

Interaction of Tn5 Transposase with the Transposon Termini

Torsten W. Wiegand† and William S. Reznikoff

Department of Biochemistry
420 Henry Mall
University of Wisconsin-Madison
Madison, WI 53706, U.S.A.

Transposition of Tn5 requires the binding of the transposase protein to the transposon outside end (OE) DNA sequences. Transposase mutants that increase the transposition frequency result in the formation of two distinct transposase/OE DNA complexes, observed by gel retardation analysis. The slower migrating complex I, also formed by wild-type transposase, contains protein oligomers of transposase and transposase related proteins. The faster migrating, novel complex II is caused by the binding of monomeric, proteolytic transposase fragments γ and δ that have lost the carboxy-terminus of the protein. Transposase γ and δ bind OE DNA with a high apparent affinity but are unable to promote transposition *in vivo*. We propose that the transposase protein is functionally unstable and can undergo a conformational change that reduces the activity but protects the protein from proteolysis. The transposase mutants favor the more active but proteolytically hypersensitive protein conformation.

Keywords: transposition; transposon Tn5; protein-DNA interaction; protein oligomerization; proteolysis

1. Introduction

The initiation of complex processes like transcription, replication, recombination, and transposition requires the assembly of proteins and DNA molecules into highly organized enzymatic machineries. For this assembly, two types of molecular interaction are critical: specific recognition of DNA sequences by proteins and specific contacts between protein surfaces. We are using the bacterial transposon Tn5 to study the construction of such protein/DNA complexes and aim to eventually elucidate the mechanism and regulation of the complete Tn5 transposition process.

Tn5 is a prokaryotic, composite transposon, consisting of two insertion sequences, IS50R and IS50L, which flank several antibiotic resistance genes (for a review, see Berg, 1989). Two features of Tn5 are essential for transposition: the 476 amino acid residue transposase protein, (Tnp), encoded on IS50R (Johnson & Reznikoff, 1984), and its DNA binding sites; the peripheral 19 bp of the insertion sequences (Makris *et al.*, 1988), called outside ends (OE‡) (Fig. 1). Tn5 transposase promotes trans-

position more efficiently if the protein is expressed from the same DNA molecule that carries the transposon termini (Johnson *et al.*, 1982). The reason for this *cis*-preference is not clear, but it may be due to functional instability of the transposase protein. Transposase has also been found to inhibit the transposition of a Tn5 element in *trans* (DeLong & Syvanen, 1991; Wiegand & Reznikoff, 1992), indicating the existence of a *cis*-activating and a *trans*-inhibiting transposase species. The second IS50R protein is defective for transposition but retains the inhibitory function. This inhibitor protein (Inh) is expressed in the same reading frame as transposase but from independent transcriptional and translational start sites, resulting in a 421 amino acid residue polypeptide which lacks the first 55 residues of transposase (Krebs & Reznikoff, 1986).

Mounting evidence indicates that the Tn5 transposase protein follows a conservative, "cut-and-paste" mechanism to carry out transposition (Berg, 1989). Conservative transposition is initiated by specific binding of transposase to the transposon termini followed by contact of these protein/DNA complexes with each other (Baker & Mizuuchi, 1992; Jilk *et al.*, 1993) and with the target DNA. Transposase then cleaves the DNA at the transposon termini and joins them covalently to the target DNA sequence.

†Present address: NeXagen, 2860 Wilderness Place, Boulder, CO 80301, U.S.A.

‡Abbreviation used: OE, outside ends.

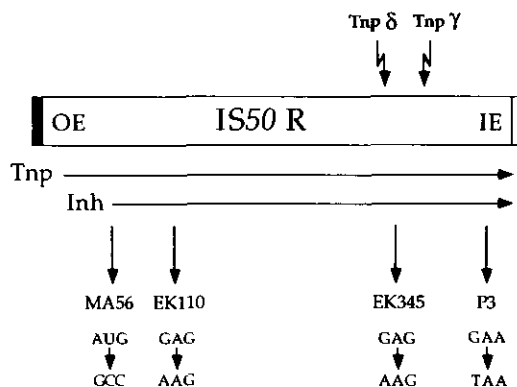


Figure 1. Structure of IS50R. The 1534 bp IS50R encodes 2 proteins involved in transposition: transposase (Tnp), 476 amino acid residues in length, and inhibitor (Inh) which is 55 residues shorter. Transposase and inhibitor share the same reading frame but use independent transcriptional and translational start sites. 19 bp of the outside end (OE) and inside end (IE) are required for the transposition of IS50R. Utilization of 2 outside ends result in transposition of the entire Tn5 element. The location and sequence changes of the two hyper-transposing EK mutations, the inhibitor-abolishing MA56 mutation, and the translation terminating P3 ochre codon are shown. The proposed locations of cleavage sites that form transposase γ and δ are estimates.

A commonly used approach to study these different transposase functions is to isolate transposase mutants that cause phenotypic alterations of the transposition process (DeLong & Syvanen, 1991; Bender & Kleckner, 1992; Leung & Harshey, 1991). We have previously reported the isolation of two separate transposase mutants that each increase the transposition frequency of Tn5 approximately tenfold and enhance the ability of transposase to act in *trans* (Wiegand & Reznikoff, 1992). These mutations are glutamate to lysine substitutions at position 110 (EK110), and 345 (EK345). The mutant transposase proteins interact with the transposon outside end DNA differently from the wild-type protein. In gel retardation assays, both wild-type and mutant transposase cause the formation of a slow migrating protein/DNA complex, called complex I. For the EK mutant proteins, however, a second, faster migrating complex, complex II, can be observed. In this study we report the molecular analysis of these two complexes. The implications of the results are presented in a model which explains the observed phenotype of the transposase mutants. Furthermore, it suggests molecular mechanisms for the *cis*-activity of the transposase protein and for the inhibitory functions of both transposase and inhibitor. Finally, we discuss the possible effects of these mutants on the transposase DNA binding and protein dimerization activities.

2. Materials and Methods

(a) Media and reagents

Bacterial strains were routinely cultured in LB medium which contains 10 g of Bacto-tryptone, 5 g of Bacto-yeast

and 10 g of NaCl per l distilled water. Antibiotics were purchased from Sigma and used in the following concentrations: kanamycin, 40 mg/ml; ampicillin, 100 mg/ml; nalidixic acid, 20 mg/ml; gentamycin, 5 mg/ml; and chloramphenicol, 30 mg/ml. Restriction endonucleases were obtained from New England Biolabs and Promega. Rabbit antitransposase antibodies were a gift from Mark Krebs. Goat antirabbit alkaline phosphatase, nitroblue tetrazolium, 5-bromo-4-chloro-3-indolyl phosphate, 1,10-phenanthroline, neocuproine, and 3-mercapto propionic acid were obtained from Sigma. Purified inhibitor protein was a gift from Norberto de la Cruz. Protein size markers were from Gibco. Radioactive isotopes were purchased from Amersham.

(b) Strains and plasmids

All bacterial strains and plasmids used in this study are described in Table 1. Details of plasmid constructions are listed in the corresponding Materials and Methods sections. DNA manipulations were performed essentially as previously described (Sambrook *et al.*, 1989).

(c) Transposase purification

Transposase and its mutant derivatives were overproduced in strain BL21 (DE3) pLysS (Studier *et al.*, 1990) from the pRZ7074 and pRZ7076 plasmid series. These plasmids are based on the T7 expression vector pET21d(+) (Novagen). pRZ7074 was constructed by ligation of the *Bsp*HI-*Bam*HI fragment from pRZ7013 (Wiegand & Reznikoff, 1992), containing IS50R, to the *Nco*I-*Bam*HI sites located in the polylinker region of pET21d(+). This cloning places transposase under the control of the T7 promoter and a strong Shine-Dalgarno sequence. To overexpress transposase point mutants, pRZ7072 MA56, pRZ7074 MA56/EK110, and pRZ7074 MA56/EK345 were constructed by analogous clonings starting with the *Bsp*HI-*Bam*HI fragments of pRZ7016, pRZ7016 EK110, and pRZ7016 EK345 (Wiegand & Reznikoff, 1992), respectively. The P3 MA56/EK345 overproducing plasmid pRZ7076 MA56/EK345 is the IS50L equivalent of pRZ7074 MA56/EK345 and was obtained by introduction of the ochre stop codon at position 451 of the transposon reading frame.

Overproduction and purification of the Tn5 transposase proteins was performed essentially as described previously (de la Cruz *et al.*, 1993). After cell lysis, the supernatants of successive 26,000 g and 100,000 g spins were precipitated with ammonium sulfate and the resuspended pellets were subjected to heparin-agarose column chromatography. To ensure maximization of complex II formation, the ammonium sulfate precipitation and heparin-agarose column chromatography steps were omitted for the EK transposase proteins. Omission of these steps during wild-type transposase purification, as a control, did not change the behavior of this protein and did not result in increased complex II formation (data not shown).

(d) Gel retardation assays

The plasmid pRZ7067 is the source of the 60 bp OE DNA probe used for gel retardation reactions. It was constructed by insertion of a Tn5 OE DNA cassette into the *Kpn*I-*Sph*I sites of pUC19 (Yanisch-Perron *et al.*, 1985). The gel retardation probe was purified from pRZ7067 following restriction digests with *Eco*RI and *Hind*III. pRZ7067 7G contains a C to G transition at

Table 1
Bacterial strains and plasmids

Strain	Relevant genotype	Reference
<i>A. Bacterial strains</i>		
RZ212	$\Delta(lac-pro)$, <i>ara</i> , <i>str</i> , <i>recA56</i> , <i>srl</i> , <i>thi</i> , pOX38-Gen	Johnson & Reznikoff, 1984
RZ221	<i>polA</i> , $\Delta(lac-pro)$, <i>ara</i> , <i>str</i> , <i>nal</i>	Johnson <i>et al.</i> , 1982
BL21 (DE 3)	F ⁻ , <i>omp T</i> , <i>r_B</i> ⁻ , <i>m_B</i> ⁻ , λ DE3	Studier <i>et al.</i> , 1990
Construct	Description	Source/reference
<i>B. Plasmids</i>		
pRZ 7067	Source of 60 bp WT OE DNA probe	John Makris
pRZ 7067 7G	Source of 60 bp 7G OE DNA probe	John Makris
pRZ 7011	Source of 266 bp WT OE DNA probe	Vincent Schulz
pRZ 7074	T7 driven overproducing plasmid for WT transposase	This study
pRZ 7074 MA56	T7 driven overproducing plasmid for MA56 transposase	This study
pRZ 7074 MA56/EK110	T7 driven overproducing plasmid for MA56/EK110 transposase	This study
pRZ7074 MA56/EK345	T7 driven overproducing plasmid for MA56/EK345 transposase	This study
pRZ 7076 MA56/EK345	T7 driven overproducing plasmid for MA56/EK345 P3 protein	This study
pRZ 4732	Transposition assay plasmid for WT transposase	Michael Weinreich
pRZ 7080	Transposition assay plasmid for EK345 transposase	This study
pRZ 7081	Transposition assay plasmid for WT P3 protein	This study
pRZ 7082	Transposition assay plasmid for EK345 P3 protein	This study

position 7 in the OE sequence, which reduces transposase binding (Jilk, unpublished results). It was constructed analogously to pRZ7067, starting with a mutant OE DNA cassette. A 266 bp OE DNA probe was isolated from pRZ7011 by digestion with *EcoRI* and *HpaI*. pRZ7011 contains IS50R on an *EcoRI*-*BamHI* fragment which was inserted into the polylinker of pTZ19U (Mead *et al.*, 1986).

The gel retardation assay was performed as described previously (Wiegand & Reznikoff, 1992). In a standard retardation reaction 5 pmol (0.3 μ g) of transposase proteins were incubated with 5 fmol of ³²P-labeled 266 bp (0.9 ng) or 60 bp (0.2 ng) DNA fragments containing the outside end sequence of Tn5.

(e) *Analysis of complexed transposase and transposase related proteins*

To obtain sufficient amounts of complexed proteins for analysis, the gel retardation reactions were scaled up to 500 fmol (20 ng) of the 60 bp OE DNA fragment and 25 pmol (1.5 μ g) of the transposase proteins. The DNA binding reactions were electrophoresed as described above, except that 0.6 mm thick polyacrylamide gels were used. After electrophoresis transposase OE DNA complexes were located in the gel by autoradiography. Gel slices containing either complex I or complex II were cut from the non-denaturing gel and loaded onto a 3%/10% discontinuous SDS/polyacrylamide gel (Laemmli, 1970). To control for transposase proteins entering the retardation gel unbound or non-specifically bound to DNA, appropriate gel slices were cut from lanes in which the specific OE DNA was omitted. Electrophoresis was carried out for 2 h 30 min at 200 volts and the proteins were electrotransferred to nitrocellulose overnight at 100 mA. The blot was incubated with a 1:500 dilution of polyclonal rabbit antitransposase antibodies and then with a 1:1000 dilution of goat antirabbit immunoglobulin G-alkaline phosphatase conjugate.

Antibody complexes were visualized with 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium.

(f) *DNA footprinting analysis*

The *in situ* DNA footprinting analysis was performed on the retarded complexes as described (Sigman *et al.*, 1991). The gel retardation reaction was scaled up to obtain retardation of approximately 100 fmol of the 266 bp DNA probe into complex I and complex II. After electrophoresis of the gel retardation reactions on a non-denaturing 8% polyacrylamide gel, the entire gel was placed in 200 ml of 50 mM Tris·HCl (pH 8.0). 20 ml of 2 mM 1,10-phenanthroline and 0.45 mM copper sulfate were added, along with 20 ml of 0.5% 3-mercapto propionic acid. The gel was incubated at room temperature for ten min and the cleavage reaction was stopped by the addition of 20 ml of 29 mM neocuprione. After 2 min, the gel was washed thoroughly with distilled water, covered with Saran wrap, and exposed to film. The developed film allowed the localization of unbound DNA and the retarded DNA complexes, which were excised from the gel. The DNA was eluted overnight into elution buffer, precipitated with ethanol, resuspended in denaturing formamide loading buffer (Sambrook *et al.*, 1989), heat denatured, and electrophoresed on an 8%, 8 M urea/polyacrylamide gel. The G and A reactions of a Maxam-Gilbert sequencing ladder (Maxam & Gilbert, 1980) of the same DNA fragment were used as size markers. The polyacrylamide gel was dried, exposed to film for autoradiography, and quantified with a Betascope 603 blot analyzer (Betagen Corporation, Waltham, MA).

(g) *Transposition assay*

The plasmids used for the mating out transposition assays are based on pRZ4732 (Michael Weinreich, personal communication). In this construct, the transposase gene,

lacking IS50 outside and inside end sequences, was placed under control of the *lacUV5* promoter and inserted into pBR322. An adjacent mini Tn5 element, consisting of 2 outside end sequences flanking the kanamycin resistance gene from Tn903, is used to measure transposition frequencies. pRZ7080, pRZ7081 and pRZ7082 are transposase point mutant variants of this plasmid, allowing us to test EK345 transposase, P3 protein and EK345 P3 protein, respectively, for transposition activity *in vivo*. These plasmids were constructed by substitution of the internal *NotI*-*BglIII* IS50 fragment of pRZ4732 with fragments carrying the appropriate point mutations.

The mating-out transposition assay, using donor-strain RZ212 and recipient-strain RZ221, was performed essentially as described previously (Yin *et al.*, 1988). The assay measures the movement of the mini Tn5 element from plasmids pRZ4732, pRZ7080, pRZ7081, and pRZ7082 to the F-factor in the cell.

3. Results

(a) Wild-type and EK345 transposase proteins form specific but different complexes with Tn5 OE DNA

Previously we have shown that wild-type and EK345 transposase proteins form different types of complexes with Tn5 OE DNA (Wiegand & Reznikoff, 1992). The gel retardation assay in Figure 2A shows that wild-type transposase retards a 266 bp DNA fragment, containing the 19 bp Tn5 outside end sequence, into complex I (lane 2). Since this transposase preparation contains a small amount of contaminating inhibitor protein, we changed the methionine codon at position 56, serving as the start codon for inhibitor, to alanine. This MA56 mutation does not allow expression of the inhibitor protein and lane 3 shows that MA56 transposase retards OE DNA identically to wild-type transposase. Therefore, formation of this transposase/OE DNA complex is not dependent on the presence of the inhibitor protein and the MA56 substitution does not alter the DNA binding characteristics of transposase. The MA56/EK345 double mutant transposase protein (lane 4), however, retards the same DNA into a novel, faster migrating complex, called complex II.

In order to test if these protein/DNA complexes are caused by the specific binding of the transposase proteins to the OE DNA sequence, we repeated the gel retardation assay with two different DNA probes: a 60 bp fragment containing the wild-type OE sequence and a mutant version in which position 7 of the outside end was changed from a G to a C. This sequence alteration abolishes transposition *in vivo* (Makris *et al.*, 1988) and has been shown to reduce transposase binding *in vitro* (Jilk, unpublished results). It, therefore, allows us to distinguish between the specific and non-specific contributions of the protein/DNA interaction. In Figure 2B, comparison of wild-type transposase binding to wild-type OE DNA (lane 2) and to 7G OE DNA (lane 3) shows that the 7G mutation causes a dramatic reduction in the formation of complex I, indicating a high degree of DNA binding specificity. DNA

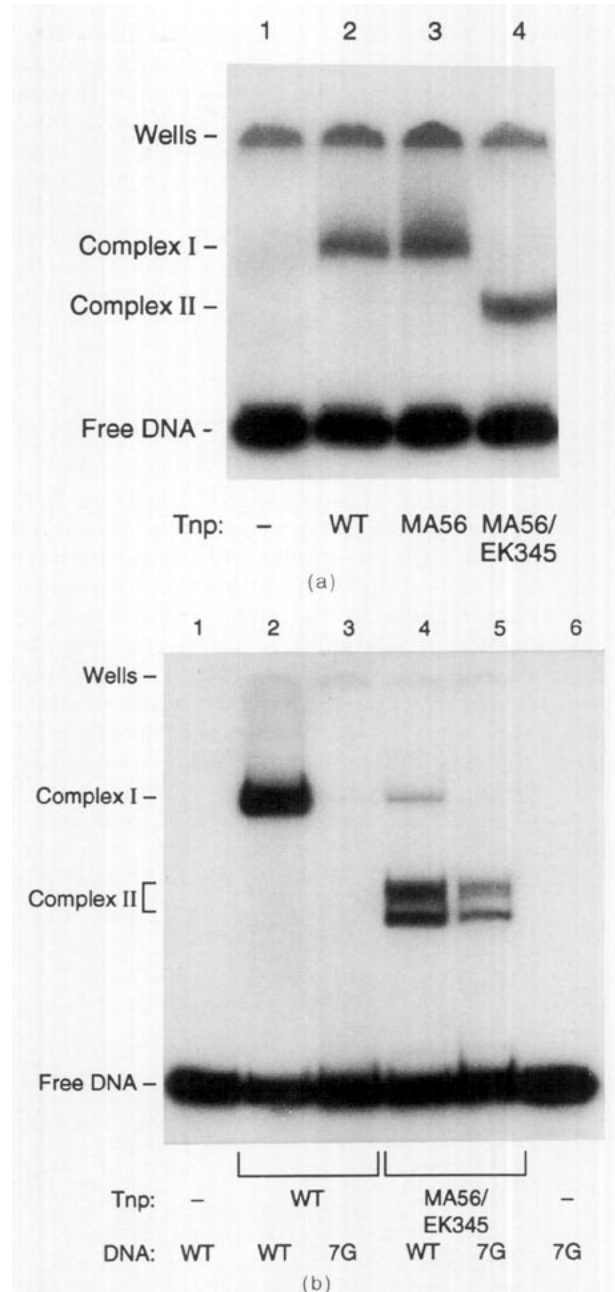


Figure 2. Gel retardation of Tn5 OE DNA fragments by transposase proteins. DNA/protein complexes were formed in 20 μ l reactions containing 5 fmol of Tn5 OE DNA fragments and 5 pmol of transposase proteins. They were analyzed by native polyacrylamide gel electrophoresis. Positions of complex I and complex II are indicated. A, Retardation of a 266 bp Tn5 OE DNA Fragment. Lane 1, no transposase protein; lane 2, wild-type transposase; lane 3, MA56 transposase; lane 4, MA56/EK345 transposase. B, Retardation of 60 bp Tn5 OE DNA Fragments. The fragments contain either the wild-type Tn5 OE sequence or a mutant sequence with a G to C transition at position 7, counting from the transposon terminus. Due to the higher resolution of the shorter DNA probe, complex II is split into two bands. Lane 1, wild-type DNA, no transposase protein; lane 2, wild-type DNA, wild-type transposase; lane 3, 7G DNA, wild-type transposase; lane 4, wild-type DNA, MA56/EK345 transposase; lane 5, 7G DNA, MA56/EK345 transposase; lane 6, 7G DNA, no transposase protein.

retardation of the 60 bp wild-type OE DNA fragment with MA56/EK345 transposase results in the formation of both complex I and complex II (lane 4). Comparison with retardation of 7G OE DNA (lane 5) shows that complex II formation is specific, but that this binding specificity is apparently lower than that observed for complex I. In addition, complex II is split into two distinct bands. We believe that this is caused by the enhanced resolution of the shorter DNA probe and we will provide a molecular explanation below.

(b) *Complex I and complex II differ in transposase size and stoichiometry*

Since binding of transposase to OE DNA leads to the formation of two distinct complexes, we wanted to determine the molecular difference between complex I and complex II. In general, this difference could be due to the size, conformation, or stoichiometry of the DNA or of the protein. Gel electrophoretic analysis of denatured DNA isolated from each of the retarded complexes indicates that both DNA strands are intact, ruling out the possibility that the differential migration is caused by DNA cutting or nicking (data not shown).

A conformational difference in DNA, resulting in altered mobility of the protein/DNA complexes, could have been due to differential DNA bending in

complex I and complex II upon transposase binding. However, both complexes were observed with a 60 bp probe (see Fig. 2B) which is too short for bending to significantly affect the complex mobility

Finally, to test if the differential migration of the two complexes is caused by a difference in DNA stoichiometry, we repeated the retardation reaction with a mixture of two OE DNA probes, differing in length. If complex I contains two DNA molecules, retardation of the DNA mixture should result in three retardation complexes: a complex containing two short DNA probes, one containing two long probes, and a third, intermediate complex containing one short and one long DNA molecule. Since this indicative, intermediate complex was not observed (data not shown), we conclude that complex I contains only a single DNA molecule. The reduced mobility of complex I, therefore, cannot be explained by a difference in DNA stoichiometry.

After ruling out that alterations in the bound DNA probe cause the mobility difference between the retarded complexes, we decided to analyze the complexed transposase proteins. This required the comparison of transposase proteins added to the DNA binding reaction with proteins present in the retarded protein/OE DNA complexes. Portions of the DNA binding reactions were separated by electrophoresis on a denaturing, discontinuous polyacrylamide gel and transposase related proteins

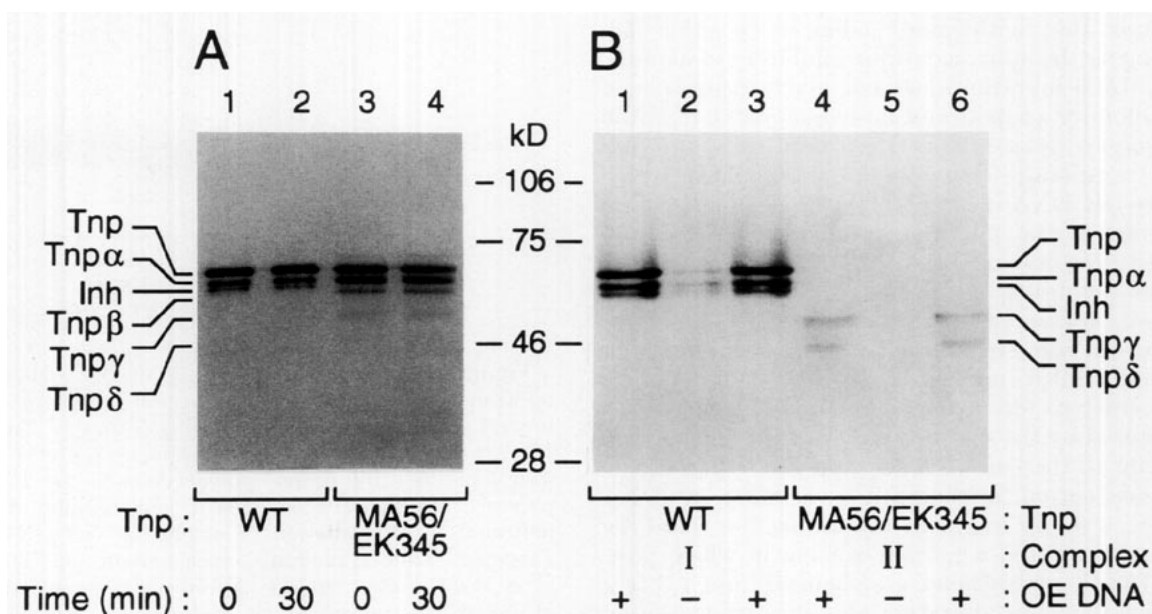


Figure 3. Immunoblotting analysis of complexed transposase proteins. Transposase proteins separated on 3%/10% discontinuous SDS/polyacrylamide gels were electrotransferred to nitrocellulose and detected with polyclonal antitransposase antibodies. A, Portions of DNA binding retardation reactions, containing 0.5 pmol transposase proteins, were taken before and after incubation and loaded directly onto the SDS/protein gel. Lane 1, wild-type transposase, 0 min; lane 2, wild-type transposase 30 min; lane 3, MA56/EK345 transposase, 0 min; lane 4, MA56/EK345 transposase, 30 min. B, Gel slices containing transposase/OE DNA complexes were cut from a non-denaturing retardation gel and then loaded onto the SDS/protein gel. Complex I gel slices contain approximately 4 times as much DNA as complex II slices as determined by ^{32}P quantitation of the bottom part of the protein gel (data not shown). Control gel slices were prepared from gel retardation reactions in which the specific OE DNA was omitted. Lanes 1 and 3, wild-type transposase related proteins in complex I; lane 2, complex I control; lanes 4 and 6, MA56/EK345 transposase related proteins in complex II; lane 5, complex II control. The locations and molecular weights of protein size markers are indicated.

were detected by immuno-blotting. The resulting Western blot is shown in Figure 3A (lane 1, wild-type transposase; lane 3, MA56/EK345 transposase). Besides the transposase protein, the preparations contain two amino-terminal proteolytic transposase fragments, transposase α and β . The wild-type transposase preparation contains, in addition, contaminating inhibitor protein.

The remainder of the DNA binding reactions were loaded onto a non-denaturing retardation gel to obtain separation of the protein/DNA complexes. Gel slices containing complex I and complex II were cut out of the native retardation gel and loaded onto a denaturing, discontinuous polyacrylamide gel. After electrophoresis the proteins were blotted onto nitrocellulose and analyzed by immuno-detection. Three proteins are present in wild-type transposase/OE DNA complex I: transposase, transposase α , and the inhibitor protein (Fig. 3B, lanes 1 and 3). To confirm that these proteins are specifically bound to OE DNA rather than entering the retardation gel unbound or non-specifically bound to DNA, a corresponding gel slice was prepared from a retardation reaction lacking the specific OE DNA (lane 2). Details about the presence of transposase α and the inhibitor protein in complex I is published elsewhere (de la Cruz *et al.*, 1993).

In contrast to complex I, the proteins from the EK345 transposase/DNA complex II are two previously unobserved proteolytic transposase fragments and are called transposase γ and δ (lanes 4 and 6, control in lane 5). These proteins apparently bind DNA in monomeric form, each causing retardation of the DNA probe into a single band within complex II (see Fig. 2B). If they were able to dimerize with each other or with full-length transposase, intermediate bands in between the complex II bands or between complex I and complex II should have been observed.

To test if the transposase fragments γ and δ are formed by proteolysis of full-length protein during the gel retardation reaction, we analyzed DNA binding reaction aliquots after a 30 minute incubation period. These samples are shown in Figure 3A, lane 2 (wild-type transposase) and lane 4 (MA56/EK345 transposase). Since the 0 and 30 minute patterns are indistinguishable, no significant protein degradation occurs during the retardation reaction. Furthermore, prolonged incubation causes an increase in complex I intensity, but complex II disappears over time, indicating that full-length transposase is more stable in the DNA binding reactions than transposase γ and δ (data not shown). Therefore, proteolysis of the full-length transposase protein, forming transposase γ and δ apparently occurs *in vivo*.

(c) *Transposase γ and δ are carboxy-terminal deletion fragments*

Because of the correlation between the appearance of complex II in gel retardation reactions and

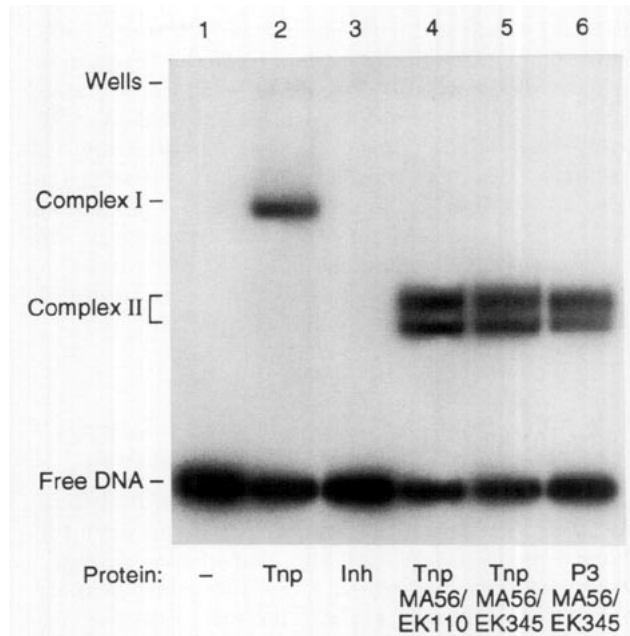


Figure 4. Gel Retardation of 60 bp Tn5 OE DNA by transposase related proteins. 5 fmol of 60 bp Tn5 OE DNA fragments and approximately 5 pmol of transposase related proteins were incubated in 20 μ l reactions. Lane 1, no transposase protein; lane 2, wild-type transposase; lane 3, inhibitor protein; lane 4, MA56/EK110 transposase; lane 5, MA56/EK345 transposase; lane 6, MA56/EK345 P3 protein.

increased transposition frequencies observed *in vivo*, we set out to further characterize transposase γ and δ . The transposase fragments could represent proteolytic cleavage at a local disturbance caused by the EK345 substitution in the transposase protein or could reflect a protease-hypersensitive form of transposase caused by the induction of a more general conformational change. To distinguish between these possibilities, we tested a second transposase mutant, EK110 transposase, which similarly increases the Tn5 transposition frequency (Wiegand & Reznikoff, 1992), for the presence of transposase γ and δ . Using the 60 bp DNA probe for the gel retardation reaction, we show in Figure 4, lane 4, that EK110 transposase causes the same complex II pattern as EK345 transposase (lane 5). The fact that the formation of transposase γ and δ is caused by two independent substitutions in the protein more than 200 residues apart must reflect a general conformational change in the transposase protein.

To determine if transposase γ and δ represent amino-terminal or carboxy-terminal deletions of the transposase protein, a stop codon was introduced into the EK345 transposase construct, resulting in the expression of a protein that lacks the last 26 amino acids of transposase. This protein is called EK345 P3. Preliminary size estimation of transposase γ and δ indicates that these fragments have lost more than 100 amino acids of the transposase protein (data not shown). Therefore, if transposase γ and δ are caused by the loss of the transposase

carboxy-terminus, the premature P3 stop codon should have no effect on the size of the transposase fragments. On the other hand, if transposase γ and δ reflect amino-terminal deletions of transposase, the stop codon should shorten them by an additional 26 residues. The complex II retardation pattern of EK345 is shown in Figure 4, lane 6, and is identical to the EK345 transposase retardation pattern. This indicates that transposase γ and δ are caused by cleavage at two individual sites, resulting in carboxy-terminal deletions of the protein.

(d) *Transposase proteins protect DNA in complex I and complex II at the Tn5 OE sequence*

To test where the transposon proteins bind to the DNA probes, and if this binding differs between full-length transposase and the transposase γ and δ fragments, we performed DNA footprinting analysis. DNA containing the Tn5 OE sequence was incubated with wild-type transposase, MA56 transposase, and MA56/EK345 transposase. The protein/DNA complexes were electrophoretically separated and the entire retardation gel was treated with phenanthroline-copper, resulting in DNA cleavage *in situ*. DNA molecules from complex I and complex II were isolated and analyzed on a high-resolution, denaturing polyacrylamide gel. The resulting DNA footprints are shown in Figure 5. Comparison of unbound DNA (lane 1) with bound DNA (lanes 2 to 4) shows that protection by all forms of the transposase proteins occurs primarily at the OE sequence and extends into the flanking DNA. The complex I protection patterns of wild-type transposase (lane 2) and MA56 transposase (lane 3) are virtually identical, confirming that this mutation does not alter the DNA binding characteristics of the transposase protein. The complex II protection pattern of MA56/EK345 transposase (lane 4), caused by the binding of monomeric transposase fragments γ and δ , is very similar to the complex I protection pattern in extent and intensity. Thus, the DNA binding domain of the transposase protein seems to be relatively unaffected by the carboxy-terminal deletion. However, subtle differences between the two protection patterns can be observed. First, the complex II footprint extends a few more bases further into the flanking DNA, and secondly the lowest protection in the OE sequence is observed at position 7C rather than at 11T. Both observations may help to explain why the 7G OE point mutation shows a reduced specificity for complex II formation.

Since complex I is caused by binding of protein-oligomers containing transposase and transposase-related proteins (de la Cruz *et al.*, 1993; and Fig. 3), the question arises how many protein subunits are in contact with the DNA. The observation that the complex I footprint does not exceed the size of the monomeric complex II footprint lets us conclude that only a single molecule, presumably transposase, contacts the DNA directly in complex I.

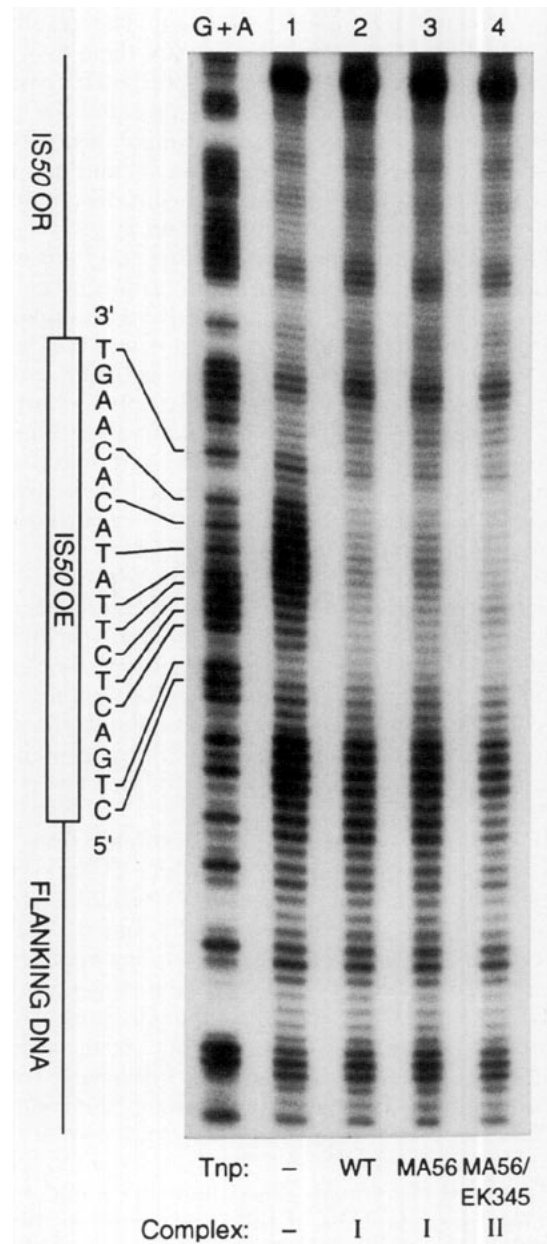


Figure 5. *In situ* phenanthroline-copper footprinting of transposase/OE DNA complexes. Autoradiograph of footprinting gel. Gels containing the transposase/OE DNA complexes described in Fig. 1 were subjected to phenanthroline-copper attack. After *in situ* cleavage the DNA was extracted and analyzed on denaturing polyacrylamide gels. Lane G+A, Maxam-Gilbert sequencing reactions "G" and "A" as size markers; lane 1, unbound DNA, no transposase protein; lane 2, complex I, wild-type transposase; lane 3, complex I, MA56 transposase; lane 4, complex II, MA56/EK345 transposase.

(e) *Transposase γ and δ are inactive for transposition in vivo*

The correlation between complex II and increased transposition frequencies can be explained by two different models. Either the EK mutations could induce a more active transposase conformation which is hypersensitive to proteolytic degradation, resulting in formation of inactive transposase γ and

Table 2
Transposition frequencies of wild-type and EK345 IS50 proteins

Plasmid	Transposase protein	Transposition frequency†
pRZ 4732	WT Tnp	$1.1 (\pm 0.3) \times 10^{-5}$
pRZ 7080	EK345 Tnp	$1.4 (\pm 0.8) \times 10^{-4}$
pRZ 7081	WT P3	$\leq 4.0 \times 10^{-8} \ddagger$
pRZ 7082	EK345 P3	$\leq 5.3 \times 10^{-8} \ddagger$

† The transposition frequencies were determined in quintuplicate by the mating-out assay. The average value and the standard error are shown.

‡ The limit of detection in this assay is approximately 5×10^{-8} events per cell per generation.

δ , or, the carboxy-terminus of transposase could act as an inhibitory domain and needs to be cleaved off to form the active transposase γ and δ fragments. In order to distinguish between these possibilities, we decided to test the transposition activity of transposase γ and δ in the absence of full-length transposase. Therefore, the EK345 mutation was moved into the P3 protein which lacks the 26 amino-terminal residues of transposase. Neither wild-type P3 ((Johnson *et al.*, 1982) and Table 2, line 3) nor EK345 P3 (Table 2, line 4) show transposition activity *in vivo*. Since EK345 P3 is a substrate for the formation of transposase γ and δ but does not promote transposition *in vivo*, we can rule out the possibility that transposase γ and/or δ are the active transposase species.

4. Discussion

An essential step of the transposition reaction is the binding of the transposase protein to the DNA sequences which form the transposon termini. We have previously reported the isolation of transposase mutants which form two distinct DNA/protein complexes and found a correlation between the appearance of the faster migrating complexes II and increased Tn5 transposition frequencies (Wiegand & Reznikoff, 1992). The slower migrating complex I, observed for wild-type and mutant transposase proteins, represents oligomers of transposase and transposase-related proteins (de la Cruz *et al.*, 1993). In this study, we show that complex II is caused by proteolytic transposase fragments featuring high apparent DNA binding affinities. These fragments, transposase γ and δ , are not active for transposition but rather seem to reflect degradation of an active and hypersensitive conformation of the transposase protein.

(a) Proteins contained in complex II

Complex II does not contain full-length transposase but instead is due to OE DNA retardation by transposase γ and δ , two proteolytic transposase fragments. In gel retardation reactions wild-type transposase forms very little complex II and, there-

fore, seems quite resistant to this proteolytic cleavage. Two transposase point mutations, EK110 and EK345, however, show significant increases in the abundance of transposase γ and δ , indicated by the appearance of complex II in retardation gels. DNA complexes formed by transposase γ and δ appear unaltered if a stop codon is introduced into the transposase gene, resulting in the lack of 26 carboxy-terminal amino acids. Therefore, transposase γ and δ are proteolytic fragments which have lost the carboxy-terminus of transposase. Size estimation and preliminary attempts to localize the cleavage sites (data not shown) indicate the loss of approximately 100 and 150 amino acids, respectively.

The enhanced resolution of the 60 bp DNA probe used for gel retardation shows that transposase γ and δ each form a distinct retarded band within complex II. This suggests that transposase γ and δ bind in monomeric form to the OE DNA sequence. If they were able to oligomerize with each other or with full-length transposase, additional bands should have been observed either in between the two complex II bands or between complex I and complex II. Preparations of EK mutant transposase show very little γ and δ fragments, judged by immunoblotting, but still give rise predominantly to complex II in gel retardation reactions. Transposase γ and δ , therefore, have higher apparent DNA binding affinities than full-length transposase. However, at this time we cannot rule out the possibility that the increased apparent binding affinities are due to higher activities of transposase γ and δ . Quantitative measurements of DNA binding affinities and determination of binding constants will only be possible after purification of the various transposase species to homogeneity. Consistent with these observations are preliminary studies in which various carboxy-terminal deletions of transposase were expressed *in vitro*. Deletion of 69 amino acids results in a protein with approximately tenfold higher affinity for OE DNA, and which is still able to oligomerize. Deletion of a further 19 amino acids causes the loss of transposase oligomerization without reduction of the increased DNA binding affinity (Mahnke, unpublished results). These results support the conclusion that the amino-terminus of transposase contains the DNA binding domain (de la Cruz *et al.*, 1993). Furthermore, they suggest that the dimerization domain is located towards the carboxy-terminus, and that the carboxy-terminus itself reduces OE DNA binding affinity.

We decided to test the ability of transposase γ and δ to carry out the transposition reaction *in vivo* because of the observed correlation of complex II formation with increased Tn5 transposition frequencies and because of a study suggesting that transposase may be activated by *recA*-mediated cleavage (Kuan *et al.*, 1991). We found that EK345 P3 produces transposase γ and δ , forming complex II in retardation reactions, but is defective for transposition *in vivo*. These results indicate that transposase γ and δ , although capable of DNA binding,

are unable to promote the complete transposition reaction. One possible explanation is that the EK transposase mutations induce a conformational change that results in an activated, protease-sensitive form of the protein. The increased activity of this conformation explains the hypertransposing phenotype, and the higher protease sensitivity causes the formation of transposase γ and δ , resulting in the appearance of complex II. Alternatively, transposase could fold into the active conformation during translation and then quickly change into the more stable but less active form. The EK mutations would increase the protein activity by slowing down this conformational inactivation.

This functional instability of transposase could explain why the protein promotes transposition preferentially in *cis*. The active transposase molecule simply does not reach a DNA binding site located on a different molecule before the conformational inactivation occurs. Since the carboxy-terminus of transposase seems to reduce DNA binding but is required for transposition, one intriguing possibility suggests that the nascent protein, tethered by the RNA molecule, binds to the transposon end sequences before completion of translation (DeLong & Syvanen, 1991; Reznikoff, 1993). Translation of the carboxy-terminus then completes the protein, resulting in functional transposase bound to its DNA binding site.

(b) *Model for the mechanism and regulation of Tn5 transposition initiation*

We have previously reported that not only the inhibitor protein but also transposase itself can inhibit the transposition reaction in *trans* (Wiegand & Reznikoff, 1992). To explain this observation, as well as the results presented in this study, we propose the model shown in Figure 6. After translation, transposase is present briefly as an active, monomeric species (ovals). This transposase protein binds to the Tn5 outside ends and brings these sequences together by protein dimerization. After addition of target DNA, strand cleavage, exchange, and DNA ligation complete the transposition reaction. This active transposase, however, is susceptible to proteolysis. The resulting transposase γ and δ fragments bind DNA tightly but have lost the ability to dimerize, and transposition cannot continue. This protein/DNA interaction is observed as complex II. The EK mutations shift the equilibrium to favor this active but unstable conformation, therefore increasing transposition frequencies and complex II formation. Wild-type transposase is present predominantly in a conformation which is inactive or less active for transposition (circles). This transposase species can interact in solution with the inhibitor protein (squares). The resulting hetero-dimers are able to bind the Tn5 outside ends by using the DNA binding domain of the transposase subunit, but transposition is blocked since

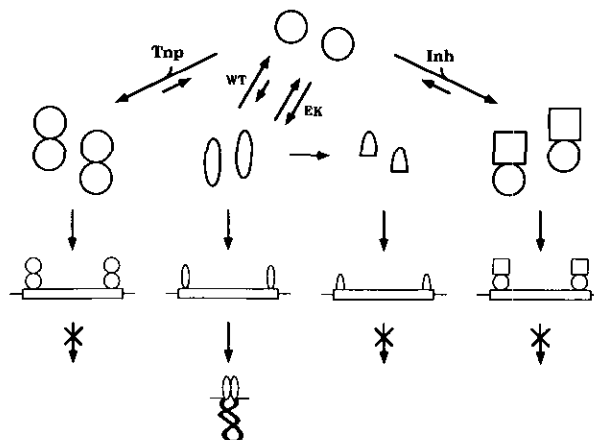


Figure 6. Model of the mechanism and regulation of Tn5 transposition initiation. Transposase is present in the cell in two conformations: one *cis*-activating form (ovals), which is hypersensitive for proteolysis, and a *trans*-inhibiting species (circles) which can dimerize with a second transposase molecule or the inhibitor (squares) protein (for details, see Discussion).

the bound complexes contain no unoccupied dimerization site to contact each other. It seems likely that the inhibiting function of transposase is caused by the same mechanism. Transposase homo-dimers, therefore, that form in solution and then bind to OE DNA may also be inactive due to occupied dimerization domains. Precedence for inactivation by protein dimerization has been reported for RepA, the initiator protein of the P1 plasmid replication system (Wickner *et al.*, 1991). Binding of the transposase homodimers and transposase/inhibitor heterodimers is observed as complex I formation. This mechanism inhibits transposition efficiently since it reduces both the amount of active transposase and occupies the necessary transposase binding sites. Theoretically, a transposase homodimer bound to one OE DNA sequence could interact with a second, unbound transposon end to form an active complex. The observed *trans*-inhibitory effect of the transposase protein, however, argues against this possibility. Therefore, either the timing of the dimerization relative to the DNA binding event must be important, or the transposase dimerization in solution somehow differs from the contact of two transposase proteins that are bound to the transposon termini.

The question that remains to be addressed is why we do not observe the binding of the active, full-length transposase monomer in the gel retardation assays. Since Tn5 transposition evolved to be a very infrequent process, it is possible that this complex is present in only minor amounts. In addition, the interaction may be too transient with the single DNA binding site we provide in our assays and may require a second OE sequence to be stabilized. This possibility will be tested in future experiments.

We thank Michael Weinreich, Patricia Kiley, and Elizabeth Goodwin for technical advice and for helpful discussions. This research was supported by American Cancer Society grant MV-64163.

References

- Baker, T. A. & Mizuuchi, K. (1992). DNA-promoted assembly of the active tetramer of the Mu transposase. *Genes Develop.* **6**, 2221–2232.
- Bender, J. & Kleckner, N. (1992). IS10 transposase mutations that specifically alter target site recognition. *EMBO J.* **11**, 741–750.
- Berg, D. E. (1989). Transposon Tn5. In *Mobile DNA* (Berg, D. E. & Howe, M. M., eds), pp. 185–210. American Society for Microbiology, Washington, DC.
- de la Cruz, N., Weinreich, M. D., Wiegand, T. W., Krebs, M. P. & Reznikoff, W. S. (1993). *J. Bacteriol.* **175**, 6932–6938.
- DeLong, A. & Syvanen, M. (1991). Trans-acting transposase mutant from Tn5. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 6072–6076.
- Jilk, R. A., Makris, J. C., Borchardt, L. & Reznikoff, W. S. (1993). Implications of Tn5 associated adjacent deletions. *J. Bacteriol.* **175**, 1264–1271.
- Johnson, R. C. & Reznikoff, W. S. (1984). Role of the IS50 R proteins in the promotion and control of Tn5 transposition. *J. Mol. Biol.* **177**, 645–661.
- Johnson, R. C., Yin, J. C. & Reznikoff, W. S. (1982). Control of Tn5 transposition in *Escherichia coli* is mediated by protein from the right repeat. *Cell*, **30**, 873–882.
- Krebs, M. P. & Reznikoff, W. S. (1986). Transcriptional and translational initiation sites of IS50. Control of transposase and inhibitor expression. *J. Mol. Biol.* **192**, 781–791.
- Kuan, C., Liu, S. & Tessman, I. (1991). Excision and transposition of Tn5 as an SOS activity in *Escherichia coli*. *Genetics*, **128**, 45–57.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)*, **227**, 680–685.
- Leung, P. C. & Harshey, R. M. (1991). Two mutations of phage Mu transposase that affect strand transfer or interactions with B protein lie in distinct polypeptide domains. *J. Mol. Biol.* **219**, 189–199.
- Makris, J. C., Nordmann, P. L. & Reznikoff, W. S. (1988). Mutational analysis of insertion sequence 50 (IS50) and transposon 5 (Tn5) ends. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 2224–2228.
- Maxam, A. M. & Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.
- Mead, D. A., Szczesna-Skorupa, E. & Kemper, B. (1986). Single-stranded DNA 'blue' T7 promoter plasmids: a versatile tandem promoter system for cloning and protein engineering. *Protein Eng.* **1**, 67–74.
- Reznikoff, W. S. (1993). The Tn5 transposon. *Annu. Rev. Microbiol.* **47**, 945–963.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edit., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sigman, D. S., Kuwabara, M. D., Chen, C. B. & Bruice, T. W. (1991). Nuclease activity of 1, 10-phenanthroline-copper in study of protein-DNA interactions. *Methods Enzymol.* **208**, 414–433.
- Studier, R. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.
- Wickner, S., Hoskins, J. & McKenney, K. (1991). Monomerization of RepA dimers by heat shock proteins activates binding to DNA replication origin. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 7903–7907.
- Wiegand, T. W. & Reznikoff, W. S. (1992). Characterization of two hypertransposing Tn5 mutants. *J. Bacteriol.* **174**, 1229–1239.
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, 103–119.
- Yin, J. C., Krebs, M. P. & Reznikoff, W. S. (1988). Effect of dam methylation on Tn5 transposition. *J. Mol. Biol.* **199**, 35–45.

Edited by N. Sternberg

(Received 5 April 1993; accepted 13 September 1993)