**Title**
Tn7 transposition in vitro



**Author**
Bainton, Roland J.

**Publication Date**
1992

Peer reviewed|Thesis/dissertation

# Tn7 Transposition *In Vitro*

by

Roland J. Bainton

**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**

I dedicate this thesis to my family, past, present and future.

## Acknowledgments

I want to thank Nancy L. Craig for her time, insights, and perseverance in my training as a graduate student. I owe Nancy a great debt for allowing me to participate in the community of scientists that is her lab. I would also like to thank Alexander Johnson for his comments on this thesis and for allowing me to crash through open doors. Bruce Alberts deserves special mention for giving me lessons on how to remain calm in tense situations and also reading my thesis with great care and thoughtfulness. I want to thank the people of the Hooper foundation, the UCSF department of Biochemistry, and the UCSF campus community who made graduate study a very interesting experience. I would like to think that we learned something from each other. In particular I want to mention my labmates Ken Kubo, Owen Hughes, Anne Stellwagen, Pascal Gamas, Renata Gallagher, Ann Hageman, and Lidia Arciszewska for tremendous support when all around was dark. I also want to thank Pat Brown for numerous helpful comments on my work. Many of my friends deserve special praise for their faith in my survival through the process of graduate education. In particular I want to thank Adam Travis, Evi Strauss, Grant Hartzog, Kevin Cahill, and Chris Chang for discussions about their various philosophies of life. I also want to thank my parents who were the best roomates a graduate student could ever have.

# Tn7 Transposition *In Vitro*

by

**Roland J. Bainton**

## Abstract

This thesis discusses the establishment and analysis of *in vitro* transposition of the *Escherichia coli* transposon Tn7 to its specific *in vivo* DNA attachment site, *attTn7*. This endeavor began with the preparation of crude extracts from *E. coli* that contained the genes required for *in vivo* site-specific Tn7 transposition, *tns*A, *tns*B, *tns*C and *tns*D. The inclusion of exogenous DNA substrates, ATP and MgAc with these extracts provoked the movement of a mini-Tn7 element from a donor molecule to an *attTn7* target molecule. These insertions were site- and orientation-specific and contained 5 base pair target site duplications, characteristics of the *in vivo* site-specific reaction.

Tn7 transposition *in vitro* uses a non-replicative transposition reaction mechanism. It produces a simple insertion of the transposon into the *attTn7* target site with stoichiometric recovery of a linear donor backbone species. During the *in vitro* reaction, staggered double-stranded breaks at the donor-transposon junctions produce an excised linear transposon intermediate with exposed 3' ends. The 3' ends of the transposon are joined to the 5' ends created by a staggered break made in the target DNA.

Highly purified fractions of TnsA, TnsB, TnsC and TnsD, and ATP and MgAc are required to promote any transposition. The addition of host extract can be stimulatory, but it is not required for efficient recombination. Supercoiling of the donor DNA substrate is stimulatory to transposition, but not essential. Several lines of evidence suggest that the transposition components promote recombination

by participating in a large nucleoprotein complex with all of the proteins and DNA substrates in communication. Novel products of the highly purified transposition system include single-strand joins of one end of the transposon to the target site.

We show evidence that suggests that TnsC forms a complex with TnsD at the attTn7 site in the presence of ATP. We also show that TnsA + TnsB + TnsC and the non-hydrolyzable analogue of ATP, AMP-PNP, promote efficient transposition to non-specific DNA sequences in the absence of TnsD. Under these conditions, target site immunity, i.e. inhibition of transposition by the presence of a transposon copy in the target DNA, is relieved. These data suggest to us that TnsC through its interaction with ATP plays a central role in communication between the target DNA and the other components of the transposition machinery.

## Table of Contents

## Figures

# Chapter 1

# Introduction

## A) A Note on Scientific Endeavor

The philosophical foundations of modern scientific endeavor come from the synthesis of two philosophical schools, Empiricism, i.e. models based on observation and Rationalism, i.e. models based on abstractions. Science joins the ideas of factual observation and abstraction into one method. This method uses collection of "facts" through careful and repeated observation within the framework of a generally accepted world view, i.e. a scientific theory or paradigm, to formulate and refute scientific models. Scientific models are an attempt to give coherent description to the interplay of facts and theories. These models are considered "acceptable" within the context of a world view by their predictive value and aesthetic beauty.

A compelling aspect of scientific models is the way they evolve. A given model is improved over time so that it accounts for more and more observations, i.e. scientific facts, thus making it more generalizable and therefore, more useful and powerful. The apparent predictive power and, technological utility of scientific models make it easy to forget that the framework supporting them is at least as important, if not more important, to the ultimate interpretation of "facts" as the "facts" are themselves. It is indeed, such a framework which allows a "fact" to show up as a "fact". In other words, to say that scientific practice is "factually" based obscures the true relationship between facts and theories.

This is essentially the nature of the scientific method as we use it today, whether one believes that scientific models are the "truest" form of intellectual experience, or whether they are merely seen as useful descriptive tools with which to order oneself within the world. Nevertheless, it is part of the human condition to come forth with opinion, i.e. moral vision, about the world and the methods for describing such opinion. Thus, as we endeavor to describe our world in a way which make sense to us, we have developed methods of description of which scientific description is one.

In order to train a scientist in the art of model building, i.e. scientific description, apprenticeship under the craftspeople of model building is a necessary experience. The expertise of the masters is in their ability to acquire "useful" factual knowledge and to assimilate it into models which are consistent with a particular scientific framework. Central to the subject of my thesis is the framework we use to describe the patterns that allow life forms to replicate--the physical and genetic models of DNA.

## B) The Models of DNA Rearrangement

### Some History

A working model for the structure of DNA was proposed by Watson and Crick in 1953 (Watson and Crick, 1953). Numerous other experiments suggested that DNA was the physical material with which living systems stored and propagated genetic information. In one such experiment, Avery and his colleagues were able to show that nucleic acids from bacteria contained a transforming principle which could alter the virulence of other bacteria (Avery et al, 1944). With a structural model of DNA, genetic observations could now be given a physical basis. This conjunction of genetic and biochemical views lead to a flowering of models describing the biological properties of DNA; these included models of how information in DNA is encoded and translated, how DNA is replicated and how the mutability, assortment and distribution of DNA coding information affects the organism. One popular view of heredity was that genes were extremely stable structures that transmitted information with great fidelity from parents to progeny and that genes mutated only infrequently such that the structure of DNA rarely changed.

An important exception to the principle that the structure of DNA was genetically constant was recognized by Barbara McClintock in her genetic and cytological analysis of transposable controlling elements in maize (McClintock, 1956). She demonstrated that

there were genetic segments that could transpose to new chromosomal locations, thereby altering the expression of nearby genes and causing chromosome breakage in a developmentally regulated way. For many years, controlling elements of maize seemed a curiosity, but then other examples of mobile elements came to light. In the 1960s and 70s, bacterial transposable elements were recognized for their ability to cause spontaneous mutation in E. coli and in the transfer of drug resistance between replicons (Malamy, 1966; Jordon et al., 1967; Adhya and Shapiro, 1969; Hedges and Jacob, 1974; Kopecko et al, 1976). Soon after, many prokaryotic and eukaryotic systems were discovered to contain mobile DNA elements (Berg and Howe, 1989). That the genome was a dynamic biological structure and not merely an inert storehouse of genetic information became strikingly apparent.

Mobile DNA elements that have the capacity to translocate from one site to another are now recognized in a wide variety of organisms including humans, bacteria, fungi, protozoa, fruit flies and plants (Berg and Howe, 1989). Within these organisms, the biological roles for mobile DNA are extremely diverse. The generation of mutations, the generation of repetitive DNA, genome evolution, developmental changes, virulence mechanisms, and the spread of antibiotic resistance in bacteria can all involve mobile DNA (Berg and Howe, 1989).

## Methods of Study

Genome rearrangement events can be studied both genetically and biochemically. From a genetic standpoint, these events can be studied when a genotypic change caused by a genome rearrangement is represented by an easily assayable phenotypic change. From a biochemical standpoint, genome sequence rearrangements can be seen as the physical breakage and joining events of DNA. The DNA sequence becomes the bridge between these two different ways of seeing the rearrangement event. It is the power of overlapping,

but distinct, detection methods that allows a deeper and broader description of the phenomenon of DNA rearrangement.

Classes of Genome Rearrangement

Classes of genome rearrangement events can be defined by the extent of required homology between the DNA sites that are undergoing rearrangement. "Homologous" rearrangement events require large regions of homology between the sites, as occurs between two parental genomes during meiotic recombination. "Non-homologous" events are not dependent on such homology. One class of "non-homologous" DNA rearrangement events includes mobile DNA elements that use specialized DNA sites. Such reactions are called site-specific recombination reactions.

## C) Models of Site-Specific DNA Recombination: Conservative Site-Specific Recombination and Transposition

Two classes of site-specific recombination reactions are conservative site-specific recombination and transposition. Conservative site specific recombination (CSSR) restructures DNA sequences by reciprocal recombination reactions that require no DNA synthesis (see Figure 1) (Craig and Kleckner, 1987; Craig, 1988). Bacteriophage lambda integration into the E. coli chromosome is the best studied example of this type of reaction (Pato, 1989). These recombination events use a single crossover event to join two DNA sites which have limited sequence homology; however, some homology between the sites is essential. Intermolecular recombination results in integration and the reverse of this reaction is excision (as shown in Figure 1). Intramolecular recombination can also result in inversion of the DNA between the two sites, depending on the relative orientation of these sites (not shown). All DNA strands are broken and rejoined without exogenous energy input, relying on covalent protein-DNA intermediates to accomplish the strand transfers. The enzymes that carry out these

reactions are specialized forms of sequence-specific DNA topoisomerases (Wang, 1985).

Transposons are discrete DNA segments that can move from one chromosomal location to another non-homologous site (see Figure 1) (Grindley and Reed, 1985; Craig and Kleckner, 1987). Transposition contrasts with CSSR in many ways. Transposons bring together three DNA sites, namely the two transposon ends and the target site. The transposon ends, although related to one another, are unrelated to the target site sequence. The DNA breaking and joining reactions do not involve covalent protein-DNA intermediates. Another distinguishing feature of transposition is that DNA synthesis is involved in this reaction. The transposon ends are joined to the target DNA at staggered positions on the top and bottom target strands such that each end is flanked by a single-stranded region. Repair of these gaps by the host replication machinery generates the target site duplications characteristic of transposition. In some cases, the entire mobile element is copied by DNA replication as an obligatory step in transposition. In addition, DNA synthesis likely plays a role in the processing the donor site from which an element moves. The resulting broken DNA at the donor site is most likely repaired by the host DNA repair machinery.

## D) Overview of Transposition

### 1) General Features of Transposition

Transposition reactions can give rise to a variety of genome rearrangements (Gindley and Reed, 1985). These rearrangements involve joining transposon ends to new sites in the genome by intra- or inter-molecular events. Intramolecular transposition reactions can produce deletion or inversion of adjacent DNA segments. Intermolecular reactions can produce either cointegrates: a joint structure between the donor molecule and the target molecule, which contains the donor and target DNA and two

copies of the transposon (see Figure 3, bottom left), or simple insertions: the integration of the transposon into a new target site (see Figure 3, bottom right).

## 2) Replicative Versus Conservative Transposition

A question that is central to understanding the mechanism of transpositional recombination is whether genetic rearrangements promote duplication of the elements undergoing recombination (Figure 2). A replicative transposition reaction copies the element as part of the process of transposition such that two copies of the element are now present (Figure 2, bottom left). A non-replicative reaction does not replicate the element as part of its reaction mechanism (Figure 2, bottom right).

With genetic transposition systems it is difficult to directly observe individual transposition events at the molecular level. Thus, distinctions between conservative and replicative transposition mechanisms are not simple to make. Early genetic experiments with the transposable element phage Mu exemplify the difficulty of interpreting genetic experiments mechanistically. Mu is a bacteriophage that uses transposition to replicate itself. Experiments with phage Mu suggested that a copy of the element was left behind after a transposition event as it began its lytic phase of growth (Ljungquist and Bukhari, 1977). This observation implied that transposition was replicative, leaving behind a copy of itself as it moved to other parts of the genome. However, insertions of Mu into the chromosome of newly infected cells appeared to be conservative, that is, no replication was required to promote insertional recombination (Liebert et al, 1982; Akroyd and Symonds, 1983; Harshey, 1984). Were these experiments contradictory or could transposition have both conservative and replicative mechanisms?

The full meaning of these apparently contrary views of the transposition reaction were difficult to appreciate because the precise nature of in vivo DNA transactions were often difficult to discern. Multiple copies of the host genome make it difficult to know

from which chromosome an element may have come. If, upon transposition, the donor site which was just vacated is lost, i.e. donor suicide, then no evidence of the conservative event remains. Moreover, when both a new insertion and the old insertion are recovered, a conservative transposition event can look replicative (Weinert et al, 1984). In addition the products of transposition are subject to manipulation by host homologous recombination systems. With such systems cointegrates can be changed into simple insertions, and vice versa. Genome manipulation by such systems can make it difficult to determine the structure of the donor site and the insertion products immediately after the transposition event and can therefore confuse the analysis of the mechanism of transposition (Berg, 1983). It was the development of in vitro transposition systems and the study of transposition intermediates that clarified much of the genetic data.

By in vitro tests the phage Mu transposition system was found to be capable of both replicative and non-replicative mechanisms of transposition (see Figure 3). Mu accomplishes this by forming a joint molecule between the donor molecule and the target molecule called a strand transfer intermediate (Shapiro, 1979; Grindley and Sherratt, 1978; Ohtsubo et al, 1981; Craigie and Mizuuchi, 1985a). This species can proceed through a non-replicative transposition reaction by cutting away the donor backbone or through a replicative reaction in which the transposon DNA is replicated to form a cointegrate (see Figure 3 and below). The molecular pathway of Mu transposition allowed the unfication of previously contradictory genetic data (discussed above), demonstrating the power of genetic and biochemical methods when used together to describe a particular biological phenomenon.

In non-replicative transposition, the element is excised from the donor backbone and inserted into a novel target site without replication and reconstitution of the donor molecule (see Figure 3). In addition to Mu, Tn10 and Tn7 in bacteria, P-elements in Drosophila and Ac elements in maize are elements that are thought to use this pathway

(Bender and Kleckner, 1986; Bainton et al, 1991; Greenblat and Brink, 1962; Engels, 1990; Craig, 1990). The excised donor backbone species is one product of the transposition event. A number of outcomes for the donor backbone are possible: it can be degraded and lost (Weinert et al, 1984; Morisato and Kleckner, 1984; Bender and Kleckner, 1991) or it could be rejoined or repaired. In Ac transposition, the donor site is imperfectly restored either by the transposon or host machinery (Federoff, 1989). In other systems, the donor site can be exactly restored by a double-strand gap repair that uses the homolog to the donor molecule, as in the case of P-elements and Tn7 (Engels et al, 1990; Craig, 1990; Ann Hageman, personal communication). This method may regenerate the donor site with or without the transposon, depending on the available homologous DNA.

In replicative transposition, the transposon is joined to the target in such a way that allows replication of the element producing a novel insertion product and reconstituting the donor molecule (see Figure 3). In this molecule, i.e. a cointegrate, the donor and target backbones are joined by two copies of the transposon (see Figure 2). The bacterial transposon Tn3 is an example of an element that seems to use an exclusively replicative pathway (Arthur and Sherratt, 1979). The Tn3 cointegrate is resolved by special recombination systems resident in the transposon (see below).

A different transposition strategy is used by retroviruses and retroelements. Their mechanism of movement relies on the production of an RNA copy of the transposon genome (Varmus and Brown, 1989). An RNA copy of the viral genome is reverse transcribed into DNA and this DNA element is then inserted into a host genome. In this reaction, the original transposing element is not broken out of the donor site, but is copied prior to transposition. Such a mechanism leaves the donor site intact.

### 3) The Structure of Elements

Although the elements that participate in transpositional recombination are very diverse and have little, if any, sequence identity with one another, they do contain common structural features (Grindley and Reed, 1985). The structural features of a some bacterial elements are discussed below.

One very diverse group of mobile elements in bacteria are known as Insertion Sequences (ISs) (Grindley and Reed, 1985; Galas and Chandler, 1989). These are small elements ranging from 700 bp to 1600 bp in length that are highly heterogeneous. They carry only the minimal sequence information necessary to move the element from place to place, i.e. cis-acting terminal sequences that participate directly in recombination and genes for transposition proteins, and lack antibiotic resistance genes. The copy number of these elements varies widely; in some bacterial strains they can make up 0.5% of the genome (Shapiro, 1983).

Composite transposons such as Tn5, Tn9, and Tn10 are composed of two related IS elements that flank an antibiotic resistance gene (Grindley and Reed, 1985). The transposase gene, a gene required for transposition (see below), is contained within the IS. The composite transposon uses the two outside ends for transposition of the resistance gene, thereby maintaining the integrity of the element. However, transposition events involving other combinations of ends can also occur.

A highly related group of elements belong to the Tn3 family of elements (Grindley, 1983). They carry antibiotic resistance genes and also encode two transposition gene products, a transposase gene that is required for transposition itself and a repressor of transposition. The repressor of transposition is also a resolvase that is involved in processing the primary transposition product. The resolvase is a sequence-specific topoisomerase that performs a conservative site-specific recombination reaction to resolve the cointegrate species generated by the transposition

reaction (Grindley and Reed, 1985). Thus, elements of this family carry two distinct site-specific recombination systems.

Another diverse class of elements include both transposing bacteriophages such as Mu (Pato, 1989) and non-composite antibiotic resistance elements such as Tn7 (Craig, 1989). Elements of this type are generally large, having complicated terminal transposition sequences with multiple repeats at each end, and encode multiple transposition genes. Phages like Mu use transposition for DNA replication during viral growth and also for integrating into the host chromosome during lysogeny. Tn7 encodes multiple antibiotic resistance determinants.

The Cis-Acting Sequences Required for Transposition

Transpositional recombination is dependent on the presence, correct orientation and integrity of the two transposon ends (Grindley and Reed, 1985; Roberts et al, 1991). The two ends of a transposon are closely related to each other and contain perfect or nearly perfect terminal inverted repeats of various lengths at their extreme termini.

The ends of transposons are the specific sites for transpositional recombination. They provide binding sites for the proteins which mediate the specific breaking and joining reactions that underlie transposition (Mizuuchi, in prep.). The complexity of the transposition protein binding sites vary widely among the families of transposons; binding sites for the transposition proteins can occur singly, as with MLV or Tn10 (Varmus and Brown, 1989; Kleckner, 1989) or in more complicated arrays, as with Mu (Pato, 1989). Essential transposon end sequences range in size from a few tens of nucleotides to several hundred base pairs.

The required end sequences are, in general, quite small compared to overall size of the mobile element. Elements range in size from seven hundred bases for some insertion sequences up to 36 kb for bacteriophage Mu. The DNA between the ends of the

transposon may carry many different kinds of genes that promote recombination. In addition some elements, such as bacterial transposons, carry antibiotic resistance genes. Other elements such as bacteriophage Mu and retroviruses encode the information essential to execute the viral life cycle.

## 4) Target Sites In Transposition Reactions

### The Consequences of Transposon Insertion Into Target Sites

Target site choice is an important aspect of transpositional recombination since transposons can cause changes in the host genome. One genome rearrangement that offers an advantage to the host is the stable acquisition of antibiotic resistance genes (Hedges and Jacob, 1974; Kopecko et al, 1976). However, not all genome changes caused by transposons are beneficial to the host. An insertion into an essential host gene can inactivate that gene; this is detrimental to the host and potentially detrimental to the element as well (Malamy, 1966; Jordon et al., 1967; Adhya and Shapiro, 1969). Furthermore, insertions can alter the expression of host genes either by juxtaposing a transposon-encoded promoter near a host gene or by creating new promoters upon insertion.

Therefore, it seems likely that a delicate balance exists between frequency and specificity of transposition, and potential deleterious effects to the host (Kleckner, 1990). The mechanisms which keep dynamic genome rearrangement events within a range that is advantageous for both the element and the host are under study. Those aspects that are relevant to target site choice are discussed below.

### Target Site Choice

Through many studies on target site specificities, it has become apparent that no element transposes in a truly random fashion. Two views of the target DNA suggest different ways of specifying the position of insertion in transposition. In one view,

target sites are specified by the nucleotide sequence at the point of insertion. In the other, the overall structure of the DNA determines where an insertion will occur. In this view, it is the overall context of the target site, determined by the constraints of higher order DNA structures and by primary sequence, that influence target site selection.

*A Target Consensus Sequence*

By analyzing many independent insertions, it is often possible to demonstrate that a particular transposon favors a particular consensus sequence. This consensus can be used to predict the frequency of insertion into a potential target site. Tn10 has a well studied consensus sequence (Halling and Kleckner, 1982). In Tn10 transposition, the frequency of insertion into a particular site often correlates with its degree of match to the consensus sequence (Lee et al, 1983). The coincidence of a consensus sequence with the point of transposon insertion suggests that these sites may be preferred target DNA binding sites for the transposition proteins which must recognize and cleave the DNA at these sites.

*Contextual Specificity*

When transposons are found only in particular regions of the chromosome, they are said to possess a regional specificity. In the case of IS1, some regions of the chromosome appear to be favored for and others forbidden to insertion (Gamas et al, 1987). One suggested explanation for this phenomenon is that the physical packaging of the genome provides a particular context in certain regions that is favorable for insertion.

In retroviral transposition, local interference with insertion can occur when proteins bind, and presumably exclude, the transposition machinery from a particular site (Pryciak, in press). Evidence in vivo suggests retroviruses prefer to insert near

DNase hypersensitive sites (Vijaya et al., 1986; Mooslehner et al., 1990). This data has been interpreted to suggest that retroviruses prefer to insert into structurally open DNA. However, a closer look reveals that the situation is not so simple. In vitro experiments suggest retroviral integration can occur in regions of DNA that are bound by protein. Local bending and shaping of the DNA seems to make particular regions bound by nucleosomes favorable targets for insertion (Pryciak et al,1992.) These data suggest that there may be multiple mechanisms that affect the preference of a particular target site for insertion.

Another aspect of target site preference in some systems is the phenomenon of preferring insertion sites that are close to the donor transposon. In Maize, Ac elements transpose more frequently into nearby chromosomal sites rather than sites that are farther away (Greenblat and Brink, 1962). This phenomenon could reflect an increased accessibility of regions close to the element.

*Specialized Targeting Mechanisms*

Transposons have the capacity to direct themselves to target sites with special features. Ty1, a yeast retrotransposon, favors the 5' ends of genes (Natsoulis et al, 1989; Wilke at al, 1989), Ty3 elements insert into regions close to the sites of initiation of Pol III transcripts (Chalker and Sandmeyer 1992), and Tn7 favors a special target site known as the attTn7 site (Lichtenstein and Brenner, 1981). One potential value of such stringent target site specificity is to provide a "safe haven" for the transposon. Transposition to special sites decreases the chance of producing an insertion that may have deleterious effects on the host or the transposon, and provides a "safe" place for elements carrying antibiotic resistance genes to reside.

The mobile elements Mu, Tn3 and Tn7 display another interesting targeting phenomenon; these elements transpose very inefficiently to targets that already contain a copy of the element (Adzuma and Mizuuchi, 1988; Robinson et al., 1977; Hauer and

Shapiro, 1984). This phenomenon is referred to as transposon target site immunity or transposition immunity. Such a phenomenon may aid in the dissemination of the element by encouraging it to transpose into target sites that do not already contain it. Also elements that transpose at high frequency, such as phage Mu during phage replication, may use this mechanism to prevent insertions into itself or into related elements thereby inactivating them.

Clearly, target site choice in transposition can have huge effects on host function and on the relationship between the host and the mobile element. Transposition can change the physical and genetic arrangement of the host genome and the mobile element and therefore is likely to have important and unforeseen consequences on the physical shape and the content of the genome. That mechanisms have developed to exclude transposition from particular areas of the genome or direct it to others suggests an intimate and ongoing relationship between mobile elements and their hosts (Kleckner, 1990).

**5) The Pathway of Transposition: The Chemical Steps**

Although many transposition systems have been characterized genetically, only a few are characterized biochemically. In vitro reconstitution of transposition systems has revealed many fundamental features about transposition reaction chemistry, intermediates, and pathways and has also helped define the substrates, proteins and cofactors involved in the transposition reaction. Kyoshi Mizuuchi's pioneering studies on phage Mu made it the first and best described in vitro transposition system. His novel use of bacteriophage targets in in vitro transposition reactions provided the field with a sensitive assay with which to study infrequent recombination events (Mizuuchi, 1983). This technique played an important role in the biochemical reconstitution of the best known in vitro transposition reactions, those of phage Mu (Mizuuchi, 1983) retroviruses (Brown et al, 1987) and retroelements (Eichinger and Boeke, 1988),

Tn10 (Morisato and Kleckner, 1985) and now Tn7 (Bainton et al, 1991). From these systems a broad overview of biochemical features common to these transposition reactions can be presented.

## Transpositional Recombination

The chemical changes in the DNA substrates define the three steps of transposition: the donor cleavage reaction, the strand transfer reaction, and product resolution. The first two steps involve breakage and joining reactions that are brought about by transposon-specific machinery and are therefore known as transpositional recombination (Mizuuchi, in prep.). The chemical changes in the DNA that underlie transpositional recombination are now understood in some detail. This has revealed remarkable similarity between the various in vitro transposition systems (discussed below). The final step in transposition, product resolution, utilizes host-encoded replication and repair functions to complete the process of transposition (also discussed below).

### *Donor Cleavage*

A transpositional recombination reaction begins with an endonucleolytic cleavage that exposes a hydroxyl group at the 3' ends of the transposon, thereby separating the ends from the donor backbone (Craigie et al, 1985; Fujiwara and Mizuuchi, 1988; Brown et al, 1989; Benjamin and Kleckner, 1989; Eichinger and Boeke, 1990; Bainton et al, 1991). In many cases, the terminal 3' nucleotide is adenosine. This cleavage is the first chemically detectable change seen in the DNA substrate and is absolutely required for the subsequent strand transfer step. Recent evidence suggests that the donor cleavage reaction results from the nucleophilic attack of $H_2O$ or another hydroxyl group on the phosphate that joins the 3' end of the transposon to the donor backbone to release the 3' end of the transposon (Mizuuchi and Adzuma, 1991; Engelman et al,

1991; Vink and Plasterk, 1991). The phosphate around which this reaction proceeds undergoes an inversion in chirality consistent with an SN2 reaction mechanism. Transposons that are artificially broken at this junction can still participate in the transposition reaction, suggesting that the energy of this bond does not have to be conserved for efficient transposition (Craigie and Mizuuchi, 1987; Craigie et al, 1990; Bainton et al, 1991). Unlike CSSR, no covalent protein-DNA intermediate has been identified that preserves the energy of the bond that joins the transposon ends and the donor backbone. Thus, at this stage of transposition, two phosphodiester bonds have been broken and none made, such that two breaks in the DNA have accumulated.

In Tn10 and Tn7, the 5' ends of the transposon are broken before the 3' end of the transposon undergoes strand transfer (Morisato and Kleckner, 1987; Hanniford et al, 1991; Bainton et al, 1991). A transposition reaction which completely separates the transposon from the donor backbone assures that the reaction will proceed in a non-replicative manner since without physical connection to donor backbone, the transposon cannot regenerate itself. If the 5' transposon end breaks are not made prior to joining the 3' ends to the target, the transposition machinery can, at a later step, decide whether to replicate the transposon or cut away the donor backbone (Mizuuchi and Craigie, 1986).

*Strand Transfer*

The strand transfer reaction, or integration step, links DNA sequences at the target site with the transposon end sequences to form a novel junction (Mizuuchi, in prep.). Although many transposition systems have yet to be characterized, a compelling model for the mechanism of the strand transfer reaction has been proposed by Mizuuchi, based on the studies of phage Mu, retrovirses, and Ty elements (Mizuuchi and Adzuma, 1991; Engelman et al, 1991, Eichinger and Boeke, 1990). The strand transfer step joins the hydroxyl group at the 3' end of the transposon to a phosphate in the target DNA.

In order to study the chemistry of recombination at the target site, specialized target DNAs were prepared that contain chiral phosphates in the DNA phosphodiester backbone. As these phosphates are potential target sites for transposon insertion, joining to these phosphates is informative about the chemistry of the reaction. The products of insertions into these phosphates show inversion of the chiral center (Mizuuchi and Adzuma, 1991; Engelman et al, 1991). A model that is harmonious with this data suggests that the 3' hydroxyl group at each of the transposon ends acts as a nucleophile and attacks the phosphate in the phosphodiester backbone of the target site. This reaction results in covalent joining of the 3' ends of the transposon to the 5' ends of the target DNA, while concomitantly displacing the alcohol group at the 3' ends of the target strand. Chiral inversion of the phosphate at the point of insertion provides strong evidence for the proposed one-step transesterification mechanism of transposition. Thus, the reaction exchanges a break at each end of the transposon for a break in each new transposon-target junction.

An interesting analogy between transpositional recombination and Group 1 self-splicing introns can be made. RNA splicing is a ribonucleic acid rearrangement reaction that proceeds by two consecutive transesterification reactions with inversion of chirality at phosphorus (Cech, 1990). The first step is initiated by guanosine which attacks the phosphorus at the 5' splice site and forms a 3' 5'-phosphodiester bond to the first nucleotide of the intron. The 5' exon, now terminating in a free 3' hydroxyl group, then attacks the phosphorus atom at the 3' splice site. The result is the ligation of the exons and the excision of the intron. In both Group I splicing and transposition, the 3' end of a nucleotide participates as a substrate in a transesterification reaction, resulting in inversion of phosphorus in the nucleic acid backbone. Such mechanistic similarities may point to common roots between DNA transposition, RNA splicing and other reactions that break and join nucleotides.

When the two 3' strands at the transposon ends join the same target site, they exchange bonds on opposite strands of the target site DNA at a defined distance from one another. The resulting breaks introduced into the target site are thus staggered. The distance between these breaks is invariant for a individual transposon, but varies between particular transposons (Berg and Howe, 1989). For example, these staggers are 5 bps with Mu and 9 bps with Tn10. These staggered breaks are the progenitors of target site duplications which flank the transposon in the final product of all transposition reactions (Shapiro, 1979; Benjamin et al, 1989).

It is important to note that for those transposition reactions studied to date, no high energy protein-DNA intermediate or hydrolysis of a high energy cofactor is necessary to promote any of the breakage and joining reactions (Mizuuchi, in prep.). Strand transfer is accomplished by a phosphoryl exchange reaction that begins with two breaks in DNA adjacent to the transposon ends and finishes with two breaks in the target DNA. Thus, the energy for formation of the new phosphodiester bonds that join the transposon ends to the target DNA derives from the original phosphodiester bonds in the target DNA. However, it should be noted that the overall transposition reaction, i.e. donor cleavage and strand transfer, does result in two leftover breaks in the DNA, which explains transposition's requirement for DNA replication to regenerate intact duplex DNA.

Of special note are mutations in transposons which affect both the efficiency of donor cleavage and strand transfer. Numerous genetic studies suggest that mutations at the very termini of the transposon can greatly diminish its ability to participate in transposition (Berg and Howe, 1989). In vitro studies in Mu, retroviruses, and Tn7 show that both donor cleavage and strand transfer reactions are dependent on a 3' terminal adenosine nucleotide at the transposon ends (Surette et al, 1991; Engelman et al, 1991; Chapter 4 this thesis). Single base changes at this nucleotide interfere with the donor cleavage in Mu and with strand transfer and reverse strand transfer reactions

in retroviruses (Surette et al, 1991; Craigie et al, 1990; Chow et al, 1992). As suggested above, both donor cleavage and strand transfer proceed through transesterification reactions that involve the hydroxyl at the 3' end of the transposon. This suggests that although these reactions produce different outcomes they may, in fact, be closely related (Engelman et al, 1991). A special positioning of the 3' end of the transposon by the transposase (discussed chapter 4) may promote attack on the phosphate by a nucleophilic moiety to release the transposon end and then position the broken end to nucleophilic attack on a target molecule. A mechanistic unification of these two reactions could explain their mutual dependence on the same 3' terminal adenosine.

*Product Resolution*

The products of transpositional recombination are two new transposon-target junctions with adjacent single stranded gap regions. Resolution and repair of these structures can involve either a small amount of DNA synthesis to repair the gaps without replication of the element (i.e. non-replicative transposition), or extensive replication through the new junctions producing two elements (i.e. replicative transposition). The latter process depends on the donor backbone being still attached to the transposon after transpositional recombination (see above). The strand transfer intermediate produced in Mu transposition is a molecule that allows both outcomes (see Figure 3). In the strand transfer intermediate, the target DNA can prime DNA synthesis such that replication through the element from both ends of the transposon produces two copies of the transposon (Mizuuchi,1983). The product of this reaction is called a cointegrate (see Figure 3). A different outcome results if the donor backbone is disconnected from the strand transfer intermediate by a second break at the donor-backbone junction prior to initiation of DNA synthesis (see Figure 3, Mizuuchi,1983). In this case, the transposition product is a simple insertion (see Figure 3). This process is accomplished by as yet unknown host functions. Tn10 and Tn7 assure a non-replicative transposition

reaction by completely releasing the donor backbone from the transposon. (Morisato and Kleckner, 1987; Benjamin et al, 1989; Bainton et al, 1991). In these cases, the transposon-backbone junctions are severed prior to joining the transposon to the target. Thus, a non-replicative reaction is assured because the donor backbone is not associated with the target DNA after strand transfer.

Further processing of non-replicative transposition products requires repair of the single-stranded gaps left over after transpositional recombination. Little is known about the processing of transposition insertion products (Mizuuchi, in prep.). Although this phase of transposition is a host-mediated event, transposon-encoded functions may play a role in the repair reaction either by recruiting the host-repair machinery, by protecting the incomplete junction from nucleases, or by preventing initiation of another transposition event. As the genetic and biochemical interplay of the element and the host becomes more fully understood, transposon mediated repair functions may come to light.

## 6) The Pathway of Transposition: Conjunction of Proteins and DNA Sites

### Transposase Function: Protein-DNA Interaction in Transposition

Transposons generally encode their own transposition functions required to accomplish transposition (Mizuuchi, in prep.). The transposition functions must specifically recognize the ends of the transposon and catalyze the breakage and joining reactions that occur between the ends of the transposon and the target site (discussed above). These functions must, therefore, be able to bind to two types of DNA sequences: the ends of the transposon, and the target site. The protein that can accomplish these functions is called a transposase. MuA, the retroviral IN protein, and Tn10 transposase, can accomplish both the donor cleavage and the target joining reactions, demonstrating that all the activities of transpositional recombination can be contained within one polypeptide (Maxwell et al 1987; Craigie et al, 1990; Benjamin et al, 1989). The

coincidence of these activities in a single protein lends credence to the suggestion that these reactions are closely related (as discussed above) or suggests that these proteins contain multiple active sites.

Protein-Protein Interactions in Transposition: The Transposase and Conjunction of the Transposon Ends

A number of lines of evidence suggest that transposon end identification, and transposon end-end conjunction are likely to occur prior to the initiation of transposition. Numerous genetic and biochemical studies suggest that transposition proteins interact directly with transposon ends (Craigie and Mizuuchi, 1986; Berg and Howe, 1989). The presence and correct orientation of the two transposon ends is necessary even for transposition reactions in which only a single end is inserted into a new site to cause an adjacent deletion (Roberts and Kleckner, 1991, Huisman and Kleckner, 1987). This suggests that the transposon ends communicate with one another and lends credence to the view that there is some sort of physical association between the ends.

Physical evidence for such an association comes from the in vitro analysis of bacteriophage Mu. The Mu transposase, MuA, is responsible for bringing the Mu transposon ends together and holding them in a manner appropriate for donor cleavage (Surette et al, 1987; see below). This process requires DNA-protein contact between the substrate and transposase, and also protein-protein contact between transposase molecules bound to each transposon end. The transposon ends are brought into precise alignment through the formation of an complex known as the transpososome (Mizuuchi, in prep.; Surette et at, 1987). The formation of such a complex is facilitated by DNA supercoiling and is assisted by host protein components (see below).

The Role of DNA Supercoiling in Transposition

DNA supercoiling influences many processes involving complex protein-DNA interactions, including DNA transcription, replication, recombination and repair (Vosberg, 1985; Wang, 1985; Wells, 1988). Two functions have been suggested for the role of supercoiling in site-specific recombination reactions. The first is demonstrated in the case of lambda site-specific recombination (Richet et al, 1986) and Mu recombination (Surette and Chaconas, 1989; Mizuuchi and Craigie, 1986), where supercoiling seems to facilitate the wrapping of distant DNA sites so that they more easily become juxtaposed. A second function is seen in Mu transposition where the ability to interwrap the transposon ends is necessary to identify their relative orientation. Because supercoiling aids in interwrapping, it is no surprise that supercoiling also mediates the proper recognition of the relative orientation of the transposon ends (Craigie and Mizuuchi, 1986).

In contrast, supercoiling of target site DNA in site-specific recombination systems does not affect its ability to act as a target (Richet et al, 1986 and Brown et al, 1987). The explanation for this may lie in the simplicity of the target sites in these reactions. Target site recognition may depend mostly on target site availability and nucleotide identification (see above) rather than the conjunction of distant DNA sites as in the donor. Because target sites appear to be relatively compact and simple, winding and wrapping of the site is less likely to be important in the formation of a competent target site recognition complex and, therefore, DNA supercoiling of the target is less likely to have a dramatic effect on the rate of recombination.

Host Proteins Can Play a Role in Transposition

Mechanistically tied to the role of supercoiling in some site-specific recombination reactions are host proteins. In Mu transposition, the host proteins HU and IHF participate in the formation of the transpososome complex (Surette et al,

1987). HU is a small histone-like protein in E. coli with pleiotropic effects on the physiology of the host (Drilca and Rouviere-Yaniv, 1987). By promoting DNA flexure, HU is thought to facilitate the assembly of complexes that require multiple protein-DNA and protein-protein interactions within short DNA sequences that are ordinarily too stiff to allow the degree of bending required for appropriate interactions (Lavoie and Chaconas, 1990). HU and proteins like HU are likely to participate in a way that facilitates the wrapping and bending of DNA, in conjunction with DNA supercoiling.

IHF is another small histone-like protein of E. coli which can bind DNA site-specifically. IHF can participate in site-specific recombination reactions as a site-specific DNA bending protein (Robertson and Nash, 1988). In Mu transposition, the formation of the transpososome complex at the Mu transpositional enhancer site benefits from the effects of IHF-mediated DNA bending that reduces the supercoiling requirement of the donor molecule (Surette et al, 1989). Although these small proteins can be highly stimulatory to transposition, their roles are generally limited to acting as accessory factors, as evidenced by the fact that transposases can accomplish transpositional recombination in reactions that lack host proteins (Maxwell et al, 1987). Tn10 is an exception; it requires the presence of Hu to promote transposition in vitro (Morisato and Kleckner, 1987; Benjamin and Kleckner, 1989) Thus, in some systems, the host factor requirement is quite substantial.

Other host proteins play additional roles in the processing and repair of the transposition intermediates and products (Craigie and Mizuuchi, 1985). Though these processes are important to the completion of the transposition reaction, their roles are not critical to the formation of the protein-DNA complexes essential to the process of transpositional recombination.

Transposon-Target Communication: Proteins that Influence Target Site Selection

Although the machinery that promotes the transposition reaction, the transposase and host factors, seems to assemble at or around the ends of the transposon in the donor molecule, an emerging view in the field of transposition is that target site recognition can influence when and how that machinery will act. If this is true, the target site must have ways of communicating with the donor site. In Mu transposition, a target-bound protein, MuB, helps provoke the donor cleavage activity by a protein-protein contact (Baker et al, 1991). Genetic and in vitro evidence in Tn7 transposition suggest that a special site, the attTn7 target site, is seen by the transposition machinery before the breakage and joining reactions are initiated (A. Stellwagen, personal communication and see below). Thus, incorporated into the transposition mechanism is an ability to sense the target site even before it is used as a substrate for insertion.

The ability of the transposition machinery to sense properties of the target site is exemplified by the phenomenon of transposition immunity. In the transposition systems of Mu, Tn3 and Tn7, when a prospective target carries transposon-related sequences that target becomes a poor substrate for insertion (Adzuma and Mizuuchi, 1988; Robinson et al, 1977; Hauer and Shapiro, 1984). Thus, the presence of one transposon in the target makes the target "immune" to further transposon insertion events. Knowledge about the target DNA is useful to the transposon, because transposition into itself or into another transposon may disable the element. Target site recognition is also a prudent way of initiating transposition. Introduction of breaks into the DNA is not of trivial consequence to the transposon or the genome; that the breaks are made when the transposon is in close proximity to an appropriate target site helps to guarantee the completion of the transposition reaction.

## 7) Established In Vitro Transposition Systems

### Mu

Bacteriophage Mu was the first transpositional recombination system well-characterized in vitro (Mizuuchi, 1983). The purified Mu system uses two transposon-encoded proteins to accomplish transposition. MuA is the transposase which binds to the ends of the transposon and executes all of the breakage and joining reactions. MuB is a protein that can associate with target sites in an ATP-dependent manner, stimulate the transposase activity of MuA, and participate in transposition immunity. ATP or a non-hydrolyzable analogue of ATP is highly stimulatory to transposition. The ATP is not required as an energy source to drive the transposition reaction, but is thought to play a role in stabilizing MuB contact with the target sites and thus promotes the ability of MuB to simulate the transpositional activity of MuA (also see below). The host proteins, HU and IHF, are highly stimulatory to Mu transposition. As discussed in the sections on the roles of supercoiling and host proteins in transposition (above), these proteins are thought to help in the formation of the protein-DNA complexes necessary to carry out transposition.

The pathway of Mu transposition, as it is understood in vitro, involves MuA binding to the transposon ends and bringing them together in a "stable synaptic complex" (SSC) (Mizuuchi et al, 1991). The formation of this complex is facilitated by HU and IHF. In the presence of $Mg^{++}$, cleavage at the 3' ends of the transposon can occur, producing a "cleaved donor complex" (CDC). Upon the addition of a target DNA, MuB and ATP in the absence of $Mg^{++}$, the CDC can associate with the target DNA to form a "strand transfer complex" (STC). With the addition of $Mg^{++}$, this complex can undergo the strand transfer reaction joining the 3' transposon ends to the target DNA. The resulting strand transfer intermediate has both donor backbone sequences and target sequences joined to the transposon (see Figure 2). However the donor backbone is not required for transposition because an artificially-made excised linear transposon species can

participate efficiently in transposition (Craigie and Mizuuchi, 1987). This species can be further processed by host extracts into a simple insertion by removing the donor backbone species or replicating through the element to generate a cointegrate. These steps are not yet understood in biochemical terms.

MuB is a targeting protein in Mu transposition. Although its presence is not required for transposition, MuB greatly stimulates the breakage and joining activity of MuA (Baker et al, 1991). MuB can be viewed as an effector of MuA that can also sense properties of the target site. An aspect of target site recognition by MuB is its ability to redistribute itself among potential targets. In the presence of MuA, MuB will selectively associate with targets that do not contain transposon end sequences (Adzuma and Mizuuchi, 1988). The binding of MuB to DNA is stimulated by ATP. MuA, when present in sufficiently high local concentration, stimulates the ATPase activity of MuB and its dissociation from DNA. When MuA and MuB are bound to the same DNA molecule, MuA causes MuB to hydrolyze ATP and come off the DNA. This causes MuB to redistribute itself onto other DNA molecules. It is this ability to redistribute MuB that causes the Mu transposition machinery to prefer target sites which do not have other resident Mu transposons or Mu end sequences (i.e. to display transposition immunity).

This data suggests that it is MuB bound to an appropriate target site which recruits the uncleaved "stable synaptic complex" to the target site. In this view, target DNA-bound MuB stimulates donor-bound MuA to initiate the transposition reaction by cleaving the transposon-backbone junction at the 3' transposon ends. This complex subsequently joins the 3' end of the transposon to the target DNA to form the strand transfer intermediate. MuB also has the ability to stimulate donor cleavage reaction when in the "stable synaptic complex", i.e. not bound to target DNA. In this case, MuB is not obviously bound to DNA (Baker et al, 1991). Thus, it appears that the donor-cleavage reaction can occur when the transpososome complex is associated with target DNA or when it is in solution. Such an observation suggests that phage Mu may be able to

use different pathways to bring the broken or unbroken donor molecule and the target molecule together.

It is important to note that in the transposition reaction promoted by MuA alone, a nucleotide target site preference is demonstrated and this preference is preserved in the MuB-stimulated transposition reaction (Mizuuchi and Craigie, 1986). This data is consistent with the view that MuB stimulates an activity in MuA, encouraging the transposition machinery to choose sites near MuB, but not specifying the exact nucleotides at which the transposon will insert.

Retroviruses and Retroelements

Retroviruses use a very different pathway from that of phage Mu to produce an integration intermediate, but use a remarkably similar mechanism to form an integrated transposition product. Upon infection, the RNA of the viral particle is converted to DNA by reverse transcription (Brown and Varmus,1989). This process produces a linear DNA species two nucleotides longer than the transposon which will integrate into the chromosome. After the production of this intermediate, retroviral integration and Mu transposition are very similar. In a step called end trimming, the viral IN protein cleaves away this dinucleotide species on one strand, revealing the 3' ends of the transposon. This reaction uses an attacking hydroxyl group (water or glycerol) as a nucleophile to generate the break at the 3' end of the transposon (Engelman et al, 1991; Vink and Plasterk, 1992). The resulting species is analogous to the "cleaved donor complex" in Mu transposition, without a donor backbone species. The integration particle seems to choose target sites with some preference for naked DNA and for certain regions of DNA bound by nucleosomes. The wrapping of the DNA around the nucleosome may make certain nucleotides accessible to the transposition machinery (Pryciak and Varmus, 1992).

The integration reaction is promoted by IN protein, which uses a mechanism similar to that of phage Mu to join the ends of the transposon to the target site (Engelman et al, 1991). As found for Mu, no protein-DNA covalent intermediate is seen in this reaction and a linear DNA species with precut ends can participate in the transposition reaction (Craigie et al, 1990). This suggests that the trimming step of retroviruses, like the donor cleavage step in Mu transposition (see above), is required only for the exposure of the 3' adenosine at the ends of the transposon. Unlike Mu, the retroviral integration reaction has no ATP requirement. In vivo, the active retroviral integration intermediate is a linear transposon found within a large nucleoprotein particle (Brown et al, 1987). This particle contains other viral protein components that play a role in holding the integration intermediate together, protecting the transposon from nucleases and may also help direct the targeting of the transposon (Bowerman et al, 1988). It is important to note the purified IN protein alone can accomplish the end trimming and integration reactions (Craigie et al, 1990).

The Ty retrotransposon from yeast is another in vitro transposition system with some similarities to retroviruses (Boeke, 1989). The Ty element is resident in the host chromosome and transposes to other parts of the genome by reverse transcribing RNA copies of its genome into DNA. The products of reverse transcription are flush with the end of the transposon, therefore no trimming step is required to uncover the transposon ends (Eichinger and Boeke, 1988). The Ty integration intermediate is also found as a large nucleoprotein complex with other proteins (Eichinger and Boeke, 1988).

## Tn10

The bacterial transposon Tn10 encodes a single transposase gene product which can promote intramolecular transposition in vitro (Morisato and Kleckner, 1987). This reaction requires HU in addition to the Tn10 transposase in order to promote

transposition (Benjamin and Kleckner, 1989). Neither ATP nor any other high energy cofactor is required for transposition in vitro. The products of transposition are intramolecular deletion/inversions (Benjamin and Kleckner, 1989). As found for Mu and retroviruses, the 3' ends of the transposon are exposed by an endonucleolytic step. However, with Tn10, this step is accompanied by cleavage of the 5' ends of the transposon as well, leaving the transposon ends flush (Benjamin and Kleckner, 1989). The donor molecule is completely released from the backbone, ensuring a conservative transposition reaction. Analysis of the products of the in vitro transposition reaction suggest that Tn10 transpositional recombination occurs by a mechanism similar to Mu and retroviruses ---that is, the 3' end of the transposon is joined to a 5' target site break (Morisato and Kleckner, 1987; Benjamin and Kleckner, 1989). Analysis of Tn10 transposition products has provided evidence that the transposon makes staggered breaks in the target DNA upon insertion (Benjamin and Kleckner, 1989).

The Tn10 in vitro reaction produces an excised transposon fragment which is suggested to represent a transposition intermediate similar to the "cleaved-donor complex" of Mu transposition and the retroviral integration intermediate (Haniford et al, 1991). This species has the expected breaks at both the 3' and 5' ends of the transposon consistent with an intermediate in transposition, but it has not yet been shown to participate directly in recombination.

Tn10 transposition can demonstrate a high degree of site selectivity. Tn10 transposes to a particular "hot spot" at high frequency when used in a intramolecular transposition reaction and in vivo in intermolecular transposition (Benjamin and Kleckner, 1989; Halling and Kleckner, 1982) . Thus, just as in the case of Mu, the Tn10 transposase demonstrates some degree of target site specificity .

## E) Tn7 Transposition

### 1) The Genetics of Tn7 transposition.

As with many bacterial transposable elements, Tn7 was found in a drug resistant strain of E. coli. Tn7 encodes several antibiotic resistance determinants including a novel dihydrofolate reductase that provides resistance to the antifolate trimethoprim and an adenyltransferase that provides resistance to the aminoglycosides spectinomycin and streptomycin (Barth et al, 1976). A unusual feature of Tn7 is its ability to transpose at very high frequency to a particular site in the E. coli chromosome known as the attTn7 site (Lichtenstein and Brenner, 1981, 1982). Site-specific insertion of Tn7 also occurs into the chromosomes of many other bacteria (Craig, 1989).

Tn7 has an usually large number of transposition genes; five genes are involved in promoting its transposition. Four of these genes are required for transposition to attTn7: tnsA, tnsB, tnsC and tnsD (Rodgers et al., 1986; Waddell and Craig, 1988). This ensemble of tns genes also promotes low-frequency Tn7 insertion into a limited number of sites related in sequence to the attTn7 site, known as pseudo-attTn7 sites (Kubo and Craig, 1990). Thus, tnsD mediates insertion into target sites of a particular sequence. A distinct, but overlapping set of genes, tnsA, tnsB, tnsC and tnsE, promotes transposition to apparently random target sites that are not obviously related to each other or to attTn7 (Kubo and Craig, 1990; Rogers et al, 1986; Waddell and Craig, 1988). Thus Tn7 can also transpose to many other sites in the chromosome and on plasmids, but at a much lower frequency.

Like all transposition reactions, both site specific and non-site specific Tn7 insertions produce target site duplications, 5 base pairs in the case of Tn7 (Lichtenstein and Brenner, 1982). This evidence of DNA synthesis strongly argues that Tn7 inserts into attTn7 via transposition and not by conservative site-specific recombination (see description above). In addition, comparison of the nucleotide sequences at the point of

insertion and at the ends of the transposon show no obvious homology (Gay et al., 1986; McKown et al., 1988)

**2) Cis acting sites and the target site**

The attTn7 target site lies at minute 84 between the phoS and glmS genes (Lichtenstein and Brenner, 1982). Insertion into this site is position-specific, that is the majority of insertions occur at a specific nucleotide junction. Insertions into this site are also orientation-specific. Tn7 inserts so that its left end is adjacent to phoS and its right end is adjacent to glmS. The central bp of the 5 bps usually duplicated upon Tn7 insertion is designated "0". Sequences to the right and left of this favored insertion site are given a positive and negative designation, respectively. Fine mapping of the attTn7 site has shown that the specific DNA sequences required for attTn7 activity lie to the right of the point of insertion (Gringauz et al., 1988; McKown et al., 1988; Qadri et al, 1989; Waddell and Craig, 1989). Sequences from +23 to +58 are capable of promoting wild type levels of insertion into the attTn7 site both in vivo and in vitro Kubo and Craig, 1990; Kubo, in prep). Thus, the sequences that are important for attTn7 target activity lie at a distance from the point of insertion.

The cis acting sequences required for Tn7 transposition lie at the ends of the transposon (Arciszewska et al,1989). 150 bps of the left end and 100 bps of the right end are capable of promoting wild-type levels movement of a mini-Tn7 element in vivo and are competent substrates in vitro (Arciszewska et al, 1989; Bainton et al 1991, respectively). The extreme termini of Tn7 consist of highly related inverted repeats of about 30 bps in length. The 30 bp elements are repeated throughout the ends of the transposon, three are found in the left end and four in the right end, thus the ends of Tn7 are asymmetric (Arciszewska et al, 1989).

**3) Biochemical Roles for Some of the Tns Proteins**

Studies, including work described in this thesis, have revealed that the Tns proteins directly mediate Tn7 transposition. We suggest that the Tn7 transposition machinery is composed of a common transpositional machine--- TnsA + TnsB + TnsC--- that can be directed by alternative target-specific proteins, TnsD and TnsE (this thesis).

Biochemical study has suggested specific roles for the individual proteins in Tn7 transposition. TnsB binds specifically to sites in the transposon ends that are included in the 30 bps repeats mentioned above; these sequences are thought to be the TnsB binding site (McKown et al, 1987; Arciszewska et al, 1991; Tang et al, 1991). TnsB has been shown to participate directly in transposition (Arciszewska and Craig, 1991). TnsB is likely to participate in end recognition and end-end conjunction required in transposition in a manner analogous to MuA in Mu transposition (discussed above). The breakage and joining activity required to promote transposition has not been identified in the TnsB polypeptide, suggesting the possibility that this activity lies in another polypeptide or is shared between TnsB and the other Tns proteins (Arciszewska et al, 1991; this thesis).

TnsD binds specifically to attTn7 and has been shown to participate directly in transposition (Kubo, in prep.) The TnsD polypeptide protects a region from +30 to +50 of the attTn7 insertion site from attack by DNaseI (Kubo, in prep). That the sequences protected by TnsD binding and the sequences required to promote Tn7 insertion lie to one side of the insertion site suggest that TnsD binding to the attTn7 site is necessary and sufficient to direct the transposition machinery to this particular site. The role of TnsD as an insertion site-specifying factor in Tn7 transposition is discussed in Chapter 3.

An in vitro transposition system has established that TnsA, TnsB, TnsC and TnsD are required to promote transposition of Tn7 to the attTn7 site. The ability of the Tns

protein components to complement one another has helped establish their roles in Tn7 transposition. TnsC was purified by its ability to complement the transposition reaction and shown to participate directly in transposition in vitro (Gamas and Craig, in press). TnsC is a non-site specific, ATP-dependent DNA binding protein. Its role in transposition is discussed in Chapter 3. Work in this thesis shows that TnsA also participates in transposition although its role is not known.

## F) Summary

The genetics and biochemistry of genome rearrangement are well-established scientific paradigms that center on the properties of DNA. Transposition is one class of genome rearrangement events. Biochemical study of the substrates, products, proteins and cofactors involved in transposition has lead to a much increased understanding of this process in light of its many biological consequences. From eukaryotic transposition systems such as retroviruses to bacterial transposition systems such as Mu and Tn10, common mechanisms of DNA breakage and joining have demonstrated underlying similarities between this seemingly diverse group of mobile elements.

My contribution to the field of transposition has been to initiate the biochemical dissection of the bacterial transposon, Tn7. I began this by establishing an in vitro system which faithfully duplicated the movement of a mini-Tn7 element from a donor molecule to an attTn7 target molecule using crude extracts derived from cells which contained the tns genes, tnsA, tnsB, tnsC and tnsD. Chapter 2 is a paper which describes the work I did in collaboration with Pascal Gamas. We developed a highly efficient, partially fractionated in vitro system in which the Tns proteins were provided in four separate extracts. Using this crude system, we were able establish the overall pathway and chemistry of the transposition reaction and comment on the minimum requirements for the reaction. In chapter 3, I describe a purified in vitro system and show the minimum components required for Tn7 transposition in vitro. In collaboration with Ken

Kubo, a model of target site selection is developed. Chapter 4 is unpublished work on some interesting species produced by the transposition reaction and some work on transposons that contain mutations in their terminal nucleotides. Chapter 5 is a short comment on the effects of DNA supercoiling on the transposition reaction. Chapter 6 provides a summary and some of my perspectives on Tn7 transposition in vitro. An appendix is included that discusses the specific activity of the purified TnsA polypeptide.

**Figure 1 Conservative Site-Specific Recombination Versus Transposition**

This figure is a schematic depiction of the fundamental differences between conservative site specific recombination (CSSR) and transposition. CSSR recombines the sequences between two related sites. The required core homology between the two sites undergoing recombination is represented by the vertical lines in the two boxes. The sites and the proteins that interact with these sites promote recombination that requires no DNA synthesis and that is often, but not always, reversible.

Transposition requires the conjunction of three DNA sites, the two ends of the transposon (boxes with triangles), and the target site (bold lines). The ends of the transposon are related to one another and contain terminal inverted repeats (triangles). Recombination is not reversible and requires DNA synthesis. The target site duplications that are the result of the new DNA synthesis are shown as arrows in the product molecule.

# Figure 1

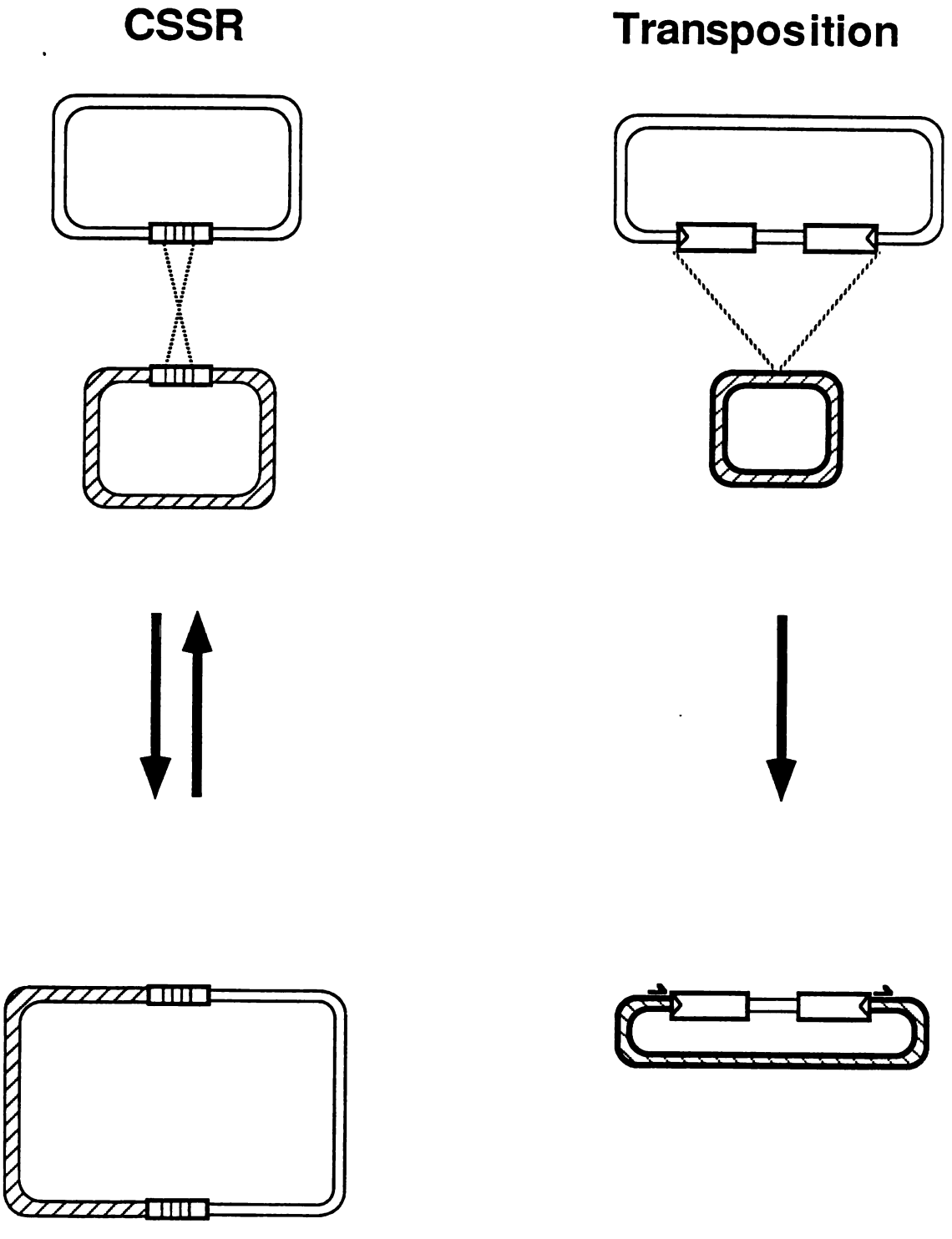


**Figure 2 Pathways of Transposition**

Transposition reactions use fundamentally different pathways to promote recombination. Replicative transposition requires that the element be replicated upon movement to a new site so that one element is found in the new insertion site and the other remains in the donor site (bottom left). Non-replicative transposition removes the element from the donor site without replicating the element, leaving behind only the donor backbone (bottom right).

# Figure 2 Pathways of Transposition

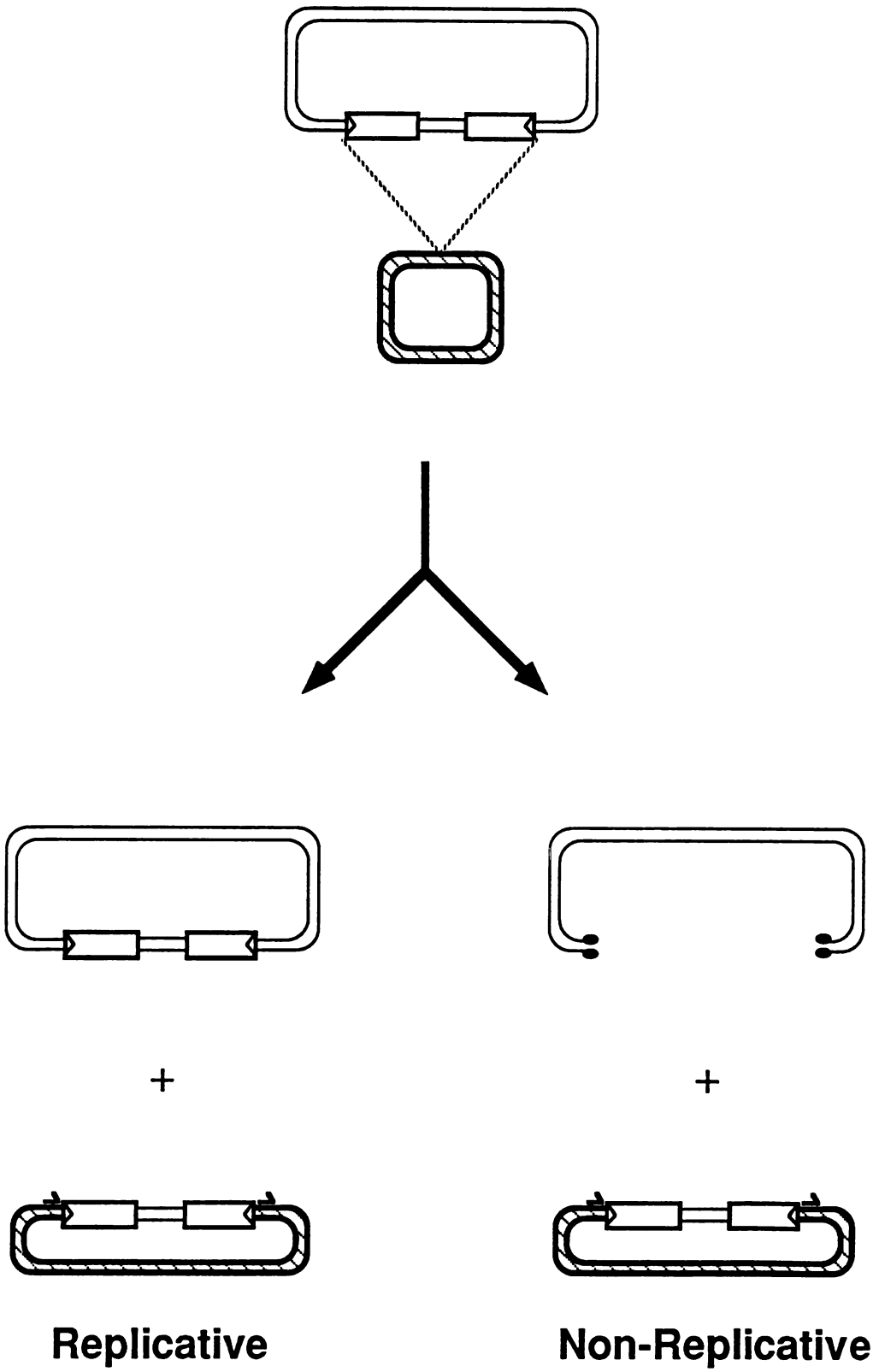
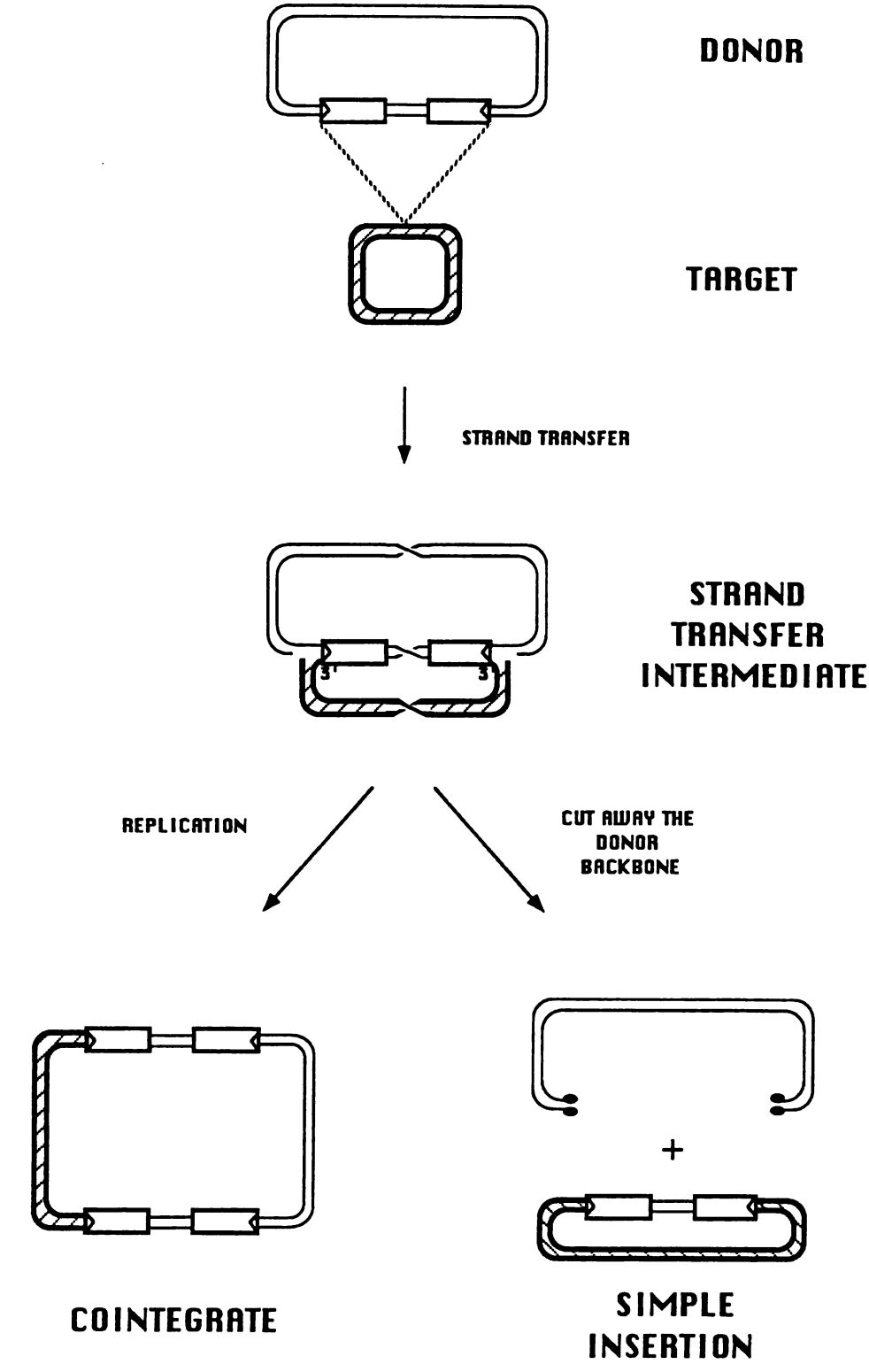


**Figure 3 The Strand Transfer Intermediate Allows for both Replicative and Non-replicative Transposition**

The strand transfer or Shapiro intermediate joins the 3' ends of the transposon to 5' target breaks while the donor backbone remains attached to the transposon at its 5' ends. This intermediate can be processed in a replicative manner producing a cointegrate (a molecule which joins the donor backbone and target DNA between two copies of the transposon) or it can be processed in a non-replicative manner producing a simple insertion and a donor backbone species. This is the intermediate the phage Mu uses to promote recombination.

# Figure 3


DONOR
TARGET
STRAND TRANSFER
3'
3'
STRAND
TRANSFER
INTERMEDIATE
REPLICATION
CUT AWAY THE
DONOR
BACKBONE
+
COINTEGRATE
SIMPLE
INSERTION

# Chapter 2

## Tn7 Transposition In Vitro Proceeds through an Excised Transposon Intermediate Generated by Staggered Breaks in DNA

Chapter 2 is a collaboration between Pascal Gamas and myself. The initial establishment of the transposition reaction was done by me. In that work I showed that ATP, $Mg^{++}$, an attTn7 site and a crude extracts made from cells containing the tnsABCD genes were required for transposition. I also demonstrated that the in vitro transposition products were site- and orientation-specific and contained 5 bp duplications. The bulk of my work in this paper focuses on the characterization of the products of transposition, the mechanism of transposition, and the breakage and joining reactions. We collaborated directly on the intermediates of transposition. Pascal concentrated on the double-strand break species and I concentrated on the excised-linear transposon species. Pascal did the work on separating the transposition proteins into different crude extracts and showing that they could complement each other in vitro. The experiment on transposition immunity is also my work.

# Tn7 Transposition In Vitro Proceeds through an Excised Transposon Intermediate Generated by Staggered Breaks in DNA

**Roland Bainton, Pascal Gamas,* and Nancy L. Craig**
Department of Biochemistry and Biophysics
and Department of Microbiology and Immunology
George W. Hooper Foundation
University of California, San Francisco
San Francisco, California 94143

## Summary

**We have developed a cell-free system in which the bacterial transposon Tn7 inserts at high frequency into its preferred target site in the Escherichia coli chromosome, *attTn7*; Tn7 transposition in vitro requires ATP and Tn7-encoded proteins. Tn7 transposes via a cut and paste mechanism in which the element is excised from the donor DNA by staggered double-strand breaks and then inserted into *attTn7* by the joining of 3′ transposon ends to 5′ target ends. Neither recombination intermediates nor products are observed in the absence of any protein component or DNA substrate. Thus, we suggest that Tn7 transposition occurs in a nucleoprotein complex containing several proteins and the substrate DNAs and that recognition of *attTn7* within this complex provokes strand cleavages at the Tn7 ends.**

## Introduction

Mobile DNA segments translocate by a variety of mechanisms. These mechanisms differ in the nature of the recombining sites, the type and order of strand breakage and joining reactions, and the degree of involvement of DNA replication. We are interested in understanding the molecular basis of transposition, a reaction in which a discrete DNA segment moves between nonhomologous DNA sites (Berg and Howe, 1989). This reaction is distinguished by the fact that it involves three distinct DNA segments, the two ends of the transposon and the target site; no sequence similarity between the transposon ends and the target site is required. Another hallmark of transposition is that element insertion is accompanied by the duplication of several base pairs of target sequence. This duplication results from the introduction and subsequent repair of a staggered break in the target DNA, the transposon being joined to the ends of the overhanging target strands at this break.

We are particularly interested in understanding the transposition of the bacterial transposon Tn7 (Barth et al., 1976; Craig, 1989), because this element displays an unusual degree of insertion specificity. Most transposable elements show relatively little insertion site selectivity when transposing to large DNA molecules (Berg and Howe, 1989). By contrast, Tn7 inserts at high frequency into a single site in the Escherichia coli chromosome called *attTn7* (Barth et al., 1976; Lichtenstein and Brenner, 1982; Craig, 1989). Tn7 is unrelated in nucleotide sequence to *attTn7* (Lichtenstein and Brenner, 1982; Gay et al., 1986; McKown et al., 1988); thus, no DNA sequence homology is involved in the selection of *attTn7* as a preferential target. This element's transposition machinery is elaborate. Tn7 encodes a surprising array of transposition genes, *tnsABCDE*. Insertion into *attTn7* in vivo requires four of these genes, *tnsABC* + *tnsD* (Rogers et al., 1986; Waddell and Craig, 1988; Kubo and Craig, 1990). Large DNA sites at both ends of Tn7 (Arciszewska et al., 1989) and at *attTn7* (McKown et al., 1988; Gringauz et al., 1988; Qadri et al., 1989) are also required. When *attTn7* is unavailable, Tn7 resembles most other transposable elements, moving at low frequency and inserting into many different sites (Rogers et al., 1986; Waddell and Craig, 1988; Kubo and Craig, 1990). This alternative transposition pathway requires a distinct, but overlapping, ensemble of four *tns* genes, *tnsABC* + *tnsE*.

Although many transposable elements have been characterized genetically, only a few have been examined biochemically (Mizuuchi, 1983; Brown et al., 1987; Morisato and Kleckner, 1987; Eichinger and Boeke, 1988). The high frequency of Tn7 insertion into *attTn7* suggested that this reaction would be amenable to biochemical dissection.

We report here the development of a cell-free system for Tn7 transposition. We find that ATP and four *tns* fractions, i.e., protein preparations derived from cells containing individual *tns* genes, are required for the formation of recombination intermediates or products; no recombination is detected when any fraction is omitted. Furthermore, the presence of the *attTn7* target site is required for the initiation of recombination, i.e., the production of recombination intermediates. These results suggest that Tn7 transposition is executed by a nucleoprotein assembly containing multiple proteins and the substrate DNAs and that the proper assembly of this entire recombination machine is required for the strand breakage reactions that initiate recombination.

We also demonstrate that Tn7 transposition in vitro is a nonreplicative reaction that proceeds by a series of double-strand breaks that disconnect the transposon from the flanking donor DNA to produce an excised transposon, which is then inserted into *attTn7*. The DNA breakage reactions at the transposon ends and at the target DNA both occur by staggered breaks that generate 5′ overhanging ends. Tn7 insertion occurs via single-strand joins between the 3′ transposon ends and the 5′ ends of target DNA.

## Results

### Establishment of a Cell-Free System for Tn7 Transposition

Tn7 transposition in vitro occurs efficiently when substrate DNAs, a donor plasmid containing a mini-Tn7 transposon,

*Current Address: Laboratoire de Biologie Moleculaire des Relations Plantes-Microorganismes, CNRS-INRA, BP27-31326 CASTANET TOLOSAN CEDEX, France

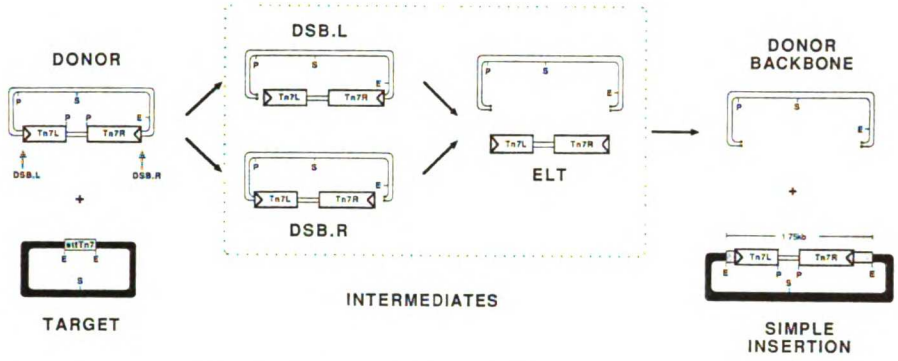



Figure 1. Tn7 Transposition Substrates, Intermediates, and Products

The 6.1 kb donor molecule contains a 1.6 kb mini-Tn7 element flanked by sequences unrelated to *attTn7*; the mini-Tn7 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-strand 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 element inserts site- and orientation-specifically into *attTn7*, producing a 4.5 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. Restriction sites: E = EcoRI, P = PstI, and S = ScaI.

and a target plasmid containing *attTn7* (Figure 1), are incubated at 30°C with *tns* protein fractions, ATP, and MgAc. **Recombination is detected by extraction of DNA from the reaction mixture, digestion with restriction enzymes, separation by gel electrophoresis, and hybridization with specific probes.**

**After incu**bation, several new species are detectable **with a mini-Tn7-specific probe (Figure 2A): a simple insertion transposition product (mini-Tn7 in *attTn7*) and several recombination intermediates, including donor DNA molecules that are broken by a double-strand break at either the** Tn7L–backbone junction (DSB.L) or the Tn7R–backbone junction (DSB.R) (see Figure 1). Insertion specificity was verified by DNA sequence analysis of cloned *attTn7*::mini-Tn7 segments. All insertions were orientation-specific, i.e., with Tn7R oriented towards the bacterial *glmS* gene, **and were accompanied by a target duplication of 5 bp.** More than 85% of the insertions occurred at either of two adjacent nucleotide positions in *attTn7*, and a few nearby insertions were also observed. This distribution closely resembles the pattern of Tn7 insertions into *attTn7* in vivo (Lichtenstein and Brenner, 1982; Gay et al., 1986; McKown et al., 1988; Gringauz et al., 1988). The efficiency of the cell-free reaction is striking, especially in view of the fact that it is conducted in crude extracts: translocation of about 25% of the Tn7 elements from donor molecules to target molecules is readily observed.

Tn7 transposition to *attTn7* in vivo requires four Tn7-

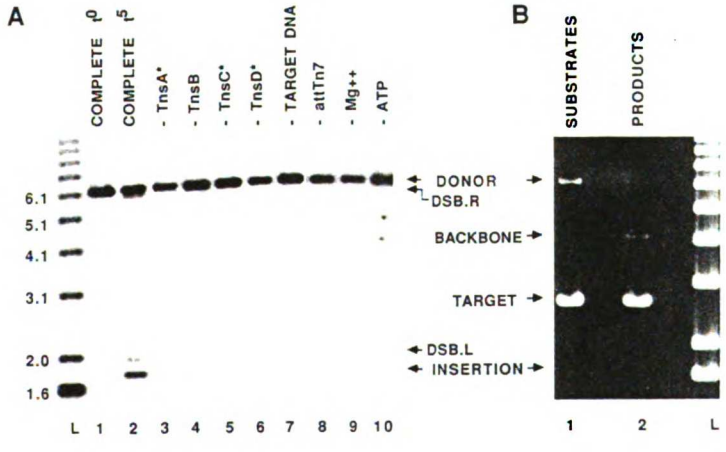



Figure 2. Tn7 Transposition In Vitro

**(A) Requirements for Tn7 transposition. Shown** is an autoradiogram that displays EcoRI-digested species detected with a mini-Tn7-specific probe from transposition reactions in which the indicated components were omitted. DNA molecules are labeled as in Figure 1; DSB.L and DSB.R appear very different in size because of the EcoRI digestion. Reactions were performed as described in the Experimental Procedures, except for lane 8, where plasmid lacking *attTn7* was used as a target instead of the usual *attTn7*-containing plasmid. Asterisks denote that the indicated Tns protein was supplied in a crude fraction that also contained host proteins. Following the preincubation step, reactions were incubated for 5 min at 30°C, except that shown in lane 1, which was stopped immediately after mixing of the reaction components. Lane L shows DNA size markers.

(B) A gapped donor backbone is a transposition product. An ethidium-stained agarose gel is shown, which displays the EcoRI-digested substrates (lane 1) and products (lane 2) of an in vitro Tn7 transposition reaction; DNA molecules are labeled as in Figure 1. After recombination, two new species are evident: an *attTn7* segment molecule containing a simple insertion of mini-Tn7 and a donor backbone from which the mini-Tn7 element has been excised. Lane L shows DNA size markers.

encoded genes, *tnsABC + tnsD* (Rogers et al., 1986; Waddell and Craig, 1988; Kubo and Craig, 1990). Tn7 transposition in vitro can occur efficiently when four protein fractions, each derived from cells containing one of these *tns* genes, are mixed (Figure 2A). In this experiment, the *tnsA* fraction is a slightly fractionated extract containing TnsA, the *tnsB* fraction is substantially purified TnsB protein, the *tnsC* fraction is a slightly fractionated extract containing TnsC, and the *tnsD* fraction is partially purified TnsD protein. It is notable that no transposition intermediates or products are detected when any one of the *tns* fractions is omitted (Figure 2A, lanes 3–6).

We also find that no recombination intermediates or products are detected when *attTn7* is omitted (Figure 2A, lane 8). This result is particularly interesting, as it suggests that the transposon ends "see" the target site before recombination actually begins, i.e., before any double-strand breaks occur at the transposon ends. We have been unable to detect any changes in the target plasmid when any *tns* fraction or the mini-Tn7 donor plasmid is omitted from the reaction (data not shown).

These experiments show that the four *tns* fractions and the *attTn7* target are required for recombination. Notably, the production of both recombination intermediates and products requires the presence of all the reaction components, i.e., no partial reactions are detected when any single component is omitted. An attractive hypothesis that explains these results is that Tn7 transposition is a highly concerted reaction carried out by a nucleoprotein assembly (machine) containing the DNA substrates and multiple recombination proteins. Because transposition in vitro requires combining four different fractions, at least four different proteins are likely to participate directly in recombination.

Tn7 transposition in vitro has several other requirements. ATP (Figure 2A, lane 10) or dATP (data not shown) must be added. Efficient recombination also requires 10 mM MgAc (lane 9). If MgAc is included at the beginning of the incubation, little recombination is observed (data not shown); substantially more (perhaps 50-fold) recombination occurs when MgAc is added after a preincubation step. Transposition is most efficient with supercoiled substrates, but considerable transposition can be observed with relaxed DNAs (data not shown). Recombination is stimulated by the presence of polyvinyl alcohol or polyethylene glycol (data not shown).

### Tn7 Transposition In Vitro Does Not Involve Replication of the Element

A key issue in understanding a transposition reaction is determining whether the transposon is copied by DNA replication or whether it is transferred from the donor molecule to the target molecule via a nonreplicative, cut and paste reaction. Several observations indicate that Tn7 transposition in vitro is nonreplicative. In addition to the simple insertion of mini-Tn7 into *attTn7*, another transposition product is a gapped donor molecule, i.e., a donor backbone from which mini-Tn7 has been excised (Figure 2B). The simple insertion product and gapped donor backbone appear with approximately the same time course and in equimolar amounts (data not shown). Production of this gapped donor backbone indicates that Tn7 translocation occurs by cutting the transposon away from the donor DNA and joining it to the target DNA without replication, a view supported by the demonstration that an excised transposon is a transposition intermediate (see below).

Another indication of the nonreplicative nature of Tn7 transposition is that recombination in vitro is not affected by the inclusion of DNA synthesis inhibitors such as dideoxynucleotides under conditions in which a strong inhibition of DNA polymerase activity is observed (data not shown).

### Double-Strand Breaks at the Ends of Tn7 Produce Recombination Intermediates

Several types of evidence indicate that the DSBs, i.e., donor DNA molecules broken by a double-strand break at a transposon end, are indeed transposition intermediates. These species have the kinetic properties of intermediates: they are most abundant at early times and disappear at late times during the incubation (Figure 3). These kinetic properties are particularly evident when the overall rate of transposition is slowed by lowering the incubation temperature to 22°C. Here, the DSB species are detected earlier and in greater amounts than are the transposition products; it is also notable that a considerable fraction of the substrate donor DNA is converted to DSBs. Disappearance of the DSBs is not due to nonspecific nuclease degradation, because a control fragment was not degraded in these reactions. The view that the DSBs are transposition intermediates is also supported by the fact that they, like the simple insertion product, are generated by staggered breaks in DNA (see below). Similar levels of DSBs are observed when the reactions are stopped by a variety of

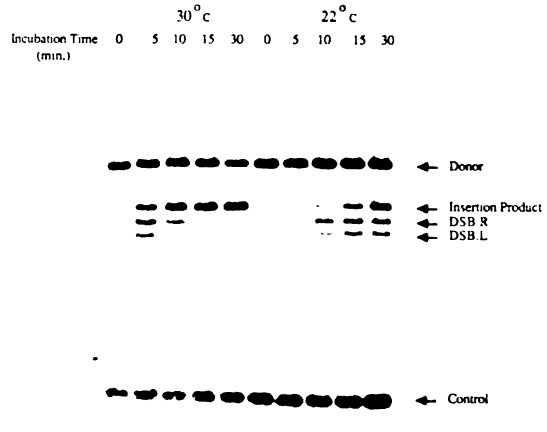



Figure 3. The Double-Strand Break Species Have the Kinetic Properties of Transposition Intermediates

Shown is an autoradiogram that displays ScaI-digested species detected by a mini-Tn7-specific probe produced in transposition reactions, after incubation for the indicated times at either 30°C or 22°C (preincubation also at 22°C). Reactions were performed as described in the Experimental Procedures, except that they also contained a 1200 bp linear fragment recognized by the probe ("Control") to evaluate nuclease degradation. DNAs are as described in Figure 1; the DSBs appear to be different in size because of the ScaI-digestion.

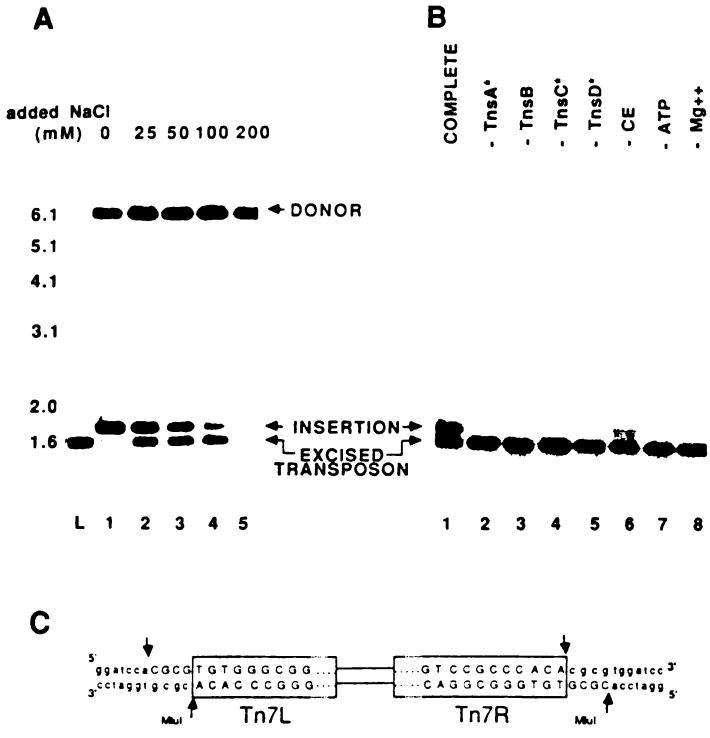



Figure 4. An Excised Transposon Is a Recombination Intermediate

(A) Production of excised transposon. Shown is an autoradiogram that displays EcoI-digested species detected with a mini-Tn7-specific probe from transposition reactions that contained various amounts of monovalent salt. Lanes 1–5: reactions supplemented with the indicated salt; the salt concentrations supplied by protein fractions were 30 mM KCl and 3 mM NaCl. Lane L: DNA size markers with a prominent 1.6 kb species. DNAs are as described in Figure 1.

(B) An "excised" transposon is an effective transposition substrate. Shown is an autoradiogram of the EcoRI-digested products detected with a mini-Tn7-specific probe of transposition reactions containing about 25 ng (0.25 nM) of MluI-generated "excised" transposon as the mini-Tn7 substrate, instead of the usual mini-Tn7-containing donor plasmid. Reactions were carried out as described in the Experimental Procedures, lacking particular components as indicated. Asterisks denote that the indicated Tns protein was supplied in a crude fraction that also contained host proteins. Addition of crude extract (CE) from cells lacking Tns proteins stimulates, but is not essential for, recombination. DNAs are labeled as in Figure 1, except that the "excised" transposon is generated by restriction instead of recombination (as in Figure 4A and elsewhere).

(C) MluI transposon. An MluI transposon–containing plasmid was produced as described in the Experimental Procedures. Digestion with MluI enzyme, which cuts at the indicated positions (arrows), releases a linear "excised" transposon, whose 3′ ends are cleanly exposed (as also with authentic excised transposon generated in recombination reactions) and whose 5′ ends have 4 nt extensions (as opposed to the 3 nt extensions of authentic excised transposon).

methods, including treatment with proteinase and phenol (data not shown); thus, the DSBs are not likely to be protein–DNA complexes, nor is DNA breakage likely to be provoked by our manipulations.

### Excised Tn7 Can Be a Transposition Substrate

In addition to DSBs, a transposon species that has been excised from the donor backbone by double-strand breaks at both transposon ends is produced in the incubation mixtures. This excised species is present in all reactions, but often at a very low level. However, high levels of excised transposon and low levels of insertion product are evident under some reaction conditions, for example, with high NaCl (Figure 4A). We have been unable to observe disappearance of the excised transposon, as might be expected for a recombination intermediate. However, high levels of excised transposon may reflect substantial impairment of the insertion step. Like the simple insertion product and the DSB species, the excised transposon is generated by staggered breaks in DNA (see below).

To determine whether an excised mini-Tn7 can participate in recombination, we asked whether exogenously supplied, "excised" mini-Tn7 could insert into *attTn7* when present as the sole transposon substrate. We introduced restriction sites at the Tn7 termini such that restriction enzyme digestion generates a mini-Tn7 species with staggered ends (Figure 4C), the 3′ ends being cut adjacent to L1 and R1 and the 5′ ends cut within the flanking DNA to generate 4 nucleotide (nt) overhangs. (As described below, we believe that the excised mini-Tn7 species generated during transposition has exposed 3′ ends and 3 nt overhangs on its 5′ ends.) The exogenously supplied mini-Tn7 element also has two sequence changes in its Tn7L segment.

The restriction-generated "excised" mini-Tn7 inserts efficiently in site- and orientation-specific fashion into *attTn7* (Figure 4B), supporting the view that the excised transposon is indeed a transposition intermediate. As with a plasmid substrate, insertion of the "excised" transposon requires the presence of all of the *tns* fractions, ATP, and MgAc. In addition to demonstrating that a transposon species that is not connected to a donor backbone can participate in transposition, these results suggest that no *tns* function is involved exclusively in the excision stage of recombination.

We also note that we have been unable to detect a strand transfer intermediate under any condition, either in complete Tn7 transposition reactions or when any *tns* component is omitted. This species, in which the transposon, the donor backbone, and the target are covalently linked, is a key intermediate in bacteriophage Mu transposition (Shapiro, 1979; Craigie and Mizuuchi, 1985; Pato, 1989).

### The Polarity of Strand Breakage and Joining

Tn7 insertion is accompanied by the duplication of 5 bp of target sequence (Lichtenstein and Brenner, 1982), pre-

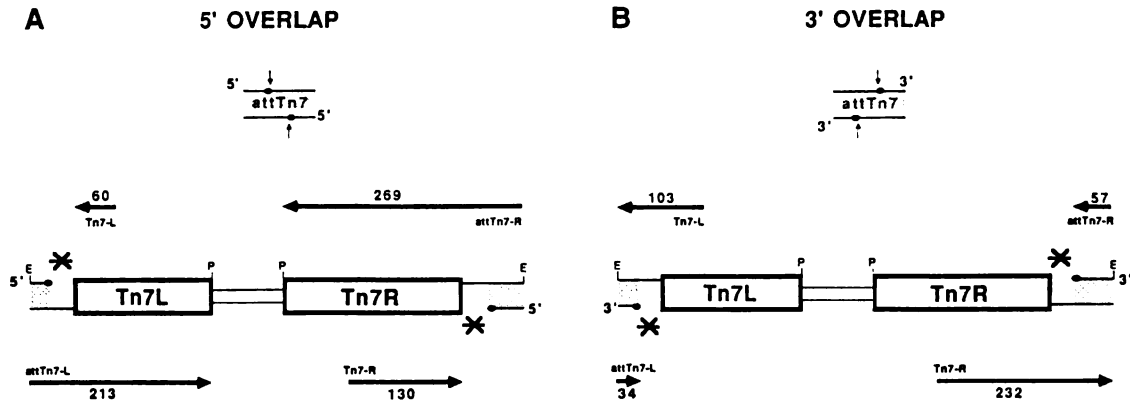



Figure 5. Transposition Can Occur via Two Different Polarities of Strand Breakage and Joining

The duplication of target sequences upon transposon insertion is presumed to result from cleavage of the target site by a staggered break and subsequent repair of the gap resulting from linkage of the transposon ends to the target site ends. Two polarities of target cleavage and transposon joining can be imagined: (A) the target is cleaved by a staggered break generating 5' overhanging ends that join to the 3' transposon ends, or (B) the target is cleaved by a staggered break generating 3' overhanging ends that join to the 5' transposon ends. The two polarities can be distinguished by identifying which target strands are covalently joined to which transposon strands in the simple insertion transposition product. Asterisks mark the specific, diagnostic gaps in the transposition product. The primers Tn7-L, attTn7-L, Tn7-R, and attTn7-R are complementary to particular strands in *attTn7* and the Tn7 ends. Extension from these primers yields products of different and defined lengths as indicated (thick arrows), depending on which polarity of target breakage and strand joining is used in Tn7 transposition. Restriction sites: E = EcoRI, P = PstI.

sumed to result from the cleavage and subsequent repair of a staggered break in the target DNA (Grindley and Sherratt, 1979; Shapiro, 1979). Two alternative types of target breaks can be imagined (Figure 5), one producing 5' overhanging ends (Figure 5A) and the other producing 3' overhanging ends (Figure 5B). These two types of target cleavage will result in two distinct types of joints between the transposon ends and the target site: either the 3' transpo-

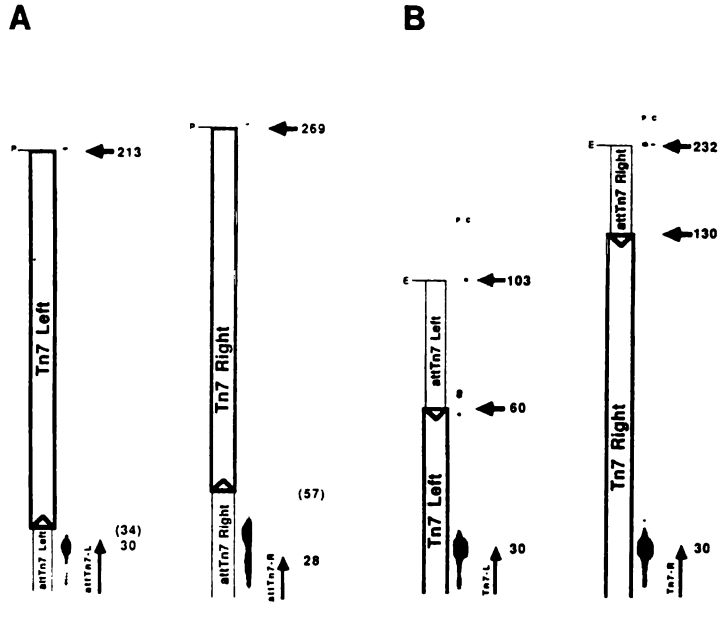



Figure 6. The 3' Ends of Tn7 Join to the 5' Target Ends

Shown are autoradiograms of DNA sequencing gels on which the products of primer extensions are displayed, in which the simple insertion transposition product was a template with the indicated primers. Also shown schematically are the relevant Tn7 end and *attTn7* regions (see also Figure 5). The lengths of the extension products were determined by comparison with dideoxy sequencing ladders. The lengths (28 or 30 nt) and mobilities of the primers are also shown.

(A) The 3' transposon ends are covalently joined to *attTn7*. Extensions from the target attTn7-L and attTn7-R primers using the EcoRI-digested and PstI-digested simple insertion product as templates are shown. The observed products (horizontal arrows), 213 nt for attTn7-L and 269 nt for attTn7-R, are the lengths expected if the 3' transposon ends are joined to *attTn7* (Figure 5A). No product is evident at the positions corresponding to free 3' transposon ends, extension from attTn7-L is 34 nt, and that from attTn7-R is 57 nt.

(B) The 5' transposon ends are not joined to *attTn7*. Extensions from the transposon end Tn7-L and Tn7-R primers using PstI-digested simple insertion product as template (P: left lane of pair) and, as a control, an intact *attTn7*::mini-Tn7 fragment as a template (C: right lane of pair). Two classes of products are observed from each primer with the simple insertion template. One class of products extends to about the 5' transposon ends (black arrows), approximately 60 nt from Tn7-L and 130 nt from Tn7-R (the exact lengths of these products are considered in Figure 7). Thus, some 5' transposon ends in the simple insertion product are not linked to *attTn7*, corresponding to a 5' staggered break in the target (Figure 5A). The other class of products (upper, gray arrows) extend to the positions expected for linkage of the 5' transposon ends to the target (103 nt from Tn7-L and 232 nt from Tn7-R), probably reflecting DNA repair.

son ends are joined to the 5′ target ends (Figure 5A), or the 5′ transposon ends are joined to the 3′ target ends (Figure 5B). In the first case (Figure 5A), gaps between the 5′ transposon ends and the 3′ target ends are expected; the second breakage and joining polarity leaves gaps adjacent to the 3′ transposon ends (Figure 5B).

To determine the polarity of Tn7 strand breakage and joining, we determined the positions of the intact and gapped strands in the simple insertion product, by using it as a template for extensions with end-labeled oligonucleotide primers specific to the transposon ends and the flanking *attTn7* DNA.

Our analysis has revealed that the target site is broken by staggered breaks with 5′ overhanging ends (Figure 5A). Primers attTn7-L and attTn7-R anneal with the strands of the target DNA that, following transposition using 5′ target breaks, will be joined with the 3′ ends of the inserted transposon. We find that the attTn7-L extension product is 213 nt and that the attTn7-R product is 269 nt (Figure 6A). These lengths correspond to the distance from the primers to restriction sites within the transposon ends (Figure 5A). Thus, in the simple insertion product, these template strands are intact, i.e., the 3′ transposon ends are covalently linked to the 5′ ends of the target. Similar primer extension analysis of the gapped donor backbone supports the view that Tn7 excision involves cleavages directly adjacent to the 3′ transposon ends, which are then joined to target DNA (data not shown).

If the 3′ ends of the transposon are joined to the target DNA, i.e., are intact, then it is expected that the other strands are gapped (Figure 5A). The state of the 5′ transposon ends was examined by extension from the Tn7-L and Tn7-R primers (Figure 6B). For both primers, two distinct classes of extension products are observed. From Tn7-L, there is a cluster of products at about 60 nt and another species at 103 nt from Tn7-R, there is a cluster of products at about 130 nt and another species at 232 nt. The shorter class of extension products from each primer (those at about 60 nt for Tn7-L and about 130 nt for Tn7-R) correspond closely to the distance to the 5′ transposon ends. (The exact lengths of these products are considered in detail below.) These results demonstrate that a considerable fraction of the 5′ transposon strands are broken and thus are not covalently linked to target DNA.

The longer extensions from the Tn7-L and Tn7-R primers do extend to lengths consistent with covalent linkage of the 5′ transposon end with the target DNA, i.e., for Tn7-L, 103 nt and for Tn7-R, 232 nt. We hypothesize that these species result from some repair in this crude system of the gap that initially flanks the 5′ ends. Repair of the right end gap appears to be more efficient than repair of the left end gap, i.e., more 232 nt Tn7-R product is observed than 103 nt Tn7-L product.

We interpret these results to indicate that during Tn7 transposition, the target DNA is cleaved by a staggered break that yields 5′ overhanging ends and that the 5′ ends of the target DNA are joined to the 3′ ends of the transposon (Figure 5A). We presume that a 5 bp gap flanks the 5′ transposon ends and is subsequently converted to duplex form through DNA repair.

**Staggered Breaks at the Transposon Ends**

What type of DNA strand cleavages release Tn7 from the donor molecule? Analysis of the simple insertion product suggests that Tn7 is cut away from the donor DNA via staggered breaks, in which one DNA strand is cleaved precisely at the 3′ transposon end, and the other strand is cleaved at a displaced position within the flanking donor DNA. We were led to this view through our characterization of the 5′ ends of the transposon in the simple insertion product by extension from the primers Tn7-L and Tn7-R (see Figure 5A). If the 5′ ends were cut flush with the 3′ ends, i.e., at exactly the end of the transposon, we would expect the Tn7-L extension product to be 60 nt and the Tn7-R product to be 130 nt. Rather, we find that the Tn7-L extension product (Figure 7A) and the Tn7-R extension product (Figure 7B) are both 3 nt longer than expected. This result would be obtained if cleavage of the 5′ strands occurs within the flanking donor DNA.

If Tn7 is excised from the donor backbone by staggered breaks involving strand cleavage within the flanking donor DNA, we would expect sequences from the flanking donor backbone to be attached to the 5′ transposon ends. Chemical sequence analysis of the 63 nt Tn7-L extension product reveals that the nucleotides attached to the 5′ terminus of the left end of Tn7 are those adjacent to this transposon end in the flanking donor DNA (data not shown). This result provides strong evidence that cleavage of the 5′ transposon strand actually occurs at a position displaced from the transposon end via a staggered break. We have also observed such staggered breaks at the transposon ends when Tn7 transposes in vitro from another unrelated donor site (data not shown); thus, these staggered breaks are not peculiar to a particular donor site.

We have performed similar primer extension analysis of the ends of the DSBs and the excised transposon generated in in vitro transposition reactions (Figure 7). In all cases, the structure of the 5′ transposon strands was identical, i.e., 3 nt of donor backbone are attached to the transposon. These results support the view that the DSBs and the excised transposon are authentic transposition intermediates and that the transposon is cut away from the donor backbone by staggered breaks in DNA (Figure 8).

The donor sequences attached to the excised transposon are not essential for transposition. An "excised" transposon species generated by the polymerase chain reaction (PCR) and primers whose 5′ strands end with the transposon termini (as in Figure 7, lane 1) inserts efficiently into *attTn7* (data not shown).

**Tn7 Target Immunity Is Active in the Cell-Free Transposition System**

Tn7 displays target immunity in vivo; that is, the presence of a copy of Tn7 in a target DNA molecule substantially reduces the frequency of a second Tn7 insertion into that target (Hauer and Shapiro, 1984; Arciszewska et al., 1989). Target immunity is fundamentally related to transposition, in that both events require the same transposon end sequences (Lee et al., 1983; Adzuma and Mizuuchi, 1988). The presence in a target molecule of the transposition sequences from the right end of Tn7 (R1–199) is suffi-

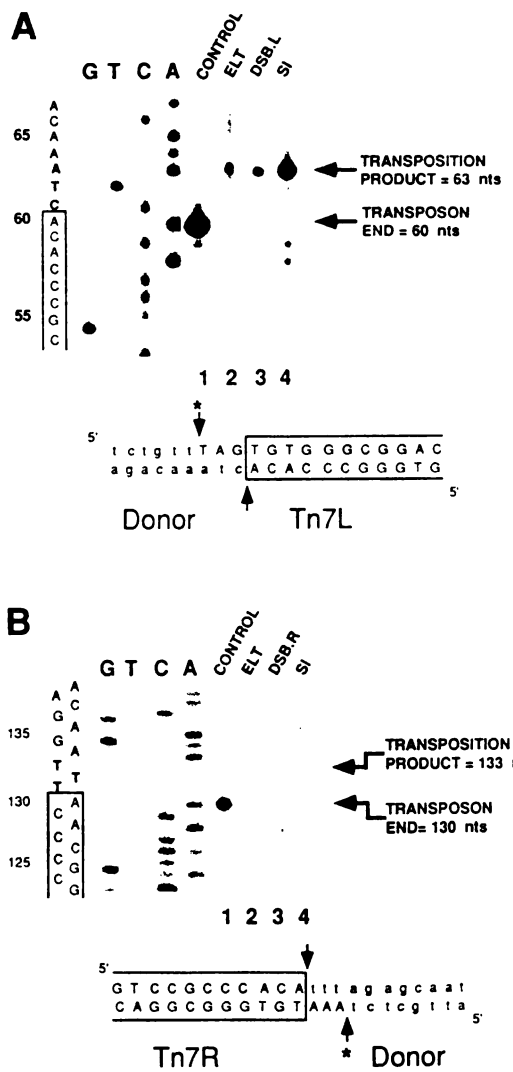



Figure 7. Donor Nucleotides Remain Attached to the 5' Transposon Strands in Transposition Intermediates and the Simple Insertion Product

Shown are autoradiograms of DNA sequencing gels that display primer extensions from Tn7 end oligonucleotides (see Figure 5A), under conditions where nontemplated addition does not occur (see Experimental Procedures). The templates were: lane 1, as a control, a PCR-generated transposon that terminates exactly at the 5' transposon ends; lane 2, excised linear transposon (ELT); lane 3, the DSB species; and lane 4, the simple insertion transposition product (SI). At the left is a dideoxynucleotide sequencing ladder generated with the same oligonucleotide on an intact *attTn7*::mini-Tn7 fragment. The nucleotide sequence of the corresponding 5' transposon strand and the sequence of the flanking donor strand in pEM are shown, along with the distance in nucleotides to the oligonucleotide primer. At the bottom, the positions of strand cleavages at the transposon termini are shown schematically; the 5' cleavage positions determined in this experiment are indicated with an asterisk.

(A) Extension from the Tn7-L primer. Extension using the simple insertion as template yields a product 3 nt longer than expected for extension to the 5' transposon end.

(B) Extension from the Tn7-R primer. Extension using the simple insertion as template yields a product 3 nt longer than expected for extension to the 5' transposon end.

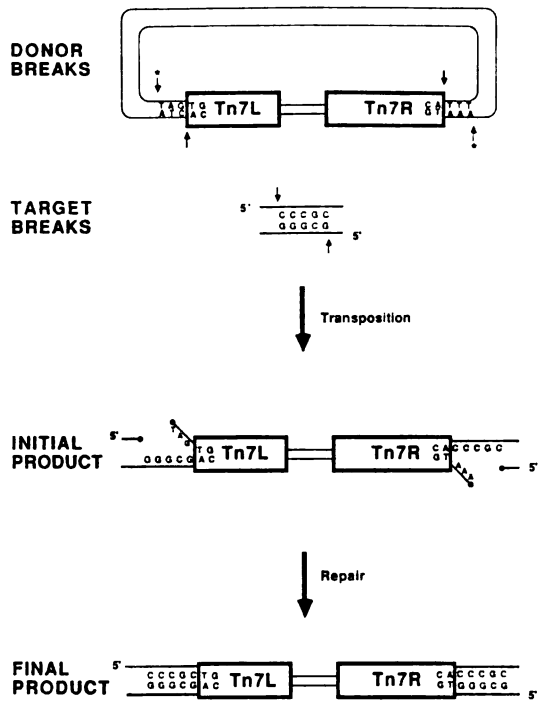



Figure 8. Strand Breakage and Joining during Tn7 Transposition

We suggest that both the ends of Tn7 and *attTn7* are broken by staggered breaks that generate 5' overhanging ends. Breaks at the target site are proposed to be 5 bp in length, and those at the ends of Tn7 are proposed to be 3 bp in length. The 3' Tn7 ends join to the 5' target site ends, generating an initial transposition product whose 5' transposon ends are attached to several nucleotides of donor DNA and are flanked by 5 bp gaps. DNA repair fills in the gaps and concomitantly removes the nonhomologous donor sequences from the 5' transposon ends, generating an intact duplex flanked by 5 bp target duplications. The target sequence shown is the 5 bp in *attTn7* usually duplicated upon Tn7 insertion.

cient to provide a high level of immunity in vivo; by contrast, a truncated right end (R1–42) cannot provide immunity and is also defective in transposition (Arciszewska et al., 1989).

We find that Tn7 target immunity is operative in the cell-free transposition system. We evaluated immunity in vitro by examining the target activity of *attTn7* plasmids that also contain Tn7R segments (Figure 9). In the experiments shown in Figure 9, lanes 2 and 3, the reactions contained a mixture of two target DNAs, one an *attTn7* plasmid (target A) and the other a plasmid containing both an *attTn7* segment and a segment from Tn7R, either R1–199 or R1–42 (targets B and C, respectively), located at a distance of about 1 kb from *attTn7*. No mini-Tn7 insertion into the *attTn7* segment of target B is detectable (Figure 9, lane 2), although insertion into the control *attTn7* molecule target A is observed. This result suggests that the presence of R1–199 in target B inactivates this molecule. In contrast, the presence of R1–42 in target C does not block insertion (Figure 9, lane 3). We note that in the presence of target B, a modest decrease in insertion into the control target A

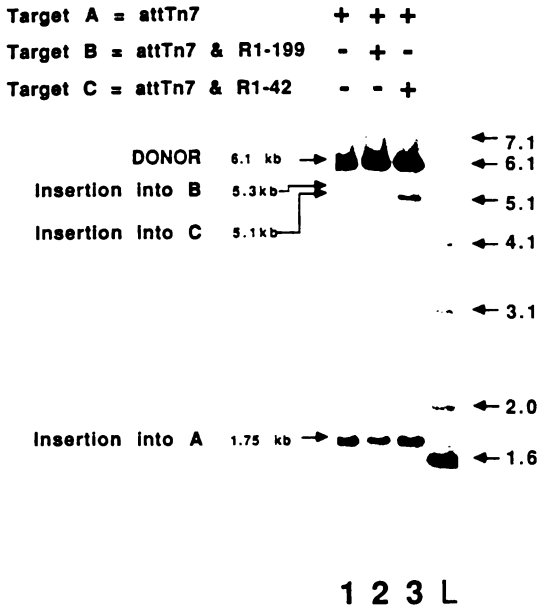



Figure 9. Tn7 Transposition In Vitro Displays Target Immunity

Shown is an autoradiogram that displays EcoRI-digested species detected with a mini-Tn7-specific probe from transposition reactions in which the *attTn7* target molecule also contained a Tn7R segment about 1 kb from *attTn7*. Target A is the *attTn7* target used in other experiments (pKAO4–3); Target B (pLA11) and Target C (pLA24) contain *attTn7* and the indicated Tn7R segment. The concentration of the target plasmids (3 nM each) is about 3-fold less than that of the standard conditions, to avoid a trans-inhibition by Tn7R-containing targets (see text). The position of a simple insertion product in each target plasmid is also shown. Lane 1, target A; lane 2: target A + target B; lane 3: target A + target C.

molecule is also observed (compare lanes 1 and 2). We suspect that this decrease results from a competition between the R1–199 segment in target B and the donor DNA for transposition proteins that interact with the transposon ends (the molar ratio of donor to target B is 1:7); such an effect has been observed in vivo (Arciszewska et al., 1989).

## Discussion

We have developed a cell-free system that faithfully reproduces many aspects of Tn7 transposition in vivo. Tn7 inserts in vitro at high frequency in site- and orientation-specific fashion into its preferred target site, *attTn7*. Transposition occurs efficiently when exogenous DNA substrates, one a donor molecule containing a mini-Tn7 element and the other an *attTn7*-containing target molecule, are incubated with extracts derived from cells containing the Tn7-encoded transposition genes *tnsABC* + *tnsD* required for insertion into *attTn7* in vivo (Rogers et al., 1986; Waddell and Craig, 1988; Kubo and Craig, 1990). The in vitro system also displays the phenomenon of transposition immunity that Tn7 displays in vivo; the presence of Tn7 in a target molecule blocks subsequent insertion of another copy of Tn7 into that target molecule (Hauer and Shapiro, 1984; Arciszewska et al., 1989). Tn7 transposition in vitro also requires the presence of ATP, which is likely to be directly involved in recombination, because TnsC is an ATP-binding protein (P. Gamas and N. L. C., unpublished data). An auxiliary role of ATP may be to facilitate supercoiling of the DNA substrates by promoting the action of DNA gyrase (Gellert et al., 1976).

### Tn7 Transposition Occurs in a Multiprotein Complex

An attractive hypothesis is that the four Tns proteins participate directly in recombination. Transposition in vitro requires the mixing of four protein fractions derived from four different bacterial strains, each containing one of the four *tns* genes, *tnsABC* + *tnsD*. Specific roles for some Tns proteins have emerged from other experiments (see below). It is likely that host proteins will also be directly involved; most transactions involving extrachromosomal DNA elements are mediated by a combination of element- and host-encoded proteins (Pato, 1989; Thompson and Landy, 1989; Kleckner, 1990). Mutations in several E. coli genes can affect Tn7 transposition (O. Hughes and N. L. C., unpublished data). Reconstitution of Tn7 transposition in vitro with purified proteins will be required to identify both the Tns- and host-encoded participants in recombination.

We propose that the Tn7 recombination proteins, together with the substrate donor and target DNAs, form a specialized nucleoprotein assembly in which transposition occurs. Elaborate protein–nucleic acid complexes also mediate other recombination reactions, DNA replication, transcription, and RNA processing (Echols, 1990). We are attracted to the view that Tn7 transposition occurs in such a complex, because no transposition intermediates or products are detected in the absence of any required protein or DNA component. Thus, we imagine that the complete transposition machinery must properly juxtapose the three DNA segments involved in recombination—the two transposon ends and the target DNA—and the multiple transposition proteins, prior to the initiation of the strand breakage and joining reactions that underlie recombination. In this scenario, *attTn7* is recognized by the rest of the recombination machinery before strand breakage at the transposon ends occurs. Thus, specific recognition of *attTn7* actually provokes transposition by stimulating the initiation of recombination.

What are the roles of the recombination proteins? Other work from this laboratory has established that two of the Tns proteins are sequence-specific DNA-binding proteins: TnsB binds to the ends of Tn7 (McKown et al., 1987; L. Arcisewska, R. McKown, and N. L. C., unpublished data), and TnsD binds to *attTn7* (Waddell and Craig, 1989; K. Kubo and N. L. C., unpublished data). Possible roles for TnsA and TnsC (and/or host proteins) include acting as linker proteins to mediate the specific juxtaposition of the TnsB-bound transposon ends and the TnsD-bound target site and interacting with the positions of strand breakage at the Tn7 termini and at *attTn7*.

### Tn7 Transposition Is Nonreplicative

Two distinct pathways for the movement of transposable elements have been identified: nonreplicative transposi-

tion, in which the transposon is cut out of the donor site and pasted into the target site; and replicative transposition, in which a copy of the transposon is made by DNA synthesis and inserted into the target site (Berg and Howe, 1989). We have established that Tn7 translocates in vitro via a nonreplicative mechanism (Figure 1). The principal evidence for this view is that the transposition products we observe are a simple insertion of Tn7 into *attTn7* and a linear donor molecule from which Tn7 has been excised. In vivo studies also provide evidence that Tn7 transposes to *attTn7* via a nonreplicative mechanism (K. Orle, R. Gallagher, and N. L. C., unpublished data).

The fate in vivo of the donor backbone from which Tn7 is excised is unknown. Tn7 transposition in vivo provokes expression of the SOS system (A. Stellwagen and N. L. C., unpublished data), consistent with the production of a gapped donor backbone (Roberts and Kleckner, 1988). We do not observe simple rejoining of the backbone in vitro or in vivo (data not shown). Perhaps the broken donor backbone is lost when transposition occurs, i.e., donor suicide occurs. Another possibility is that the broken donor molecule is repaired by a double-strand-gap repair reaction using an unbroken Tn7-containing donor molecule as a template, as has been observed in Drosophila P-element transposition (Engels et al., 1990).

### The Tn7 Transposition Pathway

Our analyses have revealed that Tn7 transposition occurs by a series of double-strand breaks and single-strand joins (Figures 1 and 8). Double-strand breaks at the transposon ends produce a linear excised transposon that is subsequently inserted by single-strand joins into an *attTn7* target site that is also broken by a double-strand break.

Double-strand cleavages of a donor molecule at both Tn7L and Tn7R generate an excised transposon intermediate. These double-strand breaks are staggered, generating 5′ overhanging ends such that the 3′ transposon termini are cleanly exposed, but several nucleotides of donor backbone remain attached to the 5′ transposon strands. We observe that three donor nucleotides are attached to the 5′ transposon ends, suggesting that these staggered cleavages are 3 bp in length. Because the species we have analyzed were generated in a crude system, however, it is difficult to exclude the possibility that a species with longer donor sequences is initially produced and then exonucleolytically reduced. Nonetheless, our studies have established that the transposon ends are released from the donor backbone via staggered breaks. The presence of donor sequences on the 5′ transposon strands is not essential for transposition.

The double-strand breaks at the termini of Tn7 need not occur simultaneously. Similar amounts of donor molecules broken at the junction of the donor backbone are observed with either Tn7L or Tn7R, which are both converted to insertion product. The ends of Tn7 are structurally and functionally distinct (Arciszewska et al., 1989). The orientation-specific insertion of Tn7 into *attTn7* also implies discrimination between the transposon ends. Our observations here suggest that discrimination between Tn7L and Tn7R occurs at a step other than formation and utilization of the double-strand break species.

The target site must also be broken by a staggered double-strand break; the joining of the ends of this break to the transposon ends and subsequent repair of the resulting gaps result in the hallmark target sequence duplication that accompanies transposon insertion (Grindley and Sherratt, 1979; Shapiro, 1979). Like the cleavages at the transposon ends, the staggered target cleavage generates 5′ overhangs; the target cleavage overhangs are 5 bp in length, as opposed to apparently 3 bp at the transposon ends. In contrast with the donor DNA, no cleavage of the target DNA is detected prior to its linkage to the transposon ends.

Upon Tn7 insertion, the 3′ transposon ends are ligated to the 5′ target ends (Figure 8). Formation of the new bonds between the target site and the transposon termini involves single-strand joins in which high energy phosphodiester bonds must be formed. The energy source for the covalent bonds between transposon ends and target strands is as yet undefined in Tn7 transposition. However, we have established that the energy of the transposon–donor bonds is not essential to the formation of the transposon–target bonds: a transposon lacking the transposon–donor bonds can join to the target site.

It will be important to identify the catalytic centers that execute strand cleavages at the transposon ends and at the target site. Because the two types of breaks have the same polarity, i.e., staggered with 5′ overlaps, both cleavages may be carried out by the same polypeptide with position specificity provided by other proteins that bind specifically to the transposon ends and to *attTn7*. Alternatively, the end and target cleavages may be carried out by different polypeptides with their own specificity determinants. Since conversion of the excised transposon to insertion product apparently requires the same *tns* fractions as a reaction using an intact donor molecule, no *tns* function appears to act exclusively in excision.

### Comparison of Tn7 with Other Transposable Elements

Tn7 transposition resembles the movement of retrotransposons, particularly in forming an elaborate nucleoprotein recombination complex. The origins of these complexes are quite different: the Tn7 DNA in its complex is formed by excision, whereas retrotransposon DNA in its complex is formed by reverse transcription of an RNA intermediate (Bowerman et al., 1989; Varmus and Brown, 1989). However, the initial joints between the element ends and insertion sites are very similar. For both elements, the 3′ transposon strands are linked to the 5′ target ends, and several nucleotides that will not appear in the final transposition product are attached to the ends of the 5′ transposon strands (Fujiwara and Mizuuchi, 1988; Brown et al., 1989; Eichinger and Boeke, 1990). With Tn7, these extra nucleotides derive from staggered breaks at the transposon termini; with retrotransposons, these nucleotides derive from single-strand cleavages of the linear retroviral DNA. These extra nucleotides are presumably removed during repair of the gaps that flank the initial insertion product.

Prior to linkage with the insertion site, the phosphodiester bonds that link Tn7 to its donor site are broken. Tn7 is excised from the donor backbone by double-strand breaks, and formation of a strand transfer intermediate is not observed. The strand transfer intermediate, i.e., a joint molecule in which the transposon, donor, and target are linked, is a key intermediate in both nonreplicative and replicative bacteriophage Mu transposition (Shapiro, 1979; Craigie and Mizuuchi, 1985; Pato, 1989). Tn7, however, appears to move exclusively by a nonreplicative reaction; excision of the transposon prior to linkage to the target effectively precludes a replicative reaction. Although distinct in temporal order, the strand cleavages at the ends of Mu and Tn7 are chemically similar, as are their target cleavages. In both systems, the 3′ transposon strands are joined to 5′ overhanging target strands exposed by a staggered double-strand break at the target (Mizuuchi, 1984). However, Tn7 is cut away from the donor site by staggered double-strand breaks, whereas cleavage of the 3′ and 5′ Mu strands occurs at very different stages, recombination being initiated by single-strand breaks at 3′ transposon ends. The exact positions of the 5′ Mu strand cleavages remain to be established.

The movement of Tn7 is, in several respects, highly reminiscent of the movement of the bacterial transposon Tn10 (Kleckner, 1989). Tn10 also translocates via a nonreplicative mechanism (Bender and Kleckner, 1986) that involves double-strand breaks at the transposon termini (Morisato and Kleckner, 1987), an excised transposon is a key recombination intermediate (Haniford et al., 1991), and chemically similar joints are formed between the transposon ends and target DNA (Benjamin and Kleckner, 1989). The ends of Tn10, however, appear to be exposed by flush cleavages (Benjamin and Kleckner, 1989), whereas those that expose the Tn7 ends are staggered.

A striking feature of Tn7 transposition revealed by this work is the highly concerted nature of this reaction. It will be interesting to determine how recognition of a single site in a bacterial chromosome is achieved and is communicated through multiple recombination proteins to trigger strand cleavages at the transposon ends and thereby provoke transposition.

## Experimental Procedures

### DNA Substrates

The donor plasmid pEM-1 contains a 1.6 kb mini-Tn7 element flanked by E. coli chromosomal sequences unrelated to *attTn7* (Arciszewska et al., 1989). The mini-Tn7 element contains a 166 bp Tn7L segment and a 199 bp Tn7R segment flanking a kanamycin resistance gene; these end segments contain the Tn7 cis-acting transposition sequences. The terminal nucleotides of Tn7 are designated L1 and R1, the numbers increasing toward the middle of the element. The *attTn7* target plasmid pKAO4–3 (McKown et al., 1988) contains a 150 bp segment that includes the sequences necessary for *attTn7* target activity. Target DNA lacking *attTn7* is Bluescript pKS+ (Stratagene, San Diego, CA). pMIM-1 contains a mini-Tn7 element whose termini directly adjoin MluI restriction sites (see Figure 4). The mini-Tn7 element with flanking MluI and BamHI sites was made by PCR amplification (below). Following digestion with BamHI, the amplified transposon was ligated into BamHI-digested Bluescript pKS+. Prior to use, the transposon was released by MluI digestion and gel purified using Gene Clean. pLA11 and pLA24 are *attTn7* plasmids that also contain Tn7R segments. To make pLA11, the EcoRI DNA fragment containing R199 from pLA1 (McKown et al., 1987) was inserted by blunt ligation into DraII-digested pRM2 (McKown et al., 1987); position R199 is adjacent to the polylinker HindIII site. To make pLA24, the EcoRI fragment containing Tn7R42 from pCW5 (Arciszewska et al., 1989) was inserted by blunt ligation into DraII-digested pRM2; R42 is adjacent to the polylinker HindIII site.

### *tns* Strains

Protein fractions were obtained from E. coli strains containing various *tns* plasmids. The bacterial strains used were: MC4100, F⁻ *araD139A* Δ(*argF-lac*)*U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR* (Casadaban, 1976); NLC51, MC4100 *val*$^R$ *recA56* (McKown et al., 1987); and SY 903.1, *recA1 srl*::Tn10, Δ(*lac-pro*) *argE*am *araD* Rif$^R$ Nal$^R$/ F′*lacI*$^{Q1}$ *lacZ*::Tn5 (Sauer et al., 1988). The *tnsA* strain was NLC51 carrying pKAO52 (Orle and Craig, 1990), the *tnsB* strain was SY 903.1 carrying ptac-*tnsB* (Arciszewska et al., 1989), the *tnsC* strain was NLC51 carrying pKAO53 (Orle and Craig, 1990), and the *tnsD* strain was MC4100 carrying pCW23 (Waddell and Craig, 1988).

### Cell Growth and Preparation of *tns* Fractions

Cells were grown at 37°C in LB broth supplemented with 100 μg/ml carbenicillin. Cells were grown to an $OD_{600}$ of approximately 0.6, harvested by centrifugation, washed in buffer A (25 mM HEPES [pH 7.5], 1 mM EDTA, 2 mM dithiothreitol [DTT], and 100 mM KCl) (Fuller et al., 1981) and frozen as cell paste. Crude cell lysates (stage I) were prepared by a freeze–thaw lysis method (Fuller et al., 1981) using buffer A. The protein concentration of all lysates was about 20 mg/ml except for that of NLC51 pKAO53, which was about 10 mg/ml. Stage II *tnsA* and *tnsC* fractions were prepared by treatment of the crude cell lysates with 1% polyethylenimine in buffer A at 500 mM KCl, collection of the resulting supernatant by centrifugation, protein precipitation with 60% saturation ammonium sulfate, and resuspension and dialysis of the resulting pellet with 25 mM HEPES (pH 7.5), 1 mM EDTA, 2 mM DTT, and 500 mM KCl. The *tnsB* fraction was fraction II, about 90% TnsB polypeptide (L. Arciszewska, R. McKown, and N. L. C., unpublished data). The *tnsD* fraction was stage III (about 1% TnsD, K. Kubo and N. L. C., unpublished data). The host extract fraction was prepared by boiling NLC51 crude cell lysate for 2 min and collecting the resulting supernatant after centrifugation in a microfuge for 2 min at room temperature.

### In Vitro Transposition Reactions

Unless otherwise indicated, all reaction mixtures (100 μl) contained 0.1 μg (0.25 nM) of pEM donor plasmid, 2.0 μg (10.0 nM) of pKAO4–3 *attTn7* target plasmid, 2 mM ATP, 26 mM HEPES (pH 8.0), 1.3 mM Tris–HCl (pH 8.0), 2.5 mM $KPO_4$, 0.1 mM EDTA, 2.2 mM DTT, 15 mM MgAc, 100 μg/ml tRNA, 50 μg/ml BSA, and 5% polyvinyl alcohol MW 8000. Reactions in Figures 2B, 6, and 7 also contained approximately 10 μg of stage II *tnsA* extract, 100 ng of TnsB (fraction II), 2.5 μg of stage II *tnsC* extract, 4 μg of stage III *tnsD* fraction, 1 μg of stage II host extract, 128 mM KCl, 15 mM NaCl, and 0.8% glycerol (v/v). Reactions in Figures 2A, 3, and 4B contained 5 μg of stage II *tnsA* fraction, 75 ng of TnsB, 2.5 μg of stage II *tnsC* fraction, 4 μg of stage III *tnsD* fraction, and 1 μg of host fraction; the reactions in Figures 2A and 4B also contained 50 mM KCl and 6 mM NaCl, and those in Figure 3 also contained 136 mM KCl and 2.8 mM NaCl. The reactions in Figure 4A contained 80 μg of stage I *tnsA* fraction, 75 ng of TnsB, 15 μg of stage I *tnsC* fraction, 4 μg of stage III *tnsD* fraction, 30 mM KCl, and 3 mM NaCl, in addition to the NaCl indicated in the figure. The reactions in Figure 9 contained a lower concentration (3 nM) of target plasmid, 40 μg of stage I *tnsA* fraction, 100 ng of TnsB, 10 μg of stage I *tnsC* fraction, 4 μg of stage III *tnsD* fraction, 128 mM KCl, 15 mM NaCl, and 0.8% (v/v) glycerol.

The order of addition of the components is unimportant, except for MgAc, which must be added after preincubation of the other components for 7 min at 30°C. After precincubation, MgAc is added, and incubation is continued for 30 min at 30°C. $MgCl_2$ can be substituted for MgAc, but neither $CaCl_2$ nor spermidine is effective. Reactions are stopped and the DNAs recovered by a modification of the Hoopes protocol (Hoopes and McClure, 1981; G. Gloor and G. Chaconas, personal communication). Four hundred microliters of 10 mM Tris–HCl (pH 7.5), 100 mM KCl, 5 mM EDTA, and 55% (w/v) urea are added to the reactions, followed by 56 μl of 100 mM spermine and then

incubation for 15 min on ice. The resulting pellet is collected by centrifugation, resuspended in 200 μl of 10 mM MgAc and 300 mM NaAc, tRNA added (150 μg/ml), and the mixtures incubated for 10 min on ice. DNA is recovered by ethanol precipitation and resuspended in 25 μl of 10 mM Tris–HCl (pH 8.0), 1 mM EDTA, and 10 μg/ml RNAase.

**Restriction and Gel Electrophoresis**

Reaction DNA (5 μl) is digested with 30–40 U of EcoRI in 10 μl reactions for 30 min at 37°C. (These rapid digestion conditions were used because of nucleases in the extracted DNAs.) After digestion, the DNAs are electrophoresed in 0.7% agarose gels in 1 x Tris–borate–EDTA buffer.

**Southern Hybridization**

DNA was electrotransferred to Nytran membranes (0.45 micron, Schleicher and Schuell), and the DNA is covalently bound by ultraviolet irradiation in a Stratalinker (Stratagene). The membrane is probed with a radioactive DNA fragment specific for the mini-Tn7 element—the 1200 bp PstI fragment— generated by random priming labeling (Feinberg and Vogelstein, 1984) using [α-$^{32}$P]dATP and the Klenow fragment of DNA polymerase I. Hybridizations are carried out at 70°C in a Hybrid-Eaze chamber (Hoefer Scientific Instruments). A 1–3 hr prehybridization (1.0 M NaCl, 1.0% SDS) was followed by overnight hybridization (10% dextran sulfate, 1.0 M NaCl, 1.0% SDS, 100 μg/ml denatured salmon sperm DNA) with probe and followed by four washes in high stringency buffer (0.1 x SSPE, 1% SDS) over a 2 hr period.

**Isolation of the Simple Insertion Transposition Product**

The in vitro reaction product species were recovered and pooled from 20 individual reactions. After EcoRI digestion, the DNA was electrophoresed on a 0.7% agarose gel in 1 x Tris–acetate-EDTA buffer. The gel was stained with 0.5 μg/ml ethidium bromide, DNA was visualized with long wave-length ultraviolet light, gel slices containing the desired species were excised, and the DNA was purified with Gene Clean.

**Oligonucleotide Primers**

The primers are complementary to sequences at the junctions of the ends of Tn7 and *attTn7* (see Figure 5; sequences described in Craig, 1989). Tn7-L is complementary to L31–60 of the top strand at the left end of Tn7; Tn7-R is complementary to R101–130 of the bottom strand at the right end of Tn7. attTn7-L is complementary to *attTn7* −25 to −5 on the bottom strand of *attTn7* and also has at its 5′ end the 5 nt complementary to the flanking polylinker region in pKAO4–3, which flanks *attTn7* −25; attTn7-R is complementary to *attTn7* +32 to +56 on the top strand of *attTn7* and also has at its 5′end TTAA. Oligonucleotides were labeled at their 5′ ends using T4 polynucleotide kinase and [γ-$^{32}$P]ATP (6000 Ci/mmol; ICN Radionucleotides).

**Primer Extensions**

Extension reactions (10 μl) contained approximately 0.2 pmol of end-labeled oligonucleotide, approximately 10 ng of isolated simple insertion transposition product, 1 U of Taq polymerase (Cetus), dNTPs as indicated below, and PCR buffer (10 mM Tris–HCl [pH 8.3], 50 mM KCl, 2.0 mM $MgCl_2$, 0.01% gelatin). In some reactions (Figure 6), dNTPs were added to 0.2 mM. Under these conditions, Taq polymerase is highly processive, but 1–2 nt are added to the 3′ end of the extension product owing to a high level of nontemplated nucleotide addition (Clark, 1988). In other reactions (Figure 7), dNTPs were added to 2 mM. Under this condition, no template-independent addition is detected on the 3′ end of the extension product (see control in Figure 7), but the extensions are not highly processive. The extension products and DNA sequencing reactions, using the same primers as markers, were run on gels containing 6% polyacrylamide (19:1 acrylamide: bisacrylamide), 42% (w/v) urea, and 1 x Tris–borate–EDTA.

**Production of the MluI Transposon by PCR Amplification**

We synthesized an R-terminus oligonucleotide primer complementary to R1–19 of the top strand at the right end of Tn7 and to 17 of 19 positions in L1–19 of the bottom strand at the left end of Tn7 (positions L14 and L16 differ from R14 and R16), plus an additional MluI site, a BamHI site, and two extra nucleotides. Using this oligonucleotide and a mini-Tn7 plasmid as template (Mullis and Faloona, 1987), PCR amplification was used to generate a linear transposon whose ends were flanked by MluI and BamHI, which was used to produce pMIM-1 (see above). Reactions contained 100 pmol of primer, 1 fmol of EcoRI-cut pEM, 2 mM dNTPs, and PCR buffer.

**Production of Control PCR Transposon**

PCR amplification and a mini-Tn7 plasmid template were used to produce a control transposon whose 5′ ends terminate at the last transposon nucleotide (Figure 7). The primer was complementary to R1–19 of the top strand at the right end of Tn7 and to 17 of 19 positions in L1–19 of the bottom strand at the left end of Tn7 (positions L14 and L16 differ from R14 and R16). The 5′ ends of this PCR transposon are provided by the primer, and the 3′ ends will be generated by polymerization; thus, the 5′ ends have a known structure. The resulting PCR transposon was purifed through two cycles of agarose gel electrophoresis with Gene Clean extraction.

**Acknowledgments**

We thank other members of the laboratory for materials and advice: Karina Orle for *tns* plasmids and anti-Tns antibodies, Lidia Arciszewska for TnsB, and Ken Kubo for TnsD fractions. Glenn Gloor, George Chaconas, and Pat Brown also provided useful methods. We are especially grateful to Rudi Grosschedl, Ann Hagemann, Donald Morisato, Anne Stellwagen, Evi Strauss, and Candy Waddell for their comments on the manuscript. R. B. was supported by funds from the Medical Scientist Training Program (NIH-MSTP-GMO 7618), Merck Sharp and Dohme Research Laboratories, the UCSF Department of Biochemistry Tompkins Memorial Fund, and the UCSF Program in Biological Sciences. P. G. was supported by funds from the Weingarten Foundation and the NIH-CNRS program, and by a NATO fellowship. The work was supported by a grant from the National Institute of Allergy and Infectious Disease to N. L. C.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with USC 18 Section 1734 solely to indicate this fact.

Received January 16, 1990; revised March 21, 1991.

**References**

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

Arciszewska, L. K., Drake, D., and Craig, N. L. (1989). Transposon Tn7 cis-acting sequences in transposition and transposition immunity. J. Mol. Biol. *207*, 35–52.

Barth, P. T., Datta, N., Hedges, R. W., and Grinter, N. J. (1976). Transposition of a deoxyribonucleic acid sequence encoding trimethoprim and streptomycin resistances from R483 to other replicons. J. Bacteriol. *125*, 800–810.

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

Benjamin, H. W., and Kleckner, N. (1989). Intramolecular transposition by Tn10. Cell *59*, 373–383.

Berg, D. E., and Howe, M. M., eds. (1989). Mobile DNA (Washington, DC: American Society for Microbiology).

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

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

Brown, P. O., Bowerman, B., Varmus, H. E., and Bishop, J. M. (1989). Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. Proc. Natl. Acad. Sci. USA *86*, 2525–2529.

Casadaban, M. J. (1976) Transposition and fusion of the *lac* genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu. J. Mol. Biol. *104*, 541–555.

Clark, J. M. (1988). Novel non-templated nucleotide addition reactions

catalyzed by procaryotic and eucaryotic DNA polymerases. Nucl. Acids Res. *16*, 9677–9686.

Craig, N. L. (1989). Transposon Tn7. In Mobile DNA, D. E. Berg and M. M. Howe, eds. (Washington, DC: American Society for Microbiology), pp. 211–225.

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

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

Eichinger, D. J., and Boeke, J. D. (1988). The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition. Cell *54*, 955–966.

Eichinger, D. J., and Boeke, J. D. (1990). A specific terminal structure is required for Ty transposition. Genes Dev. *4*, 324–330.

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

Feinberg, A. P., and Vogelstein, B. (1984). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. *137*, 266–267.

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

Fuller, R. S., Kaguni, J. M., and Kornberg, A. (1981). Enzymatic replication of the origin of the Escherichia coli chromosome. Proc. Natl. Acad. Sci. USA *78*, 7370–7374.

Gay, N. J., Tubulewicz, V. L., and Walker, J. E. (1986). Insertion of transposon Tn7 into the Escherichia coli *glmS* transcriptional terminator. Biochem. J. *234*, 111–117.

Gellert, M. K., Mizuuchi, K., O'Dea, M. H., and Nash, H. A. (1976). DNA gyrase: an enzyme that introduces superhelical turns into DNA. Proc. Natl. Acad. Sci. USA *73*, 3872–3876.

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

Gringauz, E., Orle, K. A., Waddell, C. S., and Craig, N. L. (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.

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

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

Hoopes, B. C., and McClure, W. R. (1981). Studies of the selectivity for DNA precipitation by spermine. Nucl. Acids Res. *9*, 5493–5504.

Kleckner, N. (1989). Transposon Tn10. In Mobile DNA, D. E. Berg and M. M. Howe, eds. (Washington, DC: American Society for Microbiology), pp. 227–268.

Kleckner, N. (1990). Regulation of transposition in bacteria. Annu. Rev. Cell Biol. *6*, 297–327.

Kubo, K., and Craig, N. L. (1990). Bacterial transposon Tn7 utilizes two different classes of target sites. J. Bacteriol. *182*, 2774–2778.

Lee, C. H., Bhagwat, A., and Heffron, F. (1983). Identification of a transposon Tn3 sequence required for transposition immunity. Proc. Natl. Acad. Sci. USA *80*, 6765–6759.

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

McKown, R. L., Waddell, C. S., Arciszewska, L. A., and Craig, N. L. (1987). Identification of a transposon Tn7-dependent DNA-binding activity that recognizes the ends of Tn7. Proc. Natl. Acad. Sci. USA *84*, 7807–7811.

McKown, R. L., Orle, K. A., Chen, T., and Craig, N. L. (1988). Sequence requirements of Escherichia coli *attTn7*, a specific site of transposon Tn7 insertion. J. Bacteriol. *170*, 352–358.

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*, 395–404.

Morisato, D., and Kleckner, N. (1987). Tn10 transposition and circle formation in vitro. Cell *51*, 101–111.

Mullis, K. B., and Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Meth. Enzymol. *155*, 335–350.

Orle, K., and Craig, N. L. (1990). Identification of transposition proteins encoded by the bacterial transposon Tn7. Gene *96*, 1–7.

Pato, M. L. (1989). Bacteriophage Mu. In Mobile DNA, D. E. Berg and M. M. Howe, eds. (Washington, DC: American Society for Microbiology), pp. 23–52.

Qadri, M. I., Flores. C. C., Davis, A. J., and Lichtenstein, C. P. (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.

Roberts, D., and Kleckner, N. (1988). Tn10 transposition promotes RecA-dependent induction of a λ prophage. Proc. Natl. Acad. Sci. USA *85*, 6037–6041.

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

Sauer, R. T., Smith, D., and Johnson, A. D. (1988). Flexibility of the yeast alpha-2 repressor enables it to occupy the ends of its operator, leaving the center free. Genes Dev. *2*, 807–816.

Shapiro, J. A. (1979). Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proc. Natl. Acad. Sci. USA *76*, 1933–1937.

Thompson, J. F., and Landy, A. (1989). Regulation of bacteriophage lambda site-specific recombination. In Mobile DNA, D. E. Berg and M. M. Howe, eds. (Washington, DC: American Society for Microbiology), pp. 1–22.

Varmus, H. E., and Brown, P. (1989). Retroviruses. In Mobile DNA, D. E. Berg and M. M. Howe, eds. (Washington, DC: American Society for Microbiology), pp. 53–109.

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

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

# Chapter 3

## Tn7 Transposition: Target DNA Recognition by Multiple Tn7 Encoded Proteins in a Purified *In Vitro* Transposition System

This work was done in collaboration with Ken Kubo. The reconstitution of the purified system, and the preincubation experiments are my work. The gel shift work and footprint analysis is the work of Ken Kubo. We collaborated on the experiment which uses the linear attTn7 site as a target. The TnsD-independent transposition and the immunity experiments are also my work.

**Summary**

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 transposition proteins, TnsA + TnsB + TnsC + TnsD, and ATP. Insertion without target site-selectivity is observed with TnsABC 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 selectively insert 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 reaction 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 choice of attTn7 as an insertion site (Lichtenstein and Brenner 1982; 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 (Rodgers 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 the 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 (Robinson et al. 1977; Lee 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. 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 (Rogers et al. 1986; Waddell and Craig 1988; Kubo and Craig 1990). In Tn7's "random target site" pathway, tnsABC + tnsE promote low frequency insertion into many different sites that are distinct from attTn7 and are not apparently related to each other. Thus, tnsABC provide functions common to both 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, proteins. What roles do these proteins play in recombination? Analysis of recombination reactions using the alternative cofactor AMP-PNP instead of the usual cofactor ATP reveal 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 key role in the identification of attTn7 as a specific and efficient site for Tn7 insertion (Kubo and Craig, in preparation). We also show here that TnsC, an ATP-dependent non-sequence-specific DNA binding protein (Gamas and Craig in press), interacts with attTn7 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 indicate 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 to *attTn7* *in vivo*. We established that a mini-Tn7 element can transpose from a donor plasmid into *attTn7* in a target plasmid upon incubation of these substrates in crude extracts derived form 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 form the donor and specifically inserted into *attTn7* (See Figure 1). We have reconstituted Tn7 insertion into *attTn7* using four highly purified protein fractions containing TnsA, TnsB, TnsC or TnsD (Figure 2A, lane 3 to 6).

We report here the purification of TnsA (as described in experimental protocols). We constructed a fusion protein in which TnsA is fused to the carboxyterminus of glutathione-S-transferase (GST) and purified the GST-TnsA fusion (54kd) by binding it to glutathione agarose beads (Figure 2A, lane 1) and releasing TnsA (29kd) 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 extracts lacking TnsA is comparable to authentic crude TnsA protein (see Appendix). We have previously described the purifications of TnsB, TnsC and TnsD (Figure 2A, lanes 4, 5, and 6) (Arciszewska et al. 1991; Gamas and Craig, in press; Kubo and Craig, in preparation).

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 mini-Tn7 donor and attTn7 target plasmids are incubated at 30°C with highly purified Tns protein fractions, ATP, magnesium acetate and small molecules. Using a transposon-specific probe, we observe that more than 80% of the donor substrate can be converted to simple insertion product; the other product of transposition, the donor backbone from which the element is excised, is also present in these reactions, but not detectable by this hybridization probe.

Also detectable with the reconstituted system are several DNA species shown in the crude transposition system (Bainton et al) to be intermediates in Tn7 transposition: donor molecules broken by a single double strand break between the backbone and the transposon (DSB.L = double-strand break.left end or DSB.R = double-strand break.right end) and excised linear transposons (ELTs) cut away from the donor backbone by two DSBs (see Figure 1). Time courses demonstrate the the DSBs in the purified 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 the attTn7 target; these species are called DSB-SEJ.L (= single end join of the left end of a DSB.L species) (Figure 2B, lane 1, 2.2kb; see Figure 1 for schematic representation). 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 a DSB-SEJ.R species (= single end join of the right end of a DSB.R species) is detectable in amounts equivalent to that of the SEJ.Ls. Also, ELTs which have one transposon end joined to one strand of attTn7 in the target DNA can also be observed (ELT-SEJ.L and ELT-SEJ.R) at very low levels (shown schematically in Figure 1, actual data not shown). All of these SEJ species appear late and accumulate in reaction time courses (data not shown), suggesting that they are not transposition intermediates. We emphasize that, although

these other species are present, 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 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 of recombination, 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 Tn7 transposition, no donor cleavage occurs in the absence of attTn7 (Figure 2B, lane 7).

As transposition does occur with highly purified Tns fractions, 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 increase both the rate and extent of transposition, particularly under sub-optimal conditions (data not shown, see Thesis Appendix). 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 press).

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 magnesium acetate: 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 incubation at 30°C continued after the addition of MgAc (Figure 3A, lane 3; shown schematically in panel B). 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 TnsC, TnsD, attTn7 and ATP is required in the preincubation for high level recombination; if one of them is omitted, then little recombination is observed (Figure 3A, lanes 4-7 respectively). Efficient recombination is still observed if TnsA, TnsB and the donor DNA are added after the preincubation step, i.e. when MgAc is added (Figure 3A, lane 8; shown schematically in Figure 3B). The requirement for the simultaneous presence of TnsC, TnsD, attTn7 and ATP in the preincubation step suggests that the interaction of these four components is stimulated by the low magnesium acetate 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 (Kubo and Craig, in prep.). Other studies have also revealed that TnsC is an ATP-dependent non-sequence-specific DNA binding protein (Gamas and Craig, in press). We have used gel mobility shift assays to establish that in the presence of ATP, TnsC and TnsD can form a distinct complex on attTn7.

Specific binding of TnsD to a labeled attTn7 fragment is observed by the formation of a new complex of retarded electrophoretic mobility (Figure 4A, lane 3). No binding of TnsC 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 2). Formation of this novel complex requires ATP; no "supershifted" complex is observed in the presence of ADP (lane 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 dependent on TnsD and attTn7 as no protein-DNA complexes are observed with a DNA fragment lacking attTn7 sequences (lane 8).

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 (Figure 3). This view is supported by the finding the formation of the TnsC-TnsD-attTn7 complex is blocked by the presence of MgAc in the reaction mixture (Figure 4A, lane 5), a condition the 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 reaction attTn7 is part of a supercoiled plasmid whereas the attTn7 used in analyzing complex formation is a short linear DNA fragment and also the ionic conditions of the incubations are different (see Exp. Prot.). We have therefore repeated the transposition reaction under identical conditions to those used for the gel mobility shift assay. 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). Transposition into these small linear targets also requires a preincubation step that includes TnsC (Figure 4B, lane 2) and a preincubation without MgAc (Figure 4B, 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 of attTn7.

We have explored the stability of the TnsC-TnsD complexes formed on attTn7 during the preincubation step by challenging the reaction mixtures with supercoiled attTn7 target DNA as competitor prior to or after the preincubation 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 DNA after the preincubation period (Figure 4B, lane 5), whereas competitor attTn7 DNA added from the beginning of the reaction successfully competes with linear attTn7 target DNA (Figure 4B, lane 4). This finding is consistent with the view that the complex formed during the preincubation step and detectable as TnsC-TnsD complex by mobility shift assays are the targets for Tn7 insertion.

**Footprinting analysis of the 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 by 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 nearly 55 bp in length, extending rightwards from about 7 nt. to the right of the specific point of Tn7 insertion (designated +7) to about 60 nt to the right of the Tn7 insertion point (+60), is protected from DNase I attack (Figure 5). This pattern is notably different from that observed with TnsD alone; the TnsC+TnsD pattern being more extensive and the degree of protection being much greater (Figure 5). On the bottom strand of attTn7, protection by TnsD alone extends from about +25 to +50, whereas the TnsC+TnsD protection extends from about +7 to +60 (Figure 5B). On the top strand, only a modest and limited protection is seen with TnsD alone, whereas, a striking protection is observed with TnsC+TnsD (Figure 5A). It remains to be established which positions in attTn7 may be directly contacted by TnsC and/or 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 +25) to the border of the TnsC+TnsD protected region that is closest to the point of insertion (about +7). An important feature of the TnsC+TnsD protection is that the specific point of Tn7 insertion is actually exposed to nuclease attack and thus does not appear to be directly contacted by either TnsC or TnsD (Figure 6A and B).

**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 any transposition product requires the presence of all the Tns proteins, TnsA, TnsB, TnsC and TnsD, and ATP (Figure 7, lane 2). Likely because of the high efficiency of the

reconstituted system, we do observe in these reactions low levels of the transposition intermediates DSBs and ELTs when TnsD or attTn7 are omitted form the reaction (Figure 7, lanes 1 and 3). This observation reveals that the active site for strand transfer 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 into the target backbone and high levels of recombination intermediates are generated when AMP-PNP is substituted for ATP (Figure 7, lane 4). Determination of the DNA sequence of several of these AMP-PNP dependent insertions after transformation reveals that they, like products obtained with ATP, generate 5 base pair target site duplications (data not shown), as observed for Tn7 insertions in vivo (Lichtenstein and Brenner 1982; Kubo and Craig, 1990). We suspect that TnsC, an ATP dependent binding protein is the site of action of 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 the target molecule (Figure 7A, lane 4). Another indication of the apparently random distribution of insertions is that the quantity of insertions into any particular target is proportional to the fractional size of that target (data not shown). We suggest that these insertions are for the most part random. Figure 7B shows two sets of transposition reactions done with TnsA + TnsB + TnsC. A single digestion (BglII, see Figure 1) of reactions done using ATP in the presence of TnsD (Figure 7B, lane 1) or AMP-PNP in the absence of TnsD (Figure 7B, lane 2) produce simple insertion products of 4.5kb. However, when the products of ATP + TnsD are digested with BglII and EcoRI, all the insertions are in the attTn7 site as expected with site-specific insertion (Figure 7B, lane 3). When the products of the AMP-PNP are double digested, the insertions

appear as a smear (Figure 7B, lane 4). This is the expected result of random insertion into the target.

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 data 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 and attTn7 and ATP preincubated in the absence of MgAc and the other reaction components to facilitate the formation of TnsC-TnsD complex on the target DNA. In TnsA + TnsB + TnsC reactions with AMP-PNP, recombination proceeds more efficiently when TnsC, the target DNA and nucleotide are preincubated 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 as 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 in the presence of either cofactor.

**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 includes both a standard attTn7 target plasmid and an immune attTn7 plasmid, i.e. one containing both attTn7 and a distal Tn7R segment, no insertions into the immune target (5.3kb) are observed although the standard attTn7 plasmid remains an effective target (Figure 8, lane 2, 1.75 kb). We find, however, that transposition immunity is virtually abolished when AMP-PNP is used as a cofactor instead of ATP (Figure 8, lane 4). In the presence of AMP-PNP novel insertions appear into the immune target (5.3kb) and as expected into non-immune targets, attTn7 (1.75 kb) and the target backbone (4.2 kb). A similar result is also seen in the absence of TnsD (Figure 8, lane 8).

Thus substitution of AMP-PNP for ATP dramatically changes the Tn7 transposition machinery's sensing of the target in several ways: 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, 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 a 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 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 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 insertion into attTn7 in vitro occurs efficiently in the presence of and requires four purified Tn7-encoded 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 which encode these proteins are essential to Tn7 transposition in vivo (Rodgers et al. 1986; Craig and Waddell 1988). The observation that ATP is an essential cofactor in the purified system reveals that this molecule 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 end 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 specifically interacts with the ends of Tn7 (Arciszewska et al. 1991; Tang et al. 1991), and TnsD is a sequence-specific DNA binding protein that interacts specifically with attTn7 (Kubo and Craig, in press). Because TnsC is an ATP-dependent non-sequence-specific DNA binding protein (Gamas and Craig, in press), a reasonable hypothesis is that TnsC is the site of action of the essential ATP cofactor in Tn7 transposition. As described below, we provide 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 in 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 under some conditions transposition can be stimulated by addition of host crude extract to the purified Tns proteins (data not shown, see Thesis Appendix); also, our optimization of recombination in the reconstituted system through the use of a low magnesium preincubation 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; Glascow et al. 1989; Landy 1989).

**TnsC Can Interact with the Target DNA**

Characterization of the reconstituted Tn7 transposition system has revealed that efficient transposition, albeit non-target site specific, is observed 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-PNP 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. We find 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) (data not shown). This 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 in vivo 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 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 ends, 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 transposition because TnsC cannot effectively interact with these DNAs.

The picture that has emerged for these experiments is that TnsC can interact with the target DNA during transposition, 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 TnsB and/or TnsA, perhaps thereby mediating interactions with the transposon ends.

## Comparison of Tn7 TnsC Protein and Bacteriophage Mu Transposition Protein MuB

We suggest that an ATP-dependent interaction of TnsC with the target DNA plays a central role in choosing target sites in Tn7 transposition. This model is reminiscent of the proposed role of MuB protein in bacteriophage Mu transposition. MuA protein, a Mu-end binding protein (Craigie et al. 1984), 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; 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 target. Transposition immunity is abolished in the presence of the non-hydrolyzable ATP analogue ATP-$\gamma$-S, 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 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 the interaction of TnsC with the attTn7 target DNA is promoted by TnsD during TnsD-dependent transposition. A

reasonable hypothesis is that TnsC, an ATP-dependent non-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 magnesium acetate (Mg Ac) and the initial presence of MgAc blocks the formation of the TnsC-TnsD attTn7 complex (data not shown). We hypothesize that the preincubation step without $Mg^{++}$ promotes the formation of TnsC-TnsD attTn7 complex which is relatively resistant to MgAc. The subsequent requirement for 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 (Kubo and Craig in prep.). Of course, we cannot exclude the possibility that TnsD includes a particular conformational change in attTn7 DNA which is specially 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 may be 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 the actual point of Tn7 insertion is not contacted by either TnsC or TnsD; this region is accessible to attack by DNaseI. The region of TnsC-TnsD protection is about 50 nt in length and extends rightwards from a position about 10 nt form the 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, 1989; Kubo and Craig, in prep.), lies with 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 the region surrounding the point of insertion is actually occupied by the transposon ends, TnsA and TnsB. Given the apparent spatial separation between the point of Tn7 insertion and 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 ends to the target DNA.

## Tn7 Transposition Occurs in a 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 step in transposition, i.e. cleavage of the donor DNA. Thus we argue that Tn7 transposition actually occurs with an elaborate nucleoprotein complex that contains all the Tns proteins and the three DNA

recombination substrates, i.e. the Tn7 end 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, formation of a distinct complex containing the ends of Tn7 and TnsA + TnsB + TnsC may precede the interaction of TnsC with DNA and/or TnsD.

We propose here 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 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 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 actual 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 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, thereby provoking transposition. What relationship might the mechanism of random target site TnsE-mediated transposition have to the mechanism proposed for site-specific TnsD-dependent transposition? A attractive hypothesis is that TnsE directs the basic TnsA + TnsB + TnsC machinery to

promote Tn7 insertion in 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 target DNA at many different positions and provoking transposition. Another alterative in 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 two alternative types of nucleoprotein complexes, one with TnsD and one with TnsE, that may mediate Tn7 transposition. Both complexes share the same basic machinery, i.e. TnsA + TnsB + TnsC, but a nucleoprotein complex with TnsD directs insertions to attTn7, and a TnsE-dependent complex promotes insertions to random sites. This combinatorial assembly of nucleoprotein complexes provide regulated and differential activities analogous to other processes involved in recombination, replication, transcription, and translation (Echols 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 Protocols

### In Vitro Transposition Substrates

The donor plasmid, pEM-1, contains an E. coli chromosomal insertion of the mini-Tn7 element flanked by sequences unrelated to attTn7. (Arciszewska et al., 1989). The donor plasmid pEMΔ contains the mini-Tn7 element and nearby flanking sequences (51 nt and 83 nt adjacent to the Tn7L and Tn7R end, respectively) of pEM-1. The mini-Tn7 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 form 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 ($\alpha$-$^{32}$P) dATP and Klenow enzyme (Kubo and Craig, in prep). For mobility shift assays, the 115bp 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 pKOA-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 press) and TnsD was fraction V (Kubo and Craig, in preparation).

TnsA was expressed as a glutathione-S-transferase (GST) fusion protein in Escherichia coli strain DH5-α (BRL). The gene construct encoding the GST-TnsA fusion was generated by inserting a NcoI-HindIII tnsA fragment from pKAO46 (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 1L LB broth with 100 µg/ml carbenicillin. At OD600 = 0.7, fusion protein expression was induced by the addition of 40 uM IPTG (isopropyl-B-D-thiogalactopyranoside) and then allowed to grow until OD600 = 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 1mM 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 supernatent was collected. To the crude lysate, 5M NaCl and 250 mM CHAPS (a nonionic detergent) were added to a final concentration of 500 mM and 10mM, 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 was released from the GST domain 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 fro 6 hr with glutathione agarose beads, and the supernatant, selectively enriched in TnsA polypeptide, was collected. These fractions, containing approximately 100µg/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 of SDS-PAGE with staining by Coomassie Blue R-250, using BSA as a standard.

## Tn7 Transposition In Vitro

Reaction 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), 0.6 µg (2.5 nM) of pKAO4-3 attTn7 target plasmid, 2mM ATP, 26 mM HEPES pH 8.0, 0.02 mM EDTA, 2.1 mM DTT, 15 mM MgAc, 0.1 mM $MgCl_2$, 0.01 mM $CaCl_2$, 100 µg/ml tRNA, 50 mg/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 where the reaction components were mixed and incubated for 30 min at 30°C prior to the addition of MgAc. 4µl of 375 mM MgAc was then added, and incubation was continued for an additional 30 minutes at 30°C. In Figure 3, the preincubation step was performed in reaction mixtures (96 µl) that also contained1.0 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 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 NaCl, 30 mM KCl, and 0.5% glycerol.

As described in Bainton et al. (1991), after incubation is completed, reaction DNAs were isolated by urea-spermine precipitation and EtOH precipitation, digested

with restriction enzyme, and electrophoresed through 0.6% agarose. The resolved DNAs were transferred to Nytran (Schleicher and 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 $MgCl_2$, 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, and incubation was continued for 30 min at 30°C. The attTn7 competitor, pKAO4-3 (2.0 µg) DNA was added at either at the beginning of the preincubation or just before the addition of MgAc. Reaction mixtures were 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 DNaseI Footprinting**

Mobility shift assays were performed as described in Kubo and Craig (in preparation) 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, 0.1 mM EDTA, 1.8 mM DTT, 0.1 mM MgCl2, 0.1 mM CHAPS, 0.8 mM ATP, 9% glycerol (v/v), 300 µg/ml BSA, 15 µg/ml sheared 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 Kubo and Craig (in preparation).

The DNaseI footprint analysis was performed as described in Kubo and Craig (in preparation), except that the reactions (100µl) contained 40 mM Tris-HCl (pH8.0), 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 Kubo and Craig (in preparation).

p

**Figure 1. The Transposition Substrates, Intermediates, and Products**

The 6.1 kb donor molecule contains a 1.6 kb mini-Tn7 element flanked by sequences unrelated to attTn7; the mini-Tn7 element contains segments from the left (Tn7L) and right (Tn7R) ends of Tn7 that provide the cis-acting transposition sequences. The 2.9kb 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 element inserts site- and orientation specifically into attTn7, producing a 4.5 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 single Tn7 end joins are single Tn7 end joins made by a single Tn7 end joining to a target molecule. Restriction sites: E = EcoRI, B = BglII, and S = ScaI

FIGURE 1

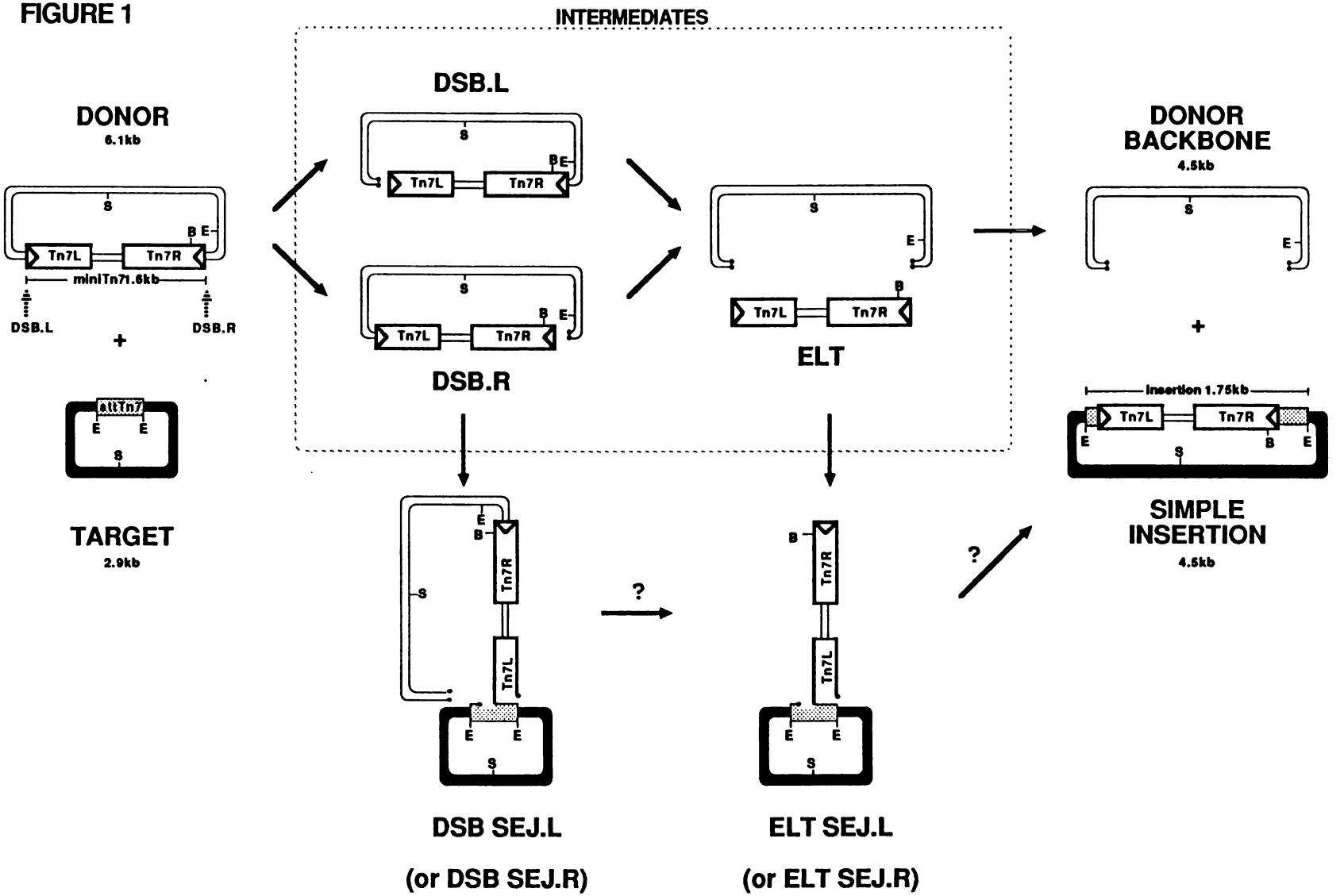

INTERMEDIATES
DONOR
6.1kb
S
B E
Tn7L
Tn7R
miniTn7l.6kb
DSB.L
DSB.R
+
altTn7
E E
S
TARGET
2.9kb
DSB.L
S
B E
Tn7L
Tn7R
DSB.R
S
E
B
Tn7L
Tn7R
ELT
DONOR
BACKBONE
4.5kb
S
E
+
Insertion 1.75kb
Tn7L
Tn7R
E
B E
S
SIMPLE
INSERTION
4.5kb
E
B
S
Tn7R
Tn7L
E E
S
?
B
Tn7R
Tn7L
E E
S
?
DSB SEJ.L
(or DSB SEJ.R)
ELT SEJ.L
(or ELT SEJ.R)

**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; 2.0 μg TnsC; lane 6, 2.0 μg TnsD.

B) Requirements for Tn7 transposition in vitro. An autoradiogram of a Southern blot 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 mini-Tn7 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, 2mM 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

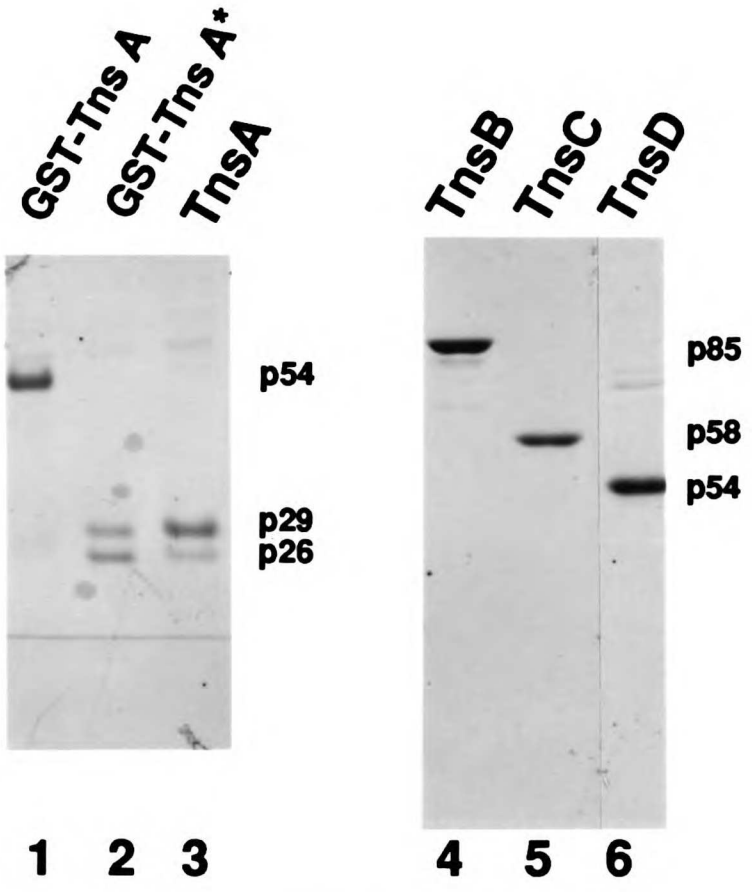

GST-Tns A
GST-Tns A*
TnsA
TnsB
TnsC
TnsD
p54
p29
p26
p85
p58
p54
1 2 3
4 5 6

Figure 2B

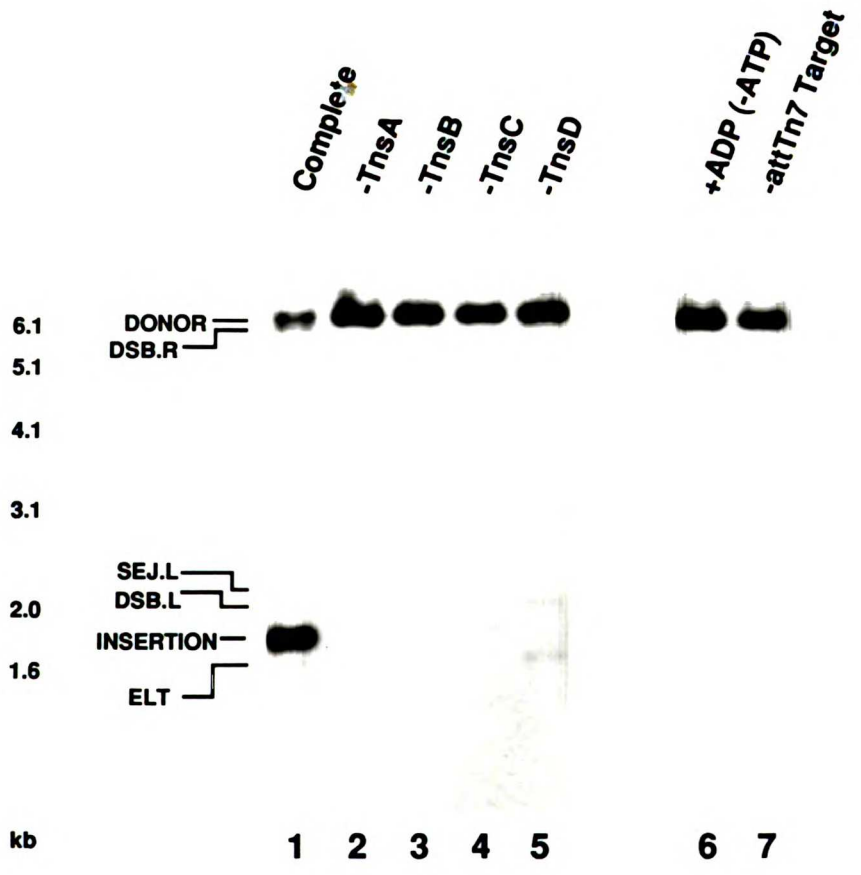
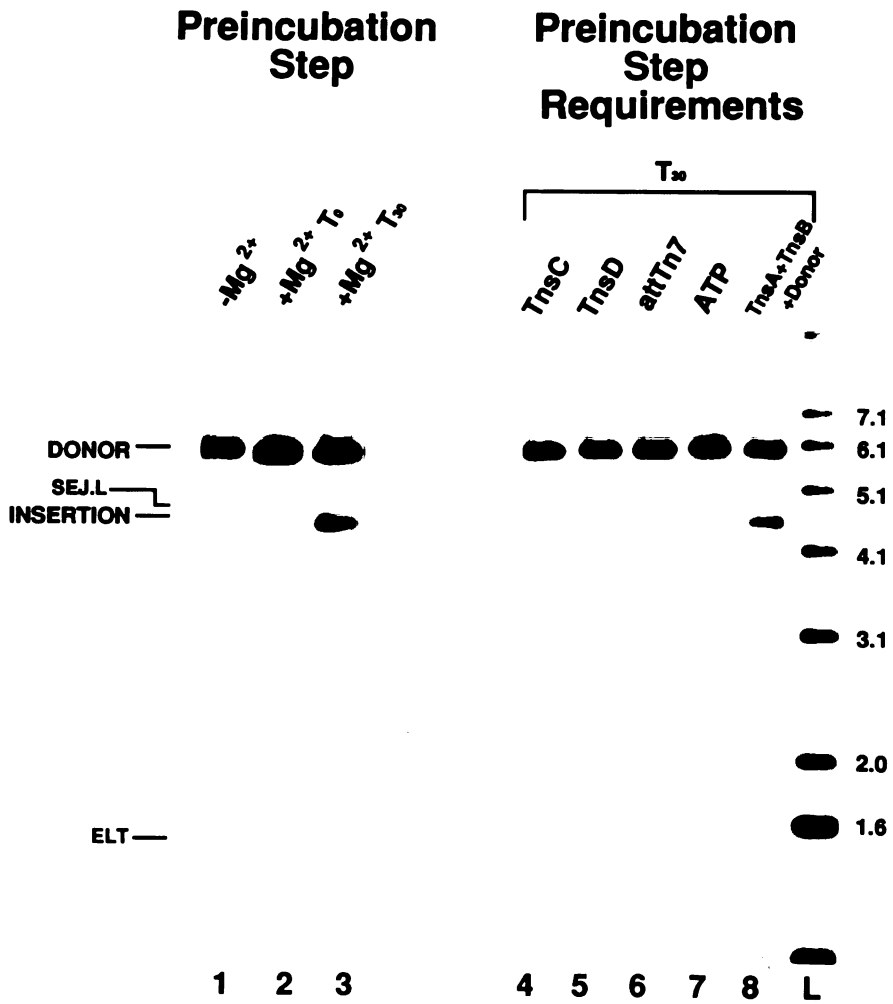

Complete
-TnsA
-TnsB
-TnsC
-TnsD
+ADP (-ATP)
-attTn7 Target
6.1
5.1
4.1
3.1
2.0
1.6
kb
DONOR
DSB.R
SEJ.L
DSB.L
INSERTION
ELT
1
2
3
4
5
6
7

**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 BglII-digested species detected with a mini-Tn7-specific probe. In lanes 1-3, MgAc is omitted entirely from the reaction (lane 1) added at the beginning of 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 (see Experimental Protocols for reaction conditions).

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

Figure 3A


Preincubation Step
Preincubation Step Requirements
$T_{30}$
-Mg 2+
+Mg 2+ $T_0$
+Mg 2+ $T_{30}$
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

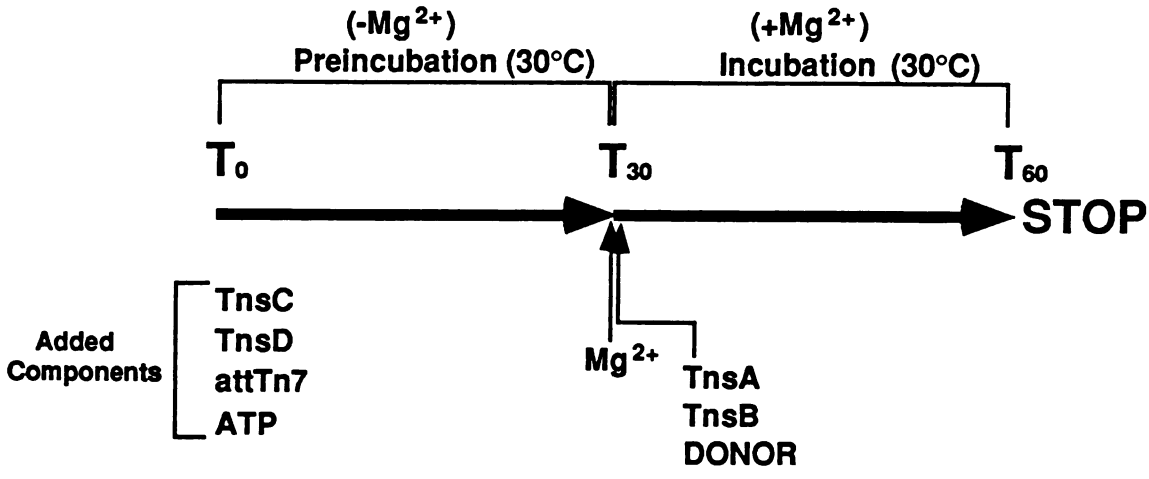

(-Mg2+)
Preincubation (30°C)
(+Mg2+)
Incubation (30°C)
T0
T30
T60
STOP
Added
Components
TnsC
TnsD
attTn7
ATP
Mg2+
TnsA
TnsB
DONOR

**Figure 4. TnsC+TnsD-Dependent Complex Formation with attTn7.**

A) Analysis of protein-DNA complexes by mobility shift assay. Lanes 1-7 contain as attTn7 segment, whereas lanes 8 and 9 contain a DNA segment lacking attTn7 sequences. Where indicated by "+", 20 ng TnsC and 20 ng TnsD are added to binding reactions. Lane 5, MgAc is added to 15 mM at the beginning of the binding reaction. Lane 7, ADP (2.0 mM) is substituted for ATP.

B) In vitro transposition into a linear attTn7 target. Particular reactions components (TnsC of MgAc) are 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, 2.0 μg of attTn7 plasmid (pKAO4-3) is provided as a competing target, either at the start or at the the end of the (-)MgAc preincubation step.

# Figure 4A

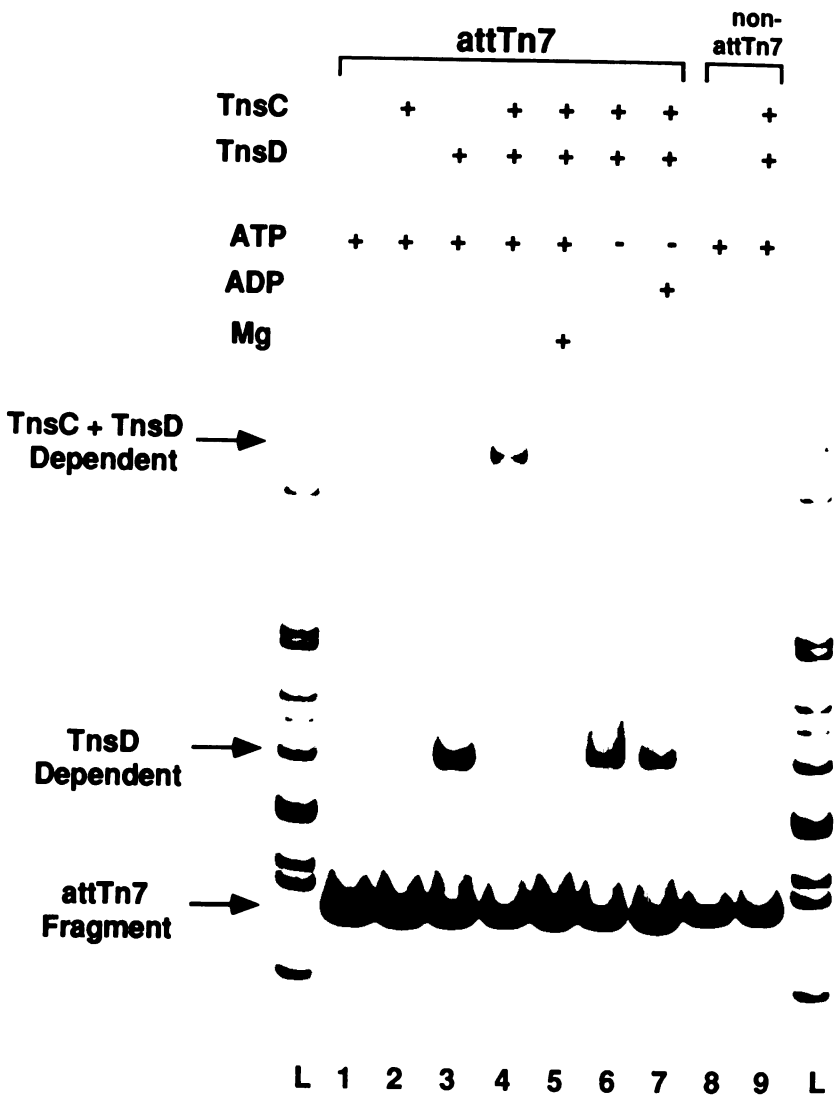

attTn7
non-attTn7
TnsC
TnsD
ATP
ADP
Mg
TnsC + TnsD Dependent
TnsD Dependent
attTn7 Fragment
L 1 2 3 4 5 6 7 8 9 L

Figure 4B

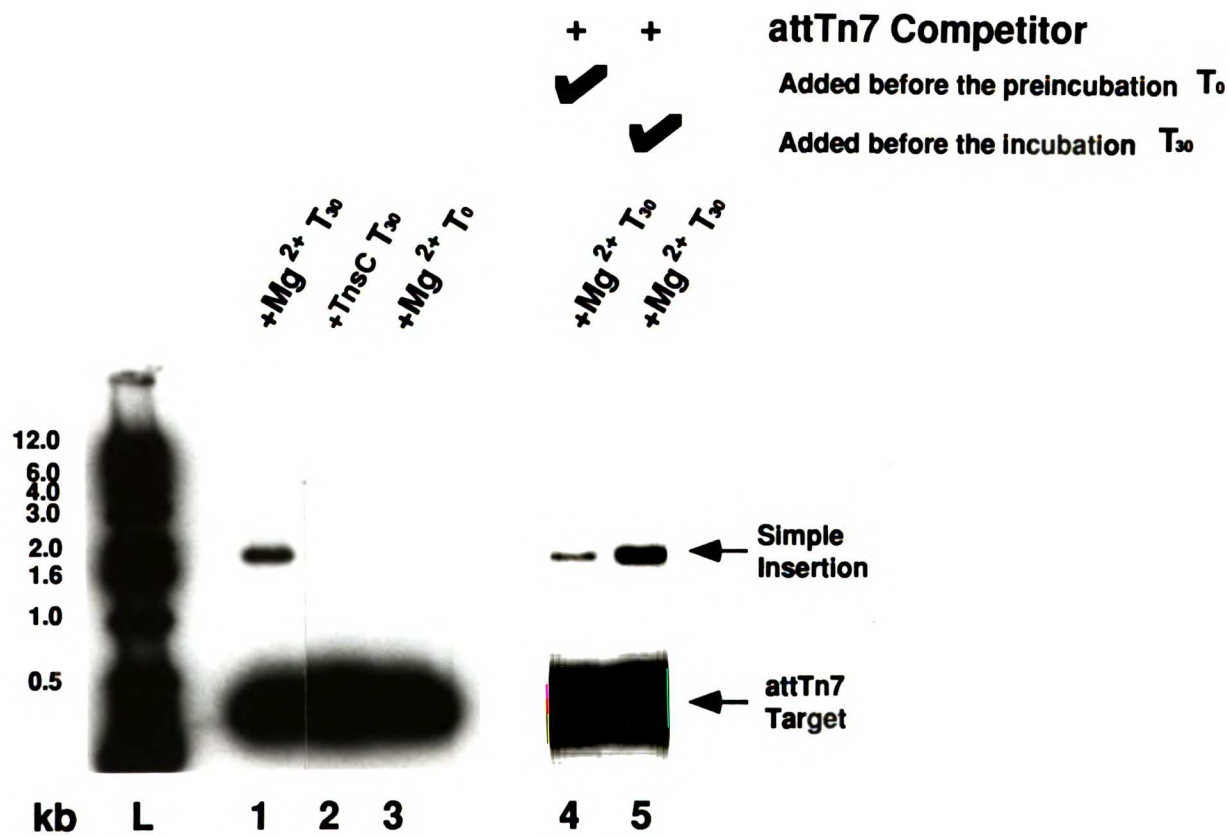

+ +
attTn7 Competitor
Added before the preincubation T0
Added before the incubation T30
+Mg 2+ T30
+TnsC T30
+Mg 2+ T0
+Mg 2+ T30
+Mg 2+ T30
12.0
6.0
4.0
3.0
2.0
1.6
1.0
0.5
Simple Insertion
attTn7 Target
kb L 1 2 3 4 5

**Figure 5. DNase I Footprint Analysis TnsC + TnsD Complexes with attTn7.**

TnsD and TnsC + TnsD complexes with attTn7 were formed, treated with DNaseI, and the DNA recovered for 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, the dotted line indicates a region of partial protection, and asterisks designate individual protected positions. Lanes marked "AG" contain Maxam-Gilbert sequence reactions of the same DNA fragment.

A) Top stand 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.

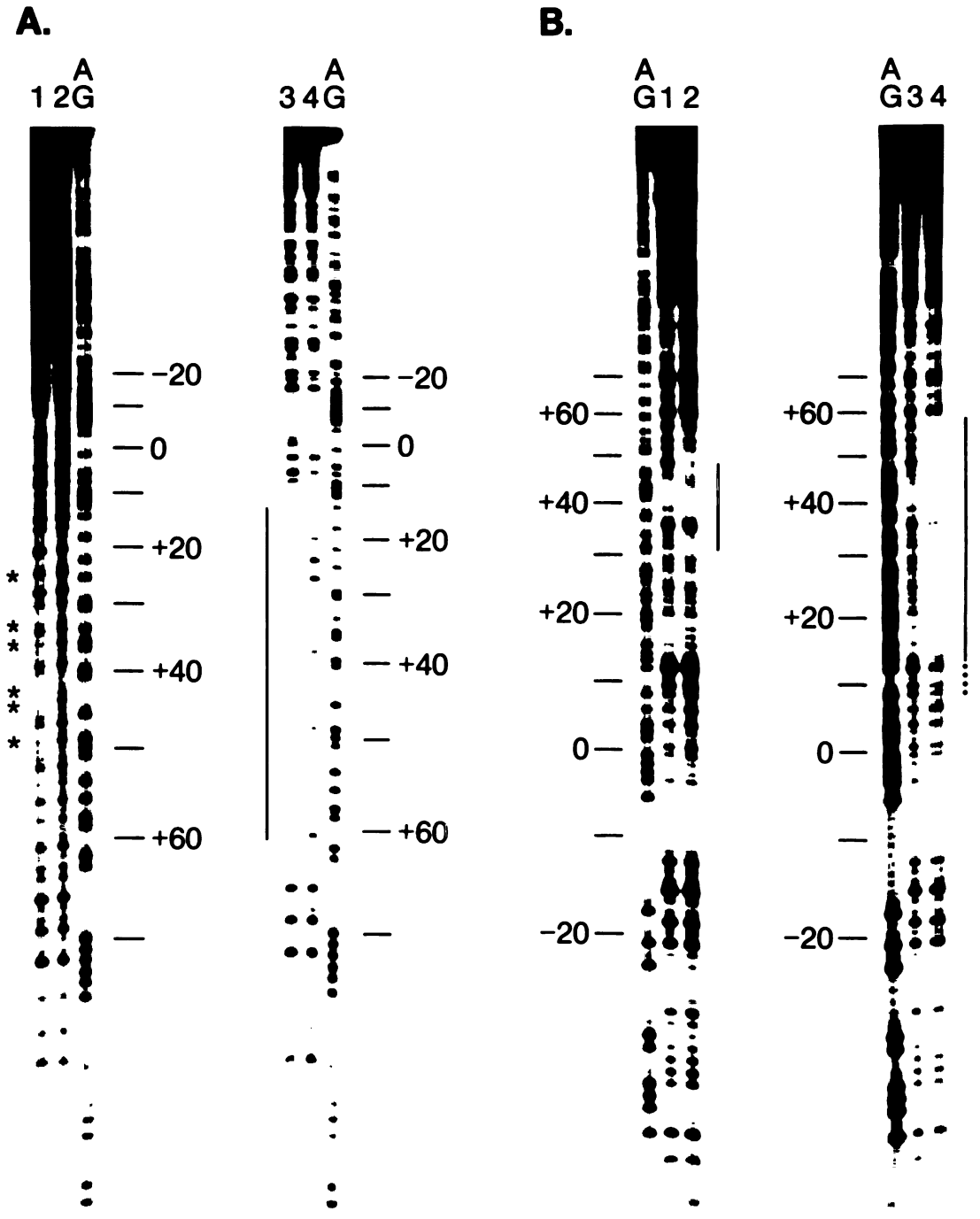
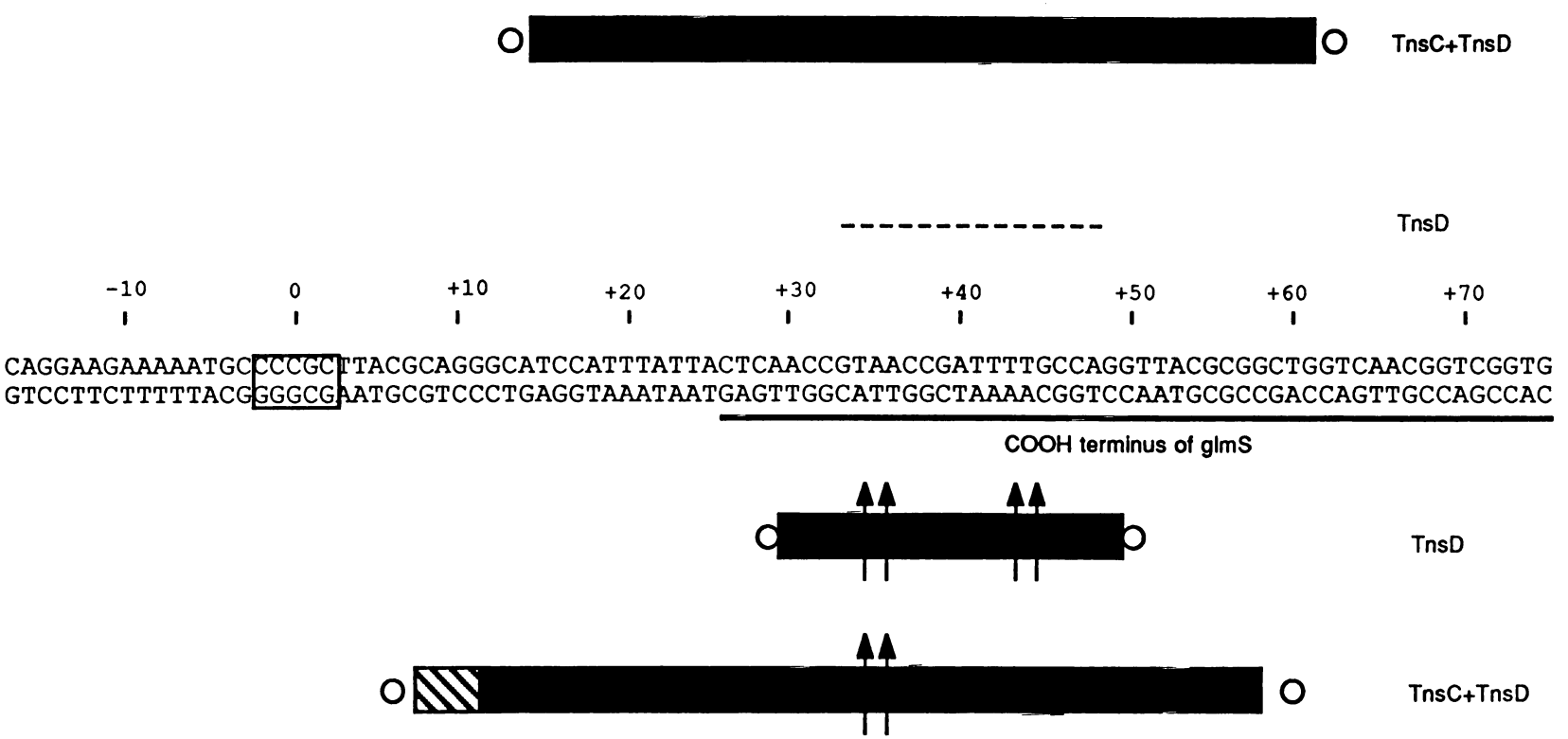

A.
A
1 2G
A
3 4G
−20
0
+20
+40
+60
B.
A
G1 2
A
G3 4
+60
+40
+20
0
−20

**Figure 6. Summary of DNAse I Protection 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 proteins. The stippled bar designates a region of partial protection. 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 remain to be determined.

# Figure 6A


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

# Figure 6B

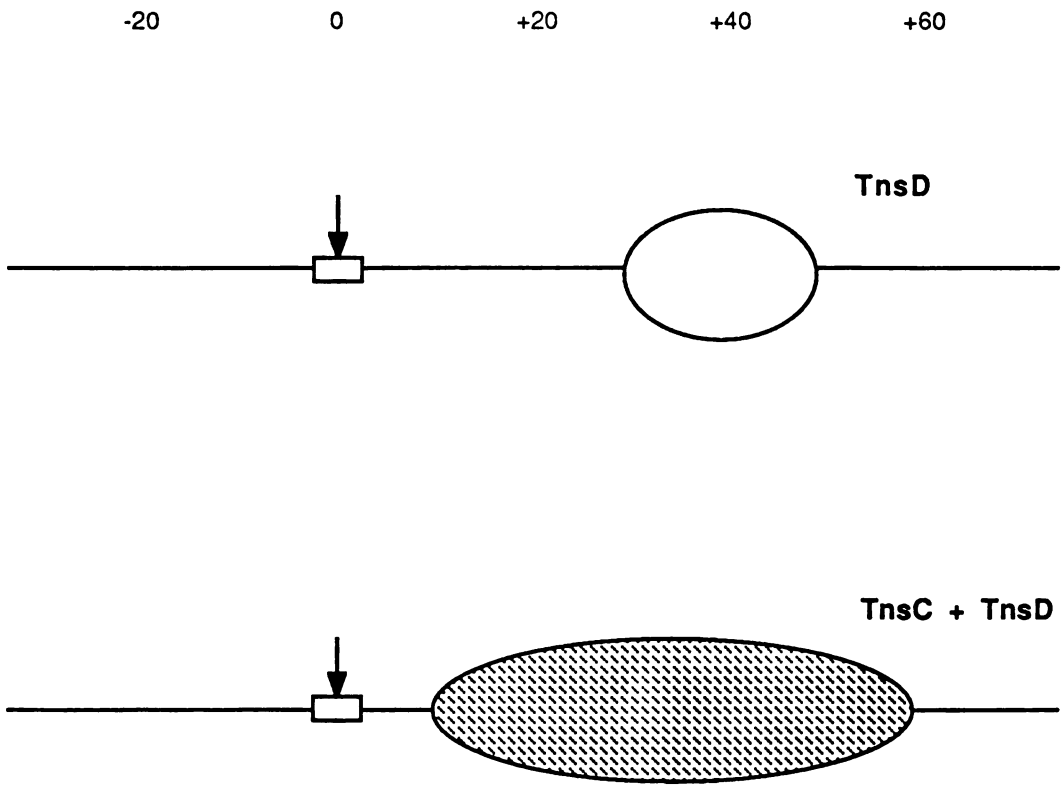


**Figure 7. Tn7 Transposition in the Presence of AMP-PNP**

A) EcoRI-digested species detected with the mini-Tn7-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 (pKAO4-3).

B) DNA species detected with a mini-Tn7-specific probe form transposition reactions were digested with BglII (lanes 1 and 2) or BglII + EcoRI (lanes 3 and 4) to evaluate attTn7 insertion specificity.

Figure 7A

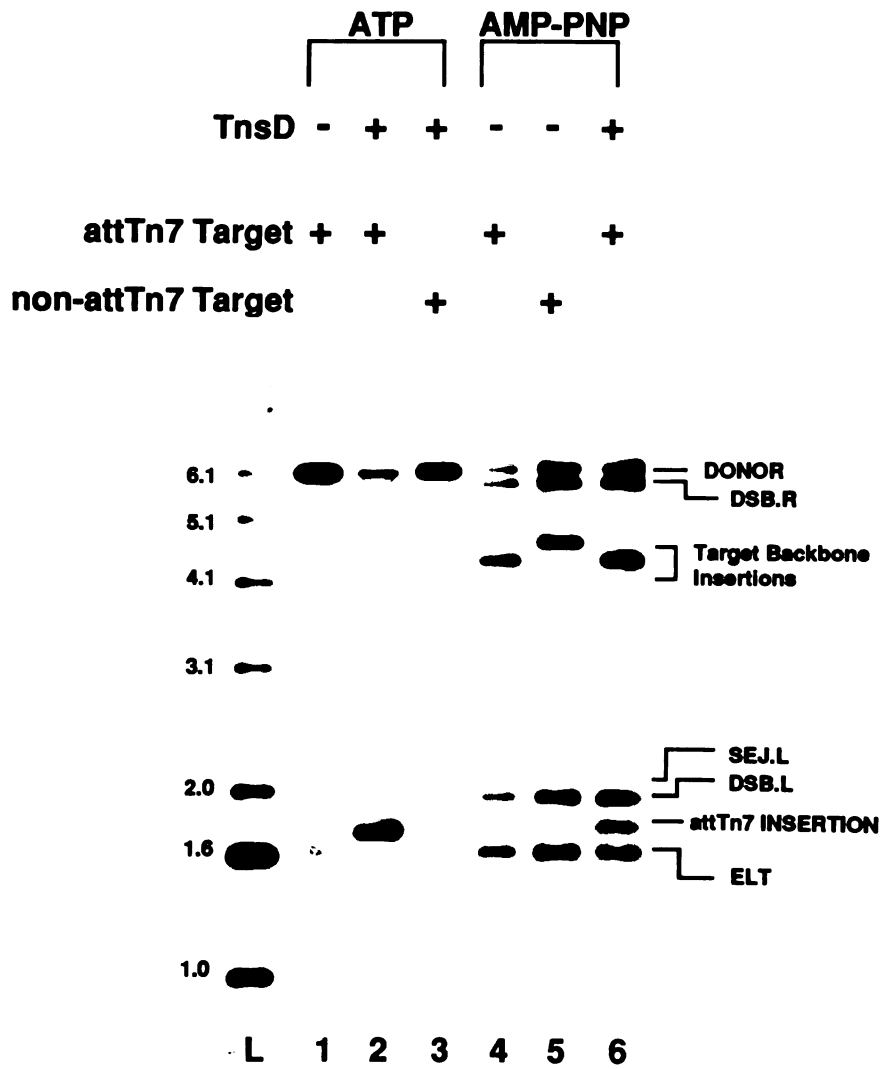

ATP
AMP-PNP
TnsD - + + - - +
attTn7 Target + + + +
non-attTn7 Target + +
6.1
5.1
4.1
3.1
2.0
1.6
1.0
DONOR
DSB.R
Target Backbone Insertions
SEJ.L
DSB.L
attTn7 INSERTION
ELT
L 1 2 3 4 5 6

Figure 7B

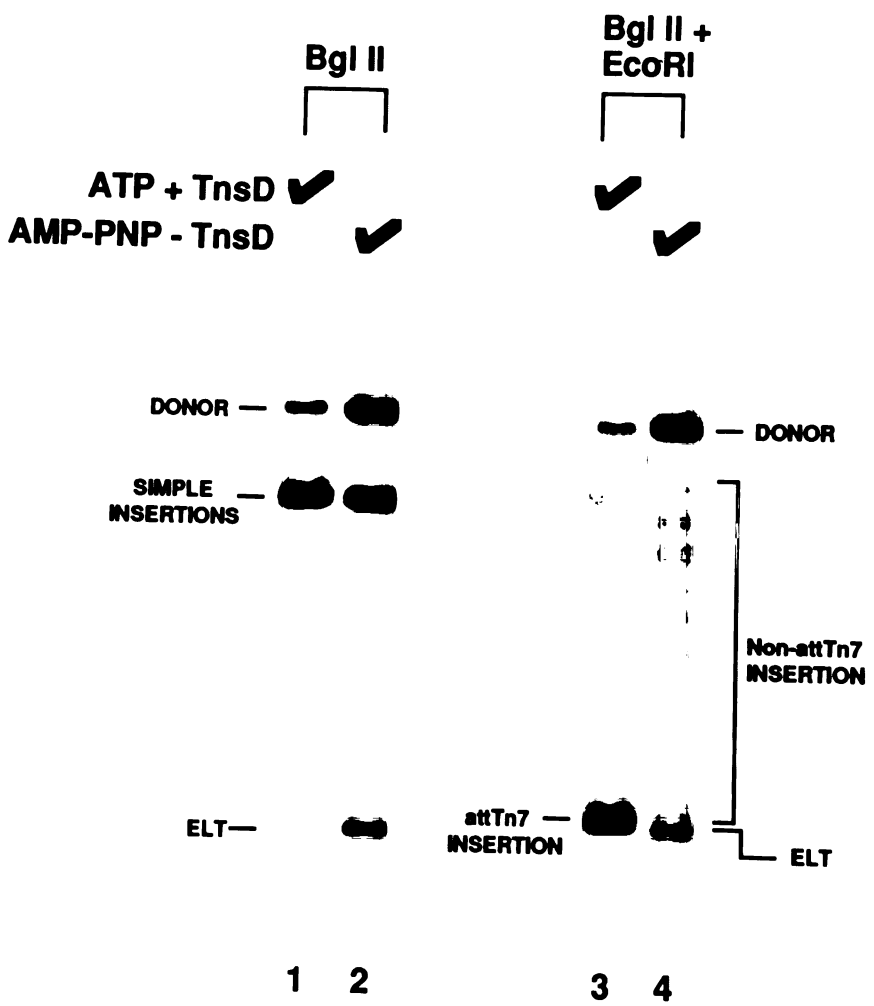
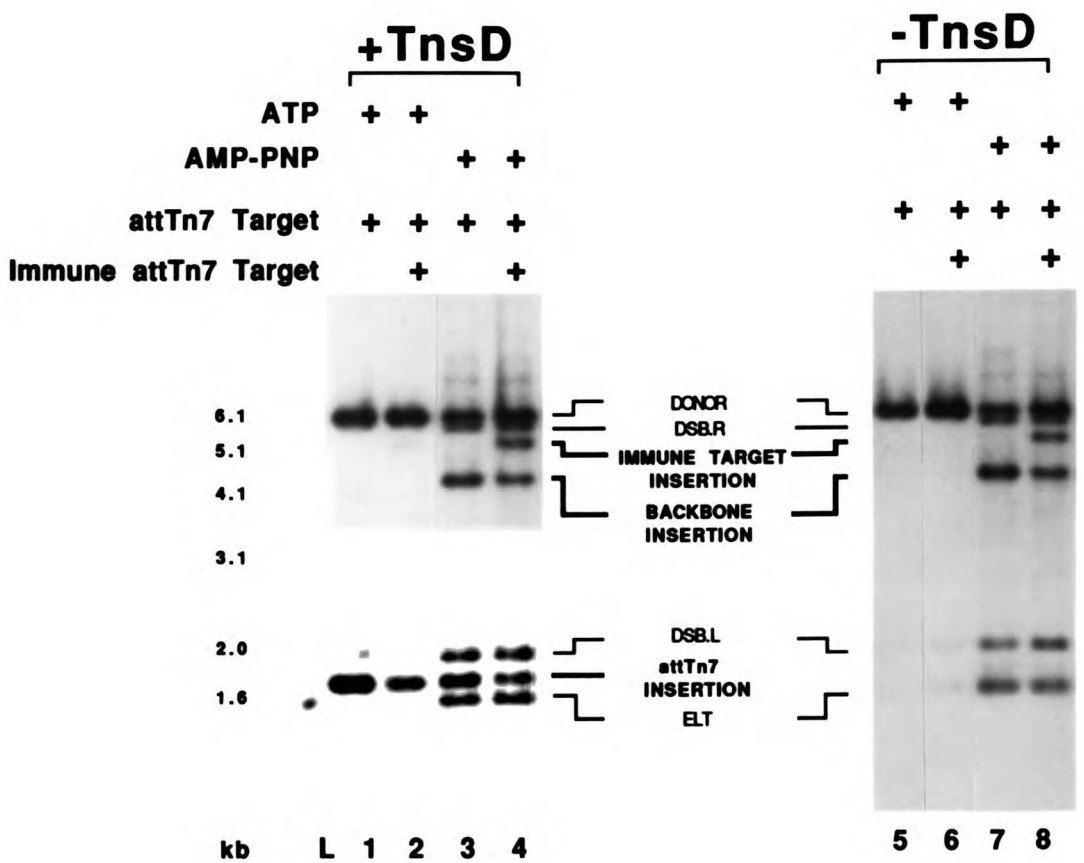

Bgl II
Bgl II + EcoRI
ATP + TnsD
AMP-PNP - TnsD
DONOR
SIMPLE INSERTIONS
ELT
DONOR
Non-attTn7 INSERTION
attTn7 INSERTION
ELT
1 2
3 4

**Figure 8. Tn7 Transposition Immunity In Vitro is Altered by the Presence of AMP-PNP**

EcoRI-digested species detected with a mini-Tn7-specific probe are shown from transposition reactions using the standard attTn7 target plasmid (pKAO4-3) 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.



# Chapter 4

## Single End Joins and Excised Linear Transposons in Tn7 Transposition *In Vitro*

## Introduction

We have proposed a pathway of Tn7 transposition, stated in Chapter 2 and reiterated in Chapter 3, which suggests that the reaction proceeds through a series of double-strand break species (DSBs) to produce an excised linear transposon species (ELT). We suggest that the ELT is an intermediate in transposition and is the direct precursor to insertion. In the highly purified reaction conditions used in Chapter 3, we find another potential intermediate species, a single-end join species (SEJ). The SEJ species is also observed at very low levels in the crude extract in vitro transposition reaction presented in Chapter 2 (data not shown). The SEJ species, shown schematically in Figure 1, is a covalent join of one end of the transposon to only one strand of the target DNA. We have observed that SEJs can be formed with DSBs (DSB-SEJ), such that the SEJ contains the transposon still attached to the donor DNA at one end and to the target DNA at the other end, and also with ELTs (ELT-SEJ). The DSB-SEJs are the only species known in Tn7 transposition in which the transposon is attached simultaneously to the donor and target DNAs.

The observation of an SEJ is important for two reasons: 1) it suggests the possibility of an alternate pathway for Tn7 transposition, i.e. that the ELT species may not be an obligate intermediate in Tn7 transposition, and 2) formation of an SEJ species reveals that each end of Tn7 can join separately to the target DNA. That the ends of Tn7 join to the target DNA in separable steps is consistent with the strand transfer model for transpositional recombination discussed in the Introduction. The strand transfer model proposes that the hydroxyl group at the 3' terminus of the transposon that is exposed during the donor cleavage phase of transposition nucleophilically attacks the phosphodiester backbone of the target, joining the transposon to a 5' target end and leaving a 3' target end unjoined in the process (Figure 2). In this model, the insertion reaction at one end of the transposon is chemically independent of the other. Thus the

concerted joining of the ends to the target that is usually observed during Tn7 transposition represents the concerted action of the overall transposition machinery rather than the obligate chemical linkage of reactions at each end of the transposon with the target DNA.

We provide evidence in this chapter that Tn7 transposition can produce SEJs. We also report that under the conditions in which SEJs are observed, we also observe transposon-dimer insertions, i.e. simple insertion transposition products made from dimer-donor molecules such that two outside cis ends, which bound two copies of the transposon and one copy of the donor backbone, are used rather than the closest cis ends which bound a single copy of the transposon. We observe greater production of both SEJs and transposon-dimer insertions in the reconstituted system described in Chapter 3 than in the crude extract system described in Chapter 2 (data not shown). This observation suggests that some step in the transposition process may be altered in the reconstituted system, mediated perhaps by more promiscuous transposition complex formation. The final section of this chapter discusses the properties of excised linear transposon substrates (ELTs) in transposition.

## Results

### SEJs are Produced in a Purified Tn7 Transposition System

Evidence for SEJs

Production of SEJs is most evident in transposition when it is carried out with purified TnsB, TnsC, TnsD and either purified TnsA (see Chapter 3) or addition of a small amount of a crude extract from a strain that overexpresses TnsA (see appendix experimental protocol). A convenient way of detecting the SEJ species is by using a small linear attTn7 fragment labeled at both ends with $^{32}P$. Under these conditions, the transposition products need not be digested with a restriction enzyme to interpret the product species. (Figure 3A). The standard donor DNA, pEM, is a mixture of monomer and dimer molecules in a 2:1 molar ratio; this donor produces a number of distinct transposition products when used in the purified reaction conditions (Figure 3). A simple insertion product of a monomer transposon (SI) is the major product of this reaction (Figure 3A, lane 1, 1.8kb). Other species are consistent with a monomer DSB-SEJ (M-SEJ, 6.0kb), a transposon-dimer simple insertion (D-SI, 7.8kb) and a DSB-SEJ from a dimer-donor (D-SEJ, 12.3kb). These species are shown schematically in Figure 3B. When a monomer donor is used, only the standard 1.8kb SI product and the 6.0kb monomer DSB-SEJ are detectable as products (data not shown), suggesting that the species we see in Figure 3A, lane 1 are indeed dimer-related transposition products.

These SEJ and transposon-dimer insertion species have also been identified and characterized in the standard in vitro reaction with supercoiled target DNA using multiple restriction digestions and Southern blot analysis (Chapter 3 and data not shown). With various restriction digestions and detection by Southern blot, the SEJ species can be further differentiated to identify which transposon end is joined to the target DNA. A DSB-SEJ can occur such that either the left or right end of the transposon is joined to the target DNA. It reasonable to suggest that these SEJ species are derived

from the double-strand break species (DSB.L and DSB.R) discussed in Chapter 2 and shown schematically in Figure 1. The DSB.L-SEJ and the DSB.R-SEJ species are found in equal quantity to one another (data not shown).

What is the nature of the joint between the transposon end and the target DNA in the SEJ and how is this species formed? Shown in Figure 4 is a schematic representation of the strand transfer reaction that shows two alternative pathways for joining the transposon end to the target. In pathway 1, the 3' end of the transposon is joined to one target strand such that one target strand is broken and one strand is intact. This species is consistent with the strand transfer model of transposition (Mizuuchi, in prep.). In pathway 2, the 3' end of the transposon is joined to one target strand of a target molecule that has been broken by a double-strand break.

The reaction products from the experiment shown in Figure 3A can help distinguish between these two models. In this reaction, the target molecule is a linear attTn7 fragment labeled at both ends with $^{32}P$. If, as in pathway 2, the target DNA is broken by a double-strand break prior to the joining of the transposon end to one target strand, an attTn7 fragment smaller than the original target fragment will be released while the rest of the target site will be attached to the transposon end, i.e. the SEJ species (see the products of Pathway 2). This small attTn7 species should be evident as a 35 bp transposition product present in roughly equal in quantity to the SEJ species. No 35bp attTn7 fragment is evident below the unused attTn7 substrate in this reaction (see Figure 3A, below the attTn7 fragment). Other experiments suggest that if such a small fragment was produced by the insertion reaction we would be able to see it in Figure 3A, lane 1 (data not shown). This data is consistent with the hypothesis that the SEJ species is formed by pathway 1 shown in Figure 4 and implied in the schematic drawing in Figure 1, i.e. that only one target strand is broken when one transposon end is joined to attTn7. Moreover, no broken attTn7 target plasmids are observed in the standard reaction using supercoiled attTn7 targets, and detected by hybdridization (data

not shown). The finding that attTn7 is not apparently broken by a double-strand break during Tn7 transposition and that the SEJ species apparently involves only one strand of target DNA is consistent with the one step transesterifcation reaction mechanism proposed for transpositional recombination by Mizuuchi (see also Introduction and Discussion).

Are SEJs intermediates or products of transposition?

The above observations reveal SEJs are produced in the Tn7 transposition reaction. Can these species be further processed to form simple insertion transposition products, i.e. are the SEJs transposition intermediates or are they aberrant products that are refractory to further manipulation by the Tn7 transposition machinery? We examined the time course of DSB-SEJ production in a standard in vitro transposition reaction, i.e. one that uses a supercoiled target DNA; the species produced are detected by Southern hybridization following restriction digestion with ScaI (Figure 5). Both SEJ.Ls and SEJ.Rs are seen as a single band at 7.7 kb. This species is not detectable at early times in the time course (lanes 1 and 2) but seems to accumulate at later times (lanes 3 and 4). The simple insertion product (4.5kb) accumulates in the same manner. In contrast, the DSB.L (3.7kb) and and DSB.R (3.9kb) species appear early and maintain a low steady-state level. (This time course was not long enough to see the DSBs actually reduce in level, i.e behave as true transposition intermediates). Their kinetic behavior, i.e. late appearance in the reaction, suggests that DSB-SEJs are not significant intermediates in our standard transposition reaction. Perhaps the DSB-SEJs are converted to simple insertions only very inefficiently.

Evidence for production of the transposon-dimer simple insertion species (D-SI) is also provided in this reaction. (Figure 5; see also Figure 3). ScaI cuts the D-SI in two places, producing 5.1kb and 5.4kb species which are not resolved here. As expected

for transposition products, these species are not found early (lanes 1 and 2) and accumulate at late times (lanes 3 and 4).

A Comparison to Crude Extract Conditions

The two experiments described above in which SEJs and the transposon-dimer insertions are evident were performed under reconstituted transposition reaction conditions. It is important to note that the SEJs and transposon-dimer insertions were observed at much lower levels in a similar experiment done in the crude transposition system shown in Chapter 2, Figure 3 (species cannot be seen). Perhaps the more purified system does not accomplish the transposition reaction with the same care as the crude system, i.e. in the crude system, the two ends of the transposon are dealt with concertedly but, under the more purified conditions, this concertedness is somewhat relieved, both producing SEJs and allowing transposon-dimers to be used as substrates. To determine whether the lack of a host factor was responsible for SEJs and transposon-dimer insertions, we looked at the appearance of SEJs and transposon-dimer insertions in reactions in which crude extract was added to reactions using the purified Tns proteins. The SEJs and transposon-dimer insertion species were both evident in these supplemented reactions, although their amount did appear diminished from that of the reactions lacking crude extract (data not shown).

It is possible that in optimizing the purified transposition system solely on the basis of maximum conversion of donor substrate to simple insertion transposition product, some step in the transposition process was disrupted. Perhaps the ratio of host proteins to Tns proteins, a condition different in the crude extract and reconstituted systems, plays a role in maintaining the proper order of breaks and joins in transposition. In addition, other differences such as alteration of the Tns proteins during purification, or differences in the absolute concentrations or ratio of the Tns proteins are possible factors for the apparent differences between these two conditions. Other

factors to consider are the salt, small molecules, and glycerol concentrations, which differ between the two types of in vitro reactions.

**Effects of Mutations in the Transposon Termini on Tn7 Transposition**

The Donor Release Plasmids

Beginning a transposition reaction with artificially excised linear transposon (ELT) substrates offers a way of bypassing the requirement of the transposition reaction to perform the donor cleavage or "donor release" step of the reaction. Using these substrates, we can test the requirements of the insertion reaction, i.e. strand transfer, independently from the steps necessary to make the ELT intermediate, i.e. donor cleavage. By changing the sequences at and/or flanking the transposon termini in a donor plasmid, we introduced sites for restriction enzyme cleavage so that the transposon could be artificially excised from the donor backbone. Digestion with a restriction enzyme can release an ELT species with a free 3' end (Figure 6, see Experimental Protocols); in some cases, the sequences of the transposon termini are actually altered whereas in some cases only the flanking DNA is changed. These constructs are called Donor-Release plasmids (DRP) (Figure 6). We assayed their activity in transposition as intact plasmids to measure their activity in donor cleavage and strand transfer and as isolated ELT species derived from these plasmids to evaluate strand transfer.

The DRP denoted as pMIM has completely wild-type transposon ends but the DNA flanking the ends has been changed. (Figure 6, row 1). In pSIM, the terminal two nucleotides are changed at both ends of the transposon (row 2); in pXIM, only the terminal nucleotide is changed at both ends (row 5). Chimeric DRPs are made with one wild-type end (from pMIM) and one mutant end (from pSIM). In pCHI1 (row 3), Tn7L is wild-type and Tn7R is mutant; in pCHI2 (row 4), Tn7L is mutant and Tn7R is wild-type.

Transposition Activity of Intact DRP Plasmids Containing Transposons with Mutated Termini

Unless otherwise indicated, the following experiments were performed in reactions containing a small amount of a crude extract from a strain that overexpresses TnsA and highly purified TnsB, TnsC and TnsD (see Experimental Protocols). When either the standard donor pEM or pMIM, a DRP substrate with wild-type transposon ends and altered flanking sequences ends, are used as intact plasmid substrates, efficient recombination is observed with the production of high amounts of simple insertion product (Figure 7: pEM - panel A, lane 1, panel B, lane 1, panel C, lane 1; pMIM, panel A, lane 2). With pEM, also evident at lower levels are the intermediate DSB and ELT species and, DSB-SEJs. Thus, as expected, both donor cleavage and strand transfer occur efficiently with wild-type transposon ends. (It should be noted that an EcoRI digestion does not allow the intermediate species to be seen with the pMIM substrate.)

By contrast, transposition intermediates or insertion products are barely detectable when intact DRP plasmids containing transposons with mutations at both termini are used as substrates, pSIM (panel A, lane 4, around 4.8kb) or pXIM (panel B, lane 3 around 4.8 kb). These constructs are equally poor at producing any transposition-related species (data not shown). Thus changing the sequence of the transposon termini can block the donor cleavage step.

pCHI1 and pCHI2 contain one wild-type and one mutant transposon end. When they are used in intact plasmid form, they produce equal levels of SEJ insertion products, but no simple insertions (panel B, lane 5 and panel B, lane 7 respectively). They are therefore active substrates for transposition, but their activity is reduced as compared to substrates with two wild-type transposon ends (compare above to panel A, lane 2).

These observations reveal that mutations at the termini can affect the donor cleavage stage of transposition.

ELT With Mutated Termini as Transposition Substrates

By digesting DRPs with the appropriate enzyme, an ELT species is produced with free 3' transposon ends containing either wild-type or mutated terminal sequences. When used as transposition substrates, the DRP ELT species bypass the strand cleavage requirement and allow independent evaluation of their ability to participate in strand transfer. In general, intact DRP plasmids and their ELTs have similar insertion efficiencies in in vitro transposition; this suggests that the particular nucleotides at the end of the transposon are critical for both donor cleavage and strand transfer.

When an ELT with wild-type ends (from pMIM) is used as a substrate, joint molecules between the ELT and attTn7 target DNA are produced; however, many species other than the simple insertion product are generated (Figure 7, panel C, lane 2). In this experiment, about 25% of joint molecules are standard simple insertions (4.5kb); another 25% of the joint molecules are ELT-SEJ.L and ELT-SEJ.Rs (4.8kb). The other species at 2.3kb and 3.4kb which are also each about 25% of the joint molecules likely result from the integration of two separate ELTs into a single target molecule. (These products could result from the concerted insertion of two ELTs that associated with each other prior to integration or the integration of a second ELT into an already formed ELT-SEJ.) The structure of these products is reminiscent of the transposon-dimer insertions (D SI) that are also seen under these conditions when using pEM as a donor (Figure 7C, lane 1; also seen in Figure 3).

In the experiment described above, although specific joining of the transposon ends to attTn7 is observed, the activity of both ends of a transposon does not appear to be highly concerted, i.e. many SEJs are made. It should be noted that the types of joint molecules formed between an ELT substrate and the target appear to be somewhat different under the more purified conditions described here (crude extract TnsA and purified TnsB, TnsC, TnsD) than under the crude extract conditions described in Chapter

2. Under the crude conditions, a much higher fraction (about 80%) of the joint molecules formed between the pMIM ELT substrate and the target DNA appear to be authentic simple insertions (data not shown). As we have already noted, the products of transposition under the purified conditions from intact donor plasmids are also different, i.e. SEJs and transposon-dimer insertions are more obvious under purified compared to crude conditions. The observation that more SEJs are produced with an ELT substrate than with an intact plasmid substrate could suggest that the nature of the substrate, i.e. on a supercoiled plasmid as opposed to an ELT, may also contribute the proper stepwise assembly of a transposition complex responsible for the insertion reaction.

It should be noted that identification of the various forms of ELT-target joint molecules is highly dependent on how the reaction products are cut prior to analysis. When the reactions are digested with Sca1, many different species are discernible (Figure 7C, lane 2); digestion with EcoRI results in a more uniform product appearance (figure 7A, lane 3). As can be seen in Figure 3B SEJ species form T-shaped molecules. EcoRI cuts relatively close to the new transposon-target junctions (see Figure 1), thus making it difficult to discern the T-shaped SEJ from a standard simple insertion with both ends inserted. A ScaI digestion allows a better discrimination between the SEJ and the standard simple insertion because it cuts much further away from the new junctions maintaining the T-shape of the SEJ.

The efficiency of forming joint molecules between ELT transposons and target DNA is much reduced when the ELT has mutated termini; joint molecules are virtually undetectable with ELTs that have two mutations at both termini (pSIM ELTs - panel A, lane 5, panel C, lane 3) or with ELTs in which only the terminal nucleotide is changed at both ends (pXIM ELTs - panel B, lanes 2 and 8, panel C, lane 6)

The formation of joint molecules between target DNA and ELTs with one wild-type end and one mutant end is particularly interesting. With these substrates, all the joint

molecules formed appear to be ELT-SEJs (panel C); the distribution between SEJ.L and SEJ.R forms has not been established. Also, there is considerable difference in the activity of ELTs with mutant right ends versus mutant left ends. A chimeric ELT with a wild-type right end (pCHI2, panel C, lane 5) is a better insertion substrate than one with a wild-type left end (pCHI 1, panel C, lane 4). This suggests that there is some discrimination between the right and left ends of Tn7 by the in vitro transposition machinery. It is interesting to note, however, that when pCHI1 and pCHI2 are provided as intact DRP plasmids, they have similar insertion activities (Figure 7B, lanes 5 and 7).

Although reduced in efficiency, ELTs with mutant ends, i.e. from pXIM, pCHI1 and pCHI2, do retain the ability to specifically recognize attTn7 (Figure 7B, lanes 8, 9,and 10). Thus these terminal mutations have not altered the insertion specificity of the in vitro reaction, although the insertions they do produce are primarily SEJs and not simple insertions.

End to End Joining

When using ELT substrates in transposition, a 3.3kb species is produced (Figure 7C, lane 3). The size and restriction digestion patterns of this species are consistent with end-to-end join (EEJ) of the ELT substrate. This species is resistant to EcoRI and ScaI digestions, and partially sensitive to BglII digestion (data not shown).

The EEJ species is sometimes present at low levels in unreacted ELT substrates (data not shown). EEJ levels are increased by incubation under in vitro reaction conditions and/or processing of in vitro reaction substrates. The presence of a crude extract that contains TnsA, a target molecule, and $Mg^{2+}$ seem to be required to increase the levels of the EEJ recovered in in vitro reactions (data not shown). EEJ production can be stimulated by, but does not seem to require the presence of TnsB, TnsC and TnsD

(data not shown). Additional crude extract has little or no effect on its production (data not shown).

All of the ELT transposons can produce EEJs. The pSIM transposon end derivatives, in particular, are the best substrates for this reaction (figure 7C, lanes 2 to 4). pMIM ELT and pXIM ELT end to end join species are present, but at much lower levels (Figure 7A, lane 3 and Figure 7C, lane 5).

## Conclusions

### SEJs in Tn7 Transposition

SEJs are found at very low levels under the crude in vitro conditions described in Chapter 2. Higher levels of SEJs are apparently produced under reaction conditions that use purified Tns proteins (Chapter 3 and this chapter). It is important to note, however, that even under the reconstituted conditions, the simple insertion species is, by far, the major product when an intact supercoiled plasmid is used as a donor. These SEJ species are single-ended insertions into one strand of the attTn7 site (Figure 3, pathway 1). These species can be produced from an intact donor molecule, i.e. pEM, or from an ELT species. We find that these species accumulate in the transposition reaction. There are two possible explanations for the genesis of SEJs: 1) they are the result of an aborted insertion reaction or 2) they are an intermediate that can precede simple insertion formation. The observation of these species has revealed that the two transposon ends can insert at different times rather than in the highly concerted manner normally observed, i.e. that one end of the transposon can be broken away from the donor molecule and inserted, prior to the processing of the other end.

Other transposition systems that can produce SEJs are the Mu and the retroviral systems (Craigie and Mizuuchi, 1987; Craigie et al, 1990). Usually, however, Mu transposon end insertions are tightly coupled (Craigie and Mizuuchi, 1985) whereas the retroviral integration systems are considerably more promiscuous (Craigie et al, 1990).

It will be interesting to determine in Tn7 transposition whether artificially constructed SEJs substrates can complete the transposition reaction. Such an experiment would differentiate between models that suggest that the strand transfers are chemically coupled and those that suggest that they are merely temporally coupled through the action of the proteins that make up the transposition complex.

The Donor Release Reaction

In Tn7 transposition, the donor cleavage step at the 3' end of the transposon appears concomitantly with a break at the 5' end of the transposon that completely releases the transposon from the donor backbone (see Figure 2); these steps together are called the Tn7 donor release reaction. With retroviruses, the chemistry of the donor cleavage step at the 3' end of the transposon is similar to the chemistry of the strand transfer reaction (Engelman et al 1991, see Thesis Introduction). Our demonstration here that mutations at the ends of the Tn7 can affect both donor release and strand transfer steps supports a mechanistic unification of these reactions in Tn7 transposition (this Chapter). For further discussion of this point, see Summary and Perspectives.

Terminal Mutations in Transposition Substrates

*Intact Donor Plasmids*

In Chapter 2, we demonstrated that ELTs are an efficient substrate for in vitro transposition. This result suggested that the breaks in the ends of the transposon need not be coupled to the strand transfer reaction. Construction of transposons with mutations at their termini that allowed the production of artificial ELT intermediates, i.e. Donor Release Plasmids, was an attempt to distinguish between the requirements for donor release and strand transfer. We found that the nature of the terminal transposon nucleotides affected both reactions; for example, changing the two terminal nucleotides at both transposon ends blocked both donor cleavage and strand transfer.

Analysis of the transposition activity of transposons with one wild-type end and one mutant end, i.e the pCHI plasmids, suggests that the transposon ends are in communication during transposition. The pCHI plasmids were better substrates for both donor cleavage (Figure 7B) and strand transfer (Figure 7C) than were transposons with two mutant ends. One interpretation of this data is that the wild-type end can compensate for deficiencies in the mutant end and allow a greater level of transposition

overall. However, the wild-type end in a chimeric plasmid is not as active as a wild-type end in a fully wild-type transposon and thus it is not merely acting autonomously to improve the reaction efficiency. The modulation of the mutant phenotype by a wild type end suggests that there is some communication between the ends of the transposon before the strands are broken and joined to the target site.

A similar phenomenon is seen with phage Mu where a wild-type end can compensate for a mutant end when the transposon is undergoing the donor cleavage reaction (Surette et al, 1991). Other genetic studies with Tn10 and Tn3 also suggest end to end communication in transposition (Roberts et al, 1991; Derbyshire and Grindley, 1986). One model to explain these observations suggests that firstly, the transposon ends are aligned within a transpososome complex. Secondly, before the breakage and joining reactions can occur, the aligned ends must undergo some sort of conformational change that requires the extreme termini of the transposon ends to participate at the same time (Surette et al 1991). Thus, if two mutant ends are paired, the conformational change can not occur and no donor release occurs. However, if one end is wild-type, then it can, in some way, compensate for the mutant end and help the complex complete the conformational change and stimulate cleavage.

*ELT Donor Substrates*

The left and right ends of Tn7 are structurally and functionally distinct (Arciszewska et al, 1989; see Thesis Introduction). In vivo experiments demonstrate transposons that contain two right ends are competent donor molecules whereas transposons that contain two left ends are not. Also, Tn7 inserts with a strong orientation-specificity (Lichtenstein and Brenner, 1982). These data suggest that the left and right ends have different abilities at recruiting Tns proteins and building the necessary machinery to accomplish an insertion reaction. Asymmetrically processed transposition species such as the right and left DSBs and the right and left SEJ species

occur in equal amounts (Chapter 2 and this Chapter). If the ends of Tn7 are functionally distinct, then why are these species apparently produced and treated equivalently by the transposition machinery? One possibility is that the distinction between the left and right ends occurs prior to the breakage and joining reactions that generate the DSBs and SEJs. This suggests that discriminations between the ends may occur during alignment of the ends and prior to initiation of recombination, as discussed in the previous section.

Special transposon substrates described in this chapter have allowed us to see the first evidence of in vitro discrimination between the Tn7 transposon ends. ELT substrates that contain mutations in their termini have revealed an interesting difference between the right and left ends of the transposon. Chimeric ELT transposons containing a wild-type right end are better substrates for transposition than transposons carrying a wild-type left end. This is the first example of a mechanistic distinction between the left and right ends of Tn7 in vitro.

Why does an ELT substrate allow us to see a difference between the left and right transposon ends? One explanation derives from the idea that in order to accomplish transposition from an intact donor plasmid, both ends of the transposon must be brought together (see Thesis Introduction and above). If the building of a specific nucleoprotein complex to accomplish the donor release step is an important constraint on the Tn7 transposition reaction, then it is possible that much of the specificity of Tn7 transposition, i.e. site-and orientation-specificity and distinction between the ends, is built into this step. If this complex is responsible for positioning all the substrates and proteins prior to recombination, then it is reasonable to suppose that within this complex distinctions are made between the left and right ends of the transposon. An ELT substrate frees the transposition reaction from the constraints of the donor release step. Perhaps ELT substrates also free the transposition machinery from a step normally necessary in the assembly of a transposition complex. One consequence of this may be to bypass a step that identifies and positions both transposon ends and the target site prior

to the initiation of any recombination. Thus, because ELT substrates are precut the left and right ends may act more autonomously and therefore display different insertion efficiencies.

The Terminal Nucleotides in Tn7 Transposition

We have shown here that the identity of nucleotides at the extreme termini of the transposon can play a key role in both the donor-release reaction and the insertion reaction in Tn7 transposition. This suggests a coupling of these steps which is likely advantageous to the transposon. If the two reactions are mechanistically linked, then it is unlikely that mutations in the required terminal sequences could allow one reaction to occur without the other. This assures that if breaks in the transposon ends are made, then it can insert them into a target. If donor cleavage could occur without strand transfer, this would be disastrous to the transposon.

Alternative Products of the Transposition Reaction

We have reported here that SEJ species and transposon-dimer insertions are generated by Tn7 transposition. These species are apparent at only very low levels in the crude extract conditions (see Chapter 2), but are relatively more abundant in more highly purified transposition conditions, although they are only a fraction of the transposition products (Chapter 3). Generation of the SEJ species suggests that the highly concerted nature, (i.e. both transposon ends inserting apparently simultaneously into the same target), of the crude system is somewhat diminished in more purified systems. The transposon-dimer insertions suggest that the mechanism that normally brings the two closest transposon together is somewhat relaxed. The coincidence of the formation of these two species under similar conditions also suggests that the transposition machinery in more purified systems may be altered in some way from that of the crude system. Both the lack of host proteins and optimization of the reaction may

play a role in allowing the production of these SEJ and transposon-dimer species, perhaps by altering requirements for the formation of the nucleoprotein complex necessary to promote transposition (see Summary and Perspectives for more discussion).

End-to-End Joins

When using ELT substrates in transposition, an end-to-end join is sometimes seen. I have not been able to clearly determine whether this is a transposition-dependent product or an artifact of the process of preparing and using the ELT species. The following paragraph suggests some interesting and not-so-interesting possibilities for these presence of these species in transposition reactions. Unfortunately, I do not know which of these possibilities is the most likely; they are merely suggested explanations.

A trivial explanation for the EEJ species could be an end-to-end annealing process caused by melting of the transposon ends. This could explain why these species are sometimes found in very low quantities even before they are added to the transposition reaction. Another explanation is that a transposition-related reaction joins the ends of transposons. We have already suggested that ELT substrates display interesting transposition properties because they are released from the constraints of the donor release reaction (see above). Perhaps the end-to-end join reaction is an attempt by the transposition machinery or a subset of the transposition machinery to reverse the donor release reaction. It is interesting to note that the best substrate for this reaction was the pSIM transposon. Perhaps the mutations in this transposon encouraged an altered transposition related reaction. Another possibility is a host protein-dependent ligation of the linear substrates; these substrates were never used in the most highly purified transposition conditions in order to test this possibility.

## Experimental Protocols

### In Vitro Reaction Conditions

Figures 3 and 5 use in vitro reaction conditions identical to Chapter 3. The linear target, used in Figure 3, is a small fragment of the attTn7 site (125 bps) labelled at both ends with $^{32}P$. The in vitro reactions are done just as with the linear attTn7 in Chapter 3, except that the reactions are stopped by urea-spermine precipitation (See Chapter 2).

The in vitro reaction conditions in Figure 7 are identical to the those used in Chapter 3 except that TnsA is added as 0.1 µl of a crude extract. The TnsA crude extract is an htpr⁻ strain (CAG 456, Baker at al, 1984) carrying the pKAO52 plasmid (Orle and Craig, 1990). The extract is prepared by freeze-thaw lysis described in Chapter 2.

### Production of the Donor Release Plasmids

We synthesized a series of R-terminus oligonucleotides primers completely complementary to R1-19 of the top strand at the right end of Tn7. These primers differ from the bottom strand of the left end of Tn7 at positions 17 to 19. In addition to the above sequences, these primers contained restriction enzyme sites at the 5' ends. The MluI primer contained the complete wild type sequence of the transposon end plus an MluI site, a BamHI and two additional adenosine nucleotides. The ScaI primer contained two changes from the wild type sequence at the transposon terminus, a ScaI site, a BamHI site and two additional adenosine nucleotides. The XbaI primer contained only a single change at the terminus, an XbaI site and two additional adenosine nucleotides. Using these primers singly and a mini-Tn7 plasmid as template, PCR amplification was done to generate linear transposons whose ends were flanked by MluI and BamHI, ScaI and BamHI and XbaI (Mullis and Faloona, 1987). The PCR reactions contained 100 pmol of

primer, 1 fmol of template,EcoRI cut pEM, 2mM dNTPs, and PCR buffer. The linear transposon species were digested with a restriction enzyme that cut the terminal restriction sites at the two ends and was cloned into the complementary site in bluescript PKS+ (Stratagene). The result of this cloning was pMIM, pSIM, and pXIM (see Figure 6).

The pMIM and pSIM plasmids were used to produce the pCHI donor molecules. The right and left end of both pMIM and pSIM were isolated separately from one another by HindIII-BamHI digestions. These ends were cloned back together such that pCHI1 contains an MluI left end and a ScaI right end and pCHI2 contains an ScaI left end and a MluI right end (see Figure 6).

Digestion with the appropriate combination of enzymes releases excised linear transposon species with the breaks at the 3' end of the transposon. These species were gel purified (Geneclean) and used as ELT intermediate species in the in vitro reaction.

**Restriction Digestion, Gel Electrophoresis and Southern Hybridization**

These steps are done as in Chapter 2.

**Figure 1. The Transposition Substrates, Intermediates, and Products.**

The 6.1 kb donor molecule contains a 1.6 kb mini-Tn7 element flanked by sequences unrelated to attTn7; the mini-Tn7 element contains segments from the left (Tn7L) and right (Tn7R) ends of Tn7 that provide the cis-acting transposition sequences. The 2.9kb 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 element inserts site- and orientation specifically into attTn7, producing a 4.5 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 single Tn7 end joins made by a single Tn7 end joining to a target molecule. Restriction sites: E = EcoRI, B = BglII, and S = ScaI.

FIGURE 1

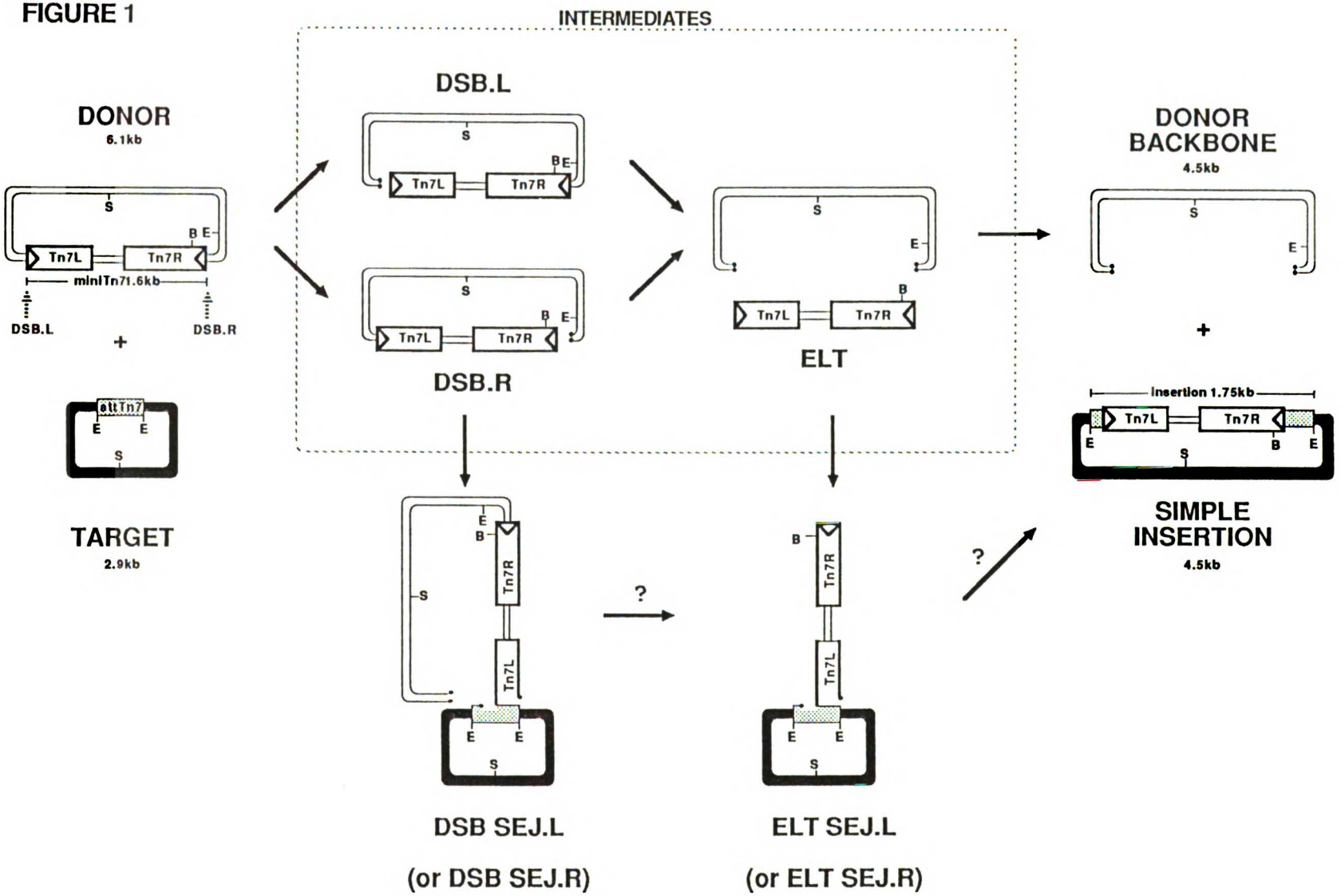

INTERMEDIATES
DONOR
6.1kb
S
B E
Tn7L
Tn7R
miniTn71.6kb
DSB.L
DSB.R
+
attTn7
E E
S
TARGET
2.9kb
DSB.L
S
B E
Tn7L
Tn7R
DSB.R
ELT
E
B
DONOR
BACKBONE
4.5kb
Insertion 1.75kb
SIMPLE
INSERTION
4.5kb
?
DSB SEJ.L
(or DSB SEJ.R)
ELT SEJ.L
(or ELT SEJ.R)

**Figure 2 The Breakage and Joining Reactions of Tn7**

Shown is a schematic representation of the breakage and joining reactions that are thought to occur during Tn7 transposition to its specific insertion site attTn7. The specific sequences at the ends of the transposon and the target site are from the pEM-1 donor and the standard attTn7 target site (see Chapter 2). The small arrows represent the breaks that occur during the donor release step that separate the transposon from the donor backbone. In Tn7 transposition, the donor release step includes both the donor cleavage reaction which breaks the 3' end of the transposon, [i.e. similar to Mu transposition (see Introduction)] and another reaction that cleaves the 5' end of the transposon (see Perspectives). The long arrows are representative of the strand transfer reactions that join the 3' ends of the transposon to the target site breaks resulting in a simple insertion. The breaks that are left over after strand transfer are at the 3' ends of the target DNA and at the 5' ends of the transposon DNA (see Chapter 2 for data).

# Figure 2

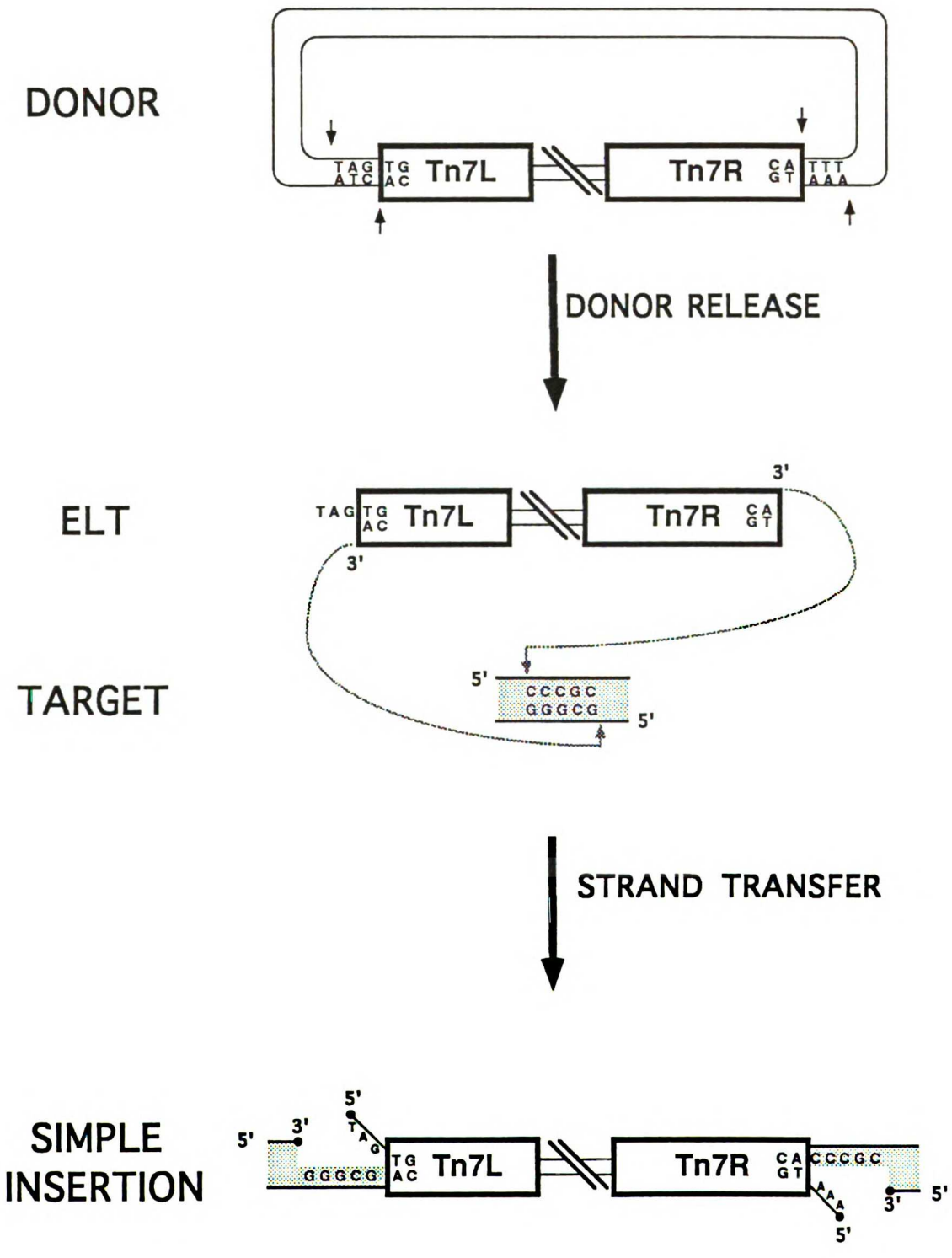
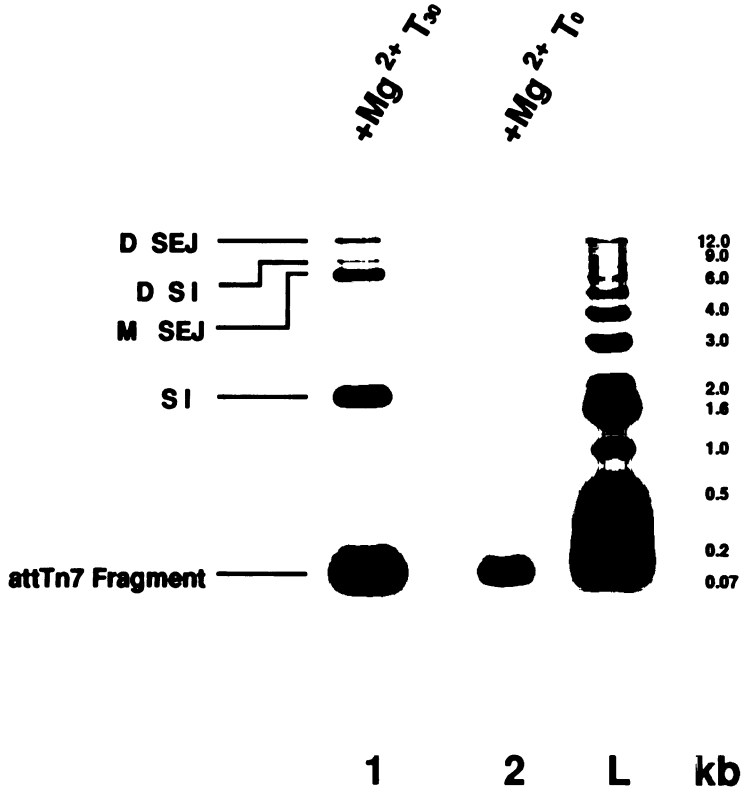

DONOR
TAG
ATC
TG
AC
Tn7L
Tn7R
CA
GT
TTT
AAA
DONOR RELEASE
ELT
3'
TAG
TG
AC
Tn7L
Tn7R
CA
GT
3'
TARGET
5'
CCCGC
GGGCG
5'
STRAND TRANSFER
SIMPLE
INSERTION
5'
3'
5'
TAG
GGGCG
TG
AC
Tn7L
Tn7R
CA
GT
CCCGC
AAA
3'
5'
5'

**Figure 3 Tn7 Transposition to a Linear AttTn7 Fragment**

A) Shown is an autoradiogram of in vitro transposition which used an attTn7 labeled at both ends with $^{32}P$ (120 base pairs in length). The donor molecule, pEM-1, is a mixture of monomer and dimer plasmids in a 2 to 1 ratio. The in vitro reaction uses the highly purified Tns proteins. The in vitro reaction conditions are the same as those used for the TnsC-TnsD complex formation in Figure 4B, Chapter 3 and see Experimental Protocols this chapter. The in vitro reaction was precipitated twice and run on a .7% agarose gel. In lane 1 ($Mg^{++}$ $T_{30}$), the reaction is done with the standard preincubation conditions, i.e. the MgAc is added after a 30' preincubation. In lane 2 ($Mg^{++}$ $T_0$), the MgAc is included from beginning of the reaction (as a negative control). Lane L is a 1kb DNA ladder. D-SEJ = dimer single-end join. D-SI = transposon-dimer insertion. M SEJ = monomer single-end join. SI = standard simple insertion.

B) A schematic drawing of the product species identified in figure 3A. The abbreviations are the same as above. The size of the species is designated in kilobases. The species are not drawn to scale relative to one another. SEJ species differ from the SI species in that only one end of the transposon is joined to the attTn7 target; these species are drawn with the attTn7 site in a vertical instead of a horizontal position.

Figure 3A


+Mg 2+ T30
+Mg 2+ T0
D SEJ
D SI
M SEJ
SI
attTn7 Fragment
12.0
9.0
6.0
4.0
3.0
2.0
1.6
1.0
0.5
0.2
0.07
1
2
L
kb

# Figure 3B

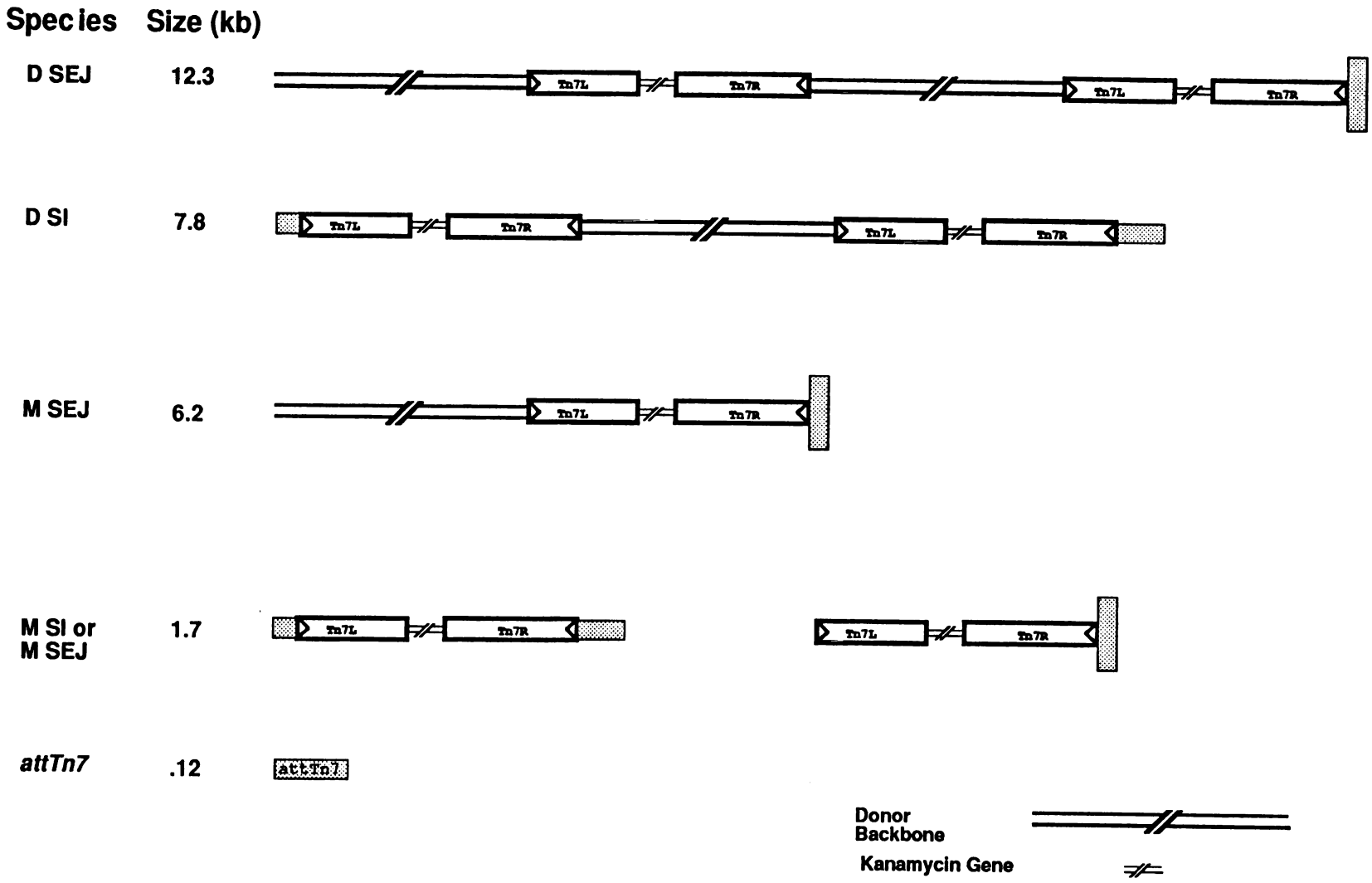

Species
Size (kb)
D SEJ
12.3
Tn7L
Tn7R
Tn7L
Tn7R
D SI
7.8
Tn7L
Tn7R
Tn7L
Tn7R
M SEJ
6.2
Tn7L
Tn7R
M SI or
M SEJ
1.7
Tn7L
Tn7R
Tn7L
Tn7R
attTn7
.12
attTn7
Donor
Backbone
Kanamycin Gene

**Figure 4 The Possible Structures of Single-End Joins**

Shown is a schematic drawing which depicts the putative structures of single end join species. In case 1, the exposed 3' end of the transposon (bold lines) attacks the target (filled lines), joining to a single target strand and leaving one break. In case 2, the transposon end joins a single target strand, but leaves a double strand break in the target. In case 2, a small attTn7 fragment is released upon insertion.

# Figure 4 The Possible Structures of Single End Joins

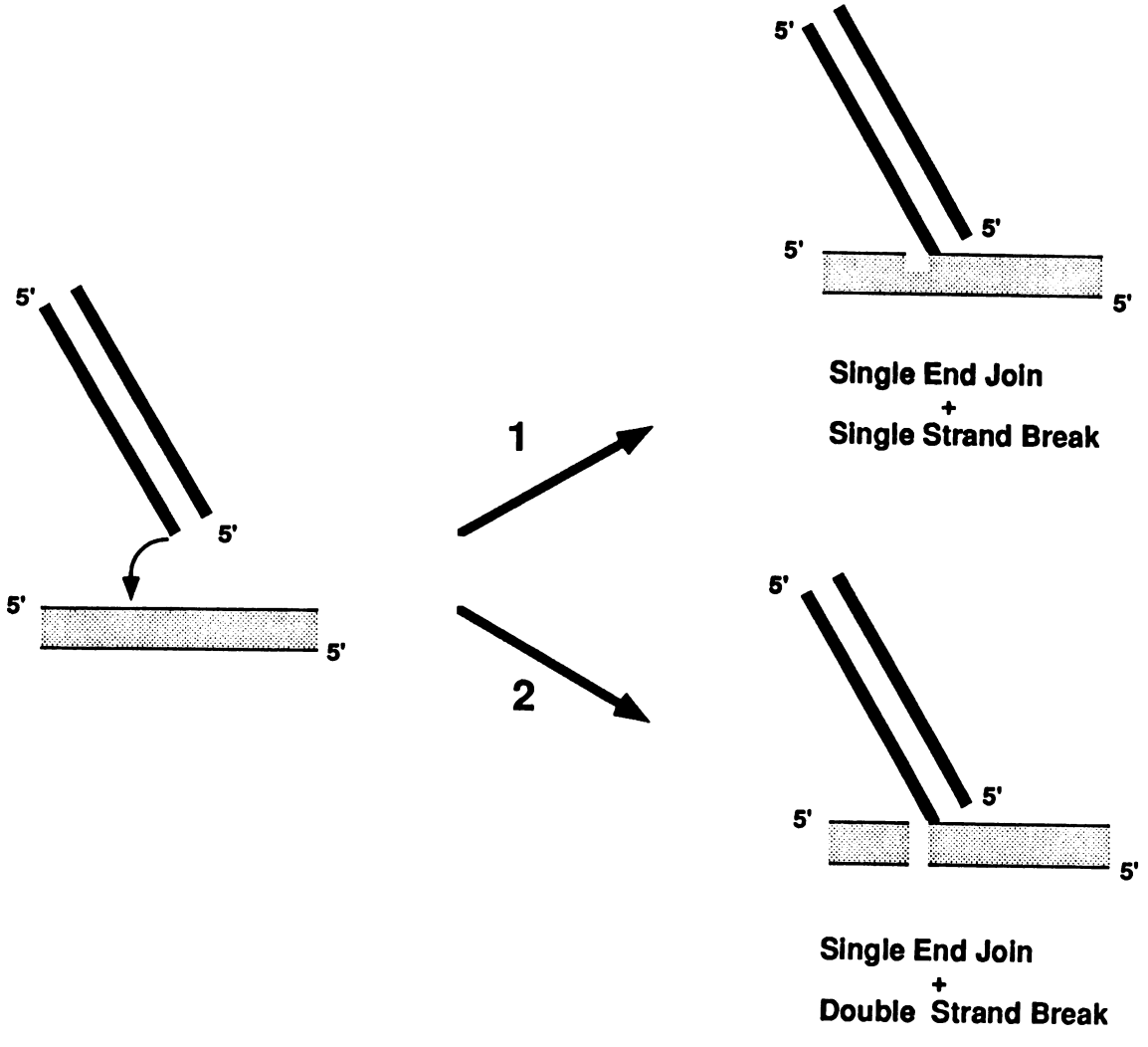


**Figure 5 An In Vitro Transposition Time Course Using the Purified Reaction Conditions**

Shown is an autoradiogram of a Southern blot using a transposon-specific probe. Aliquots of a reaction mixture were removed from the incubation step at the times indicated. The standard reaction conditions were used, DNA isolated, cut with ScaI and run on a 0.6% agarose gel (see Experimental Protocols). SEJ.L+R are the left and right single-end join species. The D SI species are the transposon-dimer simple insertions. DSB.L and DSB.R are the double strand breaks, left and right. Species which accumulate in the reaction are the (Simple) Insertion, SEJ.L+R, and D SI. The DSB.L+R appear early and maintain a steady state level as the reaction progresses, as expected for intermediates in the reaction.

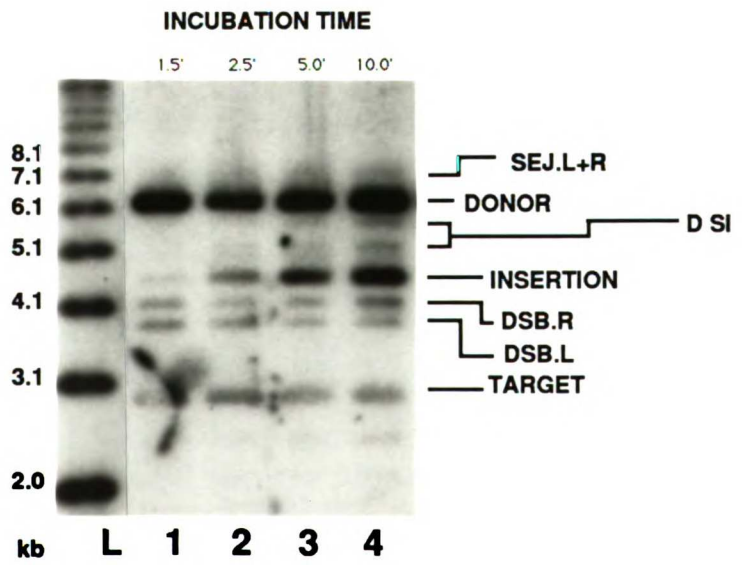


**Figure 6 The Structure of the Donor-Release Plasmids**

Shown is a schematic drawing of the structure of the Donor-Release Plasmids. The plasmid names are on the left. The restriction sites that release the transposon are shown at the ends of each element with arrows that indicate the sites of cleavage. Mutant ended transposons are indicated in two ways, verbally at the left of the elements and schematically as * in the ends of the elements.

## Figure 6 Terminal Sequences of the Donor Release Plasmids

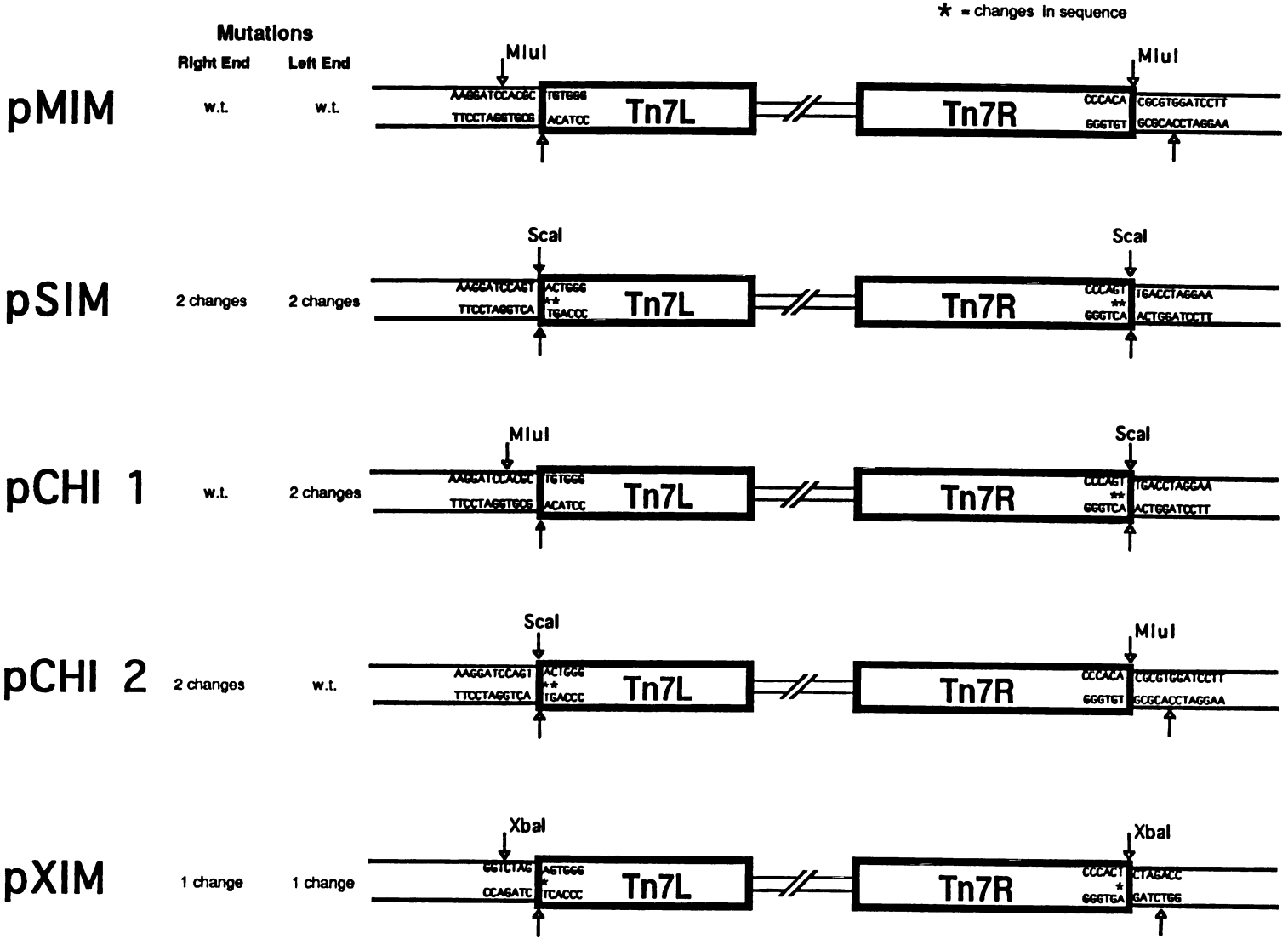


**Figure 7 The Activities of the Donor Release Plasmids (DRP) and Excised Linear Transposon (ELT) Substrates in Tn7 Transposition.**

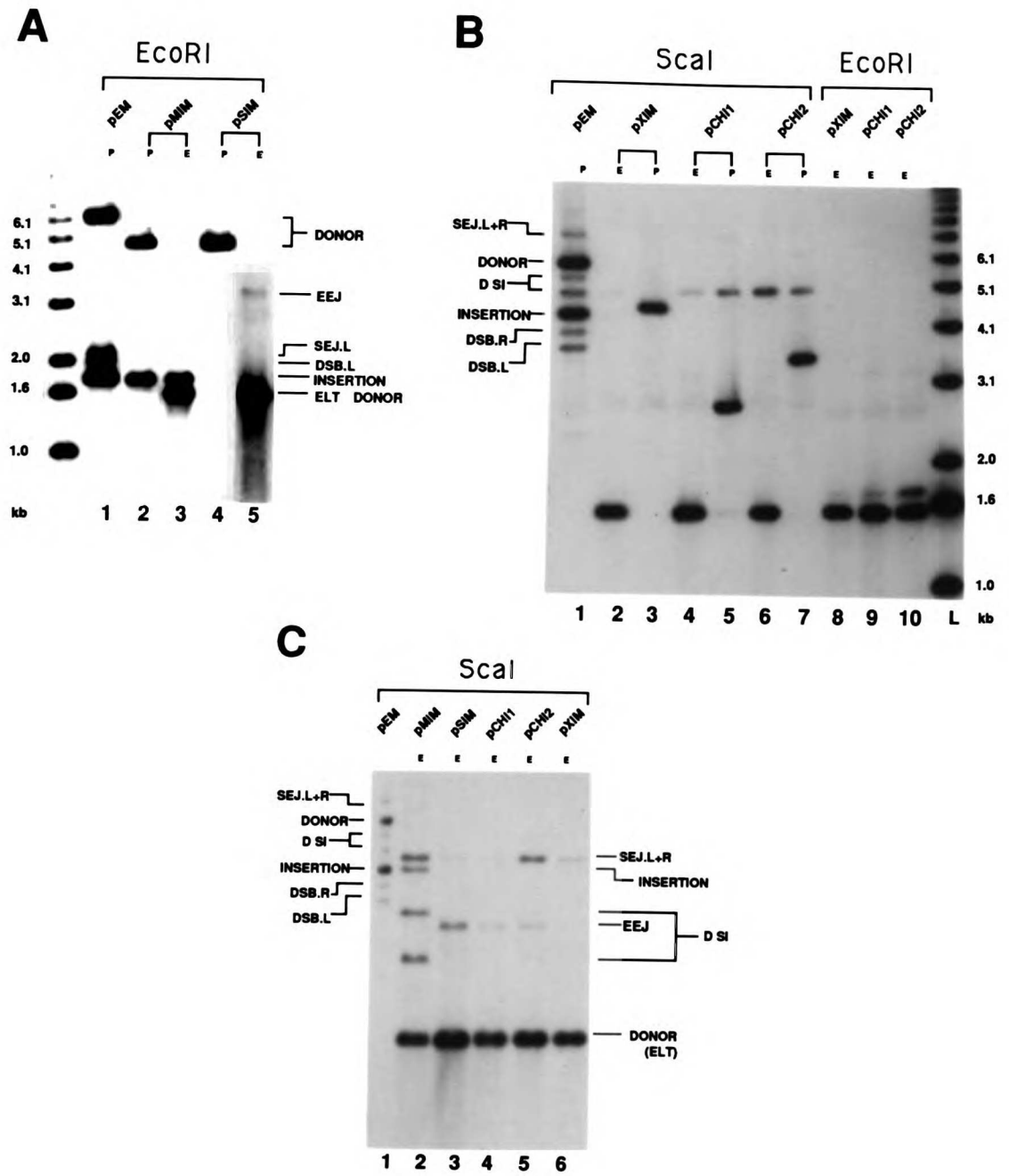
All panels are autoradiographs of the Southern blots using a transposon-specific probe. The purified Tns proteins were used except for TnsA which was provided in a crude extract (see Experimental Protocols) The restriction enzymes used to process the reaction products for analysis are shown above the panels. EcoRI digestions demonstrate whether the insertions are concentrated in the attTn7 site and ScaI digestions are best for showing intermediate species (see Chapter 2, Figure 1 for the restriction map of the product species.) The plasmids are designated by the names used in Figure 4. The intact donor-release plasmids (P) are 4.3kb in size (prior to digestion) and the excised linear transposons (E) derived from the donor-release plasmids are 1.6kb in size. The abbreviations are SEJ.L+R = single end join right of left, D SI = dimer simple insertion, DSB.L or R = double strand break left or right and EEJ = end to end join.

A) These reactions are digested with EcoRI. The attTn7 insertion product is 1.75kb. Lane 1 is positive control using the standard pEM donor. Lanes 2 and 3 use wild type DRP, pMIM, as a P and E substrate. Lane 4 and 5 use doubly mutant, pSIM, as a P and E substrate.

B) Lanes 1 to 7 are digested with ScaI. Lane 1 is positive control using pEM. The simple insertion product is 4.5Kb when digested with ScaI. The other transposition species were identified in figure 3. Lanes 2 and 3, 4 and 5, and 6 and 7, use pXIM, pCHI 1 and pCHI 2 as an E and P substrates respectively. This digestion compares the relative efficiency of the DRP plasmids in their E and P forms.

Lanes 8 to 10 are pXIM, pCHI 1, and pCHI 2 respectively digested with EcoRI. This digestion shows that all the insertions are in the attTn7 site.

C) This panel is a side by side comparison of all the ELT forms of the donor-release plasmids. All reactions are digested with ScaI. Lane 1 is a positive control using pEM; the species in this reaction are designated to the left. Lanes 2 to 6 are pXIM, pSIM, pCHI 1, pCHI 2, and pXIM, respectively. The transposition species formed in these reactions are designated to the right.


A
EcoRI
pEM
pMIM
pSIM
P
P E
P E
6.1
5.1
4.1
3.1
2.0
1.6
1.0
kb
1 2 3 4 5
DONOR
EEJ
SEJ.L
DSB.L
INSERTION
ELT DONOR
B
ScaI
EcoRI
pEM
pXIM
pCHI1
pCHI2
pXIM
pCHI1
pCHI2
P
E P
E P
E P
E
E
E
SEJ.L+R
DONOR
D SI
INSERTION
DSB.R
DSB.L
6.1
5.1
4.1
3.1
2.0
1.6
1.0
1 2 3 4 5 6 7 8 9 10 L kb
C
ScaI
pEM
pMIM
pSIM
pCHI1
pCHI2
pXIM
E
SEJ.L+R
DONOR
D SI
INSERTION
DSB.R
DSB.L
SEJ.L+R
INSERTION
EEJ
D SI
DONOR
(ELT)
1 2 3 4 5 6

# Chapter 5

# The Effects of DNA Supercoiling on Tn7 Transposition

## Introduction

Supercoiling is a fundamental property of closed circular DNA molecules. Supercoiling in biological systems is brought about by enzymes and processes which remove helical turns, thus DNA in cells is usually underwound. Supercoiling can effect many systems that manipulate DNA, such as transcription, replication, repair, and recombination (Vosberg, 1985; Wang, 1985; Wells, 1988). The mechanistic effects of negative supercoiling are thought to relate to bending, wrapping or unwinding of the DNA (Wang et al, 1983; Gellert and Nash, 1987, Wang and Giaever, 1988). Such manipulations are used to facilitate the juxtaposition and alignment of specific DNA sites that are at a distance from one another, such as with Tn3 resolvase-mediated resolution (Boocock et al, 1987) and Hin-mediated recombination (Johnson et al, 1984).

Supercoiling is stimulatory to the formation of a number of specialized nucleoprotein complexes (Richet et al, 1986; Craigie and Mizuuchi, 1986; Surette and Chaconas, 1989; Heichman and Reid, 1990). In site-specific recombination, supercoiling seems to be important for the donor molecule, i.e. the molecule in which distant sites must be aligned. In transposition, supercoiling of the target molecule has not been observed to be essential for recombination (Craigie and Mizuuchi, 1986; Brown et al, 1987).

We have performed some very simple experiments on the effects of DNA supercoiling on Tn7 transposition. Our experiments suggest that the state of supercoiling of the donor molecule is very important to the rate and kinds of products of the Tn7 transposition reaction. A surprising observation is that target supercoiling may also play a limited role in preparing the target site for insertion.

## Results

### In Vitro Transposition Reaction Done With Limiting TnsD Protein

We carried out a time course of the standard transposition reaction using the four purified Tns proteins, TnsA, TnsB, TnsC and TnsD (see Chapter 3). In order to study the effects of supercoiling on the initial rate of transposition, we reduced the amount of TnsD added to the reaction by fifty-fold and decreased the preincubation and incubation times as compared to the standard reaction conditions (see Experimental Protocols). The time course of the reaction using supercoiled substrates is shown in Figure 1A. Under these conditions, the amount of simple insertion product (SI) formed (1.75 kb species) is only about 5% of the total amount of the donor substrate as compared to greater than 80% under the standard conditions (see Chapter 3). The initial rate of SI product (1.75 kb) formed over time is linear as evaluated by quantitative analysis with a phosphorimager (data not shown). As expected, the double-strand break left, (DSB.L, 2.0kb) intermediate species is in greater abundance than the SI early in the time course; the excised linear transposon (ELT, 1.6kb) species is produced more slowly than both the SI and the DSB.L in the time course (Figure 1A).

### Donor Supercoiling

When a linear donor molecule in used in the transposition reaction instead of a supercoiled plasmid, the reaction pathway changes (Figure 1B). The intermediate DSB.L and ELT species appear to be made in much lower quantities, while the rate and extent of SI production is more moderately affected. By quantitative analysis, the amount of SI is reduced by at least half whereas the amount of the intermediates are reduced by about five-fold (data not shown).

### Target Supercoiling

When a linear target molecule is used in transposition, an interesting phenomenon takes place (Figure 1C). The amount of intermediate DSB.L and ELT species produced is similar to that observed with supercoiled donor and target (Figure 1A), but the initial rate of SI formation is two-fold lower. Ultimately, the amount of SI formation is the same whether the target is supercoiled (1A) or linear (1C) (data not shown).

### Donor and Target Supercoiling

A further diminution in both rate and extent of all transposition-related species, DSB.L, ELT and SI, occurs when both the donor and target substrates are linear (Figure 1D).

## Conclusions

### Donor Supercoiling

It is not unexpected that supercoiling of the donor molecule in Tn7 transposition would effect the rate and extent of transposition. Site-specific recombination systems are greatly dependent on the supercoiling of the donor molecule. In lambda site-specific recombination, the Tn3 resolvase system and the bacteriophage Mu systems, supercoiling seems to be involved with juxtaposition and alignment of the DNA sites and therefore lack of supercoiling can have a large effect on the efficiency of the reaction (Nash, 1981; Reed, 1981; Mizuuchi, 1983).

We suggest that the observations about the effects of supercoiling in Tn7 transposition are consistent with a larger view about the process of transposition. In Chapters 2 and 3, we suggested that a large nucleoprotein complex is formed which allows communication between all the donor elements, the attTn7 target site, and the Tns proteins of Tn7 transposition. In Chapter 4 we provided evidence suggesting that there

is end-to-end communication in Tn7 transposition. The data here support the idea that Tn7 transposition, like other site-specific recombination systems, requires the synapsis of the ends of the element within a nucleoprotein complex. Thus, because supercoiling has been implicated in just such a process in other systems, it is not surprising that supercoiling of the donor molecule can affect Tn7's ability to participate as a substrate in transposition (Figure 1B).

**Target Supercoiling**

Supercoiling of the target site has a subtle effect on the transposition reaction (Figure 1C). When the target is linear the reaction seems to stall slightly at the strand transfer step, pausing somewhat longer than under standard conditions with a supercoiled target (Figure 1A). The donor release step is quite efficient when using a linear target DNA conditions, producing large quantities of DSB.L and ELT (Figure 1C). If both donor and target supercoiling have independent effects on the overall rate of SI formation, then it is reasonable to suggest that together they would have an even greater effect on transposition, as demonstrated in Figure 1D.

However, before we can accept that target site supercoiling has some effect on the efficiency of transposition, there are some caveats that must be considered. One pertains to spurious production of transposition intermediates. Under standard conditions, the donor cleavage reaction can occur at low levels in the absence of TnsD (Chapter 3). Because low concentrations of TnsD are used in this chapter, some of the intermediates seen in Figure 1C are produced by TnsD-independent donor cleavage (data not shown). However, we feel that the donor cleavage cannot alone account for the high level of intermediates seen under the conditions in Figure 1C, and therefore that the results may instead be representative transposition reaction stalled by the lack of target supercoiling (data not shown).

Another issue to discuss is that the linearization and isolation of these linear substrates might somehow effect their ability to participate in transposition, for example either by nicking the DNA or increasing the amount of salt in the reactions. Different methods of removing the supercoils from the DNA substrates could be used to look at this question.

## What Might Target Substrate Supercoiling Be Doing In Tn7 Transposition?

If the effect of target site supercoiling on transposition is a real phenomenon, then what is its mechanism of action? The current model of Tn7 transposition suggests that a nucleoprotein complex consisting of both ends of the transposon and all the Tns proteins forms at the attTn7 site (Chapters 2, 3 and 4). We propose that the correct assembly of proteins and substrates initiates the donor-release reaction which cuts the transposon out of the donor DNA to form the ELT species. The Tn7 transposition machinery seems to wait for both ends of the transposon to be freed from the donor backbone before inserting the element (as discussed in Chapters 2, 3 and 4). How does the Tn7 transposition machinery know when both ends of the transposon have been broken? We suggest that prior to promoting strand transfer, i.e. transposon end insertion, the Tn7 transposition machinery undergoes a conformational change. In this way, the transposition machinery knows that both ends of the transposon are free and ready for insertion. What would happen if we perturb the system by using a linear target DNA? If the putative ELT-target site complex is less stable when the target is linear or if the putative conformational change is stimulated by supercoiling of the target, it is reasonable to suggest that the reaction may stall here and accumulate intermediates. If both the donor and target supercoiling are removed, then both the production of intermediates and the insertion reaction may be impaired. Implied in this

model is a communication between the target and transposon ends through the Tns proteins that encourages the conformational change at the appropriate time.

It should be noted that under the standard reaction conditions, i.e. when TnsD is not limiting (Chapters 1 and 2), the absence of target supercoiling does not obviously effect transposition (data not shown). One explanation for this is that, under the standard conditions, there are excess targets available for transposition. Thus, although marginally debilitated, a larger number of linear target complexes are effective substrates for the reaction, making it difficult to see the lower rate of strand transfer demonstrated in Figure 1C.

## Experimental Protocols

### Preparation of the In Vitro Substrates

Both the pEM and pKAO4-2a were digested with ScaI which cuts once in the donor backbone (see Chapter 2 for restriction maps). This releases the supercoils from both molecules. These species were loaded onto 0.6% 1X TAE gels, run for 4 hours at 50 volts, stained with EtBr and the appropriate linearized species were cut from the gel. The DNA was isolated from these gels by Geneclean and the concentration determined by comparison to DNA standards.

### In Vitro Reaction Conditions

The in vitro reaction conditions are the same as those used in Chapter 3, except for a few special changes. The TnsD concentration is 0.1uM instead of 10uM. This decreases the efficiency of the reaction so that the initial rate of reaction can be studied. The total volume of the reaction was 200 µl. At the appropriate time, aliquots were removed from each reaction mixture. The reaction products were recovered by the Urea-spermine method used in Chapter 2.

### Restriction Digestion, Gel Electrophoresis, and Southern Hybridization

The products were double digested with EcoRI and ScaI prior to gel electrophoresis. The gel electrophoresis and Southern hybridization steps were done as in Chapter 2.

### Autoradiographs and Phosphorimaging

The Southern Blot was exposed to the Phosporimager screen overnight and then exposed to film at -70°C for one week.

**Figure 1. The Effects of DNA Supercoiling on the Rate of Tn7 Transposition**

Shown is an autoradiogram of a Southern blot using a transposon-specific probe. The substrates are provided either as supercoiled molecules or as ScaI-cut gel-purified linear species (see Experimental Protocols). The transposition reactions are done under standard purified reaction conditions except that the level of TnsD is decreased (see Experimental Protocols). All the time course species were deproteinized and then cut with ScaI and EcoRI. The 0 time point is the start of the preincubation, the 0' time point is prior to adding MgAc to the reaction and all other time points are minutes of incubation including MgAc. The relevant transposition species are the double strand break left (DSB.L), the simple insertion (SI), and the excised linear transposon (ELT).

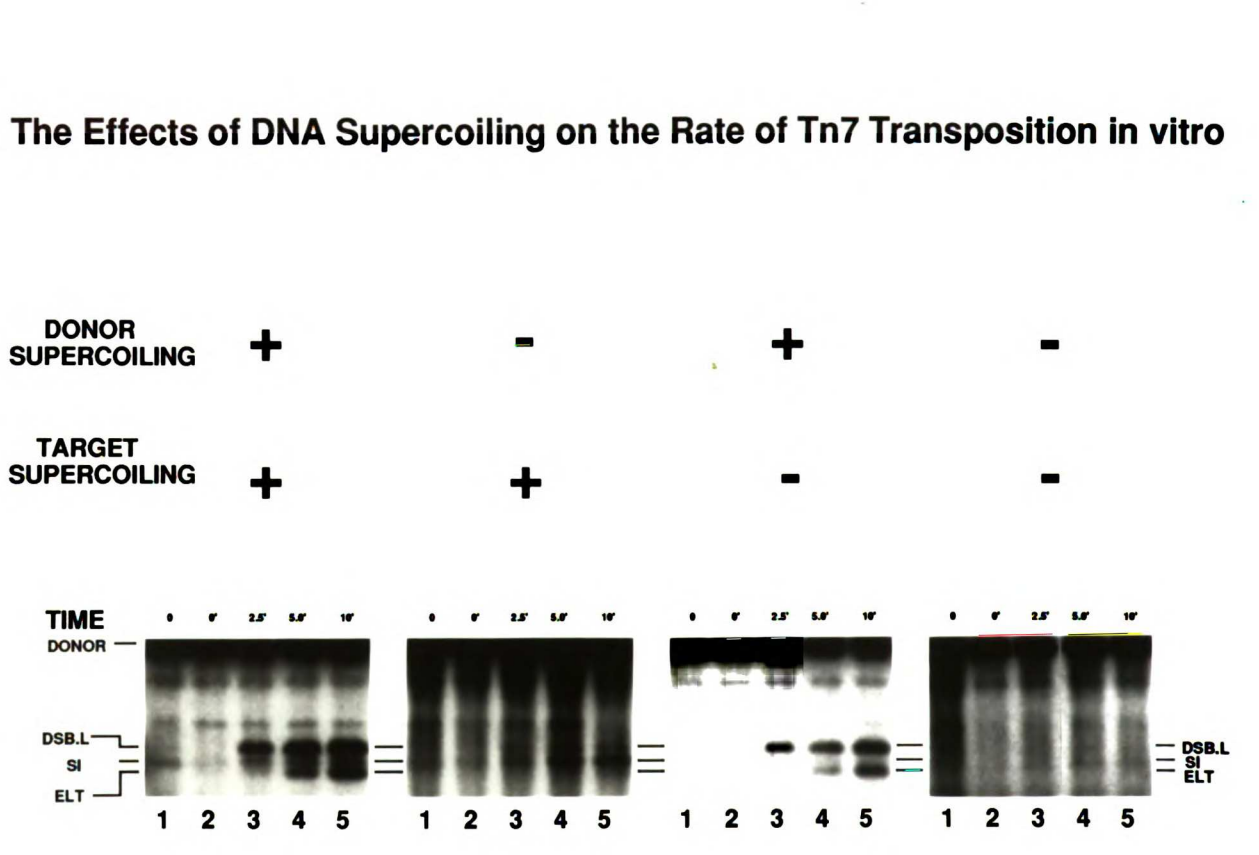
First column: Both substrates are supercoiled.

Second column: The target is supercoiled and the donor is a linear.

Third column: The target in linear and the donor is supercoiled.

Fourth column: Both substrates are linear.

# Figure 1


The Effects of DNA Supercoiling on the Rate of Tn7 Transposition in vitro
DONOR SUPERCOILING
+
-
+
-
TARGET SUPERCOILING
+
+
-
-
TIME
0
0'
2.5'
5.0'
10'
DONOR
DSB.L
SI
ELT
1 2 3 4 5
1 2 3 4 5
1 2 3 4 5
1 2 3 4 5
DSB.L
SI
ELT

# Chapter 6

# Summary and Perspectives

## Summary

This thesis is concerned with the biochemical analysis of Tn7 transposition. I established an in vitro reaction that promotes the movement of a Tn7 mini-element from a donor replicon to a target replicon. This recombination reaction has all the hallmarks of in vivo Tn7 site-specific insertion, i.e. the in vitro reaction products has the same site- and orientation- specificity as in vivo products. In vitro transposition requires a transposon donor, an attTn7 target, and extracts made from cells containing the tnsA + tnsB + tnsC+ tnsD genes. The phenomenon of transposition immunity is also active in vitro (Chapter 1). The following section summarizes our studies on Tn7 transposition in vitro by organizing the information into five overlapping categories: the products of transposition, the chemistry of transposition, the pathway of transposition, the biochemical requirements of transposition, and the roles and interactions of the substrates and proteins in transposition. My thesis concludes with some perspectives on the process of transposition and on future areas of study.

## The Products of Tn7 Transposition

Tn7 transposition in vitro promotes intermolecular movement of a mini-Tn7 donor element from a donor molecule to a target molecule containing an attTn7 site (Chapter 2). The products of this reaction are a simple insertion of the element into the attTn7 site and a donor backbone species that is recovered in stoichiometric quantities; thus, Tn7 transposition is a non-replicative, i.e. cut and paste, reaction. The donor backbone species is not resealed in the in vitro transposition reaction (Chapter 3).

## The Chemistry of Tn7 Transposition

The chemistry of Tn7 transposition is similar to that of other transposition systems (Chapter 2). Tn7 separates itself from the donor backbone DNA by cleaving its

3' ends. The 3' ends of the transposon are joined to staggered 5' target site breaks. Tn7 transposition is different from other characterized transposition reactions only in that the 5' end of the donor transposon is immediately cleaved away from the backbone by staggered cuts in the DNA. It is interesting to note that the staggered breaks that release the transposon from the donor backbone and the staggered breaks made at the target site upon insertion both produce 5' overhangs, although they are apparently different in length, 5 bps and 3 bps respectively.

## The Pathway of Transposition

In Tn7 transposition, the transposon is released from the donor backbone by staggered, double-stranded cuts, producing transient double-strand break intermediates (DSBs). Two double-strand cuts produce an excised linear transposon (ELT) (Chapter 2). The ELT species can act as direct precursor for insertion, suggesting that it is an intermediate in the transposition reaction. We suggest that, at low frequency, a single transposon end from the DSB or ELT species may undergo the strand transfer reaction by itself to produce a single-end join (SEJ, Chapter 4). The SEJ species are found at a very low level in the crude extract transposition system, but at more significant levels in the purified transposition system (Chapters 4). These species do not appear early in transposition, but accumulate as the reaction progresses, suggesting that they may be aborted transposition products rather than reaction intermediates (Chapter 4). The presence of these species however suggests that the two strand joining reactions needed to produce a simple insertion transposition product are not chemically coupled, but can occur separately.

## The Requirements of the Purified <u>In Vitro</u> Transposition System

The purified protein components required to promote efficient recombination to <u>attTn7</u> are TnsA, TnsB, TnsC, and TnsD (Chapter 3). ATP and MgAc are required

cofactors in the recombination reaction. Supercoiling of the donor molecule and, to a lesser extent of the target molecule is stimulatory to the transposition reaction (Chapter 5). In the presence of ATP, purified TnsA + TnsB + TnsC can promote the production of low levels of DSBs and ELTs, i.e. intermediate species, and, in the presence of non-hydrolyzable analogues of ATP, TnsA + TnsB + TnsC can promote high levels of transposition to non-specific target sites (Chapter 3). This suggests that TnsA, TnsB and TnsC form the central recombinational machine of Tn7 transposition and that this machine is directed to utilize the attTn7 target site by TnsD and non-specific sites by TnsE (Chapter 3). Each of the purified Tns proteins, TnsA, TnsB, TnsC and TnsD, have been shown to participate directly in the transposition reaction. Under the most optimal transposition conditions, the Tns proteins are present in close to a one-to-one ratio, suggesting that these components participate structurally in transposition, i.e. that no Tns component participates merely catalytically. It is important to note that the Tns proteins can promote transposition in the absence of any host component; however, we have observed that under some conditions host factors may stimulate recombination.

## The Roles and Interactions of the Substrates and Proteins in Tn7 Transposition

An interesting aspect of Tn7 transposition is that the presence of all the Tns proteins, the donor molecule and an attTn7 site target site, are required to efficiently initiate any recombination events (Chapter 3). The supercoiling requirement of the donor and, to a lesser extent, the target molecule, is suggestive of the need for communication between the special recombination sites in Tn7 transposition (Chapter 5). The behavior of transposons with mutations at their termini is consistent with the suggestion that the transposon ends communicate during transposition (Chapter 4). A model which synthesizes these observations is that a large nucleoprotein complex, consisting of the Tns proteins, the target site and both transposon ends in communication

with one another and with the target site, is formed during the process of transposition. This model suggests that it is within such a complex that transpositional recombination is initiated and completed.

We suggest that Tn7 nucleoprotein complex assembly may be obligatory to initiate recombination at the attTn7 target site (Chapter 3). We have presented evidence to suggest that a complex consisting of TnsC, TnsD and the attTn7 site is highly stimulatory to both donor cleavage and strand transfer under standard transposition reaction conditions. In this complex, TnsD acts a site-specificity factor which directs TnsC to the attTn7 target site. We further suggest that this complex can recruit the remaining transposition machinery, TnsA, TnsB and the donor molecule, to this site, thereby promoting transposon insertions at attTn7.

Several other observations suggest that TnsC's interaction with target DNA is a key to promoting transpositional recombination (Chapter 3). Biochemical evidence indicates that TnsC can bind DNA in an ATP-dependent manner (Gamas and Craig, in press). TnsC + TnsA + TnsB can promote efficient transposition to non-specific DNA sites in the presence of AMP-PNP (Chapter 3). We also suggest that TnsC plays a role in target site immunity through control of its interaction with ATP (Chapter 3). These data all suggest that TnsC plays a important role in promoting communication between target sites and the remainder of the transposition machinery.

## Perspectives on Transpositional Recombination: The Breakage and Joining Reactions in Transposition

The donor cleavage and strand transfer reactions occur sequentially to cause transpositional recombination (Mizuuchi, in prep.). Transposases catalyze these two recombination reactions at the 3' terminus of the transposon (Mizuuchi, in prep.). Both reactions use direct hydroxyl nucleophilic attack with phosphoryl transfer to accomplish their respective strand breakage and strand joining reactions (Mizuuchi and

Adzuma, 1991; Engelman et al, 1991; Vink and Plasterk,1991; Mizuuchi, in prep.). These observations suggest that, although donor cleavage and strand transfer accomplish markedly different DNA manipulations, they are the result of closely related chemical reactions (Engelman et al, 1991; see Thesis Introduction).

Tn7 transpositional recombination is similar to that of Mu, retroviruses, retroelements, and Tn10, i.e. Tn7 joins the 3' end of the transposon to a 5' target site break (Chapter 2, also see Thesis Introduction). The transposition pathway of Tn7 allows a unique opportunity to delve into the chemistry of transpositional recombination (Chapter 2). Excised-linear transposons are intermediates in Tn7 transposition. With the ability to introduce specific sequence changes in both intact donor molecules and excised linear transposon substrates, it is possible to study the sequence requirements of the donor cleavage and strand transfer reactions separately. What we find, however, is that ELT substrates with mutations at their termini cannot participate efficiently in either the donor cleavage or strand transfer reactions (Chapter 4). This result is consistent with a mechanistic linkage of these reactions as suggested by the studies done on Mu and retroviruses (see Chapters 1 and 4). We also believe that the SEJ species seen in Tn7 transposition are consistent with the strand joining mechanism proposed for Mu and retroviruses (Chapter 4). These data suggest that the mechanism of Tn7 transposition is in many ways similar to that in the Mu and retroviral transposition systems.

Not all aspects of Tn7 transpositional recombination are similar to that of the purified Mu and retroviral systems. The purified Mu transposition system does not make breaks at the 5' end of the transposon; it stops at the strand transfer intermediate stage (see Thesis Introduction). Purified IN protein, the transposase of the retroviral transposition system, also does not process the transposon 5' end; instead only the 3' transposon end is processed (see Thesis Introduction). However, in the purified Tn7 and Tn10 transposition systems, a break is made early on at the 5' donor-transposon

junction to release the transposon from the donor backbone (Chapter 3; Benjamin et al, 1989). The Tn10 transposition system makes flush cuts to release the transposon. In Tn7 transposition, the donor release cuts are staggered, (i.e. the 3' cut is flush with the end of the transposon), and the 5' cut is three bases out into the donor DNA. In vivo observations suggest that Tn7 can transpose from many different donor sites, i.e. sites with different sequence flanking the transposon, with no apparent effect on the efficiency of transposition (Arciszewska et al, 1989). This suggests that the 5' breaks can be made without regard to primary sequence. Thus, the Tn7 transposition machinery can identify and break both specific DNA sequences at the 3' end and apparently non-specific DNA sequences at the 5' end of the transposon.

That Tn7 requires two different breakage activities to release it from the donor backbone must be incorporated into any unified transpositional recombination mechanism. Is the mechanism of the 5' breakage related to the adenosine-centered transpositional recombination at the 3' end? If the nucleotide sequence at the 5' end is unimportant, how is the break specified so precisely? Are the 3' and 5' breaks made together by a single mechanism or can they made at different times by different mechanisms? Increased resolution of the order and structure of these associated breaks should reveal the mechanistic relationship of Tn7 donor release to transpositional recombination. These questions are now within the reach of the purified Tn7 transposition system.

## Perspectives on the Nucleoprotein Complex Model of Tn7 Transposition

The role and structure of the proposed Tn7 nucleoprotein complex that contains both the donor and target DNA is still an open question. I discuss below several ways in which a nucleoprotein complex that contains all of the Tns proteins and substrates may influence the transposition reaction. These include ways of influencing the initiation of

recombination, the target-site specificity of recombination, the role of ATP in nucleoprotein complex assembly, and how the transposition machinery differentiates between the transposon ends.

The Nucleoprotein Complex and the Initiation of Transposition: A Role for Target Site Recognition

The initiation of transpositional recombination begins with the donor cleavage reaction (see Thesis Introduction). Is the initiation of this step under the influence of the target site? Retroviruses need not have any association with target DNA to initiate the trimming reaction, i.e. the donor cleavage reaction (Fujiwara and Mizuuchi, 1988; Brown et al, 1989; Craigie et, 1990). However, by its very nature, the retroviral integration intermediate is free entity, untethered to a donor molecule, and therefore is not likely to have the same constraints on the initiation of recombination as systems that must cut the transposon out of a particular site in the chromosome.

In Mu transposition in vitro, the initiation of transpositional recombination can occur prior to the recognition of a target site, i.e. the formation of the cleaved-donor complex (Craigie et al, 1985; Surette et al, 1987). However, the initiation of recombination in the Mu system can be influenced by the target site. MuB preferentially associates with DNA that does not contain transposon end sequences (Adzuma and Mizuuchi, 1988). MuB can stimulate the donor cleavage reaction by interacting with transposon-bound MuA (Baker et al, 1991). Thus, it reasonable to suggest that although recognition of a target site is not absolutely required for the initiation of the donor cleavage reaction, it can have some influence over when initiation occurs in Mu recombination (Mizuuchi, in prep.).

Tn7 is quite strict about the initiation of transpositional recombination. All the Tns proteins and both DNA substrates must be together before any reaction is begun (Chapter 3). Under some in vitro conditions, a TnsC-TnsD attTn7 site complex must be

formed and recognized before the process of transposition can begin (Chapter 3). In vivo experiments also suggest that Tn7 needs the attTn7 site to initiate the transposition reaction (Anne Stellwagen, pers. com.). These data demonstrate that target site recognition can predominate as a controlling mechanism for initiation of some transposition reactions. We suggest that in Tn7's case, this occurs by the conjunction of the target site and transposon ends in a nucleoprotein complex. Thus, it seems that target site recognition can have wide spectrum of influence as an effector on the initiation of transposition and that Tn7 is at one extreme of that spectrum.

A Nucleoprotein Complex Provides Specificity to the Reaction

The above view of Tn7 nucleoprotein complex assembly suggests that the formation of a specific target site complex could be a necessary step on the pathway of transposition and thus be a way of controlling where and when transposition events will occur. We know that a very specific complex can be formed at the attTn7 site, dependent on ATP, and consisting of TnsD and TnsC (Chapter 3). We suggest that this complex can recruit other transposition components to the target site largely due to the action of TnsC (discussed above and in Chapter 3). If, as we suggest TnsC, is the pivotal communication protein between the target and the donor site machinery, and target site recognition is required before the breakage and joining reactions can begin, then the formation of a specific target site complex is the crucial step in determining where insertions are going to occur in Tn7 transposition.

TnsC may function as site-selectivity factor in some ways analogous to MuB in Mu transposition. MuB displays two important properties in choosing target sites: 1) it can stimulate donor cleavage in a complex that contains the Mu transposons ends held together by MuA (see Thesis Introduction) and 2) in the presence of MuA, MuB associates selectively with DNA molecules that do not contain Mu end sequences, i.e. causing transposition immunity (see Thesis Introduction). These two properties

together explain why MuB plays such an important role in determining where Mu insertions are going to occur, even though MuB itself does not contain the machinery necessary to carry out transpositional recombination.

We suggest that TnsC may play a similar role in Tn7 transposition because TnsC has several similar phenomonological properties to MuB: TnsC is an ATP-dependent DNA binding protein (Gamas and Craig, in press.), and TnsC seems to be involved in transposition immunity (Chapter 3), TnsC is part of the central transposition machinery (Chapter 3), and TnsC may stimulate the transpositional machinery to accomplish the Tn7 donor cleavage and strand transfer reactions, i.e. the formation of double-strand breaks, particularly in the presence of AMP-PNP (Chapter 3). TnsC also has an additional property: it associates with the attTn7 site in a TnsD-dependent manner (Chapter 3). Whether by associating directly with TnsD or by binding to a conformational change in the DNA brought about by TnsD, TnsC can site-selectively associate with the attTn7 target site. That TnsC appears to be associated with the target DNA suggests a model for Tn7 transposition target site-specificity. If TnsC is the protein that recruits the catalytic machinery necessary for transpositional recombination to the target site, then it is reasonable to suggest the binding of TnsC to target DNA will in large measure determine where insertions are going to occur. If the Tn7 transposition machinery can control where TnsC binds, then it can control where transposition will occur.

The interaction of protein complexes involved with the initiation of transcription at RNA polymerase II promoters makes an interesting comparison to target site complexes in Tn7 transposition (Greenblatt, 1991). TFIID is a site-specific DNA binding complex that binds the TATA box at RNA pol. II promoters (Greenblatt, 1991). When TFIID binds to a promoter, it is recognized and bound by TFIIB, a non-specific DNA-binding protein complex (Buratowski, 1989). The recruitment of TFIIB to this site is also under the influence of other enhancer-related transcription factors, allowing

multiple inputs to influence stable complex formation. With TFIIB now stably bound, the complex can recruit RNA polymerase II, the molecule responsible for the catalytic activity of transcription (Greenblat, 1991). Perhaps a target-associated nucleoprotein complex is built in a similar manner in Tn7 transposition. The site-specificity of the complex is designated by a protein that binds DNA site-specifically, i.e. TnsD. A non-site specific DNA binding protein can associate with this specific protein and become bound to a particular site, i.e. TnsC. Then the recruitment of the catalytic machinery responsible for transposition is accomplished by the stable association of TnsC with target DNA. In such a model the machinery can be directed to or away from particular sites by controlling where the non-specific DNA protein is bound to DNA.

The Role of ATP in the Formation of a Nucleoprotein Complex

We suggest that ATP plays a role in the formation of a target site complex in Tn7 transposition. TnsC is an ATP-dependent DNA binding protein whose DNA binding is inhibited by the presence of ADP (Gamas and Craig, in press). We have already suggested that TnsC communicates with both target sites and the rest of the transposition machinery (Chapter 3). Perhaps TnsC uses ATP hydrolysis as a way of coordinating information about a target site. If so, TnsC uses this information in two distinct ways: 1) to prevent its interaction with inappropriate target sites, i.e. transposition immunity, and 2) to direct the transposition machinery to particular sites, i.e. its interaction with TnsD. A scenario which unifies these two properties is as follows.

The best analogy for the role of TnsC in transposition immunity is that of MuB in the Mu system (Adzuma and Mizuuchi, 1988; Adzuma and Mizuuchi, 1991). MuB + ATP, MuB-ATP, has high affinity for DNA; MuB-ADP has low affinity for DNA. When in high local concentration, MuA stimulates the hydrolysis activity of MuB, causing MuB to fall off the DNA. This causes MuB to selectively associate with DNA molecules that do not have Mu end sequences and, in turn, directs the Mu transposition machinery to sites

which do not contain Mu ends (see Thesis Introduction and Chapter 3, discussion). A parallel model can be proposed for Tn7: TnsC + ATP has high affinity for DNA and TnsC + ADP has low affinity for DNA (Gamas and Craig, in press). Perhaps TnsB and/or TnsA bound to the ends of Tn7 stimulates the hydrolysis of ATP by TnsC-ATP producing TnsC-ADP, which then comes off the DNA. In this way, TnsC would tend to associate with targets that do not contain transposon end sequences, i.e. immune targets, and thus TnsC would direct the transposition machinery only to those site bound by it (as described above and see Chapter 3).

TnsC's DNA binding properties can also be used in a model describing target site selectivity. In addition to the binding properties discussed above, we suggest that TnsC might possess an ATP hydrolysis activity which can be used to coordinate its binding to special DNA sites. This hydrolysis could be spontaneous, i.e. inherent to the TnsC polypeptide, or it could be stimulated by another Tns protein, perhaps TnsB or TnsA. The important point in this model is that TnsC hydrolyzes ATP at a rate high enough to be used as part of the mechanism that directs it to particular DNA sites. Thus, in the absence of TnsD, TnsC-ATP binds to DNA, it hydrolyzes ATP to become TnsC-ADP, and it comes off the DNA. Under these conditions, TnsC cannot stably bind a target and can not therefore promote recombination. However, if TnsC's hydrolysis is inhibited by a cognate interaction with TnsD, then it may stably bind to the DNA, but only in places where TnsD is stably bound, presumably the attTn7 site. In this manner, a hydrolysis activity built into the transposition system may encourage TnsC to bind DNA stably only in the presence of TnsD, causing it to associate with the attTn7 site.

Evidence supportive of the above model comes from the use of non-hydrolyzable analogues of ATP in transposition. ATP analogues disrupt the phenomenon of target site immunity in Mu and Tn7 transposition (Adzuma and Mizuuchi, 1988; Chapter 3). In addition, ATP analogues promote Tn7 transposition in the absence of TnsD (Chapter 3). One explanation for these phenomena is that TnsC can now stably bind to the DNA and

promote target complex formation without the assistance of TnsD. That the Tn7 transposition reaction does not require ATP hydrolysis to promote recombination is not surprising. Tn10 and retroviruses require no nucleotide cofactor to promote transposition (Brown et al, 1987; Benjamin et al, 1989), and Mu transposition is also efficient in the presence of non-hydrolyzable analogues of ATP (Adzuma and Mizuuchi, 1988). Thus, ATP is not necessary for the breakage and joining reactions involved in transposition (see Thesis Introduction; Mizuuchi review). We therefore suggest that ATP and ATP hydrolysis allows TnsC to play a central role in target site choice, but that ATP does not participate directly in the recombinational mechanism. In summary, this model suggests that TnsC chooses targets by associating with the target DNA. This association is influenced in a number of different ways: 1) by interacting with TnsD, 2) by participating in transposition immunity, 3) by interacting with the donor associated transposition machinery, and 4) perhaps by interacting with TnsE (see below). Thus a target-associated, nucleoprotein complex assembly suggests that TnsC's use of ATP in Tn7 transposition may be a focal point for the gathering of information about the target sites and may play a critical role in promoting a high degree of target site selectivity in the Tn7 transposition reaction.

A Nucleoprotein Complex Differentiates Between the Ends of the Transposon

Although it is clear that the ends of Tn7 "see" the target site in some way, perhaps in a nucleoprotein complex, it is not clear how they do so. The Tn7 ends might find each other before seeing the target, or they might find the target site in a particular ordered manner. However, at some point in the Tn7 donor-target recognition process, there must be a differentiation between the right and left ends of Tn7 because its orientation of insertion is highly specific (Chapter 1). Under the in vitro transposition reaction conditions with purified proteins, 98% of the insertions are in the most favored orientation with only 2% in the opposite orientation (unpublished observation).

Opposite orientation insertions are uncommon but not forbidden in vivo (Owen Hughes, pers. com.). We suggest that orientation-specificity is established before any of the breakage and joining reactions are accomplished. We base this suggestion on the observation that in all cases where a single end of the transposon is recombinationally-processed independent of the other end, i.e. double strand breaks and single end joins, both right and left ended species are found in equal quantities (Chapter 2 and Chapter 4). In other words, it seems that the transposon ends are broken and joined with the same frequency, once the decision to initiate transposition has been made. Thus, the left and right end distinctions must be made prior to the initiation of transposition. This idea is corroborated by experiments with ELT substrates that contain mutations at their termini (Chapter 4). The products of transposition when using these substrates show directly that the transposition machinery can make a distinction between the left and right transposon ends that is normally concealed when the transposon is unbroken (Chapter 4).

How then is this distinction between the ends made with an intact donor? We suggest that steps leading up to the donor release reaction are where transposition reaction orientation decisions are made. We further suggest that this function is accomplished by a nucleoprotein complex that contains all the substrates, particularly both ends of the transposon, in a specific relative configuration. If while assembling the complex, a configuration that positions one end of the transposon in a particular orientation relative to the target site is favored over the other, then a larger proportion of complexes will form in one orientation over the other. We propose that as the complex proceeds towards the donor release step, the ends of the transposon communicate with one another and that this communication readies the complex for the donor release reaction (Chapter 4). However, once the complex is ready for the donor release step, it no longer discriminates between the ends and thus breaks them with equal frequency, producing the ELT. The integrity of the nucleoprotein complex prevents

the dissociation of the excised transposon and thereby assures that it will maintain the same orientation with which it began the process.

The assembly of this complex into its proper relative configuration may be in some ways analogous to the stepwise formation of recombination complexes in the Tn3 resolvase reaction (Boocock et al, 1987). The proper relative configuration of the res sites is essential to the recombination reaction that promotes resolution of cointegrate species rather than inversions of the intervening DNA. The stepwise assembly of the resolvase complex allows the reaction to increase its precision because at multiple places in the process of assembly formation of the opposite orientation complex is so energetically unfavorable that it never gets to the stage necessary to begin the recombination reaction (Craigie and Mizuuchi, 1986; Gellert and Nash, 1987). Tn7 nucleoprotein assembly is clearly not as precise as that of the Tn3 resolvase system, but it may use a similar stepwise assembly to maintain a high level of precision in the orientation of insertion and in site-selection.

## Generalized and Specialized Roles For the Tns Proteins Allow Generalized and Specialized Pathways of Transposition

### Roles for the Tns Proteins

*The Central Transpositional Machine*

In the well-characterized in vitro transposition systems, Mu, retroviruses and Tn10, the transposase activity is clearly contained within a single polypeptide (see Thesis Introduction). Tn7 has a large number of transposition proteins, and transposes in a highly concerted manner, i.e. it requires all the transposition components to promote the reaction. Indeed Tn7 transposition is so concerted that it has been difficult to produce any recombination events when either TnsA, TnsB, TnsC or TnsD is absent (Chapter 3). However, we feel confident that TnsD does not contain or share the catalytic site for recombination because it is dispensable for transposition when AMP-

PNP is used in the reaction (Chapter 3). This observation is consistent with in vivo data that suggests that tnsA + tnsB + tnsC are required for both transposition pathways (Thesis Introduction). Thus, the designation of TnsA + TnsB + TnsC as the central transpositional machine seems reasonable (see above).

In this view, TnsA + TnsB + TnsC play generalized roles in Tn7 transposition and TnsD and TnsE play roles as activation and specificity determinants. The TnsA + TnsB + TnsC + TnsE pathway promotes transposition to many different sites with apparently random specificity (Waddell and Craig, 1988; Rogers et al, 1986; Kubo and Craig, 1990). This pathway may require a target site complex similar to that suggested for the TnsA + TnsB + TnsC + TnsD pathway. Two models emerge as possible explanations for the alternative interactions of TnsA + TnsB + TnsC with TnsD and TnsE and with their substrates. Model 1 suggests that TnsE is analogous to TnsD, in that TnsE is a target DNA binding protein which directs TnsC to sit down on the DNA. Perhaps TnsE binds the DNA non-specifically and stabilizes the association of TnsC with the target. In such a model, TnsC might interact with TnsD and TnsE in a similar manner, suggesting a conserved domain in the two target specific proteins. No such domain has yet been observed however (Flores et al, 1990).

Model 2 suggests that instead of interacting directly with the target, TnsE stimulates TnsC or a complex of Tns proteins including TnsC to bind the target DNA. TnsC has already been suggested to be a protein that communicates between target sites and donor sites (Chapter 3 and see above). Perhaps TnsE can alter the binding affinity of TnsC for DNA without itself touching the DNA. In this model, TnsE allosterically affects TnsC's DNA binding, thereby promoting insertions to non-specific DNA sites through TnsC.

*Which Tns Protein Contains the Catalytic Activity Necessary for Transposition?*

Attempts to define which Tns proteins contain the transpositional recombination activity have been difficult. Neither TnsA, TnsB nor TnsC, singly or in pairwise combination can promote any noticeable recombination (unpublished observation). Further experimentation is needed to discover in which protein or proteins the catalytic site resides. The properties of the individual Tns proteins and comparison to other transposition systems can suggest ways of thinking about this problem.

TnsB is known to bind repeated sequence elements in the ends of Tn7 (Arciszewska et al 1991; Tang et al, 1991). We suggest that this protein is likely to play a role in bringing the ends of the transposon together, analogous to MuA in Mu transposition (see Thesis Introduction). TnsB has not been demonstrated to contain any nucleolytic activity (Arciszewska et al, 1991). However, its role as a transposon end binding protein suggests that, at the very least, it plays a role in positioning the transposase on the ends of the transposon and perhaps participates catalytically in the recombination reaction.

TnsC plays a targeting role similar in some ways to MuB in Mu transposition (see above). A direct comparison to the Mu system would suggest that the recombination activity of Tn7 lies not in TnsC, but in TnsB, an end-binding protein, and/or perhaps TnsA, a protein for which no specific function has yet been found. A potentially illuminating experiment would be to expose ELT donors to subsets of purified Tns proteins. An ELT donor simplifies the recombination reaction by asking the Tns proteins to accomplish only strand transfer and not donor cleavage. Perhaps such an approach could uncover the transpositional recombination activity among or between TnsA, TnsB, and TnsC.

Other roles for the Tns proteins are numerous. One possibility is that a Tns protein could act as a linker, bridging a target site complex, such as TnsC-TnsD-attTn7, and a donor complex, such as TnsB bound to the ends of Tn7. Perhaps TnsA plays the role

of linking these two complexes in manner analogous to adaptor proteins in transcription complex formation (Martin, 1991). These proteins are thought to act as accessory factors that help protein complexes bound at different DNA sites communicate with one another, thus promoting transcription complex assembly. The conjunction of multiple specific DNA sites in Tn7 transposition and transcription initiation have similar mechanistic problems to overcome, thus if their components and those of other nucleic acid manipulation reactions share similar functions; it would not be surprising.

The Resolution of the Transposition Event

The predominant products of Tn7 transposition in vitro are simple insertions, i.e. the transposon is inserted into the target site without any connection to the donor backbone (Chapter 2). An insertion of a Tn7 element leaves single-stranded gaps at the transposon-target junctions and three bases of donor backbone DNA hanging off the 5' ends of the transposon (Chapter 2). Under the crude transposition conditions, these breaks and overhangs are repaired in vitro at high frequency; about 95% of the right and about 50% of the left transposon-attTn7 site junctions are repaired to duplex DNA (Chapter 2). That the right junction is repaired at higher frequency than the left junction suggests that the sequences around the junction have influence over the efficiency of repair (Chapter 2). Whether this is due to the transposon end sequences, the attTn7 target sequences, or a combination of both is unclear.

In the Mu in vitro system with added crude extract, repair of the strand transfer intermediate occurs either by completely replicating the element to form a cointegrate or by cutting away the donor backbone to form a simple insertion product (Mizuuchi, 1984). How Mu makes the decision to choose one pathway or the other, and whether one end is favored over the other for repair, is unknown. Perhaps the transposition machinery bound to the transposon ends and/or target sites plays some role in recruiting the replication machinery. The site-specificity of Tn7 transposition offers a

unique opportunity to study the effects of both transposon end sequences and specific target site sequences on the repair process in transposition. Future studies may demonstrate important roles and interactions of host proteins and Tns proteins in the process of repair subsequent to transpositional recombination.

# Appendix

## Notes on Purified TnsA: Specific Activity and the Effect of the Addition of Crude Extract to the Reconstituted System

## Introduction

The following is a description of experiments concerning TnsA. The purified TnsA fraction used in the experiments in Chapters 3, 4, and 5 was derived from a GST-TnsA fusion protein by isolation of the fusion protein by affinity chromatography using glutathione agarose, release of TnsA by cleavage with thrombin and removal of the released GST domain as described in detail in Chapter 3, Experimental Procedures. How does the activity of this released TnsA compare with the activity of authentic TnsA? Although this question has not yet been examined in detail, the preliminary experiments described below suggest that the specific activity in transposition of released TnsA is not substantially different from, (i.e. is within several fold of), the specific activity of authentic TnsA present in crude extracts.

## Results

### The Amount of TnsA in a Crude Extract

Crude TnsA is overproduced by the pKAO52 plasmid (Orle and Craig, 1990) in an htpr$^-$ background (CAG 456, Baker et al 1984) . Based on several observations, we estimate that the concentration of TnsA present in crude extracts made from this strain to be about 30 to 100 µg/ml in 20 mg/ml total protein. In one series of experiments, we evaluated the amount of TnsA in crude extracts by comparing with immunoblots of these extracts to known amounts of purified GST-TnsA. These experiments suggest that the concentration of TnsA is between 50-100ug/ml. In another evaluation, we compared by immunoblotting extracts containing either TnsA or a GST-TnsA fusion containing about the amino terminal two thirds of TnsA, GST-2/3TnsA. Each extract displayed similar immunoreactivity. We estimated the actual amount of GST-2/3TnsA to be between 30 to 100ug/ml based on comparison of this species with known amounts of BSA after Coomassie blue staining. These observations suggest that TnsA is about 60ug/ml,

i.e. about 0.3% of the total protein in our crude extracts, a range consistent with the fact that TnsA is not discernable in crude extracts evaluated by Coomassie blue staining.

**Comparison of the Specific Activity in Transposition of Authentic TnsA and Released TnsA in the Presence of Crude Extract**

How does the activity of the purified, released TnsA compare to the activity of authentic TnsA? In particular, has the transposition activity of released TnsA been impaired by its synthesis as a fusion protein or by its subsequent purification? To address this question, we compared the amount of transposition promoted by crude extracts containing authentic TnsA and reconstructed crude extracts to which released TnsA was added. In both cases, the total amount of crude extract in the reactions was held constant and the amount of TnsA varied by addition to each reaction of a constant amount of TnsA crude extract diluted with crude extract prepared from cells lacking TnsA; each 100µl reaction contained 1µl of crude extract. Reactions were performed as described in Chapter 3 in the absence of PVA using a 7 minute preincubation prior to the addition of MgAc; after MgAc addition, the incubations were continued for 30 minutes. The reaction products were electrophoresed without restriction enzyme digestion and without treatment to remove proteins on a 0.6% agarose gel, the gel stained with EtBr and photographed.

Shown in Figure 1 are titrations with crude extracts containing released TnsA (lanes 1-4) and authentic TnsA (lanes 5-8). The transposition product evident in these reactions is a nicked circle simple insertion (4.5kb) which is fuzzy in appearance because the samples were not deproteinized; the donor backbone (4.5kb) is not obvious here, likely because it is degraded in the presence of crude extract. With released TnsA, transposition product is readily detectable with 15 ng TnsA (lane 1) and barely so with 5 ng (lane 2); 15 ng released TnsA appears to saturate the reaction, as no increase in the amount of transposition product is observed with increasing TnsA (data not shown).

The amount of recombination observed with 15ng released TnsA (lane 1) is comparable to that observed with 18ng authentic TnsA (lane 5), estimated by assuming that the crude extract contains 60ug/ml authentic TnsA. The very similar amounts of transposition observed with comparable amounts of released and authentic TnsA suggest that the transposition activity of the released TnsA has not been significantly impaired by its preparation and purification.

**Is Transposition Stimulated by the Presence of Crude Extract?**

An important issue in Tn7 transposition is to establish whether host proteins participate directly in recombination. We have compared the amount of recombination promoted by TnsA, TnsB, TnsC and TnsD in the presence of crude extract (as in Figure 1) with recombination promoted by Tns proteins in the absence of host extract (Figure 2). Transposition reactions containing varying amounts of TnsA and constant amounts of TnsB, TnsC and TnsD were performed as described in Chapter 3 in the absence of PVA using a 7 minute preincubation prior to the addition of MgAc; after MgAc addition, the incubations were continued for 30 minutes. The reaction products were deproteinized, digested with ScaI, electrophoresed on a 0.6% agarose gel, the gel stained with EtBr and photographed. The simple insertion product is a 4.5 kb linear species and the donor backbone is present as two species, one 2.1kb and the other 2.4kb. Recombination is readily detectable with 25ng released TnsA (lane 1); when released TnsA is reduced to 9ng (lane 2), the products are barely detectable using Southern hybridization (data not shown). The observation that only slightly higher amounts of released TnsA are required in the absence of crude extract (Figure 2) to give comparable recombination to that observed in the presence of crude extract (Figure 1) suggests that recombination is not severely limited by the lack of a critical host protein in the reconstituted system that uses four highly purified Tns fractions. If a host protein is essential to Tn7 recombination, our observations suggest that it must be present in non-limiting

amounts in the Tns fractions. Although significant amounts of host proteins are not obvious in these fractions, the possibility that they contain a host protein not readily detectable by the methods we have used cannot be dismissed. A reasonable and not improbable view is that the Tns proteins themselves participate directly in recombination, contain the specificity determinants for recombination as well as the catalytic site that execute breakage and joining, and that host proteins may, under some conditions, play accessory, stimulatory roles.

It is important to note that although Tn7 does not appear to require a host protein for recombination, in vitro transposition reactions are stimulated by the addition of crude extract particularly under short preincubation times (data not shown). Longer preincubation times (> 30 minutes) can compensate for the lack of host extract in the in vitro reaction, i.e. in Chapters 3 and 4 the in vitro transposition reactions are highly efficient. This may suggest that host proteins can facilitate the formation of the TnsC + TnsD + attTn7 complex (see Chapter 3), but that their participation in the formation of such a complex is not mandatory.

## Experimental Protocol Appendix (TnsA Specific Activity) In Vitro Transposition Reactions

The reaction conditions of the experiments in this appendix are the same as in Chapter 3 except that the purified TnsA is diluted in crude extract. In Figure 1A, the purified TnsA is diluted at 1/3 intervals in an NLC51 crude extract. The NLC51 crude extract is prepared as in Chapter 2 and its protein concentration is 20mg/ml. 1 µl of each dilution is added to each reaction, thus the total amount of crude extract added to the reaction is 1 µl. In figure 1B, pTacA* htpr⁻ (pKAO52 in a CAG 456 background) (Orle and Craig, 1990; Baker et al, 1984) is diluted in the NLC51 crude extract as described above and 1 ml of the dilution is added to the reaction.

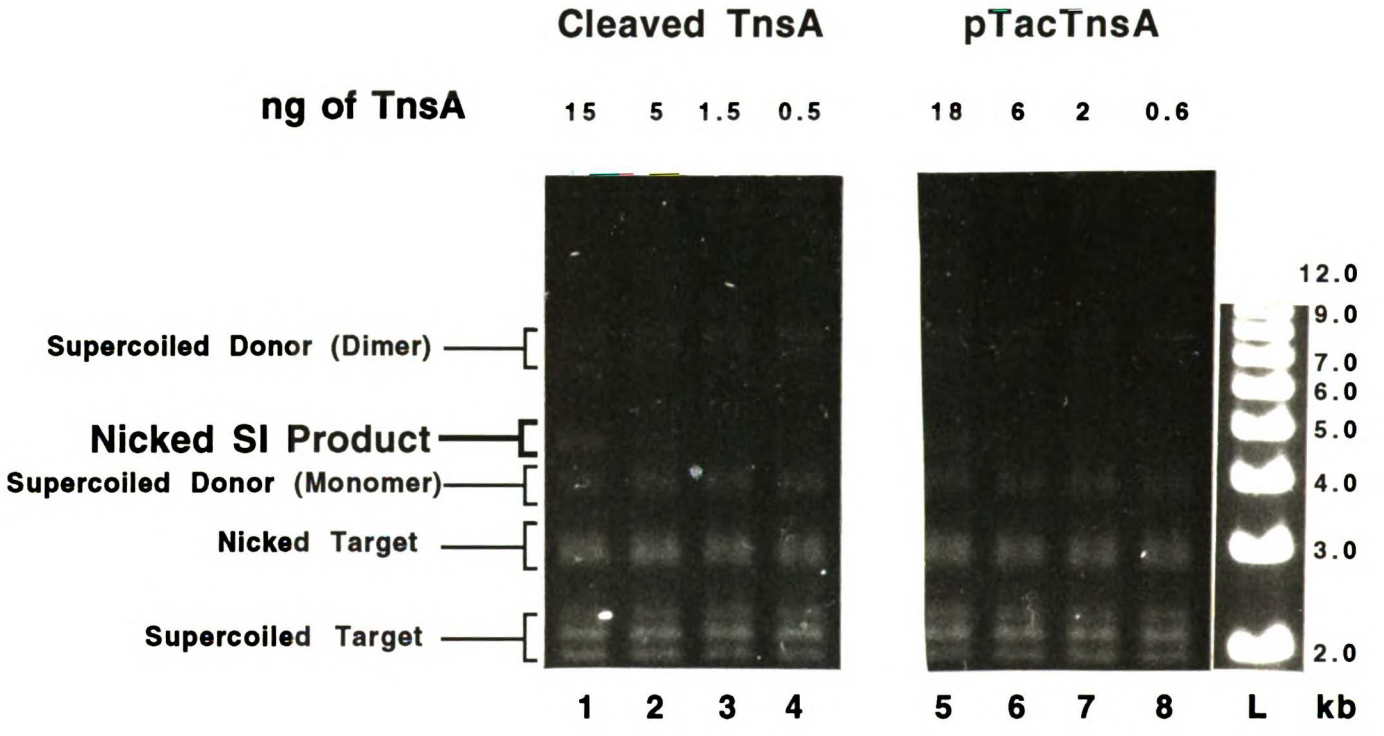
The reaction conditions in Figure 2 are the same as in Chapter 3. In this case, the TnsA is diluted in TnsA dialysis buffer plus 50 µg/ml BSA (Chapter 3).

### Processing the In Vitro Transposition Reaction Products

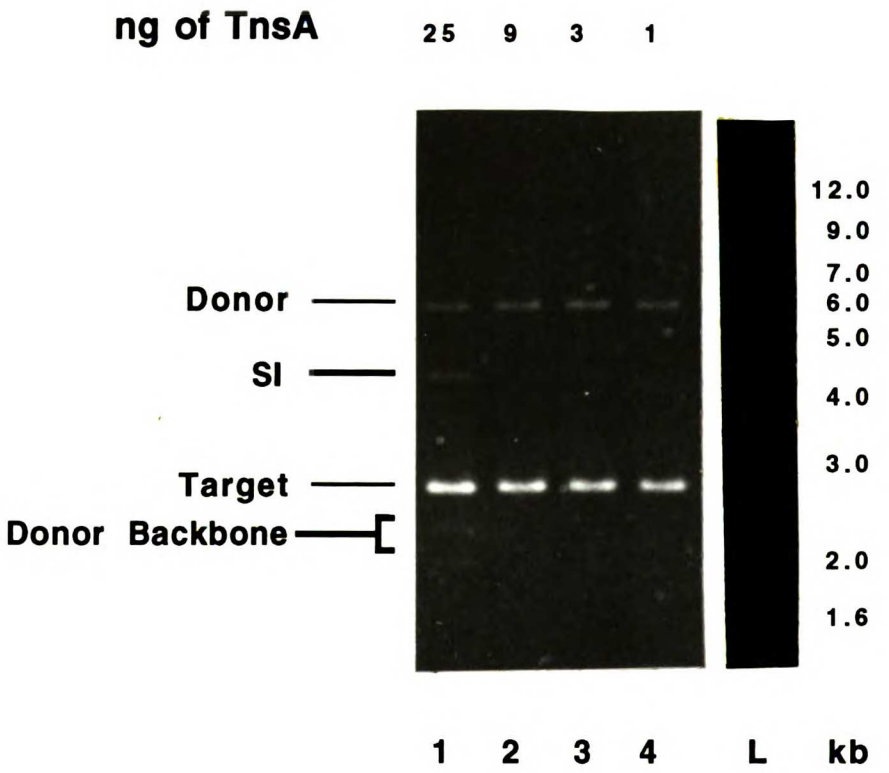
In Figure 1 the transposition reactions are stopped by adding EDTA to 20 mM and cooling the reactions to 4°C. 20 µl of each reaction is mixed with 4µl 6X agarose gel loading buffer (Chapter 2). The samples are loaded directly onto an 0.6% agarose-1X TBE gel and run for 600 volt-hrs. The gel was stained with EtBr and photographed.

In Figure 2 the transposition reactions were stopped by the Urea-spermine precipitation procedure (Chapter 2). One quarter of the sample was digested with 10 units of ScaI and loaded onto a 0.6% agarose-1X TBE gel and run for 600 volt-hrs. The gel was stained with EtBr and photographed.

# Figure 1 Titration of TnsA in C.E.



# Figure 2 Titration of TnsA Without C.E.



# References

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

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

Adzuma, K., & Mizuuchi, K. (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., & Mizuuchi, K. (1991). Steady-state Kinetic Analysis of ATP Hydrolysis by the B Protein of Bacteriophage Mu. J.B.C., 266: 6159-6167.

Akroyd, J., & Symonds, N. (1983). Evidence for a Conservative Pathway of Transposition of Bacteriophage Mu. Nature, 303: 84-86.

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

Arciszewska, L. K., & Craig, N. L. (1991). Interaction of the Tn7-encoded transposition protein TnsB with the ends of the transposon. N.A.R., 19: 5021-5029.

Arciszewska, L. K., McKown, R. L., & Craig, N. L. (1991). Purification of TnsB, a transposition Protein that binds to the ends of Tn7. J.B.C., 266: 21736-21744.

Arthur, A., & Sherratt, D. (1979). Dissection of the Transpositional Process: a transposon-encoded site-specific recombination system. Mol. Gen. Genet., 175: 267-274.

Avery, O. T., MacCleod, C. M., & McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a deoxyribonucleic acid fraction isolated form pneumococcus type III. J. Exp. Med., 79: 137-158.

Bainton, R. B., Gamas, P., & Craig, N. L. (1991). Tn7 Transposition In Vitro Proceeds through an Excised Transposon Intermediate Generated by Staggered Breaks in DNA. Cell, 65: 805-816.

Baker, T. A., Grossman, A. D., & Gross, C. A. (1984). A gene regulating the heat shock response in Escherichia coli also affects proteolysis. P.N.A.S., 81: 6779-6783.

Baker, T. A., Mizuuchi, M., & Mizuuchi, K. (1991). MuB Protein Allosterically Activiates Strand transfer by the Transposase of Phage Mu. Cell, 65: 1003-1013.

Barth, P. T., Datta, N., Hedges, R. W., & Grinter, N. J. (1976). Transposition of a Deoxynucleotide Acid Sequence Encoding Trimethroprim and Streptomycin Resistances from R483 to Other Replicons. J. Bacteriol., 125: 800-810.

Bender, J., & Kleckner, N. (1986). Genetic evidence that Tn10 transposes by a nonreplicative mechanism. Cell, 45: 801-15.

Bender, J., Kuo, J., & Kleckner, N. (1991). Genetic Evidence Against Intramolecular Rejoining of the Donor DNA Molecule Following IS10 Transposition. Genetics, 128: 687-694.

Benjamin, H. W., & Kleckner, N. (1989). Intramolecular Transposition by Tn10. Cell, 59: 373-383.

Berg, D. E. (1983). Structural requirements for IS50-mediated gene transposition. PNAS, 80: 792-96.

Berg, C., & Berg, D. (1987). Uses of transposable elements and maps of known insertions. Escheria coli and Salmonella typhimurium: Cellular and Molecular Biology (pp. p.1071-1109). In F. Neidhardt, J. Ingraham, K. Low, B. Magasanik, M. Schaechter, and H. Umbarger (ed.),.

Berg, D. E., & Howe, M. M. (1989). Mobile DNA . Washington D.C.: American Society for Microbiology.

Boeke, J. D. (1989). Transposable elements in Saccharomyces cerervisiae. In D. Berg, & M. M. Howe (Ed.), Mobile DNA (pp. 335-374). Washington, D.C.: American Society for Microbiology.

Boocock, M. R., Brown, J. L., & Sherratt, D. J. (1987). Topological specificty in Tn3 resolvase catalysis. In R. McMacken and T.J. Kelly (Ed.), In DNA Replication and Recombination. UCLA Symposia of Molecular and Cellular Biology New Series (pp. 703-718). (New York: Alan R. Liss, Inc.).

Bowerman, B., Brown, P. O., Bishop, J. M., & Varmus, H. E. (1989). A nucleoprotein complex mediates the integration of retroviral DNA. Genes Dev, 3: 469-78.

Brown, P. O., Bowerman, B., Varmus, H. E., & Bishop, J. M. (1987). Correct Integration of Retroviral DNA In Vitro. Cell, 49: 347-356.

Brown, P. O., Bowerman, B., Varmus, H. E., & Bishop, J. M. (1989). Retroviral integration: Structure of the initial covalent product and its precursor, and a role for the viral IN protein. PNAS, 86: 2525-2529.

Buratowski, S., Hahn, S., Guarente, L., & Sharp, P. A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. Cell, 56: 549-561.

Casadaban, M. J. (1976). Transposition and fusion of the lac genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu. J Mol Biol, 104: 541-55.

Cech, T. R. (1990). Self-Splicing of Group I Introns. Annu. Rev. Biochem., 59: 543-68.

Chaconas, G., Gloor, G., & Miller, J. L. (1985). Amplification and purification of the bacteriophage Mu encoded B transposition protein. JBC, 260: 2662-9.

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

Chow, S. A., Vincient, K. A., Ellison, V., & Brown, P. O. (1992). Reversal of Integrationand DNA Splicing Mediated by Integrase of Human Immunodeficiency Virus. Science, 255: 723.

Craig, N., & Kleckner, N. (1987). Transposition and site-specific recombination. In J. I. In F. Neidhardt K. Low, B. Magasanik, M. Schaechter, and H. Umbarger (Ed.), Escheria coli and Salmonella typhimurium: Cellular and Molecular Biology (pp. p.1054-1070).

Craig, N. L. (1988). The Mechanism of Conservative Site-Specific Recombination. Annu. Rev. Genet., 22:: 77-105.

Craig, N. L. (1989). Tn7 Transposition. In D. Berg, & M. M. Howe (Ed.), Mobile DNA (pp. 211-225). Washington, D.C.: American Society for Microbiology.

Craig, N. L. (1990). P Element Transposition. Cell, 62: 399-402.

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

Craigie, R., Arndt-Jovin, D. J., & Mizuuchi, K. (1985). A defined system for the DNA strand-transfer reaction at the initiation of bacteriophage Mu transposition: protein and DNA substrate requirements. P.N.A.S., 82: 7570-4.

Craigie, R., & Mizuuchi, K. (1985). Mechanism of transposition of bacteriophage Mu: structure of a transposition intermediate. Cell, 41: 867-76.

Craigie, R., & Mizuuchi, K. (1986). Role of DNA Topology in Mu Transposition: Mechanism of Sensing the Relative Orientation of Two DNA Segments. Cell, 45: 793-800.

Craigie, R., & Mizuuchi, K. (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., Fujiwara, T., & Bushman, F. (1990). The IN Protein of Moloney Murine Leukemia Virus Processes the Viral DNA Ends and Accomplishes Their Integration In Vitro. Cell, 62: 829-837.

Derbyshire, & Grindley, N. (1986). Replicative and Conservative Tansposition in Bacteria. Cell, 47: 325.

Drlica, K., & Rouviere-Yaniv, J. (1987). Histone-like proteins of bacteria. Microbiol. Rev., 51: 301-319.

Echols, H. (1990). Nucleoprotein structures initiating DNA replication, transcription, and site-specific recombination. J.B.C., 265: 14697-14700.

Eichinger, D. J., & Boeke, J. D. (1988). The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition. Cell, 54: 955-66.

Eichinger, D. J., & Boeke, J. D. (1990). A specific terminal structure is required for Ty1 transposition. Genes Dev, 4: 324-30.

Engelman, A., Mizuuchi, K., & Craigie, R. (1991). HIV-1 DNA Integration: Mechanism of Viral DNA Cleavage and DNA Strand Transfer. Cell, 67: 1211-1221.

Engels, W. R., Johnson-Schlitz, D. M., Eggleston, W. B., & Sved, J. (1990). High-Frequency P Element Loss in Drosophila Is Homology Dependent. Cell, 62: 515-525.

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

Feinberg, A. P., & Vogelstein, B. (1984). A Technique for Radiolabelling DNA Restriction Endonuclease fragments to High Specific Activity. Anal. Biochem., 137: 266-267.

Flores, C., Qadri, M. I., & Lichtenstein, C. (1990). DNA sequence analysis of five genes; tnsA, B, C, D and E, required for Tn7 transposition. N.A.R., 18: 901-11.

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

Fuller, R. S., Kaguni, J. M., & Kornberg, A. (1981). Enzymatic replication of the origin of the Escherichia coli chromosome. P.N.A.S., 78: 7370-7374.

Galas, D. J., & Chandler, M. (1989). Bacterial Insertions Sequences. In D. Berg, & M. M. Howe (Ed.), Mobile DNA (pp. 109-162). Washington, D.C.: American Society for Microbiology.

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

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

Gellert, M. K., Mizuuchi, K., O'Dea, M. H., & Nash, H. A. (1976). DNA gyrase: an enzyme that introduces superhelical turns into DNA. P.N.A.S., 73: 3872-3876.

Gellert, M., & Nash, N. (1987). Communication between segments of DNA during site-specific recombination. Nature, 325: 401-404.

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

Greenblatt, J. (1991). Roles of TFIID in Transcriptional Initiation by RNA Polymerase II. Cell, 66: 1067-1070.

Grindley, N. D., & Sherratt, D. J. (1979). Sequence analysis of IS1 insertion sites: Models for transposition. Cold Spring Harbor Symp. Quant. Biol., 432:

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

Grindley, N., & Reed, R. (1985). Transpositional recombination in prokaryotes. Annu. Rev. Biochem., 54: 863-896.

Gringauz, E., Orle, K. A., Waddell, C. S., & Craig, N. L. (1988). Recognition of Escherichia coli attTn7 by transposon Tn7: lack of specific sequence requirements at the point of Tn7 insertion [published erratum appears in J Bacteriol 1988 Aug;170(8):3792]. J Bacteriol, 170: 2832-40.

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

Hanniford, D. B., Chelouche, A. R., & Kleckner, N. (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.

Hanniford, D. B., Benjamin, H. W., & Kleckner, N. (1991). Kinetic and Structural Analysis of a Cleaved Donor Intermediate and a Strand Transfer Intermediate in Tn10 Transposition. Cell, 64: 171-179.

Harshey, R. (1984). Transposition Without Duplication of Infecting Bacteriophage DNA. Nature, 311: 580-581.

Hauer, B., & Shapiro, J. A. (1984). Control of Tn7 transposition. Mol Gen Genet, 194: 149-58.

Hedges, R., & Jacob, A. (1974). Transposition of ampicillin resistance form RP4 to other replicons. Mol. Gen. Genet., 132: 31-40.

Heichman, K. A., & Johson, R. C. (1990). The Hin Invertasome: Protein-Mediated Joining of Distant Recombination Sits at the Enhancer. Science, 249: 511-517.

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

Hoopes, B. C., & McClure, W. R. (1981). Studies of the selectivity of DNA precipitation by spermine. NAR, 9: 5493-5504.

Huisman, O., & Kleckner, N. (1987). A new generalizable test for detection of mutations affecting Tn10 transposition. Genetics, 116: 185-189.

Johnson, R. C., Bruist, M. F., Glaccam, M. B., & Simon, M. I. (1984). In vitro analysis of Hin-mediated site-specific recombination. CSHSQB, 49: 751-760.

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

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

Kleckner, N. (1989). Tn10. In D. Berg, & M. M. Howe (Ed.), Mobile DNA (pp. 227-268). Washington, D.C.: American Society for Microbiology.

Kleckner, N. (1990). Regulation of transposition in bacteria. Annu. Rev. Genet., 6: 297-327.

Kopecko, D. J., Brevet, J., & Cohen, S. N. (1976). Involvement of multiple translocating DNA segments and recombination hot spots in the structural evolution of bacterial plasmids. J. Mol. Biol., 108: 333-360.

Kubo, K., & Craig, N. L. (1990). Bacterial transposon Tn7 utilizes two different classes of target sites. J. Bacteriol., 172: 2774-2778.

Lavoie, B. D., & Chaconas, G. (1990). Immunoelectron microscopic analysis of the A, B, and HU protein content of bacteriophage Mu transpososomes. JBC, 265: 1623-7.

Lee, C. H., Bhagwat, A., & Heffron, F. (1983). Identification of a transposon Tn3 sequence required for transposition immunity. PNAS, 80: 6765-9.

Lichtenstein, C., & Brenner, S. (1981). Site-specific properties of Tn7 transposition into the E. coli chromosome. Mol. Gen. Genet., 183: 380-7.

Lichtenstein, C., & Brenner, S. (1982). Unique insertion site of Tn7 in the E. coli chromosome. Nature, 297: 601-3.

Liebert, J., Ghelardini, P., & Paolozzi, L. (1982). Conservative Integration of Bacteriophage Mu DNA into pBR322 Plasmid. PNAS, 79: 4362-66.

Ljungquist, E., & Bukhari, A. (1977). State of Prophage Mu upon Induction. PNAS, 74: 3143.

Malamy, M. H. (1966). Frameshift mutations in the lactose operon of E. coli. CSHSQB, 31: 189-201.

Martin, K. J. (1991). The Interactions of Transcription factors and Their Adaptors, Coactivators and Accessory Proteins. BioEssays, 13: 499-503.

Maxwell, A., Craigie, R., & Mizuuchi, K. (1987). B protein of bacteriophage mu is an ATPase that preferentially stimulates intermolecular DNA strand transfer. Proc Natl Acad Sci U S A, 84: 699-703.

McClintock, B. (1956). Controlling elements and the gene. CSHSQB, 21: 197-216.

McKown, R. L., Waddell, C. S., Arciszewska, L. K., & Craig, N. L. (1987). Identification of a transposon Tn7-dependent DNA-binding activity that recognizes the ends of Tn7. PNAS, 84: 7807-11.

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

Miller, J. L., Anderson, S. K., Fujita, D. J., Chaconas, G., Baldwin, D. L., & Harshey, R. M. (1984). The nucleotide sequence of the B gene of bacteriophage Mu. NAR, 12: 8627-38.

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

Mizuuchi, K., Mizuuchi, M., & Craigie, R. (1984). The mechanism of transposition of bacteriophage Mu. CSHSQB, 49: 835-8.

Mizuuchi, K. (1984). Mechanism of transposition of bacteriophage Mu: polarity of the strand transfer reaction at the initiation of transposition. Cell, 39: 395-404.

Mizuuchi, K., & Craigie, R. (1986). Mechanism of bacteriophage mu transposition. Annu. Rev. Genet., 20: 385-429.

Mizuuchi, K., & Adzuma, K. (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., Baker, T. B., & Mizuuchi, K. (1991). DNase protection analysis of the stable synaptic complexes involved in Mu transposition. PNAS, 88: 9031-9035.

Mooslehner, K., Karls, U., & Habers, K. (1990). Retroviral Integration Sites in ransgenic MOV Mice Frequently Map in the Vicinity of Transcribed DNA Regions. J. Virol., 64: 3056-3058.

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

Morisato, D., & Kleckner, N. (1987). Tn10 transposition and circle formation in vitro. Cell, 51: 101-11.

Mullis, K. B., & Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol, 155: 335-50.

Nash, H. A. (1981). Integration and excision of bacteriophage lambda: the mechanisms of conservative site specific recombination. Annu. Rev. Genet., 15: 143-167.

Natsoulis, G., Thomas, W., Roughman, M. C., Winston, F., & Boeke, J. D. (1989). Ty1 transposition in Saccharomyces cerevisiae is nonrandom. Genetics, 123: 269-279.

Ohtsubo, E., Zenilman, M., Ohtsubo, H., McCormick, M., Machida, C., & Machida, Y. (1981). Mechanism of insertion and cointegration mediated by IS1 and Tn3. CSHSQB, 45: 283-295.

Orle, K. A., & Craig, N. L. (1990). Identification of transposition proteins encoded by the bacterial transposon Tn7. Gene, 96: 1-7.

Pato, M. L. (1989). Bacteriophage Mu. In D. Berg, & M. M. Howe (Ed.), Mobile DNA (pp. 23-52). Washington, D.C.: American Society for Microbiology.

Pryciak, P. M., Sil, A., & Varmus, H. E. (1992). Retroviral integration into minichromosomes in vitro. EMBO, 11: 291-303.

Qadri, M. I., Flores, C. C., Davis, A. J., & Lichtenstein, C. P. (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.

Reed, R. R. (1981). Transposon-mediated site-specific recombination: a defined in vitro system. Cell, 25: 713-719.

Richet, G., Abcarian, P., & Nash, H. A. (1986). The interaction of recombination proteins with supercoiled DNA defining the role of supercoiling in lambda integrative recombination. Cell, 46: 1011-1021.

Roberts, D. E., Ascherman, D., & Kleckner, N. (1991). IS10 Promotes Adjacent Deletions at Low Frequency. Genetics, 128: 37-43.

Robertson, C. A., & Nash, H. A. (1988). Bending of the bacteriophage lambda attachment site by Escherichia coli integration host factor. JBC, 263: 3554-7.

Robinson, M. K., Bennet, P. M., Grinstead, J., & Richmond, M. H. (1977). Inhibition of TnA translocation by TnA. J. Bact., 129: 407-414.

Rogers, M., Ekaterinaki, N., Nimmo, E., & Sherratt, D. (1986). Analysis of Tn7 transposition. Mol Gen Genet, 205: 550-6.

Sauer, R. T., Smith, D., & Johnson, A. D. (1988). Flexibility of the yeast alpha-2 repressor enables it to ocuppy the ends of its operator, leaving the center free. Genes and Develop., 2: 807-816.

Shapiro, J. A. (1979). Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. PNAS, 76: 1933-7.

Shapiro, J. (1983). Mobile Genetic Elements. Orlando Fla.: Academic Press, Inc.

Smith, D. B., & Johnson, K. S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene, 31-40.

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

Surette, M. G., & Chaconas, G. (1989). A Protein Factor Which Reduces the Negative Supercoiling Requirement in the Mu DNA Strand Transfer Reaction Is E. coli Integration Host Factor. JBC, 264: 3028-3034.

Surette, M. G., Lavoie, B. D., & Chaconas, G. (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: 3483-9.

Surette, M. G., Harkness, T., & Chaconas, G. w. (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 1 Transposome Formation. J.B.C., 266: pp 3118-3124.

Tang, Y., Lichenstein, C., & Cotterill, S. (1991). Purification and characterization of the TnsB protein of Tn7: a transposition protein that binds to the ends of Tn7. N.A.R., 19: 3395-3402.

Varmus, H., & Brown, P. (1989). Retroviruses. In D. Berg, & M. M. Howe (Ed.), Mobile DNA (pp. 53-108). Washington, D.C.: American Society for Microbiology.

Vijaya, S., Steffen, D. L., & Robinson, H. L. (1986). Acceptor sites for retroviral integrations map near DNaseI-hypersensitive sites in chromatin. J. Virol., 60: 683-692.

Vink, C., Yeheskiely, E., Vandermarel, G. A., Vanboom, J. H., & Plasterk, R. H. A. (1991). Site-Specific Hydrolysis and Alcoholysis of Human Immunodeficiency Virus DNA Termini Mediated by the Viral Integrase Protein. NAR, 19: 24.

Vosberg, H.-P. (1985). DNA topoisomerases: enzymes that control DNA conformation. Curr. Topics Microbiol. Immunol., 114: 19-102.

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

Waddell, C. S., & Craig, N. L. (1989). Tn7 transposition: recognition of the attTn7 target sequence. PNAS, 86: 3958-62.

Wang, J. C., Peck, L. J., & Becherer, K. (1983). DNA supercoiling and its effects on DNA structure and function. Cold Spring Harb Symp Quant Biol, 1: 85-91.

Wang, J. (1985). DNA topoisomerases. Annu. Rev. Biochem., 54: 665-697.

Wang, J. C., & Giaever, G. N. (1988). Action at a Distance Along DNA. Science, 240: 300-304.

Watson, J. D., & Crick, F. H. (1953). Genetic Implications of the Structure of Deoxyribonucleic Acid. Nature, 177: 964.

Weinert, T., Derbyshire, K., & Grindley, N. (1984). Replicative and Conservative Transpositional Recombination of the Insertion Sequences. CSHSQB, 49: 251-260.

Wells, R. D. (1988). Unusal DNA structures. J.B.C., 263: 1095-1098.

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