations were as described: TnsB was fraction IV (Arciszewska et al. 1991), TnsC was fraction III (Gamas and Craig, in preparation) and TnsD was fraction V (Chapter 3).

TnsA was expressed as a glutathione-S-transferase (GST) fusion protein in Escherichia coli strain DH5α (BRL). The gene construct encoding the GST-TnsA fusion was generated by inserting a NcoI-HindIII tnsA fragment from pKA046 (Orle and Craig 1991), after partial digestion with HindIII, between the NcoI and HindIII sites of pGEX2 (Smith and Johnson 1988). Fusion protein expression and purification were done as described by Smith and Johnson (1988).

Cells used for purification were grown at 30°C in 1 L LB broth (Miller 1972) with 100 μg/ml carbenicillin. At OD_{600} = 0.7, fusion protein expression was induced by the addition of 40 μM IPTG (isopropyl-B-D-thiogalactopyranoside) and then allowed to grow until OD_{600} = 1.2-1.4. Cells were harvested by centrifugation, washed in buffer A (25 mM HEPES pH 7.5, 1 mM EDTA), and pelleted.

Subsequent steps, unless otherwise noted, were done at 4°C. 2 ml/g cells of buffer A with 100 mM KCl, 2 mM DTT, and 1 mM phenylmethyl-sulfonyl fluoride (PMSF) was added to cell pellet. Cells were lysed by sonication, cell debris was removed by centrifugation for 30 min at

16,800 x g, and supernatant was collected. To the crude lysate, 5 M NaCl and 250 mM CHAPS (a nonionic detergent) were added to a final concentration of 500 mM and 10 mM, respectively, and the lysate was incubated with glutathione agarose beads (Sigma) for 1 hr, followed by extensive washing in buffer A with 500 mM NaCl, 10 mM CHAPS, and 2 mM DTT. TnsA fusion protein was eluted from beads with 10 mM glutathione, and the GST carrier was clipped off by incubation for 30 min at 25°C with 30 units human thrombin (Sigma) per mg TnsA fusion protein. Reaction was stopped by addition of 100 mM PMSF (in isopropanol) to final concentration of 0.25 mM. Cleavage of GST-TnsA with thrombin results in a wild-type TnsA polypeptide with three additional amino acids (Gly-Ser-Pro) at its N-terminus. Cleavage products were incubated for 6 hr with glutathione agarose beads, and the supernatant, selectively enriched in TnsA polypeptide, was collected. These fractions, containing approximately 100 μg/ml TnsA, were stored at -80°C and were stable for at least several months.

SDS-PAGE using 12.5% gels was carried out by the method of Laemmli (1970). Protein concentration was determined by Bio-Rad protein assay or SDS-PAGE with staining by Coomassie Blue R-250, using BSA as a standard.

Tn7 transposition *in vitro*

Reactions were performed essentially as described in Bainton et al. (1991). Reaction mixtures (100 μl) contained 0.1 μg (0.25 nM) of the donor plasmid (pEM-1 or pEMΔ), 0.3 μg (2.5 nM) of pKAO4-3 *attTn7* target plasmid, 2 mM ATP, 26 mM HEPES pH 8.0, 0.02 mM EDTA, 2.1 mM DTT, 15 mM magnesium acetate (MgAc), 0.1 mM $MgCl_2$, 0.01 mM $CaCl_2$, 100 μg/ml tRNA,

50 μg/ml BSA, 0.1 mM CHAPS, 50 ng TnsA, 50 ng TnsB, 50 ng TnsC, and 40 ng TnsD. Reactions in Figures 2B, 7, and 8 also contained 1.0 mM Tris-HCl (pH 8.0), 5.8 mM NaCl, 15 mM KCl, and 0.25% glycerol. The reactions in Fig. 8 contained a lower concentration (1 nM each) of target plasmids pKAO4-3 and pLA11. A preincubation step was done where the reaction components were mixed and incubated for 30 min at 30°C prior to the addition of MgAc. 4 μl of 375 μM MgAc was then added, and the incubation was continued for an additional 30 min at 30°C. In Figure 3, the preincubation step was performed in reaction mixtures (96 μl) that also contained 1.0 mM Tris-HCl (pH 8.0), 5.8 mM NaCl, 15 mM KCl, and 0.25% glycerol, but were omitted for the indicated Tns proteins; after the preincubation, the omitted reaction components and MgAc were added, resulting in final reaction mixtures (104 μl) that also contained 2.0 mM Tris-HCl (pH 8.0), 11.5 mM NaCl, 30 mM KCl, and 0.5% glycerol.

As described in Bainton et al. (1991), after the incubation was completed, the reaction DNAs were recovered, digested with restriction enzyme, and electrophoresed through 0.6% agarose. The resolved DNAs were transferred to Nytran (Schleicher & Schuell) and analyzed by Southern hybridization with a radioactive DNA probe specific for the miniTn7 element. Blots were examined by autoradiography, and reaction products were quantified with a Phosphorimager (Molecular Dynamics).

In vitro transposition reactions in Figure 4 were done using a 260 bp attTn7 DNA fragment as a linear target. Reactions (100 μl) contained 34.5 mM Tris-HCl (pH 8.0), 44.5 mM KCl, 0.35 mM NaCl, 0.02 mM EDTA, 1.4 mM DTT, 15 mM MgAc, 0.02 mM $MgCl_2$, 0.03 mM $CaCl_2$, 0.02 mM CHAPS, 6.7% glycerol (v/v), 260 μg/ml BSA, 20 μg/ml sheared salmon sperm DNA, and

approximately 0.1 pmol end-labeled attTn7 fragment. 20 ng TnsC and 30 ng TnsD were initially incubated in reaction mixtures for 30 min at 30°C to allow TnsC-TnsD binding to attTn7. 50 ng TnsA, 50 ng TnsB, 50 ng donor plasmid pEMΔ, and MgAc were added; as indicated, 2 μg of the attTn7 plasmid pKAO4-3 was also added. Incubations were continued for 30 min at 30°C, treated with 0.25% SDS, 20 mM EDTA, and 0.3 mg/ml Proteinase K for 30 min at 37°C, and electrophoresed through a 1% agarose gel in 1X Tris-borate-EDTA buffer. The gel was dried and exposed to X-ray film.

Mobility shift assays and DNase I footprinting

Mobility shift assays were performed as described in Chapter 3, except that the DNA binding reactions (20 μl) contained 50 mM Tris-HCl (pH 8.0), 0.25 mM HEPES (pH 7.5), 120 mM KCl, 10 mM NaCl, 0.1 mM EDTA, 1.8 mM DTT, 0.1 mM $MgCl_2$, 0.1 mM CHAPS, 0.8 mM ATP, 9% glycerol (v/v), 300 μg/ml BSA, 15 μg/ml sheared salmon sperm DNA, and approximately 0.01 pmol end-labeled DNA. Purified TnsC and TnsD were incubated in reaction mixtures for 20 min at room temperature and were analyzed on 5% polyacrylamide as described in Chapter 3.

The DNase I footprint analysis was performed as described in Chapter 3, except that the reactions (100 μl) contained 40 mM Tris-HCl (pH 8.0), 1.3 mM HEPES (pH 7.5), 85 mM KCl, 50 mM NaCl, 0.1 mM EDTA, 1.7 mM DTT, 0.5 mM $MgCl_2$, 0.5 mM CHAPS, 1 mM ATP, 8.5% glycerol (v/v), 350 μg/ml BSA, 20 μg/ml sheared salmon sperm DNA, with approximately 0.15 pmol 3'end-labeled DNA. TnsD or TnsC + TnsD was incubated in reaction mixtures for 15 min at 30°C and treated with DNase I as described in Chapter 3.

Figure 1. Tn7 transposition substrates, intermediates, and products.

The 6.1 kb donor molecule contains a 1.6 kb miniTn7 element flanked by sequences unrelated to attTn7; the miniTn7 element contains segments from the left (Tn7L) and right (Tn7R) ends of Tn7 that provide the cis-acting transposition sequences. The 2.9 kb target molecule contains a 0.15 kb attTn7 segment flanked by EcoRI sites. Double-stranded breaks at the junction of either Tn7L or Tn7R with the donor backbone (arrows) produce the DSB.L and DSB.R species, which are transposition intermediates. An excised linear transposon (ELT) is generated by a second double-strand break in the DSBs. The excised mini-Tn7 elements inserts site- and orientation-specifically into attTn7, producing a 1.6 kb simple insertion product and a 4.5 kb gapped donor backbone. EcoRI digestion of the simple insertion product releases a diagnostic 1.75 kb fragment. Other transposition species are made by a single Tn7 end joining to a target molecule (see text).

Figure 1

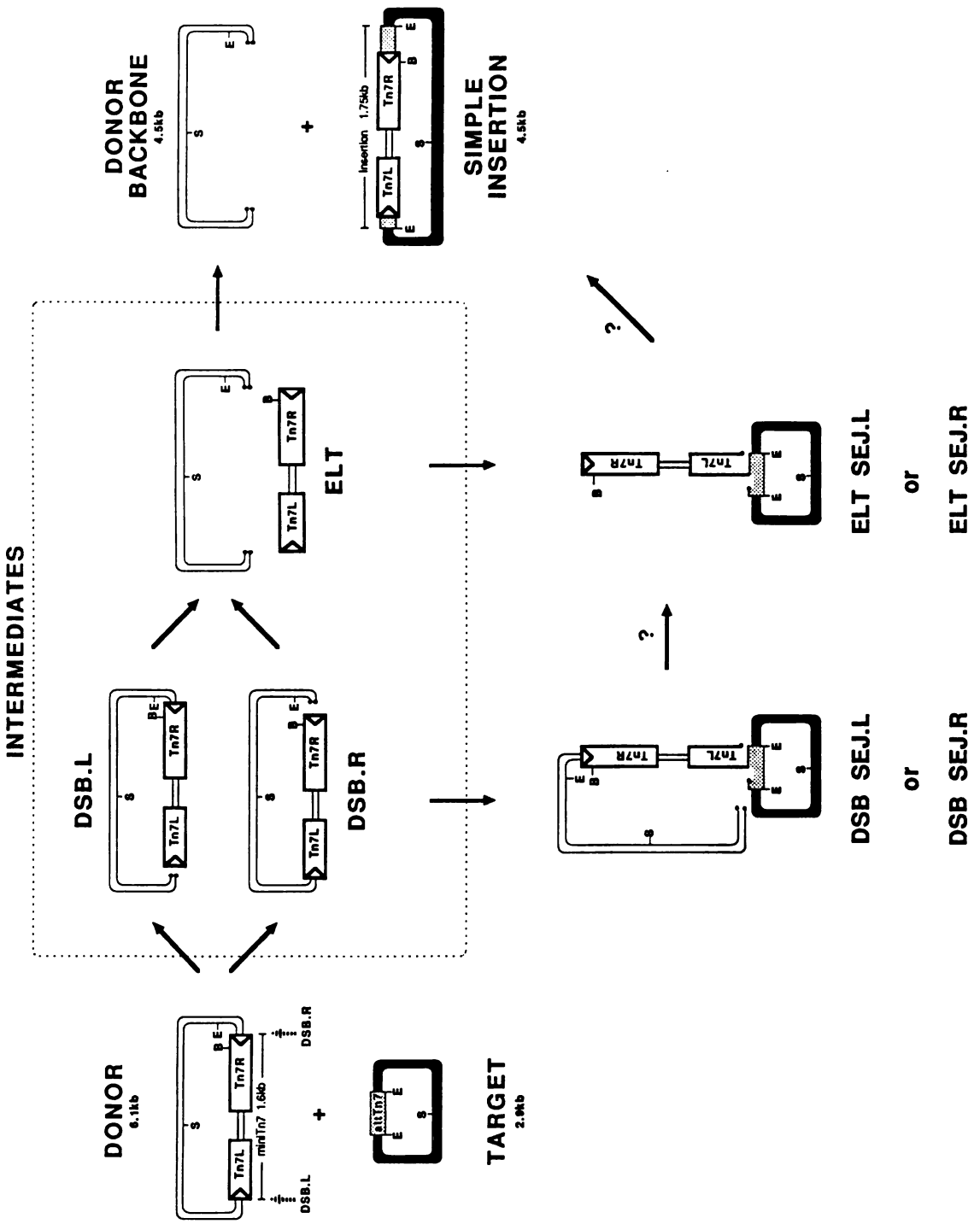

DONOR
6.1kb
DSB.L
DSB.R
Tn7L
Tn7R
attTn7
TARGET
2.9kb
INTERMEDIATES
DSB.L
DSB.R
ELT
DSB SEJ.L
or
DSB SEJ.R
ELT SEJ.L
or
ELT SEJ.R
DONOR
BACKBONE
4.5kb
SIMPLE
INSERTION
4.5kb
Insertion 1.75kb
?
?

Figure 2. Purified Tns proteins reconstitute Tn7 transposition *in vitro*.

A) Purified Tns protein fractions used in Tn7 transposition. The Tns protein fractions were resolved by 12.5% SDS-PAGE and stained with Coomassie Blue R-250. Lanes 1-3 show TnsA fractions during purification: lanes 1 and 2, 1.0 μg affinity purified GST-TnsA fusion protein, before and after thrombin cleavage, respectively (TnsA = 29 kd and GST domain = 26 kd); lane 3, 2.0 μg cleaved TnsA protein after further purification. Lanes 4-6 show other purified Tns protein fractions, which were concentrated by 10% TCA precipitation prior to loading: lane 4, 2.0 μg TnsB; lane 5, 2.0 μg TnsC; lane 6, 2.0 μg TnsD.

B) Requirements for Tn7 transposition *in vitro*. An autoradiogram of a gel is shown of transposition reactions digested with EcoRI in which the indicated components were omitted. The DNA species, which were detected by Southern hybridization with a miniTn7-specific probe, are labeled as in Figure 1. SEJ.L refers to the DSB-SEJ.L species. Reactions were performed as described in Experimental Procedures, with the following exceptions: lane 6, ADP was substituted for exogenously added ATP; lane 7, a target plasmid lacking *attTn7* sequences was used as a target instead of the usual *attTn7*-containing plasmid.

Figure 2A

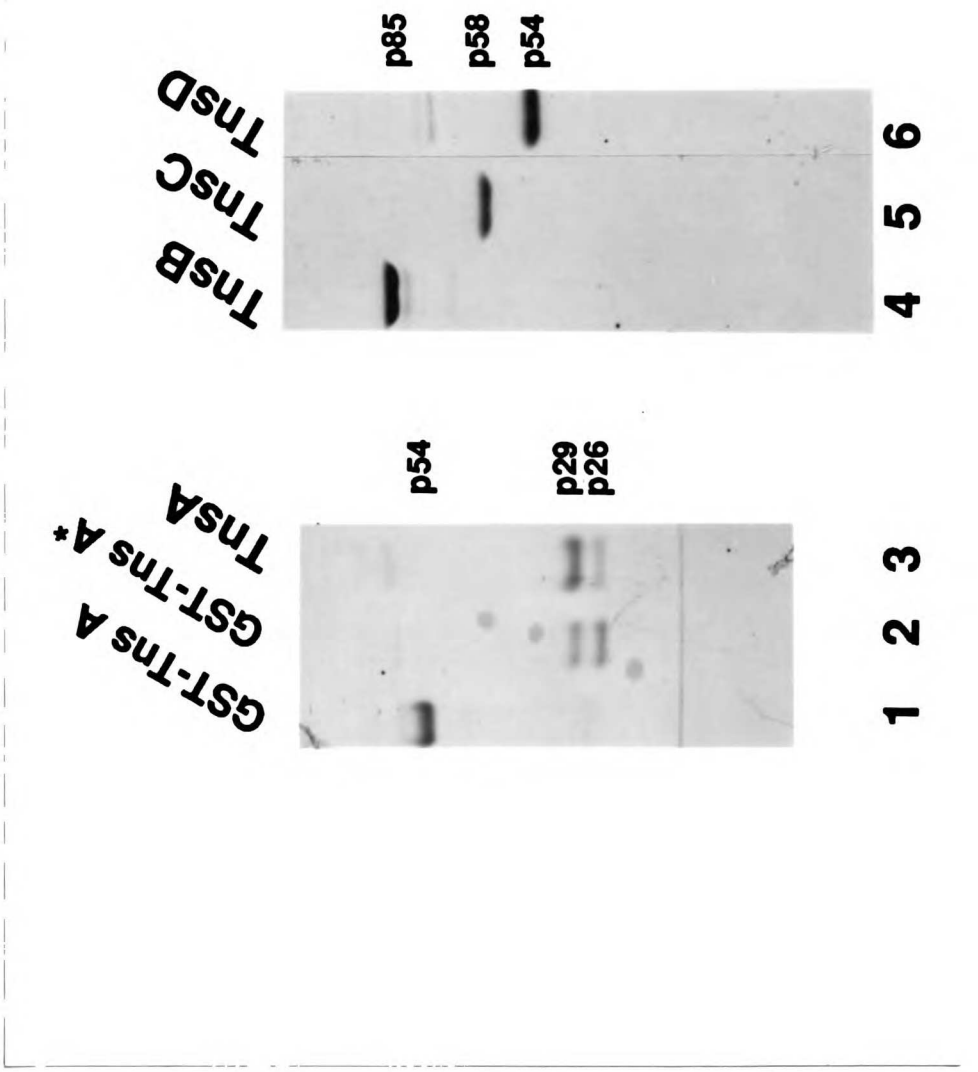

p85
p58
p54
TnsD
TnsC
TnsB
6
5
4
p54
p29
p26
TnsA
GST-Tns A*
GST-Tns A
3
2
1

Figure 2B

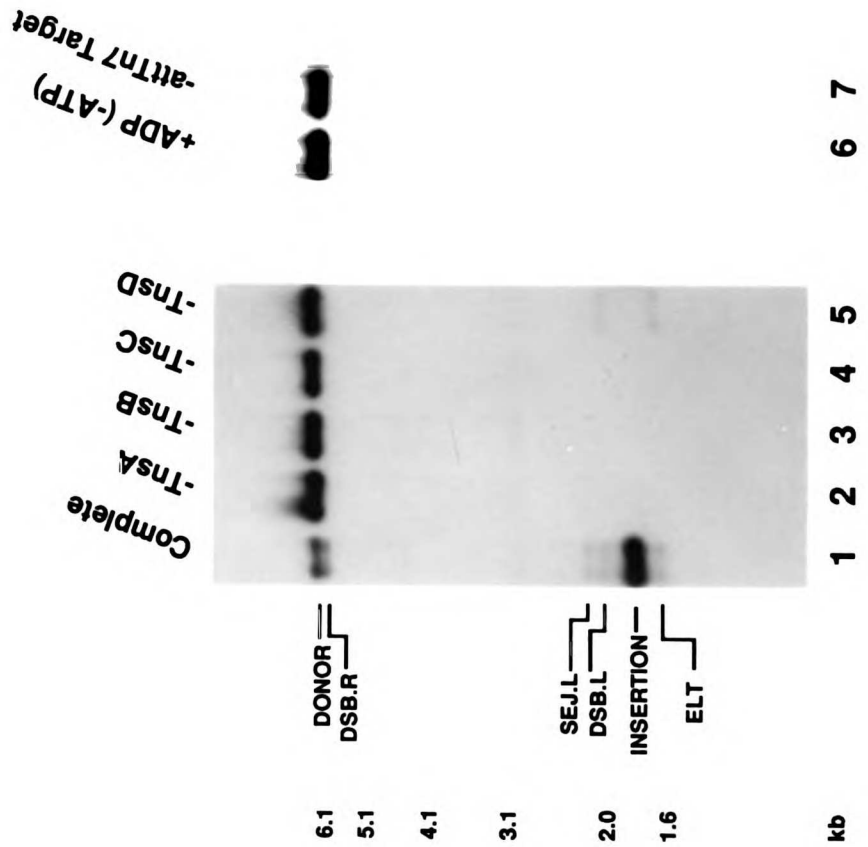

-attTn7 Target
+ADP (-ATP)
-TnsD
-TnsC
-TnsB
-TnsA
Complete
1 2 3 4 5 6 7
DONOR
DSB.R
SEJ.L
DSB.L
INSERTION
ELT
6.1
5.1
4.1
3.1
2.0
1.6
kb

Figure 3. Tn7 transposition *in vitro* is stimulated by a preincubation step that lacks MgAc.

A) Shown is an autoradiogram from transposition reactions which display Bgl II-digested species detected with a miniTn7-specific probe. In lanes 1-3, MgAc was omitted entirely from the reaction (lane 1), added initially to the reaction (lane 2), or added after a 30 min preincubation step done in the absence of MgAc (lane 3). In lanes 4-8, indicated reaction components were initially omitted and added to reactions after the 30 min (-)MgAc preincubation step.

B) Shown is a schematic representation of the (-)MgAc preincubation step and the required reaction components. TnsC, TnsD, *attTn7*, and ATP were added at the start (T_0) of the reaction in the absence of MgAc. After a preincubation step at 30°C for 30 min, MgAc and the remaining reaction components (TnsA, TnsB, and donor DNA) were added (at T_{30}), and incubation continued at 30°C for 30 min (until T_{60}).

Figure 3A

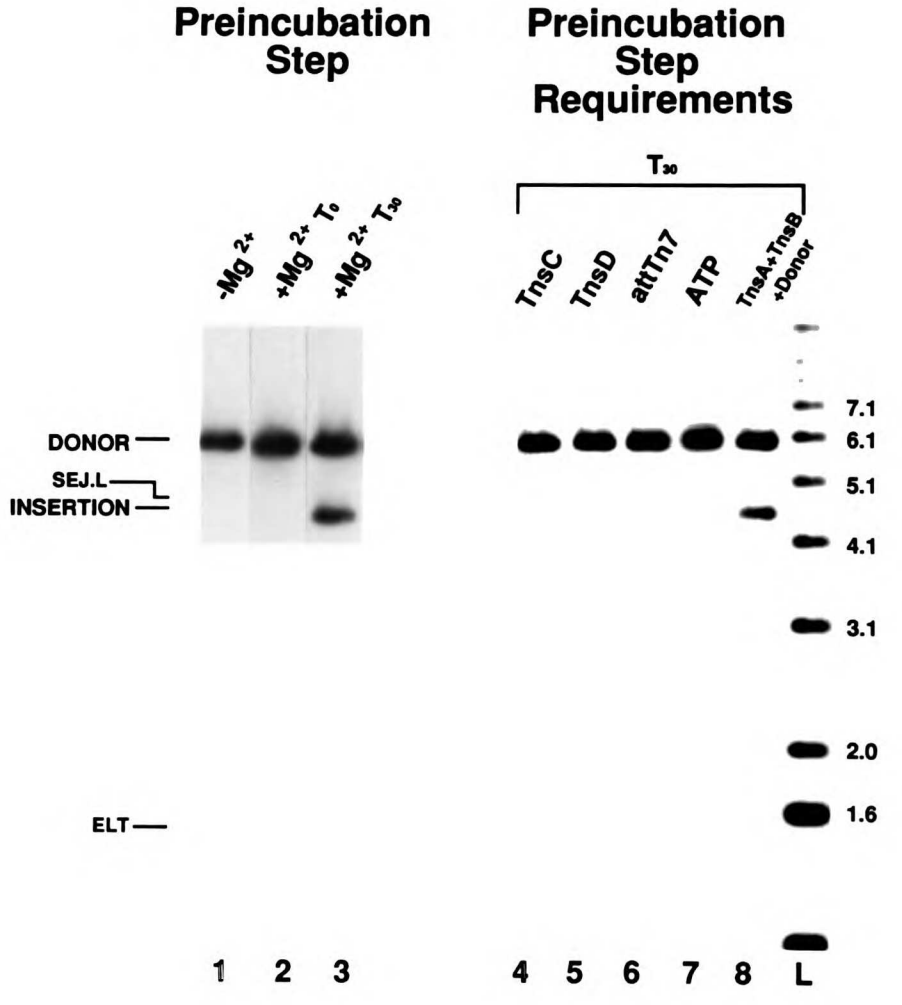

Preincubation Step
Preincubation Step Requirements
T30
-Mg 2+
+Mg 2+ T0
+Mg 2+ T30
TnsC
TnsD
attTn7
ATP
TnsA+TnsB +Donor
DONOR
SEJ.L
INSERTION
ELT
7.1
6.1
5.1
4.1
3.1
2.0
1.6
1 2 3 4 5 6 7 8 L

Figure 3B

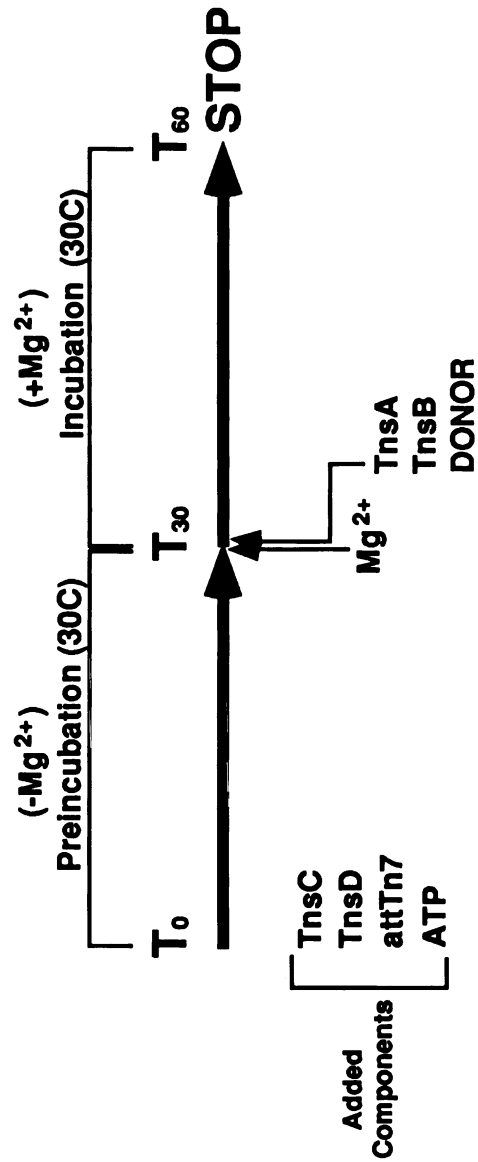

STOP
T60
T30
T0
(+Mg2+)
Incubation (30C)
(-Mg2+)
Preincubation (30C)
Mg2+
TnsA
TnsB
DONOR
TnsC
TnsD
attTn7
ATP
Added
Components

Figure 4. Formation of TnsC + TnsD complexes with attTn7.

A) Analysis of protein-DNA complexes by mobility shift assay. Lanes 1-7 contain an attTn7 segment, whereas lanes 8 and 9 contain a DNA segment lacking attTn7 sequences. Where indicated by "+", 20 ng TnsC and 20 ng TnsD were added to binding reactions. Lane 5, MgAc was added to 15 mM at the beginning of the binding reaction. Lane 7, ADP (0.8 mM) was added instead of ATP (a small amount of ATP is present in the reaction from the TnsC fraction).

B) In vitro transposition into a linear attTn7 target. Particular reaction components (TnsC or MgAc) were added at indicated reaction times T_0 or T_{30} (see Figure 3B), to determine the effects of TnsC + TnsD binding to attTn7 on Tn7 transposition. Lanes 4 and 5, 2 μg of the attTn7 plasmid (pKA04-3) was provided as a competing target, either at the start or at the end of the (-)MgAc preincubation step.

Figure 4A

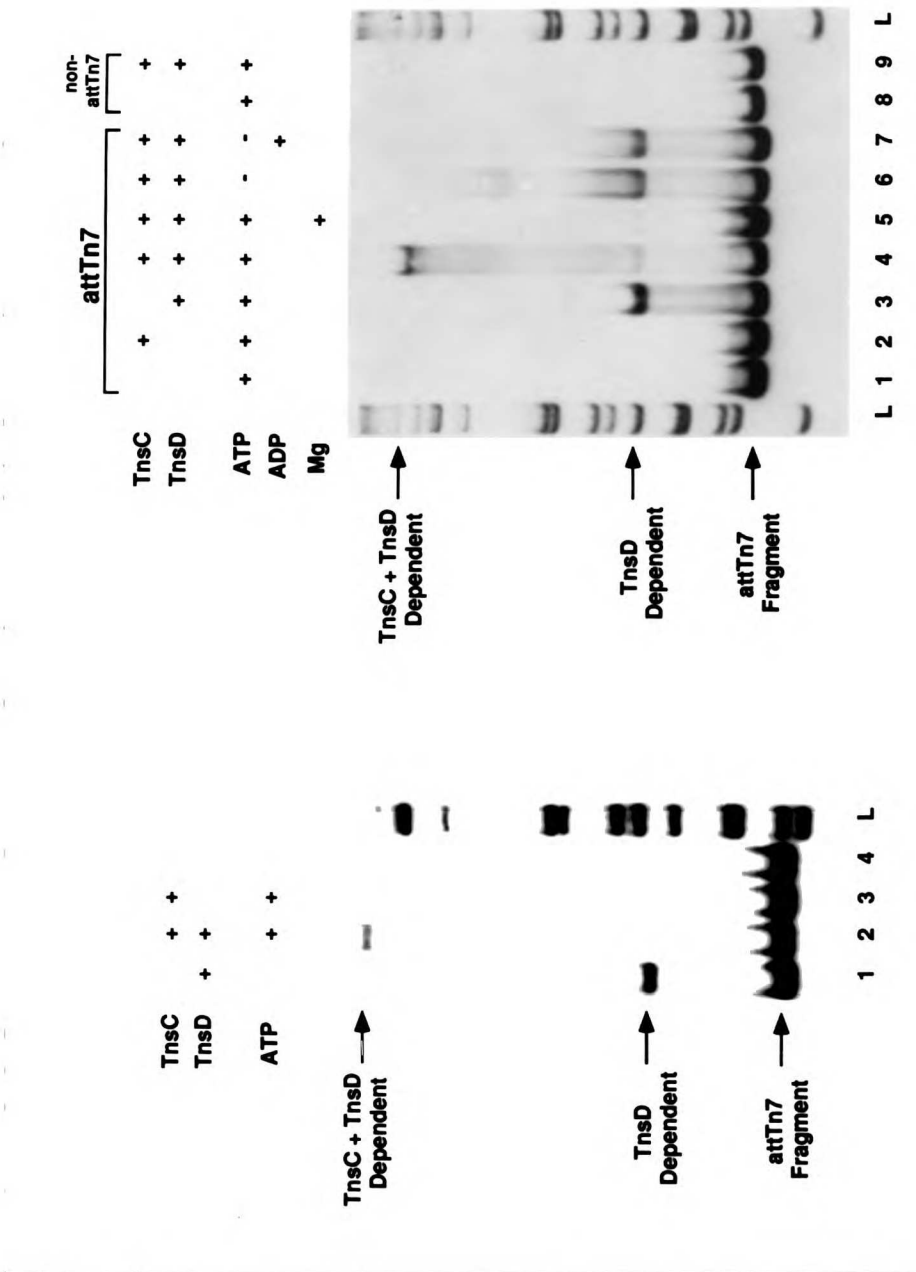

attTn7
non-attTn7
TnsC
TnsD
ATP
ADP
Mg
L 1 2 3 4 5 6 7 8 9 L
TnsC + TnsD Dependent
TnsD Dependent
attTn7 Fragment
TnsC
TnsD
ATP
1 2 3 4 L
TnsC + TnsD Dependent
TnsD Dependent
attTn7 Fragment

Figure 4B

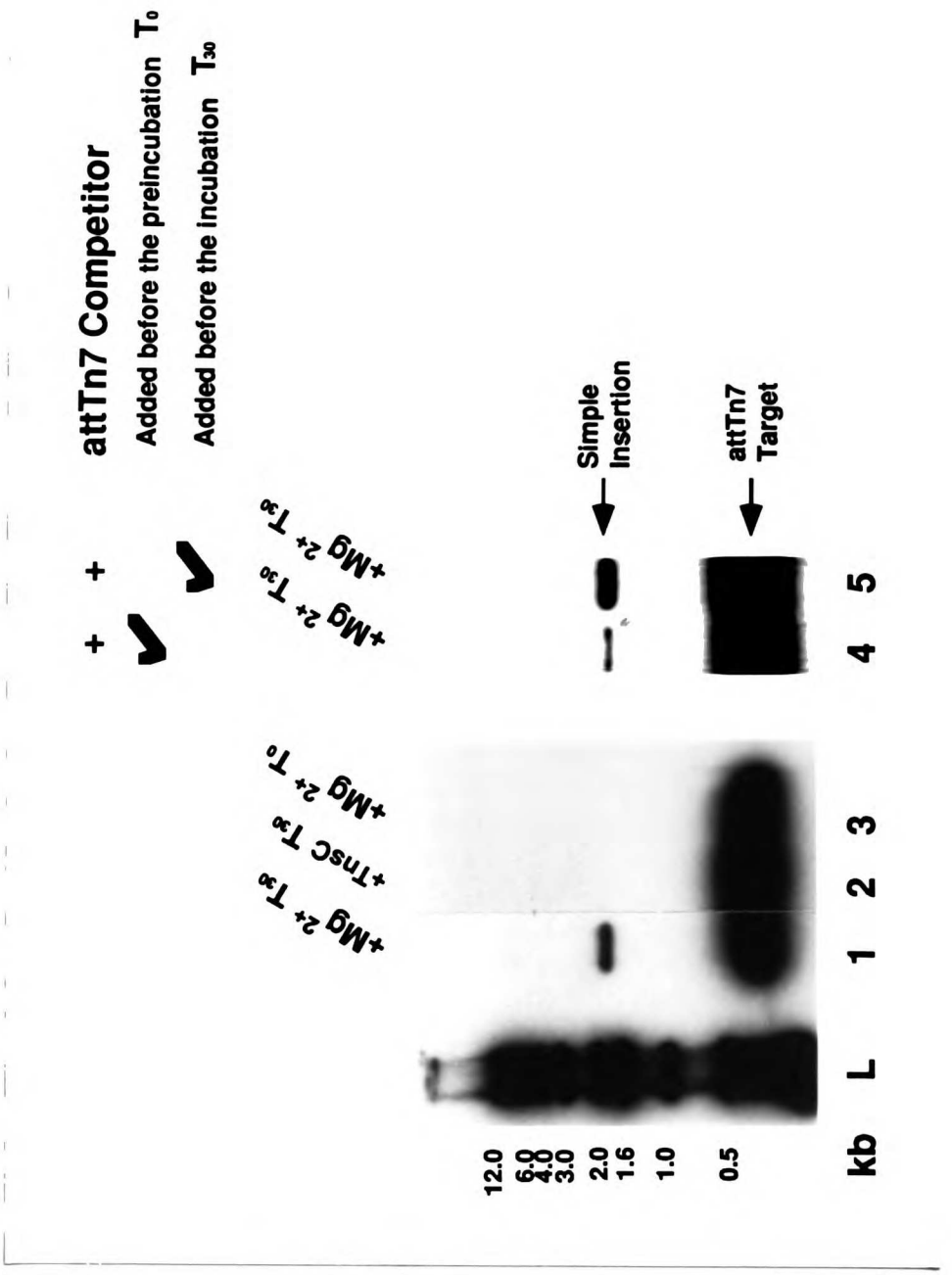

attTn7 Competitor
Added before the preincubation T0
Added before the incubation T30
+
+
+
+Mg 2+ T30
+Mg 2+ T30
+Mg 2+ T30
+Tnsc T30
+Mg 2+ T0
Simple Insertion
attTn7 Target
kb
L
1
2
3
4
5
12.0
6.0
4.0
3.0
2.0
1.6
1.0
0.5

Figure 5. DNase I footprint analysis of TnsD and TnsC + TnsD complexes with attTn7.

TnsD and TnsC + TnsD complexes with attTn7 were formed, treated with DNase I, and the DNA recovered from the isolated complexes. Top and bottom strands of attTn7 are as shown in Figure 6, and the nucleotide positions are numbered as described in Figure 6. The vertical lines indicate the extent of protection, and dashed line indicates the region of partial protection. Lanes marked "AG" contain Maxam-Gilbert A+G sequence reaction of same DNA fragment.

A) Top strand of attTn7: lane 1, TnsD at 2 μg/ml; lane 2, no Tns proteins added; lane 3, TnsC at 5 μg/ml and TnsD at 2 μg/ml; lane 4, no Tns proteins added.

B) Bottom strand of attTn7: lane 1, no Tns proteins added; lane 2, TnsD at 2 μg/ml; lane 3, no Tns proteins added; lane 4, TnsC at 5 μg/ml and TnsD at 2 μg/ml.

Figure 5

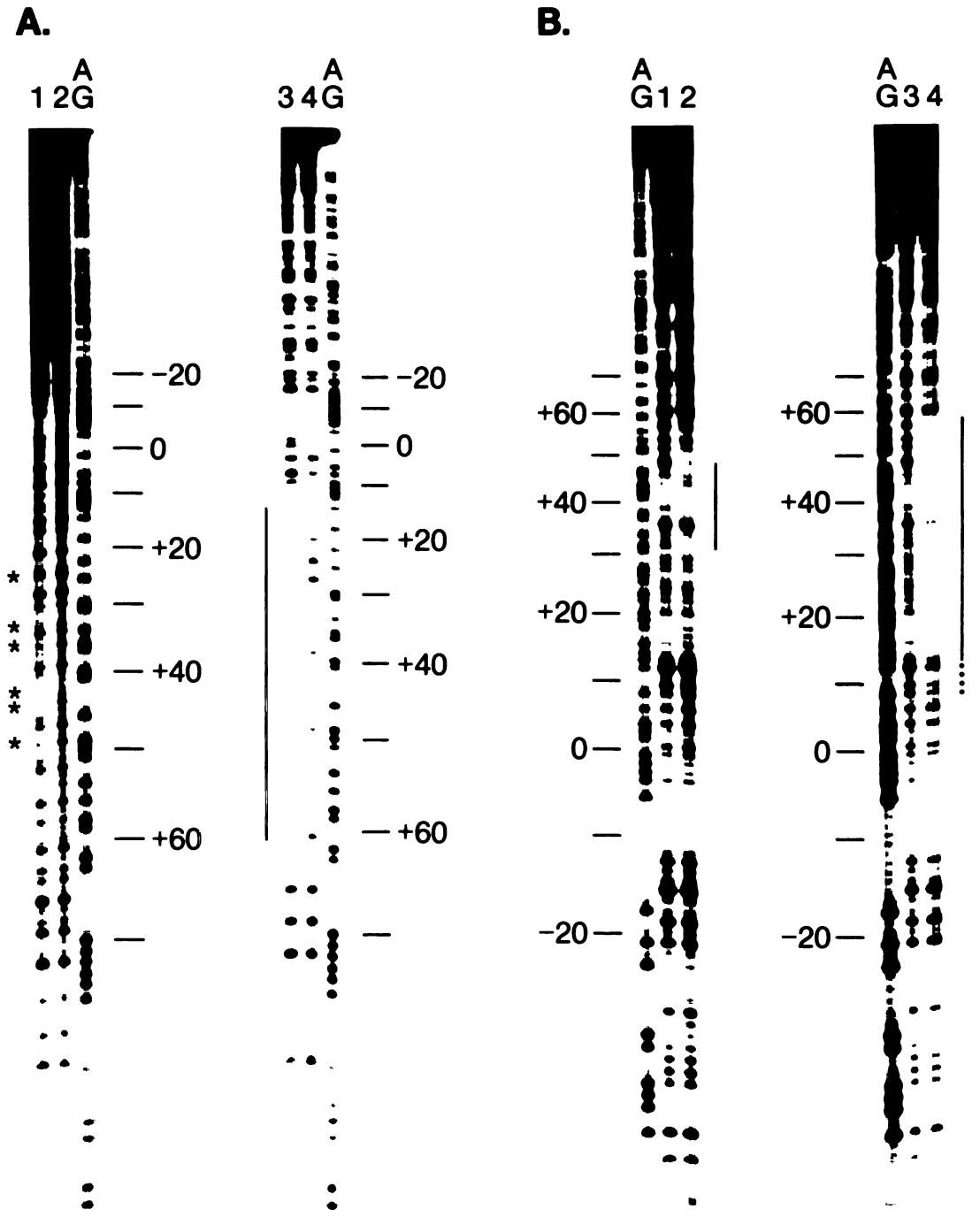


Figure 6. Summary of DNase I footprint analysis.

A) The attTn7 sequence is shown, with boxed nucleotides indicating the 5 bp sequence duplicated upon Tn7 insertion (Lichtenstein and Brenner 1982; Gay et al. 1986). The central base pair of the target duplication is designated "0"; sequences extending leftward are given negative values, whereas sequences extending rightward are given positive values. Solid bars above and below the sequence show the extent of protection on the top and bottom strands, respectively, in the presence of the indicated Tns proteins. The stippled bar designates a region of partial protection, and arrows indicate accessible positions within the protected region. The bars extend to the outermost positions that are inaccessible to DNase I. Open circles mark the positions nearest the protected regions that are accessible. The dotted line indicates the modest protection observed for TnsD on the top strand.

B) Diagram of TnsD and TnsC + TnsD complexes formed on attTn7. TnsD binds to sequences (approximately +30 to +50) which are positioned at a distance from the specific point of Tn7 insertion (indicated by arrow). In contrast, TnsC + TnsD interacts with a more extensive region of attTn7 (approximately +10 to +60). The actual placement the TnsC and TnsD proteins within the complex remains to be determined.

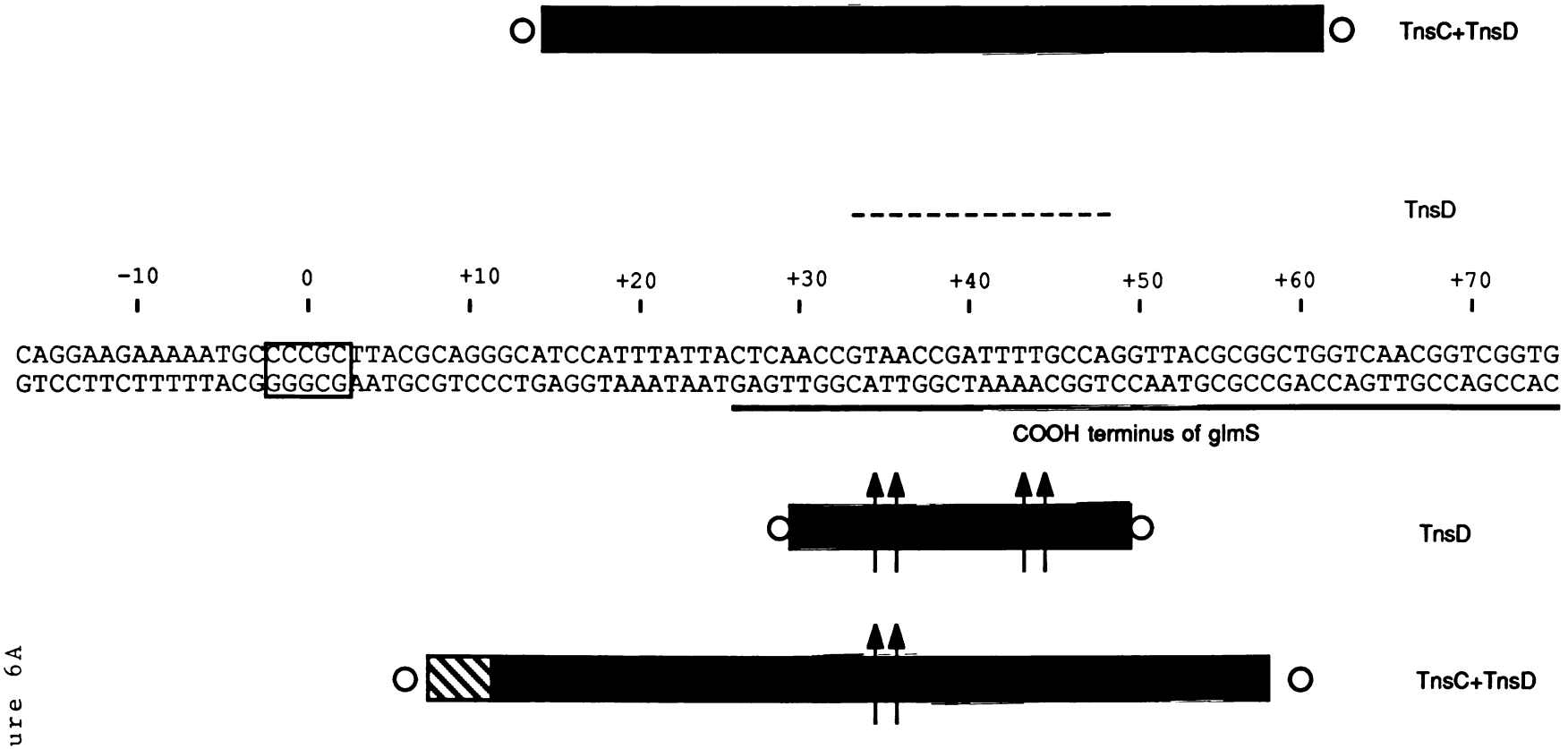

TnsC+TnsD
TnsD
-10
0
+10
+20
+30
+40
+50
+60
+70
-146-
CAGGAAGAAAAATGCCCCGCTTACGCAGGGCATCCATTTATTACTCAACCGTAACCGATTTTGCCAGGTTACGCGGCTGGTCAACGGTCGGTG
GTCCTTCTTTTTACGGGGCGAATGCGTCCCTGAGGTAAATAATGAGTTGGCATTGGCTAAAACGGTCCAATGCGCCGACCAGTTGCCAGCCAC
COOH terminus of glmS
TnsD
TnsC+TnsD

Figure 6A

Figure 6B

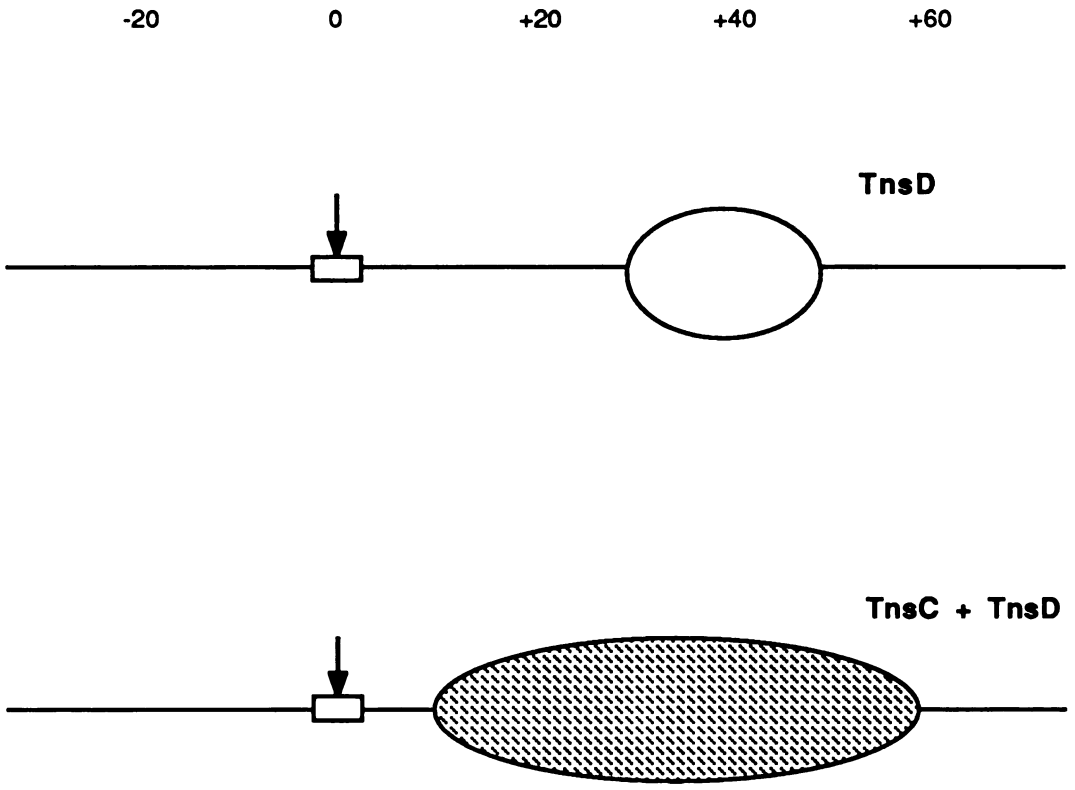


Figure 7. Tn7 transposition in the presence of AMP-PNP.

A) EcoRI-digested species detected with the miniTn7-specific probe are shown from transposition reactions done in the presence of ATP (lanes 1-3) or AMP-PNP (lanes 4-6). Lanes 4 and 6, a target plasmid (pKS+) that lacked attTn7 sequences was used instead of the standard attTn7 plasmid (pKA04-3).

B) DNA species detected with a miniTn7-specific probe from transposition reactions were digested with Bgl II (lanes 1 and 2) or Bgl II + EcoRI (lanes 3 and 4) to evaluate attTn7 insertion specificity.

Figure 7A

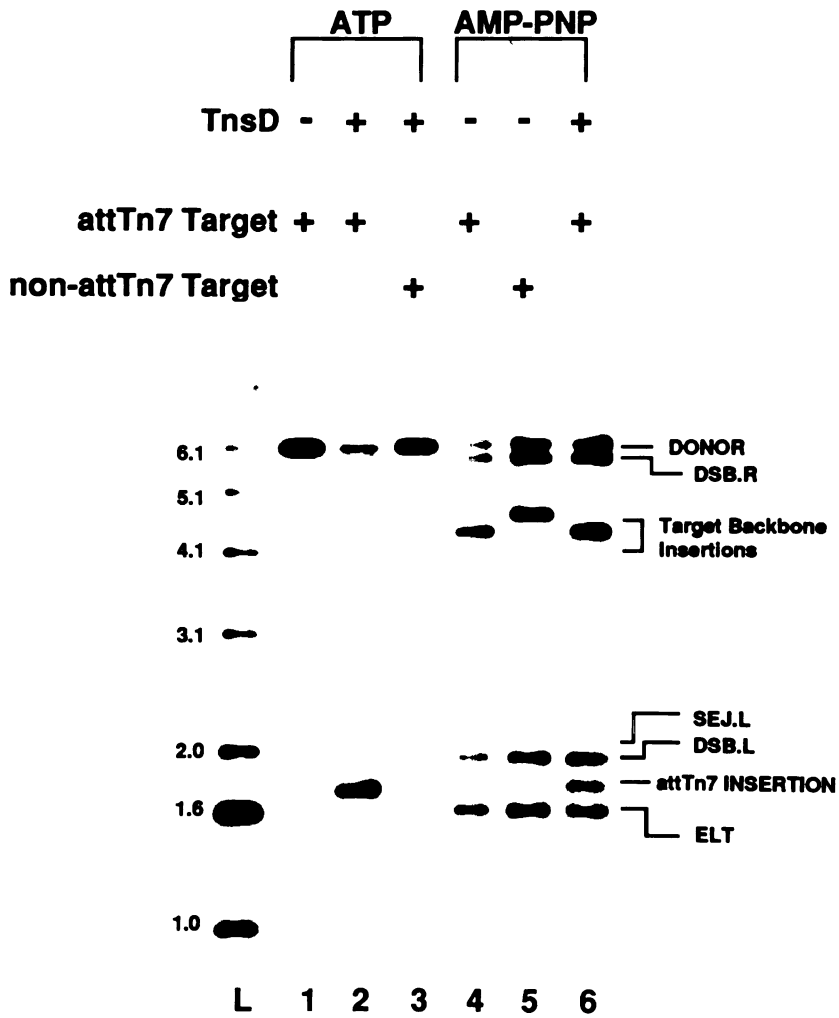

ATP
AMP-PNP
TnsD - + + - - +
attTn7 Target + + + +
non-attTn7 Target + +
6.1
5.1
4.1
3.1
2.0
1.6
1.0
DONOR
DSB.R
Target Backbone Insertions
SEJ.L
DSB.L
attTn7 INSERTION
ELT
L 1 2 3 4 5 6

Figure 7B

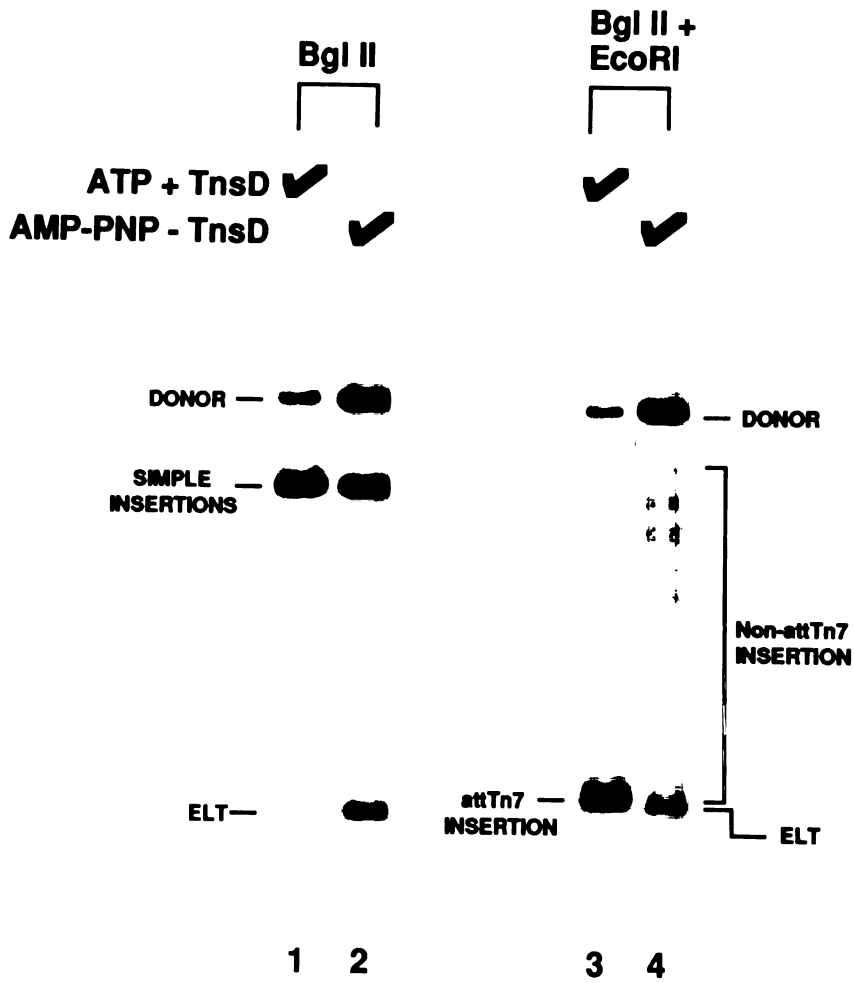
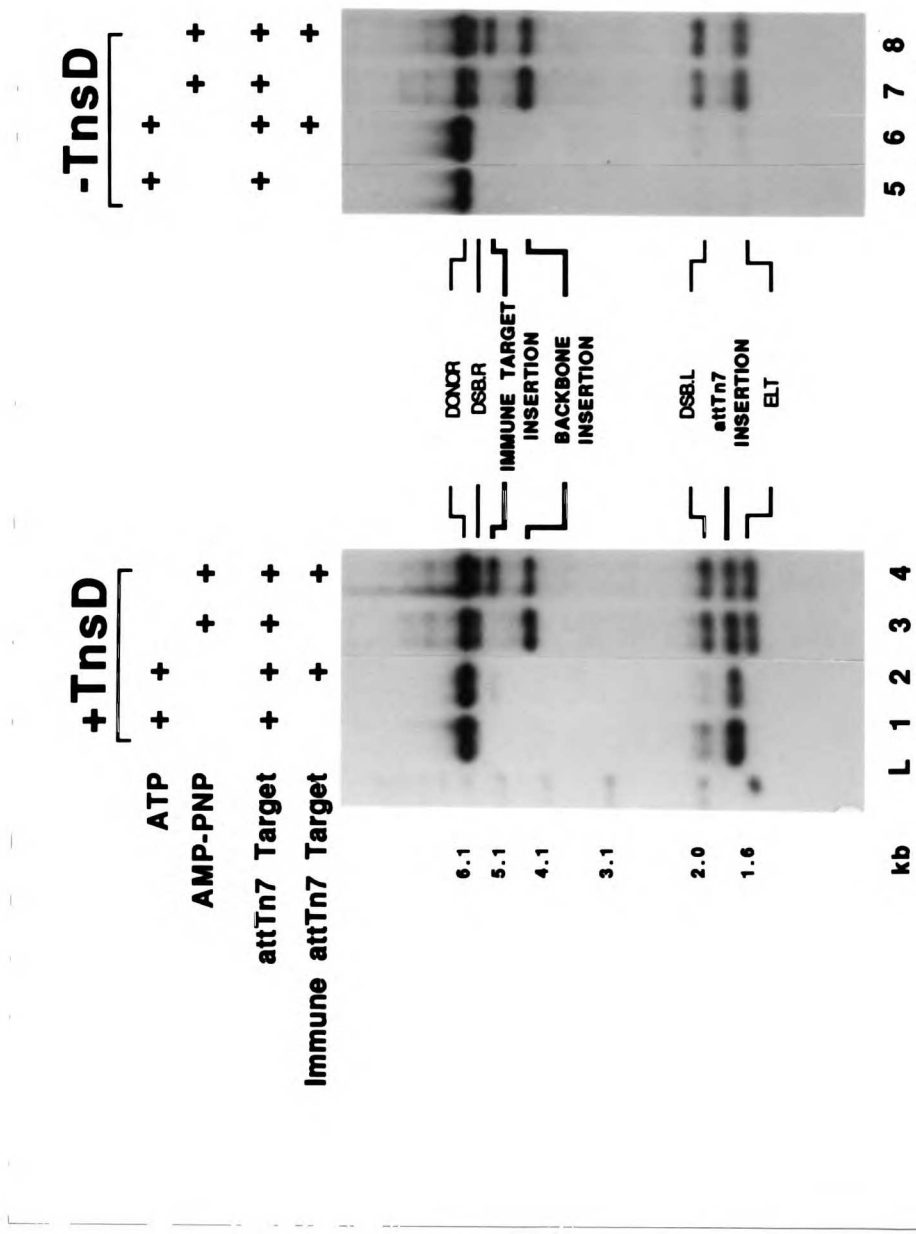

Bgl II
Bgl II + EcoRI
ATP + TnsD
AMP-PNP - TnsD
DONOR
SIMPLE INSERTIONS
ELT
DONOR
Non-attTn7 INSERTION
attTn7 INSERTION
ELT
1 2 3 4

Figure 8. Tn7 transposition immunity in vitro is altered by the presence of AMP-PNP.

EcoRI-digested species detected with a miniTn7-specific probe are shown from transposition reactions using the standard attTn7 target plasmid (pKA04-3) and an attTn7 plasmid (pLA11) which also contained a Tn7R segment. The concentration of the target plasmids is 3-fold less than that of the standard conditions. Reactions were done in the presence (lanes 1-4) or the absence (lanes 5-8) of TnsD.

Figure 8


+TnsD
-TnsD
ATP
AMP-PNP
attTn7 Target
Immune attTn7 Target
DONOR DSB.R
IMMUNE TARGET INSERTION
BACKBONE INSERTION
DSB.L
attTn7 INSERTION ELT
6.1
5.1
4.1
3.1
2.0
1.6
kb
L 1 2 3 4
5 6 7 8

Chapter 5

Summary and Perspectives

This thesis has extended earlier studies which established that Tn7 utilizes five transposition genes to direct recombination along two distinct pathways, i.e. tnsABC + tnsD and tnsABC + tnsE (Rogers et al. 1986; Waddell and Craig 1988). In the work presented here, I have analyzed Tn7 insertions promoted by these two ensembles of tns genes, which has revealed the sequence information important for target site selection. I have also examined the transposition proteins which interact with the target DNA during Tn7 insertion into attTn7, which has also provided a mechanistic understanding of Tn7 target site selection. These studies have revealed the complexity of protein-DNA and protein-protein interactions that mediate Tn7 target site selection.

Tn7 directs transposition to two distinct classes of target sites

I have demonstrated that the target sites used by the tnsABC + tnsD and tnsABC + tnsE pathways differ significantly in their nucleotide sequences. tnsABC + tnsD is a target sequence-specific pathway that directs insertion into attTn7 and chromosomal pseudo-attTn7 sites, which are related in sequence to attTn7. In contrast, tnsABC + tnsE is an apparently random target sequence pathway that promotes insertion into many different sites.

Insertion into attTn7 and the related pseudo-attTn7 sites at a reduced transposition efficiency is reminiscent of target site usage by bacteriophage λ during integration into the E. coli chromosome. λ inserts efficiently into a specific chromosomal site, called attB, but can also insert at reduced frequency into another set of secondary sites

in the chromosome (Shimada et al. 1972, 1973). Examination of these secondary sites revealed sequence homology to attB, and the lower frequency of utilization presumably reflects the differences in nucleotide sequence between attB and secondary sites (reviewed in Weisberg and Landy 1983). A distinguishing feature, however, concerning the similarity between these sites is that the homology between attB and secondary sites is confined to sequences immediately near the insertion site. The phage-encoded Int recombinase binds specifically to these sequences, where DNA cleavage and strand exchange occur (Ross and Landy 1983; Craig and Nash 1983). In contrast, the attTn7 and pseudo-attTn7 sites display a similarity in sequence that is positioned at a distance from the point of Tn7 insertion. As discussed below, this region is important for specific recognition by the transposition protein TnsD. The insertion specificity exhibited by Tn7 contrasts with that of several other transposable elements, such as Tn10, which show sequence specificity at the site of target duplication (Halling and Kleckner 1982).

The tnsABC + tnsE pathway is analogous to many other transposable elements which can insert into a variety of target sites. Although we have not to date observed any similarity in sequence between tnsE-dependent sites, it would not be unreasonable to find insertion specificity operating at some level. Examination of a large number of insertions into a defined target DNA segment may reveal a subtle similarity in sequence or a more generalized regional specificity.

Role of two transposition pathways in the maintenance of Tn7 within genomes

Tn7's ability to transpose into two distinct classes of target sites may play an important role in its establishment and proliferation within bacterial genomes. tnsABC + tnsD transposition into attTn7 assures Tn7 of an insertion site that can be tolerated by the host cell, as no E. coli gene is disrupted. Hence, high frequency insertion into a specific site can promote the stable establishment of Tn7 within a genome. Insertion into attTn7 also positions Tn7 adjacent to the E. coli glmS gene, which is involved in cell wall biosynthesis (Lichtenstein and Brenner 1982; Walker et al. 1984). The location of Tn7 downstream of glmS could provide a means for Tn7 to respond to the physiological state of the cell (Gay et al. 1986). Transcription from glmS, which is more abundant during logarithmic growth, extends into the Tn7 right end across the tnsAB promoter; thus, glmS transcription may occlude the tnsAB promoter and inhibit expression of tnsA and tnsB. Under these robust growth conditions, rapid replication of the genome generates mulitple copies of Tn7; hence, in this case transposition functions are not needed for Tn7 to proliferate.

In contrast to Tn7 insertion into attTn7, tnsABC + tnsE transposition allows the proliferation of Tn7 into other chromosomal and plasmid target sites. In particular, insertion of Tn7 into conjugable plasmids allows the element to disseminate readily in a bacterial population via plasmid transfer from cell to cell (Barth and Grinter 1977; Barth et al. 1978; Moore and Krishnapillai 1982). The identification of Tn7 in conjugable plasmids in natural populations of

enteric bacteria may explain the rapid spread of bacterial trimethoprim resistance from the heavy use of antibiotics (Goldstein et al. 1986).

Tn7 utilizes TnsD, a specific target DNA binding protein, to direct insertion into attTn7

A particularly interesting question is how Tn7 directs site-specific insertion into attTn7. I have provided biochemical evidence that TnsD binds directly to attTn7, and the sequence information important for TnsD binding correlates with that required for transposition to attTn7. Moreover, TnsD also directly participates in and is essential for in vitro transposition to attTn7, which explains the requirement for the tnsD gene for Tn7 insertion into attTn7 in vivo. Even under conditions of TnsA + TnsB + TnsC with AMP-PNP, where Tn7 insertions occur with random specificity, the addition of TnsD still promotes site-specific insertion into attTn7. Taken together, these results argue that TnsD recognition of attTn7 is a critical determinant for target activity. It is tempting to speculate that TnsD is the only specific determinant for attTn7 insertion.

Notably, TnsD specifically recognizes sequences in attTn7 that are entirely displaced from the specific point of Tn7 insertion. The region of homology between pseudo-attTn7 and attTn7 also coincides with the sequences important for TnsD recognition. We suggest that TnsD may also bind to pseudo-attTn7 sites, with the differences in nucleotide sequence between attTn7 and pseudo-attTn7 sites accounting for the disparity in target activity. It will be interesting to compare the extent to which these specific sequences affect TnsD binding and target activity.

A multiprotein-DNA complex can form with attTn7

A notable feature about the TnsD interaction with attTn7 is that the recognition site is located in a region about 30 to 50 nt away from the point of Tn7 insertion, where the target DNA breakage and rejoining events occur. TnsD acts at a distance to mediate these target DNA transactions by positioning the rest of the transposition apparatus toward the point of insertion. In support of this view, biochemical studies have shown that TnsD recruits another Tn7-encoded protein, TnsC, to the attTn7 target site. Also, the region of attTn7 that interacts with TnsC + TnsD extends close to the point of Tn7 insertion. TnsC, an ATP-dependent non-sequence-specific DNA binding protein, may also directly contact the target DNA through its interaction with TnsD. A plausible view is that the TnsC + TnsD interaction with attTn7 directs to the point of insertion the other transposition components, in particular, the transposase, which mediates DNA cleavage and strand transfer.

Although a target complex with TnsC + TnsD is competent to promote Tn7 insertion, such a discrete intermediate need not be an obligatory step for Tn7 recombination. For example Tn7 insertion into attTn7 may proceed in a manner similar to bacteriophage Mu DNA strand transfer, where multiple pathways of assembly can produce the strand transfer product (Baker et al. 1991). Even a "naked" attTn7 target, i.e. non-complexed with protein, may still efficiently promote recombination. This situation occurs during bacteriophage λ integration into the attB target site, where it was found that the λ integration complex,

consisting of Int and the E. coli IHF protein, assembles entirely on the phage DNA and encounters a naked attB target (Richet et al. 1988). Nonetheless, the studies of the TnsC + TnsD interaction with attTn7 have been valuable in dissecting the protein-DNA and protein-protein interactions that occur during Tn7 recombination.

An interesting question concerns how the two Tn7 ends communicate with the attTn7 target site during transposition. One attractive view is that TnsC, associated with the target DNA through TnsD, mediates communication between the target site and the Tn7 ends, perhaps through an interaction with TnsB and/or TnsA. These elaborate interactions likely require a specialized nucleoprotein assembly in which the transposon and target DNAs are held together by the Tn7 transposition proteins. In support of this view, the efficient production of both recombination intermediates and products in vitro requires the presence of the donor and target DNAs with all the Tn7 recombination proteins (Bainton et al. 1991; this work). This observation suggests that the transposition components, including the transposon and target DNAs, assemble together before any transposition reaction occurs.

Target specificity in other specialized protein-DNA transactions

Multicomponent protein-DNA interactions that mediate other processes, such as replication, transcription, and DNA rearrangements, also utilize target DNA binding proteins that position the multicomponent assembly onto the appropriate target site. These target DNA binding proteins recognize specific DNA sites and initiate the assembly of the nucleoprotein complex.

In E. coli, DNA replication initiation from the chromosomal origin (oriC) involves the recognition of oriC by the DnaA protein, which binds to multiple sites within oriC to form a higher-order DNA-wrapped complex (reviewed in Bramhill and Kornberg 1988, Kornberg 1988). The DnaA-oriC interaction also leads to a local unwinding of the duplex DNA. In addition, DnaA directly contacts another initiation protein, DnaC; in turn, DnaC is directly associated with the DnaB helicase (Wahle et al. 1989). This network of protein-protein interactions localizes the DnaB helicase onto the oriC region, leading ultimately to the unwinding of the DNA duplex and assembly of the replication machinery onto the available template. A similar strategy is used by bacteriophage λ to replicate its genome (reviewed in Echols 1990). The λ replication origin, oriλ, is specifically recognized by the phage encoded λO protein, which recruits the E. coli DnaB helicase through protein-protein interactions with a second phage protein λP. The λP protein is analogous to DnaC in that λP forms a complex with DnaB helicase, and the λP-DnaB complex is loaded onto the λO-oriλ complex.

Transcriptional initiation also involves sequence-specific DNA binding proteins which direct the assembly of multicomponent complexes at the promoter. Transcription by RNA polymerase III requires recognition of the target gene by the assembly factors TFIIIA, which is specific for 5S genes, and TFIIIC (reviewed in Geiduschek and Tocchini-Valentini 1988). These factors bind specifically to sequences within the target gene and position the transcription factor TFIIIB upstream of the transcription start site (Braun et al. 1989; Kassavetis et al. 1990).

Once positioned on the DNA template, TFIIIB is sufficient to load RNA pol III onto the proper initiation site.

RNA polymerase II-transcribed genes also use sequence-specific DNA binding proteins which act in concert with other transcription factors to assemble the transcription machinery. The general transcription factor TFIID binds to the TATA box a short distance upstream of the transcription start site and promotes the assembly of the transcription initiation complex, consisting of RNA polymerase II and other general initiation factors (reviewed in Greenblatt 1991). The assembly of this multicomponent complex is also stimulated by the action of activator proteins, which are positioned at specific enhancer sites upstream of the promoter core sequences (reviewed in Martin 1991). In one case, the viral activator protein VP16 through its localization to specific enhancer elements stimulates the transcription of herpes simplex virus (HSV) early genes (reviewed in McKnight et al. 1986). VP16 is positioned at these elements largely through its interaction with the Oct-1 protein, which directly binds to a core sequence within the enhancer sequences (Kristie and Sharp 1990). From this location, VP16 can activate transcription initiation, probably through a direct interaction with TFIID and TFIIB, another general initiation factor (Stringer et al. 1990; Lin and Green 1991). An attractive view is that these interactions increase the rate of assembly of the transcription complex.

The properties of these target DNA binding proteins raise other possible functions for TnsD and/or TnsC in Tn7 transposition. These Tns proteins may, for example, induce conformational changes in the target DNA, which may be important for the positioning or perhaps activity of

the other transposition components. As an example of such an effect, the transcriptional regulator MerR, which controls expression of the bacterial mercury-resistance genes, activates gene expression by inducing a structural alteration in the promoter DNA. Upon addition of Hg, the MerR protein stimulates transcription by locally unwinding the duplex DNA in the promoter region, which is proposed to activate the RNA polymerase complex (Ansari et al. 1992).

TnsC and TnsD, in addition to positioning the the other transposition components onto attTn7, may also activate the transposition machinery. In bacteriophage Mu transposition, the phage-encoded MuB protein interacts with the target DNA and recruits the MuA transposase to the target DNA (Maxwell et al. 1987; Adzuma and Mizuuchi 1988). MuB also activates the Mu end cleavage and DNA strand transfer activities mediated by the MuA transposase (Baker et al. 1991; Surette et al. 1991; Surette and Chaconas 1991). It is tempting to speculate that TnsC and TnsD may play a similar role in stimulating the transposase activity, possibly attributed to TnsA and/or TnsB.

A model for Tn7 target site selection: Alternative multiprotein complexes direct insertion into two classes of target sites

Our results suggest that TnsD, bound to attTn7, can stably position TnsC onto the target DNA. In addition, TnsA + TnsB + TnsC contain the essential active site(s) for Tn7 insertion, i.e., these components with AMP-PNP can promote cleavage at the Tn7 ends and the joining of the ends to the target DNA. Thus, during Tn7 insertion into attTn7, TnsC is likely to mediate the interaction between the

transposition machinery and TnsD, which provides target site specificity for attTn7.

The interactions between TnsC and TnsD with attTn7 that promote site-specific Tn7 insertion suggest a related mechanism for TnsE-mediated transposition to random target sites. An attractive model is that TnsE, like TnsD, positions TnsC onto the target DNA to promote transposition. In this view, TnsE is a target DNA binding protein, but displays little target site selectivity in contrast to the extreme specificity of TnsD. Through this TnsC-TnsE interaction, the transposition machinery is directed to the target DNA. Another alternative is that TnsE does not directly contact the target DNA, but instead causes TnsC to bind stably to the target DNA. In this view, target selection is determined by the non-sequence-specific DNA binding of TnsC, which directs insertion into many different target sites.

Alternative mulitprotein-DNA complexes operate in other specialized DNA transactions

Alternative multicomponent complexes are also used to regulate DNA replication and transcription. The combinatorial assembly of different protein components can generate multicomponent complexes with varied functions, particularly in DNA binding specificity.

In E. coli, several plasmid and phage replicons encode their own origin-binding proteins to recruit the DnaB-DnaC protein complex (reviewed in Bramhill and Kornberg 1988). Once assembled onto the origin, the host replication machinery can then load onto the replicon. In several cases, plasmid replicons also utilize the E. coli DnaA

protein in concert with their own origin-binding proteins, called Rep proteins, to direct replication initiation. The combination of host-encoded and phage or plasmid-encoded replication proteins generates initiation complexes with different origin-binding specificities.

Transcription complexes can also alter their DNA binding specificity by varying their modular arrangement of DNA binding proteins. An extensively studied example of this DNA binding specificity occurs in the control of yeast mating type. In yeast cells that display the α mating type, **a**-specific genes are repressed, whereas α-specific genes are expressed (reviewed in Nasmyth and Shore, 1987). The regulatory sites of the α-specific and **a**-specific genes are recognized by the combinatorial interactions of the pheromone receptor transcription factor (PRTF) with one of the α-specific factors, MAT-α1 or MAT-α2. PRTF and MAT-α1 bind cooperatively to the upstream activation sequences of α-specific genes to stimulate expression of these genes (Bender and Sprague 1987; Tan et al. 1988; Tan and Richmond 1990). In a similar manner, PRTF and MAT-α2 interact cooperatively with the operators of **a**-specific genes; in this case, binding to the operator represses expression of these genes (Keleher et al. 1988, 1989). Thus, these two protein complexes, PRTF + MAT-α1 or PRTF + MAT-α2, produce distinct DNA binding specificities that culminate in either the activation of α-specific genes or the repression of **a**-specific genes.

Concluding remarks

It is clear from the studies in this thesis that the target DNA information for Tn7 insertion is mediated by an elaborate set of

protein-DNA and protein-protein interactions. Future dissection of these interactions will provide further insights into the communications between the ensemble of proteins and DNAs in Tn7 transposition and, more generally, contribute to our understanding of other complex protein-DNA transactions.

References

Adhya, S.L., and J.A. Shapiro. 1969. The galactose operon of E. coli K-12. I. Structural and pleiotropic mutations of the operon. *Genetics* 62:231-247.

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

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

Adzuma, K., and K. Mizuuchi. 1991. Steady-state kinetic analysis of ATP hydrolysis by the B protein of bacteriophage Mu. *J. Biol. Chem.* 266:6159-6167.

Ajioka, J.W., and D.L. Hartl. 1989. Population dynamics of transposable elements. In: *Mobile DNA*. (ed. D.E. Berg and M.M. Howe), pp. 939-958. American Society for Microbiology, Washington, D.C.

Akroyd, J., and N. Symonds. 1983. Evidence for a conservative pathway of transposition of bacteriophage Mu. *Nature* 303:84-86.

Amemura, J., H. Ichikawa, and E. Ohtsubo. 1990. Tn3 transposition immunity is conferred by the transposase-binding domain in the terminal inverted-repeat sequence of Tn3. *Gene* 88:21-24.

Andrews, B.J., L.G. Beatty, and P.D. Sadowski. 1987. Isolation of intermediates in the binding of the FLP recombinase of the yeast plasmid 2-micron circle to its target sequence. *J. Mol. Biol.* 193:345-358.

Ansari, A.Z., M.L. Chael, and T.V. O'Halloran. 1992. Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR. *Nature* 355:87-89.

Arciszewska, L.K., and N.L. Craig. 1991. Interaction of the Tn7-encoded transposition protein TnsB with the ends of the transposon. *Nucl. Acids Res.* 19:5021-5029.

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.

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.* 266:21736-21744.

Arthur, A., and D. Sherratt. 1979. Dissection of the transposition process: a transposon-encoded site-specific recombination system. *Mol. Gen. Genet.* 175:267-274.

Baker, T.A., M. Mizuuchi, and K. Mizuuchi. 1991. MuB protein allosterically activates strand transfer by the transposase of phage Mu. *Cell* 65:1003-1013.

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.* 125:800-810.

Barth, P.T., and N. Grinter. 1977. Map of plasmid RP4 derived by insertion of transposon C. *J. Mol. Biol.* 113:455-474.

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.

Bender, A., and G.F. Sprague, Jr. 1987. MATα1 Protein, a yeast transcription activator, binds synergistically with a second protein to a set of cell-type-specific genes. *Cell* 50:681-691.

Bender, J., and N. Kleckner. 1986. Genetic evidence that Tn10 transposes by a nonreplicative mechanism. *Cell* 45:801-815.

Bender, J., and N. Kleckner. 1992. IS10 transposase mutations that specifically alter target site recognition. *EMBO J.* 11:741-750.

Bender, J., J. Kuo, and N. Kleckner. 1991. Genetic evidence against intramolecular rejoining of the donor DNA molecule following IS10 transposition. *Genetics* 128:687-694.

Benjamin, H.W., and N. Kleckner. 1989. Intramolecular transposition by Tn10. *Cell* 59:373-383.

Berg, C.A., and A.C. Spradling. 1991. Studies on the rate and site-specificity of P element transposition. *Genetics* 127:515-524.

Berg, C.M., D.E. Berg, and E.A. Groisman. 1989. Transposable elements and the genetic engineering of bacteria. In: *Mobile DNA* (D.E. Berg and M.M. Howe, eds.), pp. 879-926. American Society for Microbiology, Washington, D.C.

Berg, D.E. Insertion and excision of the transposable kanamycin-resistance determinant Tn5. In: *DNA insertion elements, plasmids, and episomes* (ed. A.I. Bukhari et al.), pp. 205-212. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Berg, D.E., and M.M. Howe, ed. 1989. *Mobile DNA*. American Society for Microbiology, Washington, D.C.

Berg, D.E., M.A. Schmandt, and J.B. Lowe. 1983. Specificity of transposon Tn5 insertion. *Genetics* 105:813-828

Bingham, P.M., and Z. Zachar. 1989. Retrotransposons and the FB transposon from Drosophila melanogaster. In: *Mobile DNA*. (ed. D.E. Berg and M.M. Howe), pp. 485-502. American Society for Microbiology, Washington, D.C.

Birkenbihl, R.P., and W. Vielmetter. 1989. Complete maps of IS1, IS2, IS3, IS4, IS5, IS30 and IS150 locations in Escherichia coli K12. *Mol. Gen. Genet.* 220:147-153.

Boeke, J.D. 1989. Transposable elements in Saccharomyces cerevesiae. In: *Mobile DNA*. (ed. D.E. Berg and M.M. Howe), pp. 335-374. American Society for Microbiology, Washington, D.C.

Bowerman, B., P.O. Brown, J.M. Bishop, and H.E. Varmus. 1989. A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev.* 3:469-478.

Bramhill, D., and A. Kornberg. 1988. A model for initiation of origins of DNA replication. *Cell* 54:915-918.

Braun, B.R., D.L. Riggs, G.A. Kassavetis, and E.P. Geiduschek. 1989. Multiple states of protein-DNA interaction in the assembly of transcription complexes on Saccharomyces cerevisiae 5S ribosomal RNA genes. *Proc. Natl. Acad. Sci. USA* 86:2530-2534.

Brown, P.O., B. Bowerman, H.E. Varmus, and J.M. Bishop. 1987. Correct integration of retroviral DNA in vitro. *Cell* 49:347-356.

Bushman, F.D., and R. Craigie. 1991. Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. *Proc. Natl. Acad. Sci. USA* 88:1339-1343.

Bushman, F.D., T. Fujiwara, and R. Craigie. 1990. Retroviral DNA integration directed by HIV integration protein in vitro. *Science* 249:1555-1558.

Castilho, B.A. and M.J. Casadaban. 1991. Specificity of mini-Mu bacteriophage insertions in a small plasmid. *J. Bacteriol.* 173:1339-1343.

Chaconas, G., G. Gloor, and J.L. Miller. 1985a. Amplification and purification of the bacteriophage Mu encoded B transposition protein. *J. Biol. Chem.* 260:2662-2669.

Chaconas, G., E.B. Giddens, J.L. Miller, and G. Gloor. 1985b. A truncated form of the bacteriophage Mu B protein promotes conservative integration, but not replicative transposition, of Mu DNA. *Cell* 41:857-865.

Chalker, D.L. and S.B. Sandmeyer. 1992. Ty3 integrates within the region of RNA polymerase III transcription initiation. *Genes Dev.* 6:117-128.

Chao, L., E. Vargas, B.B. Spear, and E.C. Cox. 1983. Transposable elements as mutator genes in evolution. *Nature* 303:633-635.

Chen, J., E.M. Greenblatt, and S.L. Dellaporta. 1987. Transposition of Ac from the P locus of maize into unreplicated chromosomal sites. *Genetics* 117:109-116.

Chen, J., I.M. Greenblatt, and S.L. Dellaporta. 1992. Molecular analysis of Ac transposition and DNA replication. *Genetics* 130:665-676.

Chou, S.A., K.A. Vincent, V. Ellison, and P.O. Brown. 1992. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science* 255:723-726.

Coen, E.S., T.P. Robbins, J. Almeida, A. Hudson, and R. Carpenter. 1989. Consequences and mechanisms of transposition in Antirrhinum majus. In: *Mobile DNA*. (ed. D.E. Berg and M.M. Howe), pp. 413-436. American Society for Microbiology, Washington, D.C.

Cortay, J.C., F. Bleicher, C. Rieul, H.C. Reeves, and A.J. Cozzone. 1988. Nucleotide sequence and expression of the aceK gene coding for isocitrate dehydrogenase kinase/phosphatase in Escherichia coli. *J. Bacteriol.* 170:89-97.

Craig, N.L. 1988. The mechanism of conservative site-specific recombination. *Annu. Rev. Genet.* 22:77-105.

Craig, N.L. 1989. Transposon Tn7, p. 211-225. In: *Mobile DNA* (ed. D.E. Berg and M.M Howe). American Society for Microbiology, Washington, D.C.

Craig, N.L., and N. Kleckner. 1987. Transposition and site-specific recombination. In *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology* (F.C. Neidhardt ed.), pp.1054-1070. American Society for Microbiology, Washington, D.C.),

Craig, N.L., and H.A. Nash. 1983. The mechanism of phage λ site-specific recombination: site-specific breakage of DNA by Int topoisomerase. *Cell* 35:795-803.

Craigie, R., and K. Mizuuchi. 1985. Mechanism of transposition of bacteriophage Mu: structure of a transposition intermediate. *Cell* 41:867-876.

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.

Craigie, R., D.J. Arndt-Jovin, and K. Mizuuchi. 1985. A defined system for the DNA strand-transfer reaction at the initiation of bacteriophage Mu transposition: Protein and DNA substrate requirements. *Proc. Natl. Acad. Sci. USA* 82.:7670-7574.

Craigie, R., T. Fujiwara, and F. Bushman. 1990. The IN protein of moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. *Cell* 62:829-837.

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

Derbyshire, K.M., and N.D.F. Grindley. 1986. Replicative and conservative transposition in bacteria. *Cell* 47:325-327.

Drlica, K., and J. Rouviere-Yaniv. 1987. Histonelike proteins of bacteria. *Microbiol. Rev.* 51:301-319.

Dombroski, B.A., S.L. Mathias, E. Nanthakumar, A.F. Scott, and H.H. Kazazian, Jr. Isolation of an active human transposable element. *Science* 254:1805-1807.

Doolittle, W.F. and C. Sapienza. 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284:601-603.

Doolittle, W.F., T.B.L. Kirkwood, and M.A.H. Dempster. 1984. Selfish DNAs with self-restraint. *Nature* 307:501-502.

Echols, H. 1990. Nucleoprotein structures initiating DNA replication, transcription, and site-specific recombination. *J. Biol. Chem.* 265:14697-14700.

Ellison, V., H. Abrams, T. Roe, J. Lifson, and P. Brown. 1990. Human immunodeficiency virus integration in a cell-free system. *J. Virol.* 64:2711-2715.

Engels, W.R. 1989. P elements in Drosophila melanogaster. In: *Mobile DNA*. (ed. D.E. Berg and M.M. Howe), pp. 437-484. American Society for Microbiology, Washington, D.C.

Engels, W.R., D.M. Johnson-Schlitz, W.B. Eggleston, and J.Sved. 1990. High-frequency P element loss in Drosophila is homolog dependent. *Cell* 62:515-525.

Elgin. S.C. 1988. The formation and function of DNase I hypersensitive sites in the process of gene activation. *J. Biol. Chem.* 263:19259-19262.

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

Escoubas, J.M., M.F. Prere, O. Fayet, I. Salvignol, D. Galas, D. Zerbib, and M. Chandler. 1991. Translational control of transposition activity of the bacterial insertion sequence IS1. *EMBO J.* 10:705-712.

Fanning, T.G., and M.F. Singer. 1987. LINE-1: a mammalian transposable element. *Biochim. Biophys. Acta* 910:203-212.

Federoff, N. 1989. Maize transposable elements. In: *Mobile DNA*. (ed. D.E. Berg and M.M. Howe), pp.375-412. American Society for Microbiology, Washington, D.C.

Finnegan, D.J. 1989a. Eukayotic transposable elements and genome evolution. *Trends Genet*. 5:103-107.

Finnegan, D.J. 1989b. The I factor and I-R hybrid dysgenesis in Drosophila melanogaster. In: *Mobile DNA*. (ed. D.E. Berg and M.M. Howe), pp. 503-518. American Society for Microbiology, Washington, D.C.

Finnegan, D.J., and D.H. Fawcett. 1986. Transposable elements in Drosophila melanogaster. *Oxf. Surv. Eukaryotic Genes* 3:1-62.

Fling, M., J. Kopf, and C. Richards. 1985. Nucleotide sequence of the transposon Tn7 gene encoding an aminoglycoside-modifying enzyme, 3"(9)-O-nucleotidyltransferase. *Nucleic Acids Res*. 13:7095-7106.

Fling, M., and C. Richards. 1983. The nucleotide sequence of the trimethoprim-resistant dihydrofolate reductase gene harbored by Tn7. *Nucleic Acids Res*. 11:5147-5158.

Fried, M., and D.M. Crothers. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucl. Acids Res*. 9:6505-25.

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

Fujiwara, T. and K. Mizuuchi. 1988. Retroviral DNA integration: structure of an integration intermediate. *Cell* 54:497-504.

Galas, D.J., M.P. Calos, and J.H. Miller. 1980. Sequence analysis of Tn9 insertions in the lacZ gene. *J. Mol. Biol*. 144:19-41.

Galas, D.J., and M. Chandler. 1989. Bacterial insertion sequences. In: *Mobile DNA*. (ed. D.E. Berg and M.M. Howe), pp.109-162. American Society for Microbiology, Washington, D.C.

Gamas, P., M.G. Chandler, P. Prentki, and D.J. Galas. 1987. Escherichia coli integration host factor binds specifically to the ends of the insertion sequence IS1 and to its major insertion hot-spot in pBR322. *J. Mol. Biol*. 195:261-272.

Gamas, P., and N.L. Craig. 1992. Purification and characterization of TnsC, a Tn7 transposition protein that binds ATP and DNA. *Nucl. Acids Res*. (in press).

Garner, M.M., and A. Revzin. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. *Nucl. Acids Res.* 9:3047-3060.

Gay, N.J., V.L.J. Tybulewicz, and J.E. Walker. 1986. Insertion of transposon Tn7 into the Escherichia coli glmS transcriptional terminator. *Biochem J.* 234:111-117.

Geiduschek, E.P., and G.P. Tocchini-Valentini. 1988. Transcription by RNA polymerase III. *Annu. Rev. Biochem.* 57:873-914.

Goldstein, F.W., B. Papadopoulou, and J.F. Acar. 1986. The changing pattern of trimethoprim resistance in Paris, with a review of worldwide experience. *Rev. Infect. Dis.* 8:725-737.

Ghrayeb, J., H. Kimura, M. Takahara, H. Hsiung, Y. Masui, and M. Inouye. 1984. Secretion cloning vectors in Escherichia coli. *EMBO J.* 3:2437-2442.

Glasgow, A.C., K.T. Hughes, and M.I. Simon. 1989. Bacterial DNA inversion systems. In: *Mobile DNA.* (ed. D.E. Berg and M.M. Howe), pp. 637-660. American Society for Microbiology, Washington, D.C.

Greenblatt, J. 1991. RNA polymerase-associated transcription factors. *Trends Biochem Sci.* 16:408-411.

Greenblatt, I.M. 1984. A chromosomal replication pattern deduced from pericarp phenotypes resulting from movements of the transposable element, Modulator, in maize. *Genetics* 108:471-485.

Greenblatt, J. 1991. Roles of TFIID in transcriptional initiation by RNA polymerase II. *Cell* 66:1067-1070.

Greenblatt, I.M., and R.A. Brink. 1962. Twin mutations in medium variegated pericarp maize. *Genetics* 47:489-501.

Grindley, N.D.F. 1983. Transposition of Tn3 and related transposons. *Cell* 32:3-5.

Grindley, N.D.F., and R.R. Reed. 1985. Transpositional recombination in prokaryotes. *Ann. Rev. Biochem.* 54:863-896.

Grindley, N.D.F., and D.J. Sherratt. 1978. Sequence analysis of IS1 insertion sites: models for transposition. *Cold Spring Harbor Symp. Quant. Biol.* 43:1257-1261.

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.* 170:2832-2840.

Gross, D.S. and W.T. Garrard. 1988. Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* 57:159-97.

Habermann, P., R. Klaer, S. Kuhn, and P. Starlinger. IS4 is found between eleven or twelve base pair duplications. *Mol. Gen. Genet.* 175:369-373.

Halling, S.M. and N. Kleckner. 1982. A symmetrical six-base-pair target site sequence determines Tn10 insertion specificity. *Cell* 28:155-163.

Haniford, D.B., H.W. Benjamin, and N. Kleckner. 1991. Kinetic and structural analysis of a cleaved donor intermediate and a strand transfer intermediate in Tn10 transposition. *Cell* 64:171-179.

Haniford, D.B., A.R. Chelouche, and N. Kleckner. 1989. A specific class of IS10 transposase mutants are blocked for target site interactions and promote formation of an excised transposon fragment. *Cell* 59:385-394.

Harshey, R. 1984. Transposition without duplication of infecting bacteriophage Mu DNA. *Nature* 311:580-581.

Harshey, R.M., and S.D. Cuneo. 1986. Carboxyl-terminal mutants of phage Mu transposase. *J. Genet.* 65:159-174.

Hauer, B. and J.A. Shapiro. 1984. Control of Tn7 transposition. *Mol. Gen. Genet.* 194:149-158.

Hawley, R.S., R.A. Steuber, C.H. Marcus, R. Sohn, D.M. Baronas, M.L. Cameron, A.E. Zitron, and J.W. Chase. 1988. Molecular analysis of an unstable P element insertion at the singed locus of Drosophila melanogaster: evidence for intracistronic transposition of a P element. *Genetics* 119:85-94.

Hedges, R., N. Datta, and M.P. Fleming. 1972. R factors conferring resistance to trimethoprim but not sulphonamides. *J. Gen. Microbiol.* 73:573-575.

Hershey, J.W.B. 1991. Translational control in mammalian cells. *Annu. Rev. Biochem.* 60:717-755.

Hutchison, C.A., S.C. Hardies, D.D. Loeb, W.R. Shehee, and M.H. Edgell. 1989. LINEs and related retroposons: long interspersed repeated sequences in the eucaryotic genome. In: *Mobile DNA.* (ed. D.E. Berg and M.M. Howe), pp. 593-618. American Society for Microbiology, Washington, D.C.

Isberg, R.R,. and M. Syvanen. 1982. DNA gyrase is a host factor required for transposition of Tn5. *Cell* 30:9-18.

Jones, J.D.G., F. Carland, E. Lim, E. Ralston, and H.K. Dooner. 1990. Preferential transposition of the maize element Activator to linked chromosomal locations in tobacco. *Plant Cell* 2:701-707.

Jordan, E., H. Saedler, and P. Starlinger. 1967. Strong-polar mutations in the transferase gene of the galactose operon in E. coli. *Mol. Gen. Genet.* 100:296-306.

Kamp, D., and R. Kahmann. 1981. Two pathways in bacteriophage Mu transposition? *Cold Spring Harbor Symp. Quant. Biol.* 45:329-336.

Kassavetis, G.A., B.R. Braun, L.H. Nguyen, and E.P. Geiduschek. 1990. S. cerevisiae TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell* 60:235-245.

Keleher, C.A., C. Goutte, and A.D. Johnson. 1988. The yeast cell-type-specific repressor α2 acts cooperatively with a non-cell-type-specific protein. *Cell* 53:927-936.

Keleher, C.A., S. Passmore, and A.D. Johnson. 1989. Yeast repressor α2 binds to its operator cooperatively with yeast protein Mcm1. *Mol. Cell. Biol.* 9:5228-5230.

Klaer, R., and P. Starlinger. 1980. IS4 at its chromosomal site in E. coli K12. *Mol. Gen. Genet.* 178:285-291.

Klaer, R., S. Kuhn, E. Tillmann, H.J. Frits, and P. Starlinger. The sequence of IS4. *Mol. Gen. Genet.* 181:169-175.

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

Kleckner, N. 1990. Regulation of transposition in bacteria. *Annu. Rev. Cell Biol.* 6:297-327.

Kleckner, N., J. Bender, and S. Gottesman. 1991. Uses of transposons with emphasis on Tn10. *Methods Enzymol.* 204:139-180.

Klumpp, D.J., D.W. Plank, L.J. Bowdin, C.S. Stueland, T. Chung, D.C. LaPorte. 1988. Nucleotide sequence of aceK, the gene encoding isocitrate dehydrogenase kinase/phosphatase. *J. Bacteriol.* 170:2763-2769.

Kopecko, D.J. 1980. Specialized genetic recombination systems in bacteria: their involvement in gene expression and evolution. *Prog. Mol. Subcell. Biol.* 7:135-234.

Kornberg, A. 1988. DNA replication. *J. Biol. Chem.* 263:1-4.

Krishnapillai, V., M. Wexler, J. Nash, and D.H. Figurski. 1987. Genetic basis of a Tn7 insertion mutation in the trfA region of the promiscuous IncP-1 plasmid R18 which affects its host range. *Plasmid* 17:164-166.

Kristie, T.M., and P.A. Sharp. 1990. Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV α-trans-activator protein. *Genes Dev.* 4:2383-2396.

Landy, A. 1989. Dynamic, structural, and regulatory aspects of λ site-specific recombination. *Annu. Rev. Biochem.* 58:913-949.

Lavoie, B.D., and G. Chaconas. 1989. Immunoelectron microscopic analysis of the A, B, and HU protein content of bacteriophage Mu transpososomes. *J. Biol Chem.* 265:1623-1627.

Lee, C.-H., A. Bhagwat, and F. Heffron. 1983. Identification of a transposon Tn3 sequence required for transposition immunity. *Proc. Natl. Acad. Sci. USA* 80:6765-6769.

Lee, S.Y., D. Butler, and N. Kleckner. 1987. Efficient Tn10 transposition into a DNA insertion hot spot in vivo requires the 5'methyl groups of symmetrically disposed thymines within the hot-spot consensus sequence. *Proc. Natl. Acad. Sci. USA* 84:7876-7880.

Lichtenstein, C., and S. Brenner. 1981. Site-specific properties of Tn7 transposition into the E. coli chromosome. *Mol. Gen. Genet.* 183:380-387.

Lichtenstein, C., and S. Brenner. 1982. Unique insertion site of Tn7 in the E. coli chromosome. *Nature (London)* 297:601-603.

Lin, Y.-S., and M.R. Green. 1991. Mechanism of action of an acidic transcriptional activator in vitro. *Cell* 64:971-981.

Lodge, J.K., K. Weston-Hafer, and D.E. Berg. 1988. Transposon Tn5 target specificity: preference for insertion at G/C pairs. *Genetics* 120:645-650.

Lodge, J.K., and D.E. Berg. 1990. Mutations that affect Tn5 insertion into pBR322: importance of local DNA supercoiling. *J. Bacteriol.* 172:5956-5960.

Malamy, M.H. 1966. Frameshift mutations in the lactose operon of E. coli. *Cold Spring Harbor Symp. Quant. Biol.* 31:189-201.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Martin, K.J. 1991. The interactions of transcription factors and their adaptors, coactivators and accessory proteins. *Bioessays* 13:499-503.

Mathias, S.L., A.F. Scott, H.H. Kazazian, Jr., J.D. Boeke, and A. Gabriel. 1992. Reverse transcriptase encoded by a human transposable element. *Science* 254:1808-1810.

Maxam, A.M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Meth. Enzymol.* 65:499-599.

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

Mayaux, J.-F., M. Springer, M. Graffe, M. Fromant, and G. Fayat. 1984. IS4 transposon in the attenuator region of the Escherichia coli pheS,T operon. *Gene* 30:137-146.

McCarthy, J.E.G., and C. Gualerzi. 1990. Translational control of prokaryotic gene expression. *Trends Genet.* 6:78-85.

McClintock, B. 1951. Chromosome organization and genic expression. *Cold Spring Harbor Symp. Quant. Biol.* 16:13-47.

McClintock, B. 1956. Controlling elements and the gene. *Cold Spring Harbor Symp. Quant. Biol.* 21:197-216.

McKnight, J.L.C., T.M. Kristie, S. Silver, P.E. Pellett, P. Mavromara-Nazos, G. Campadelli-Fiume, M. Arsenakis, and B. Roizman. 1986. Regulation of herpes simplex virus 1 gene expression: The effect of genomic environments and its implications for model systems. *Cancer Cells* 4:163-173.

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.* 170:352-358.

Miller, J. 1972. *Experiments in molecular genetics.* Cold Spring Harbor Laboratory, Cold Spring harbor, New York.

Miller, J.H., M.P. Calos, D. Galas, and M. Hofer. 1980. Genetic analysis of transpositions in the lac region of Escherichia coli. *J. Mol. Biol.* 144:1-18.

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

Mizuuchi, K. 1984. Mechanism of transposition of bacteriophage Mu: Polarity of the strand transfer reaction at the initiation of transposition. *Cell* 39:396-404.

Mizuuchi, K., and R. Craigie. 1986. Mechanism of bacteriophage Mu transposition. *Annu. Rev. Genet.* 20:385-429.

Mizuuchi, K., and K. Adzuma. 1991. Inversion of the phosphate chirality at the target site of Mu DNA strand transfer: evidence for a one-step transesterification mechanism. *Cell* 66:129-140.

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

Mizuuchi, M., T.A. Baker, and K. Mizuuchi. 1991. DNase protection analysis of the stable synaptic complexes involved in Mu transposition. *Proc. Natl. Acad. Sci. USA* 88:9031-9135.

Moore, R.J., and V. Krishnapillai. 1982. Tn7 and Tn501 insertions into the Pseudomonas aeruginosa plasmid R91-5: mapping of two transfer regions. *J. Bacteriol.* 149:276-283.

Mooslehner, K., U. Karls, and K. Harbers. 1990. Retroviral integration sites in transgenic Mov mice frequently map in the vicinity of transcribed DNA regions. *J. Virol.* 64:3056-3058.

Morisato, D., and N. Kleckner. 1984. Transposase promotes double strand breaks and single strand joints at Tn10 termini in vivo. *Cell* 39:181-190.

Morisato, D., and N. Kleckner. 1987. Tn10 transposition and circle formation in vitro. *Cell* 51:101-111.

Nash, J., and V. Krishnapillai. 1987. DNA sequence analysis of host range mutants of the promiscuous IncP-1 plasmids R-18 and R68 with Tn7 insertions in oriV. *Plasmid* 18:35-45.

Nasmyth, K., and D. Shore. 1987. Transcriptional regulation in the yeast life cycle. *Science* 237:1162-1170.

Natsoulis, G., W. Thomas, M.C. Roghmann, F. Winston, and J.D. Boeke. 1989. Tyl transposition in Saccharomyces cerevisiae is nonrandom. *Genetics* 123:269-279.

Nowick, E.M., and P.A. Peterson. 1981. Transposition of the enhancer controlling element system in maize. *Mol. Gen. Genet.* 183:440-448.

O'Hare, K., and G.M. Rubin. 1983. Structure of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. *Cell* 34:25-35.

Ohtsubo, E., M. Zenilman, H. Ohtsubo, M. McCormick, C. Machida, and Y. Machida. 1981. Mechanism of insertion and cointegration mediated by IS1 and Tn3. *Cold Spring Harbor Symp. Quant. Biol.* 45:283-295.

Orle, K.A., and N.L. Craig. 1991. Identification of transposition proteins encoded by the bacterial transposon Tn7. *Gene* 104:125-131.

Orgel, L.E., and F.H.C. Crick. 1980. Selfish DNA: the ultimate parasite. *Nature* 284:604-607.

Pato, M., and B. Waggoner. 1987. Replication of Mu DNA in vivo. In: *Phage Mu* (ed. N. Symonds, A. Toussaint, P. van de Putte, and M. Howe), pp. 177-189. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Plasterk, R.H.A. 1991. Frameshift control of IS1 transposition. *Trends Genet.* 7:203-204.

Prentki, P., M. Chandler, and D.J. Galas. 1987. Escherichia coli integration host factor bends the DNA at the ends of IS1 and in an insertion hotspot with multiple IHF binding sites. *EMBO J.* 6:2479-2487.

Pryciak, P.M., A. Sil, and H. Varmus. 1992. Retroviral integration into minichromosomes in vitro. *EMBO J.* 11:291-303.

Qadri, M.I., C.C. Flores, A.J. Davis, and C.P. Lichenstein. 1989. Genetic analysis of attTn7, the transposon Tn7 attachment site in Escherichia coli, using a novel M13-based transduction assay. *J. Mol. Biol.* 207:85-98.

Raleigh, E.A., and N. Kleckner. 1984. Multiple IS10 rearrangements in E. coli. *J. Mol. Biol.* 173:437-461.

Raymond, J.D., and M.J. Simmons. 1981. An increase in the X-linked lethal mutation rate associated with an unstable locus in Drosophila melanogaster. *Genetics* 98:291-302.

Richet, E., P. Abcarian, and H.A. Nash. 1988. Synapsis of attachment sites during lambda integrative recombination involves capture of a naked DNA by a protein-DNA complex. *Cell* 52:9-17.

Rogers, M., N. Ekaterinaki, E. Nimmo, and D. Sherratt. 1986. Analysis of Tn7 transposition. *Mol. Gen. Genet.* 205:550-556.

Rohdewohld, H., H. Weiher, W. Reik, R. Jaenisch, and M. Breindl. 1987. Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNase I-hypersensitive sites. *J. Virol.* 61:336-343.

Roiha, H., G.M. Rubin, and K. O'Hare. 1988. P element insertions and rearrangements at the singed locus of Drosophila melanogaster. *Genetics* 119:75-83.

Rosenberg, J.M., J.A. McClarin, C.A. Frederick, B.-C. Wang, J. Grable, H.W. Boyer, and P. Greene. 1987. Structure and recognition mechanism of EcoRI endonuclease. *Trends Biochem. Sci.* 12:395-8.

Ross, W., and A. Landy. 1983. Patterns of λ Int recognition in the regions of strand exchange. *Cell* 33:261-272.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.

Scherdin, U., K. Rhodes, and M. Breindl. 1990. Transcriptionally active genome regions are preferred targets for retrovirus integration. *J. Virol.* 64:907-912.

Sekine, Y., and E. Ohtsubo. 1989. Frameshifting is required for the transposase encoded by insertion sequence 1. *Proc. Natl. Acad. Sci. USA* 86:4609-4613.

Shand, R.F., L.J.W. Miercke, A.K. Mitra, S.K. Fong, R.M. Stroud, and M.C. Betlach. 1991. Wild-type and mutant bacterioopsins D85N, D96N, and R82Q: high-level expression in Escherichia coli. *Biochemistry* 30:3082-3088.

Shapiro, J.A., and S.L. Adhya. 1969. The galactose operon of E. coli K-12. II. A deletion analysis of operon structure and polarity. *Genetics* 62:249-264.

Sherratt, D. 1991. Jumping genes. *Current Biol.* 1:192-194.

Shih, C.-C., J.P. Stoye, and J.M. Coffin. 1988. Highly preferred targets for retroviral integration. *Cell* 53:531-537.

Shimada, K., R.A. Weisberg, and M.E. Gottesman. 1972. Prophage lambda at unusual chromosomal locations I. Location of the secondary attachment sites and the properties of the lysogens. *J. Mol. Biol.* 63:483-503.

Shimada, K., R.A. Weisberg, and M.E. Gottesman. 1973. Prophage lambda at unusual chromosomal locations II. Mutations induced by bacteriophage lambda in Escherichia coli K12. *J. Mol. Biol.* 80:297-314.

Simmons, M.J., J.D. Raymond, N.A. Johnson, and T.M. Fahey. 1984. A comparison of mutation rates for specific loci and chromosome regions in dysgenic hybrid males of Drosophila melanogaster. *Genetics* 106:85-94.

Simonsen, C.C., E.Y. Chen, and A.D. Levinson. 1983. Identification of the type 1 trimethoprim-resistant dihydrofolate reductase specified by the Escherichia coli R-plasmid R483: comparison with procaryotic and eucaryotic dihydrofolate reductases. *J. Bacteriol.* 155:1001-1008.

Smith, D.B., and K.S. Johnson. 1988. Single-step purification of polypeptides in Escherichia coli as fusions with glutathione-S-transferase. *Gene* 67:31-40.

Soriano, P., T. Gridley, and R. Jaenisch. 1989. Retroviral tagging in mammalian development and genetics. In: *Mobile DNA* (D.E. Berg and M.M. Howe, eds.), pp. 927-937. American Society for Microbiology, Washington, D.C.

Stringer, K.F., C.J. Ingles, and J. Greenblatt. 1990. Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature* 345:783-786.

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

Surette, M.G. and G. Chaconas. 1991. Stimulation of the Mu DNA strand cleavage and intramolecular strand transfer reactions by the Mu B protein is independent of stable binding of the Mu B Protein to DNA. J. *Biol. Chem.* 266:17306-17313.

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* 49:253-262.

Surette, M.G., T. Harkness, and G. Chaconas. 1991. Stimulation of the Mu A protein-mediated strand cleavage reaction by the Mu B Protein, and the requirement of DNA nicking for stable type I transpososome formation. *J. Biol. Chem.* 266:3118-3124.

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

Syvanen, M. 1984. The evolutionary implications of mobile genetic elements. *Annu. Rev. Genet.* 18:271-293.

Tan, S., G. Ammerer, and T.J. Richmond. 1988. Interactions of purified transcription factors: binding of yeast MATa1 and PRTF to cell type-specific, upstream activating sequences. *EMBO J.* 7:4255-4264.

Tan, S., and T.J. Richmond. 1990. DNA binding-induced conformational change of the yeast transcriptional activator PRTF. *Cell* 62:367-377.

Tang, Y., C. Lichtenstein, and S. Cotterill. 1991. Purification and characterisation of the TnsB protein of Tn7: a transposition protein that binds to the ends of Tn7. *Nucl. Acids Res.* 19:3395-3402.

Vijaya, S., D.L. Steffen, and H.L. Robinson. 1986. Acceptor sites for retroviral integrations map near DNase I-hypersensitive sites in chromatin. *J. Virol.* 60:683-692.

Waddell, C.S., and N.L. Craig. 1988. Tn7 transposition: two transposition pathways directed by five Tn7-encoded genes. *Genes. Dev.* 2:137-149.

Waddell, C.S., and N.L. Craig. 1989. Tn7 transposition: recognition of the attTn7 target sequence. *Proc. Natl. Acad. Sci. USA* 86:3958-3962.

Wahle, E., R.S. Lasken, and A. Kornberg. 1989a. The dnaB-dnaC replication protein complex in Escherichia coli. I. Formation and properties. *J. Biol Chem.* 264:2463-2468.

Wahle, E., R.S. Lasken, and A. Kornberg. 1989b. The dnaB-dnaC replication protein complex of Escherichia coli. II. Role of the comples in mobilizing dnaB functions. *J. Biol. Chem.* 264:2469-2475.

Walker, J., N. Gay, M. Saraste, and A. Eberle. 1984. DNA sequence around the Escherichia coli unc operon. *Biochem J.* 224:799-815.

Weinert, T.A., K.M. Derbyshire, F.M. Hughson, and N.D.F. Grindley. 1984. Replicative and conservative transpositional recombination of insertion sequences. *Cold Spring Harbor Symp. Quant. Biol.* 49:251-260.

Wiater, L.A., and N.D.F. Grindley. 1988. γδ transposase and integration host factor bind cooperatively at both ends of γδ. *EMBO J.* 7:1907-1911.

Wiater, L.A., and N.D.F. Grindley. 1990. Uncoupling of transpositional immunity from γδ transposition by a mutation at the end of γδ. *J. Bacteriol.* 172:4959-4963.

Wiater, L.A., and N.D.F. Grindley. 1990. Integration host factor increases the transpositional immunity conferred by γδ ends. *J. Bacteriol.* 172:4951-4958.

Wiater, L.A., and N.D.F. Grindley. 1991. γδ transposase: purification and analysis of its interaction with a transposon end. *J. Biol. Chem.* 266:1841-1849.

Wilke, C.M., S.H. Heidler, N. Brown, and S.W. Liebman. 1989. Analysis of yeast retrotransposon Ty insertions at the CAN1 locus. *Genetics* 123:655-665.

Yin, J., and W. Reznikoff. 1988. p2 and inhibition of Tn5 transposition. *J. Bacteriol.* 170:3008-3015.

Zerbib, D., P. Gamas, M. Chandler, P. Prentki, S. Bass, and D. Galas. 1985. Specificity of insertion of IS1. *J. Mol. Biol.* 185:517-524.