J. Mol. Biol. (1997) **271**, 362–373

JMB

Tn*5* Transposase Mutants that Alter DNA Binding Specificity

Maggie Zhou and William S. Reznikoff*

Department of Biochemistry University of Wisconsin-Madison, 420 Henry Mall Madison, WI 53706, USA

*Corresponding author

Tn*5* transposase (Tnp) binds to Tn*5* and IS*50* end inverted repeats, the outside end (OE) and the inside end (IE), to initiate transposition. We report the isolation of four Tnp mutants (YH41, TP47, EK54 and EV54) that increase the OE-mediated transposition frequency and enhance the binding affinity of Tnp for OE DNA. In addition, two of the Tnp mutants (TP47 and EK54) appear to be change-of-specificity mutants, since they alter the recognition of OE *versus* IE relative to the wild-type Tnp. EK54 enhances OE recognition but decreases IE recognition. TP47 enhances both OE and IE recognition but with a much greater enhancement for IE than for OE. This change-of-specificity effect of TP47 is observed only when TP47 Tnp is synthesized *in cis* to the DNA that contains the ends. We propose that Lys54 makes a favorable interaction with an OE-specific nucleotide pair(s), while Pro47 may cause a more favorable interaction with an IE-specific nucleotide pair(s) than it does with the corresponding OE-specific nucleotide pair(s). A model to explain the preference of TP47 Tnp for the IE *in cis* but not *in trans* is proposed.

Keywords: DNA binding specificity; gain of function; protein-DNA interaction; Tn*5* transposase; transposition

Introduction

Transposition is an illegitimate recombination process mediated by transposable elements such as bacterial transposons. Transposition can give rise to a variety of genome rearrangements such as insertions, deletions, inversions and chromosome fusions. The fundamental macromolecules involved in transposition are an element-encoded protein called transposase (Tnp) that catalyzes the multiple sequential steps in transposition and the specific DNA sequences that define the ends of the element. Because of the importance and complexity of this process, it is of interest to understand the structure/function aspects of Tnp, including how it recognizes the specific end DNA sequences. Since transposition has a destabilizing effect on the genome, it is to be expected that many transposable elements would have evolved various strategies to minimize its frequency. A variety of mechanisms have been discovered that decrease the transposition frequency for various elements. In some cases the Tnp is synthesized infrequently (Tn*10*, Kleckner, 1990), in others the Tnp is unstable (IS*903*, Derbyshire *et al.*, 1990) or its activity is downregulated by some other macromolecule (Tn*5*, Reznikoff, 1993). One intriguing mechanism for decreasing the frequency of transposition is for the Tnp to have evolved with suboptimal properties such as being ineffective at recognizing the specific end sequences that define the transposable element.

Tn*5* is a composite prokaryotic transposon consisting of two IS*50* insertion sequences in opposite orientation flanking a central unique region encoding three antibiotic resistances (Figure 1(a); for a review of Tn*5*, see Reznikoff, 1993). IS*50*R (Right) is itself an autonomous transposable element encoding two proteins: a 476 amino acid residue transposase protein (Tnp) that is required for transposition, and an inhibitor protein (Inh) that is translated from the same open reading frame as Tnp but lacks the N-terminal 55 amino acid residues of Tnp. Inh inhibits transposition efficiently both *in cis* and *in trans* (Johnson *et al.*, 1982); Tnp, however, activates transposition much more efficiently *in cis* than *in trans*. When it is produced *in trans* to another Tn*5* element that also encodes its own Tnp, Tnp acts primarily to inhibit transposition of that element (Delong & Syvanen, 1991; Wiegand & Reznikoff, 1992).

Non-productive premature multimerization is thought to be at least one of the causes of the *cis*

Abbreviations used: Tnp, Tn*5* transposase; OE, outside end; IE, inside end; wt, wild-type.

0022–2836/97/330362–12 $25.00/0/mb971188

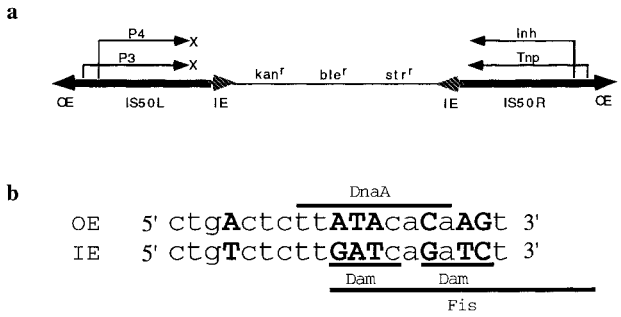


Figure 1. (a) Structure of Tn*5*. Filled arrows indicate OEs and striped arrows indicate IEs. IS*50*L differs from IS*50*R at only one base-pair, resulting in the premature termination of translation of both transposase (Tnp) and inhibitor (Inh). *kan*, Kanamycin; *ble*, bleomycin; *str*, streptomycin. (b) Sequence comparison of OE and IE, and some of the host factor binding/modification sites that they overlap. The uppercase letters indicate non-identical positions between OE and IE.

preference of Tnp in activating transposition (Weinreich *et al*., 1994). Other possible reasons for preferential *cis* activity of a protein include but are not limited to: physical and/or conformational instability of the active form of the protein (e.g. Tn*903* Tnp, Derbyshire *et al*., 1990); tight, non-specific DNA binding, which favors its association to the nearest DNA molecule (e.g. λ Q protein, Echols *et al*., 1976; Burt & Brammar, 1982); or a temporal coupling of the translation of the protein to its function (e.g. the C terminus of the protein, when fully translated, might interfere with the DNA-binding function of the N terminus: Tn*10* Tnp, Jain & Kleckner, 1993).

Transposition of Tn*5* and IS*50* requires the two 19 bp DNA sequences located at the ends of the respective transposable element: two outside ends (OE) for Tn*5*, and one OE and one inside end (IE) for IS*50* (Figure 1(a): Johnson & Reznikoff, 1983; Sasakawa *et al*., 1983). OE and IE contain 12 identical and seven non-identical base-pairs (see Figure 1(b)). Transposition of Tn*5* derivatives with different OE and IE combinations occur at different frequencies, implying that OE and IE are utilized with different efficiencies (Sasakawa *et al*., 1983; Makris *et al*., 1988), and possibly form protein-DNA complexes of somewhat different components, since OE and IE overlap with different host protein recognition sites. DnaA protein stimulates Tn*5* transposition through direct contact with a DnaA box located at positions 8 to 16 of OE (Yin & Reznikoff, 1987; Fuller *et al*., 1984; A. Bhasin, unpublished results). IE contains two *dam* methylation sites, and *dam* methylation inhibits IE-mediated transposition (Yin *et al*., 1988). In wild-type (wt) Tn*5*, IE partially overlaps with a Fis binding consensus, which was shown *in vitro* to be bound by Fis protein in footprinting and gel retardation assays (Weinreich & Reznikoff, 1992).

Generally, the first step of transposition involves Tnp binding to the ends of the transposable element. Tn*5* Tnp binds to OE and IE sequences *in vitro* (de la Cruz *et al*., 1993; Jilk *et al*., 1996). The role of specific base-pairs of the end sequences in transposition and in Tnp binding in particular has been studied through genetic analyses (Makris *et al*., 1988; Dodson & Berg, 1989; Tomcsanyi & Berg, 1989; Jilk *et al*., 1993) and chemical modification/inhibition analyses (Jilk *et al*., 1996). However, the role of different Tnp residues in end sequence recognition has not been studied extensively. Interestingly, Inh, which lacks only the N-terminal 55 amino acid residues of Tnp, does not bind to OE *in vitro* (de la Cruz *et al*., 1993; Weinreich *et al*., 1994). This indicates that the N terminus of Tnp is important for binding to OE (and possibly IE). Through deletion analysis of Tn*5* Tnp, Weinreich *et al*. (1993) found that deletion of 11 amino acid residues from the N terminus completely abolishes Tnp binding to OE, while deletion of three residues from the N terminus does not affect this binding. Both deletion derivatives still retain, at least partially, the normal folded conformation of wt Tnp, since they both possess a *trans*-inhibitory activity similar to wt Tnp. This suggests that the impaired OE binding seen with Δ11 Tnp is due to a defect in the DNA-binding domain, rather than to a gross misfolding of the mutant Tnp protein. These authors also found that point mutations AT20, DN24, RH30 and AV36, and an internal six residue deletion Δ30-35 all result in significantly reduced or undetectable OE binding activity. All this evidence is consistent with the N terminus of Tnp being important for DNA binding.

Here, we set out to characterize which residues of Tnp are involved in binding DNA, and how binding specificity is determined. While structural studies of Tnp and Inh are being carried out by other members of our group, we decided to take a combined genetic and biochemical approach to study the Tnp DNA-binding domain. Such an approach had been used successfully to study CAP, Lac repressor and Trp repressor-DNA interactions (for a review, see Ebright, 1991). In order to perform this study, we assumed that Tnp activity might be limited by a low binding affinity for the OE and IE, and that we could identify residues in the DNA-binding domain by isolating mutations with an enhanced transposition phenotype. Second, we utilized the unique feature of Tnp, that it recognizes and binds to two somewhat different sequences, OE and IE. We reasoned that a mutant that manifests an increased affinity for the OE sequence might do so by virtue of an additional interaction with a nucleotide pair(s) that differs between the OE and IE and thus might manifest a change of OE/IE specificity. We began by randomly mutagenizing the N-terminal one-third of Tnp (since the N-terminal part of Tnp most likely contains the DNA-binding domain, see above), and screening for mutants causing an increased or decreased transposition frequency of a Tn*5*-like

construct defined by two OE sequences. Among the mutants causing an increased transposition frequency (which may be due to a new specificity for OE), we then asked if any changed the differentiation of OE *versus* IE recognition displayed by wt Tnp. We found two mutants that appeared to meet these requirements: EK54 and TP47. After we characterized their OE *versus* IE preferences in transposition *in vivo*, we studied their *in vitro* DNA-binding properties with gel shift assays. The results indicate that both Tnp mutants have an altered interaction with DNA.

Results

Isolation of Tnp N-terminal mutants

Tnp catalyzes Tn*5* and IS*50* transposition. The first step in this process is Tnp binding to the OE and/or IE DNA. Inh is not required for transposition (it inhibits transposition *in vivo*), and it does not bind to OE *in vitro*. Since Tnp differs from Inh only at its N terminus, it is likely that the N terminus of Tnp is important for the OE/IE binding reaction. Consistent with this, several single amino acid substitutions as well as short deletions of the N terminus of Tnp have been found to result in significantly reduced or undetectable OE binding activity (Weinreich *et al.*, 1993). To investigate the possible DNA-binding function of the Tnp N-terminal region, we conducted a PCR random mutagenesis of the first 166 amino acid residues using Taq DNA polymerase, as described in Materials and Methods. The population of mutagenized plasmid pRZ5412 (Figure 2 and Table 1), encoding the MA56 Tnp (and not encoding Inh, see Materials and Methods), was electroporated into strain MW320, which contains the papillation factor pOXGen386 (Weinreich *et al.*, 1993; and see Materials and Methods). The papillation frequency in this strain provides an estimate of the relative

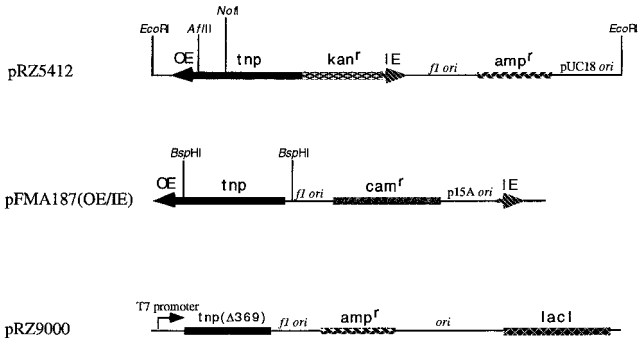


Figure 2. Structure of three plasmids used in this study. The MA56 mutation is present in all of these constructs. When representing OE or IE, the arrow points towards the outside of the transposable element (from base-pair 19 to base-pair 1). The *Bsp*HI sites in pFMA187 indicate where the deletions were subsequently made to generate pRZ5452, pRZ5453 and pRZ5454, resulting in the complete loss of the *tnp* gene. *kan*, Kanamycin; *amp*, ampicillin; *cam*, chloramphenicol.

trans-activated OE/OE-defined transposition frequency caused by the newly introduced *tnp* gene. The resulting colonies were screened for papillation levels and compared to that of the unmutagenized pRZ5412. We screened about 1000 colonies resulting from each of nine independent PCR reactions. Overall, about 30% of the colonies demonstrated a decrease in papillation (hypopapillation), whereas only six mutants showed an increase in papillation (hyperpapillation). The six mutants with a hyperpapillation phenotype and 20 mutants with a hypopapillation phenotype were isolated, and their plasmid DNA purified and sequenced. The mutants that resulted in hypopapillation phenotype are listed by Zhou (1997). Of the 26 mutants isolated, 11 had a single point mutation, and 15 had two or more point mutations. The PCR conditions suggested by Zhou *et al.* (1991), which uses low dNTP concentrations (50 μM each) and no manganese, yielded fewer double and multiple mutations and more single mutations in our hands than the PCR conditions using standard dNTP concentrations (200 μM each) and 250 μM $MnCl_2$.

We were particularly interested in the six mutants that resulted in hyperpapillation, since they are more likely to directly change OE recognition than down mutants, which could simply be due to loss of function. Because some of them contained double mutations, we subcloned each point mutation, then tested the papillation phenotype of each point mutant. We found five point mutations that resulted in hyperpapillation: YH41, TP47, EK54, EV54 and EK110. EK110 was isolated and studied previously, and was found not to increase DNA-binding activity (Wiegand & Reznikoff, 1992, 1994). It was not further studied here. Both YH41 and EK54 were isolated twice in this study, and EK54 and EV54 are mutations of the same position, indicating that the possible kinds of mutations in the region mutagenized that will result in hyperpapillation under the conditions used is not much more than what has been isolated so far.

In vivo OE/IE preferences of Tnp mutants

In order to determine quantitatively if any of the four hyperpapillating *tnp* mutations (YH41, TP47, EK54 and EV54) changed the differential recognition of OE *versus* IE displayed by wt Tnp *in vivo*, we utilized a mating-out assay in which the frequency of transposition onto an F factor was determined (see Materials and Methods). We tested transposition frequencies of wt and mutant Tnp with different end sequence combinations (OE/OE, OE/IE or IE/IE). Since the transposition frequency mediated by wt Tnp is known to be much higher when Tnp is produced by the same DNA molecule (*in cis*) than when it is produced by a different molecule (*in trans*) as the one on which the end sequences are located (Johnson *et al.*, 1982; Isberg *et al.*, 1982), we carried out the mating-out assays both *in cis* and *in trans*. The results are shown in Table 2. All the assays in this Table were per-

Table 1. Bacterial strains and plasmids

Strains, plasmids	Description	Source
A. Strains		
BL21(DE3) pLysS	Strain for T7 promoter overexpression	Studier *et al.* (1990)
DH5α	Δ*(lac)U169 endA1 gyrA96 hsdR17 recA1 relA1 supE44 thi-1* ϕ80*lacZ*Δ*M15*	Hanahan (1985)
CJ236	*dut1 ung1 thi-1 relA1*; pCJ105 (Cam^r F′)	Kunkel *et al.* (1987)
MDW320	Δ*(lac-pro) ara strs* Δ *recA thi* /pOXGen386	Weinreich *et al.* (1993)
JCM101 − pOX38-Gen	Δ*lacZX74 rpsL dam-3*/pOX38-Gen	I. Goryshin; M. G. Marinus
14R525	F$^-$Nal^r prototroph	Goryshin *et al.* (1994)
B. Plasmids		
pRZ3271-26G	IS*50*R in pTZ18U vector, *tnp* gene under natural promoter and inducible T7 promoter control	Schulz & Reznikoff (1991)
pRZ7055	IS*50*R with Kan^r gene inserted between the end of *tnp* gene and IE	Wiegand, (1993)
pRZ7067OE	pUC19 vector carrying the 53 bp OE DNA fragment	J. Makris
pRZ7067IE	pUC19 vector carrying the 56 bp IE DNA fragment	Jilk *et al.* (1996)
pRZ5412	Full-length *tnp* gene and Kan^r gene flanked by OE and IE, MA56 Tnp	This study
pFMA187(OE/IE)	Full-length *tnp* gene and Cam^r gene flanked by OE and IE, MA56 Tnp	Goryshin *et al.* (1994)
pFMA187(IE/IE)	Full-length *tnp* gene and Cam^r gene flanked by IE and IE, MA56 Tnp	Goryshin *et al.* (1994)
pFMA187(OE/OE)	Full-length *tnp* gene and Cam^r gene flanked by OE and OE, MA56 Tnp	This study
pRZ5452	pFMA187(OE/IE) with complete deletion of *tnp* gene	This study
pRZ5453	pFMA187(IE/IE) with complete deletion of *tnp* gene	This study
pRZ5454	pFMA187(OE/OE) with complete deletion of *tnp* gene	This study
pRZ9000	T7 promoter driven overexpression plasmid for TnpΔ369	York & Reznikoff (1996)

formed in a *dam* strain, since transposition involving the IE is known to be inhibited by *dam* methylation (Yin *et al.*, 1988; unpublished results) (there are two *dam* methylation sites within the 19 bp IE sequence, Figure 1(b)).

All four *tnp* mutants increased the OE/OE transposition frequency, as expected from the screening strategy. YH41 and EV54 appeared to promote a general increase in transposition, since they increased the transposition frequency with all end combinations. These mutations may (a) interact more favorably with the DNA, either by creating a more favorable interaction with a specific nucleotide identical in OE and IE, or by introducing a more positive charge in the DNA binding motif so that it can better interact with the negatively charged DNA phosphate backbone, (b) interact more favorably with a host factor involved in both OE and IE transposition complexes, or (c) alter some other step in transposition. Model (a), DNA contact enhancement, is supported by our *in vitro* DNA-binding results reported below.

On the other hand, EK54 Tnp was found to differentiate OE *versus* IE more strongly, compared to wt Tnp. When OE/OE was used, EK54 increased transposition eight- to ninefold over wt

Table 2. *In vivo* transposition frequencies of *tnp* mutants

Tnp mutation	End combination	*Cis* transposition freq[a]	(Relative freq)[b]	*Trans* transposition freq[a]	(Relative freq)[b]
Wild-type	OE/OE	$2.22(\pm0.48) \times 10^{-4}$	(1)	$1.77(\pm0.27) \times 10^{-5}$	(1)
YH41	OE/OE	$14.9(\pm4.25) \times 10^{-4}$	(6.7 ± 1.9)	$3.63(\pm0.54) \times 10^{-5}$	(2.1 ± .31)
TP47	OE/OE	$24.4(\pm4.51) \times 10^{-4}$	(11 ± 2.0)	$21.2(\pm2.07) \times 10^{-5}$	(12 ± 1.2)
EK54	OE/OE	$19.6(\pm5.30) \times 10^{-4}$	(8.8 ± 2.4)	$13.8(\pm2.40) \times 10^{-5}$	(7.8 ± 1.4)
EV54	OE/OE	$5.16(\pm0.21) \times 10^{-4}$	(2.3 ± .09)	$3.84(\pm0.59) \times 10^{-5}$	(2.2 ± .33)
Wild-type	OE/IE	$6.01(\pm0.89) \times 10^{-4}$	(1)	$6.02(\pm0.70) \times 10^{-5}$	(1)
YH41	OE/IE	$67.5(\pm16.0) \times 10^{-4}$	(11 ± 2.7)	$28.1(\pm3.23) \times 10^{-5}$	(4.7 ± .54)
TP47	OE/IE	$324(\pm137) \times 10^{-4}$	(54 ± 23)	$101(\pm6.6) \times 10^{-5}$	(17 ± 1.1)
EK54	OE/IE	$0.99(\pm0.07) \times 10^{-4}$	(0.16 ± .01)	$1.58(\pm0.49) \times 10^{-5}$	(0.26 ± .08)
EV54	OE/IE	$29.9(\pm2.37) \times 10^{-4}$	(5.0 ± .39)	$27.4(\pm3.80) \times 10^{-5}$	(4.6 ± .63)
Wild-type	IE/IE	$0.95(\pm0.07) \times 10^{-4}$	(1)	$2.65(\pm0.35) \times 10^{-5}$	(1)
YH41	IE/IE	$10.3(\pm0.26) \times 10^{-4}$	(11 ± .27)	$8.14(\pm0.94) \times 10^{-5}$	(3.1 ± .35)
TP47	IE/IE	$244(\pm89.6) \times 10^{-4}$	(257 ± 94)	$29.9(\pm11.3) \times 10^{-5}$	(11 ± 4.3)
EK54	IE/IE	$0.154(\pm0.029) \times 10^{-4}$	(0.16 ± .03)	$0.58(\pm0.09) \times 10^{-5}$	(0.22 ± .03)
EV54	IE/IE	$9.54(\pm2.30) \times 10^{-4}$	(10 ± 2.4)	$11.3(\pm2.77) \times 10^{-5}$	(4.3 ± 1.0)

[a] Transposition frequencies were measured by mating-out assays. Wild-type or mutant Tnp was supplied *in cis* or *in trans* to a DNA containing the indicated end combinations. Each frequency is the average of the results obtained from five independent initial transformants. Each error is the standard error of the mean, which is calculated as the sample standard deviation divided by the square-root of 5.

[b] Relative frequencies are calculated by taking the wild-type transposition frequency in each group as 1. The errors in the relative frequencies are calculated by taking the errors in the corresponding absolute frequencies divided by the absolute frequency of wild-type Tnp in that group.

Tnp; but when OE/IE, or IE/IE was used, EK54 decreased transposition four- to sixfold relative to wt Tnp (Table 2). EK54 is thus a change-of-specificity mutant, because it displayed an altered OE *versus* IE differentiation phenotype compared to wt Tnp. These results suggest that either (a) Lys54 is able to make a more favorable interaction with one or more OE specific base(s), or with the OE DNA backbone, compared to Glu54, or (b) EK54 Tnp interacts more favorably with an OE specific host factor (e.g. DnaA protein?) during transposition, than wt Tnp. Again, our *in vitro* DNA-binding data, presented in the next section, support model (a) and argue against model (b).

TP47 Tnp also presents an interesting case, in that it changes the OE/IE preference, but this change is manifested only *in cis*. When provided *in cis*, TP47 Tnp appears to favor IE utilization dramatically (257-fold higher than wt Tnp when using IE/IE, 54-fold higher than wt Tnp when using OE/IE, and only 11-fold higher than wt Tnp when using OE/OE). No such IE preference is observed when it is provided *in trans*. *In trans*, the frequencies of IE/IE, OE/IE and OE/OE transposition are all increased 11 to 17-fold. A model to explain this is proposed later.

In Table 2, the transposition frequency of wt and mutant Tnp *in cis* is generally much greater than *in trans*, consistent with previous findings reported by Johnson *et al*. (1982) and Isberg *et al*. (1982).

In vitro OE/IE binding properties of Tnp mutants

To answer whether the changes in transposition frequency by these mutant Tnp proteins were, at least in part, correlated with changes in their binding affinity to OE and IE sequences, we carried out an *in vitro* gel retardation assay. To measure the DNA-binding properties of the Tnp N-terminal mutants, we used derivatives of Tnp that contained a 108 amino acid residue truncation of the C terminus (Δ369), which are transpositionally inactive. The C-terminal region of Tnp was shown to be involved in Tnp dimerization (Weinreich *et al*., 1993; York & Reznikoff, 1996), which appears to influence Tnp binding to OE (de la Cruz *et al*., 1993; York & Reznikoff, 1996). TnpΔ369 binds to OE as a monomer (York & Reznikoff, 1996), so the abundance of Tnp-OE complexes is a simple reflection of the Tnp-OE affinity, and is not complicated by the dimerization reaction. Moreover, protection assays and DNA-bending assays suggest that the full-length Tnp-OE complexes and TnpΔ369-OE complexes have similar properties despite differences in protein stoichiometry (Wiegand & Reznikoff, 1994; York & Reznikoff, 1996, 1997).

We used pRZ9000 (York & Reznikoff, 1996; Figure 2 and Table 1) which encodes wt Tnp truncated after residue 369 (designated Δ369). Mutations YH41, TP47, EK54 and EV54 were each subcloned into pRZ9000 to yield the truncated version of each mutant Tnp.

First, crude cell extracts containing wt, YH41, TP47, EK54 and EV54 TnpΔ369 were used in a gel-retardation assay with a labeled 53 bp DNA fragment containing the OE or a 56 bp DNA fragment containing the IE. The DNA containing the IE was prepared from a *dam* strain. An approximately equal amount of each TnpΔ369 was used, according to quantification of crude extracts by densitometric scanning of Coomassie blue-stained SDS/polyacrylamide gels. The results are shown in Figure 3. IE binding of TnpΔ369 is complicated by the presence of Fis protein in crude cell extracts, which also binds unmethylated IE (Weinreich & Reznikoff, 1992) and thus competes with TnpΔ369 for IE binding. Nevertheless, it is clear that YH41, TP47 and EV54 TnpΔ369 are all significantly enhanced for both OE and IE binding, whereas EK54 TnpΔ369 is greatly enhanced for OE binding but only slightly enhanced for IE binding. In fact, in the gel shown (Figure 3(a)), EK54 TnpΔ369 has shifted virtually all the labeled OE into the TnpΔ369-OE complex, thus the full extent of enhanced OE binding by EK54 TnpΔ369 is not reflected in this gel. Note also that TP47 TnpΔ369 enhances IE binding more than it enhances OE binding.

In order to study OE and IE binding of TP47 and EK54 TnpΔ369 more accurately, we then purified wt, TP47 and EK54 TnpΔ369. All three proteins were found to be overexpressed to a similar level and purified to a similar yield and purity (≈ 85% pure), indicating that the three proteins have similar biochemical properties. Gel retardation assays were again carried out with fixed

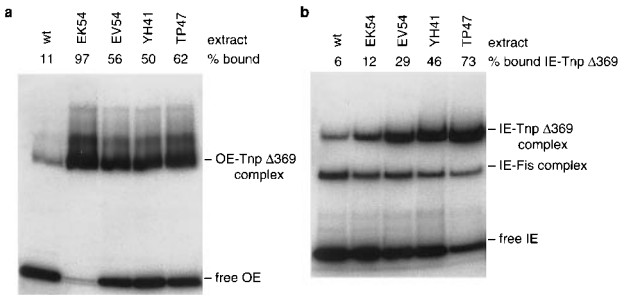


Figure 3. Gel retardation assays with crude extracts prepared from cells containing wt or mutant TnpΔ369. (a) OE binding; 2 nM 53 bp labeled DNA fragment containing OE was used in a binding reaction with crude cell extracts containing equivalent amount of wt or mutant TnpΔ369. (b) IE binding; same as OE binding except that the labeled DNA used was a 56 bp fragment containing IE, and that the amount of each TnpΔ369 used was ten times more than that in OE binding. The percentage of the total labeled DNA of each lane that is found in the OE-TnpΔ369 and the IE-TnpΔ369 complexes was determined with a Betascope (see Materials and Methods). IE is also bound by the Fis protein present in the crude extracts, and the IE-Fis complex nearly disappeared when the DNA containing IE was purified from a *dam*$^+$ strain (data not shown), consistent with results reported by Weinreich & Reznikoff (1992).

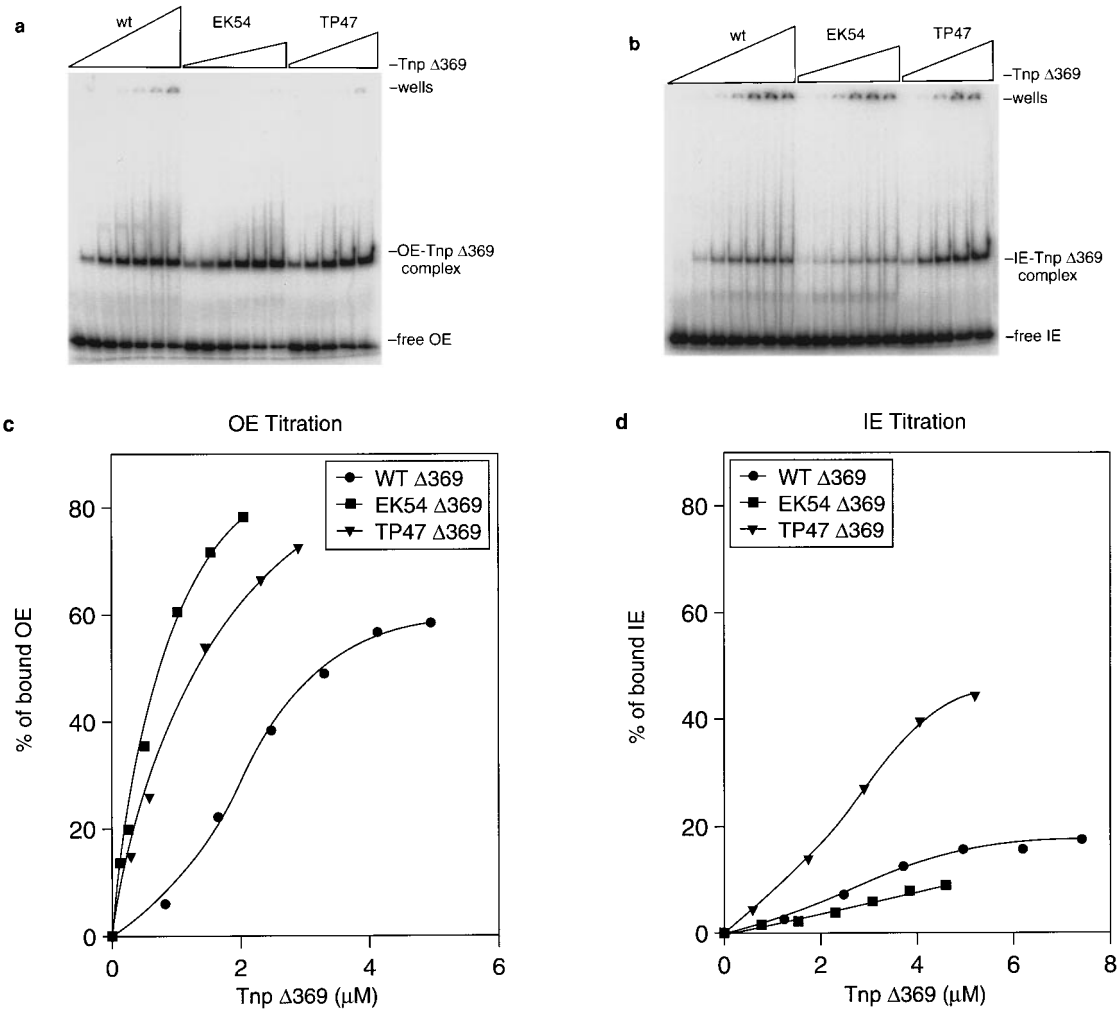


Figure 4. Gel retardation assays with purified wt, EK54 and TP47 TnpΔ369. (a) and (b) Titration of 11 fmol (0.4 nM) of the labeled OE and IE DNA, respectively, with increasing amounts of the purified TnpΔ369 preparations. The triangles above the lanes indicate the relative amount of purified TnpΔ369 added in each reaction. The percentage of the total labeled DNA of each lane that is found in the OE-TnpΔ369 and the IE-TnpΔ369 complexes was determined with a Phosphorimager (Molecular Dynamics), and shown as OE and IE titration curves in (c) and (d), respectively. Each experiment was performed at least twice with each of the two groups of purified protein preparations. Shown here are the data of one representative experiment, for OE and IE, respectively. The relative positions and trends of the curves from each experiment is the same, although the absolute activity of the proteins varied with (1) the method of the protein preparation (which group of protein preparation was used), and (2) the age of the protein preparation. This is why the results from different experiments were not averaged.

concentrations (~0.4 nM) of the same labeled OE or IE fragments mentioned above and varying concentrations of each TnpΔ369 to determine a titration curve for each protein and DNA combination. Figure 4(a) and (b) show an example of such a titration experiment with OE and IE, respectively. Figure 4(c) and (d) show the corresponding titration curves for each TnpΔ369. All three proteins were purified in parallel on two independent occasions. The gel retardation assay was performed at least twice with each group of purified protein preparations. The relative positions and trends of the curves from each experiment are the same, although the absolute activity of the proteins varied (see the legend to Figure 4).

The OE and IE binding experiments using purified TnpΔ369 proteins demonstrated that the altered *in vivo* (*trans*) transposition properties of the mutants were a result of altered OE and IE binding affinities. EK54 TnpΔ369 binds to OE better than wt TnpΔ369, but it binds to IE poorer than wt TnpΔ369. Wild-type and EK54 TnpΔ369 binding to OE has also been studied extensively by York & Reznikoff (1996, 1997) with qualitatively the same results. TP47 TnpΔ369 enhances both OE and IE binding compared to wt TnpΔ369. Since the position of the "bound" complex in these gel shift assays was already determined to correspond to TnpΔ369 monomer-OE or IE complexes (York & Reznikoff, 1996), these results indicate that the increased or decreased *in vivo* transposition frequency of these mutants observed in the previous section correlates with their increased or decreased binding affinity to the respective ends, and is un-

likely to be due to an altered ability to interact with another protein factor involved in transposition. For EK54 TnpΔ369, the increased binding to OE and decreased binding to IE suggests that Lys54 either makes a more favorable interaction with one or more OE-specific base(s) than Glu54, but a less favorable interaction with the corresponding IE-specific base(s), and/or interacts more favorably with the OE DNA backbone, and less favorably with the IE DNA backbone, than Glu54. To strictly differentiate between these two possibilities, three-dimensional structural analysis will need to be performed. In subsequent work we shall investigate this question with more genetic and biochemical experiments.

We also purified EV54 TnpΔ369 in one experiment and tested DNA binding in parallel with wt and EK54 TnpΔ369. We found that EV54 increased both OE and IE binding, but to a small extent (data not shown). This is consistent with the *in vivo* transposition data of EV54 Tnp.

Discussion

The N terminus of Tn*5* Tnp was proposed from deletion and null point mutation studies to be important in Tnp binding to OE and IE sequences of Tn*5* and IS*50*. Here, we tried a novel approach to search for the amino acid residues in the Tnp N terminus that may be in close contact with the OE/IE sequences in a Tnp-DNA complex. We sought Tnp mutants that enhanced OE/OE-mediated transposition. We reasoned that a mutant Tnp with an enhanced recognition of the OE might do so by virtue of contacts with nucleotides that differ between OE and IE (OE and IE differ at seven out of 19 positions), and thus might manifest a change-of-specificity compared to wt Tnp. Of the four mutants isolated with an enhanced OE/OE transposition phenotype, one mutant (EK54) has the desired change-of-specificity phenotype. It increased OE/OE transposition, but decreased OE/IE and IE/IE transposition. A second mutant (TP47) has a more complicated change-of-specificity phenotype. It increased IE/IE and OE/IE transposition much more dramatically, as opposed to OE/OE transposition, but this effect is seen only *in cis*. When TP47 Tnp was provided *in trans* to the end sequences, transposition involving all end combinations was enhanced to a similar, more modest degree.

It was possible that the properties of the Tnp mutants could be explained by altered interactions with host factors. This was ruled out by performing gel shift assays using wt and the mutant forms of Tnp, with OE and IE DNA. These assays demonstrated that the four mutant Tnp have altered affinities for OE and IE DNA, consistent with their *in vivo* phenotypes.

As far as we are aware, such a genetic/biochemical approach is unique in studying protein-DNA interactions. Sakai & Kleckner (1996) and N. Tavakoli & K. Derbyshire (personal communication) have studied Tnp-end sequence interactions in Tn*10* and Tn*903*, respectively, by isolating transposase mutants that suppressed defective mutant end sequences. Several interesting mutants were isolated; however, neither group has identified a transposase mutant that is specifically altered in end sequence recognition. A possible reason for this lack of success is that the base-pairs in the end sequences responsible for specific binding by transposase may act cooperatively, such that single mutations of the ends may not be easily rescued by a single mutation in the transposase. Our approach is different and possibly more advantageous, since we started off with two functional end sequences, which differ at more than one base-pair. The results of our studies are discussed in more detail below.

EK54 TnpΔ369 in crude extracts greatly enhanced OE binding but only slightly enhanced IE binding, compared to wt TnpΔ369. When purified, EK54 TnpΔ369 enhanced OE binding compared to wt TnpΔ369, while decreasing IE binding relative to wt. The reason for the different results obtained with crude extract and with purified TnpΔ369 is not clear, but it may be due to an imprecise measurement of TnpΔ369 quantities in crude extracts, or the influence of other factors present in the crude extracts. But the trend in the two experiments is the same, in that OE binding is much more preferred than IE binding when using EK54 compared to wt TnpΔ369. The results with purified TnpΔ369 are also in good agreement with the *in vivo* transposition data, indicating that the EK54 phenotype observed *in vivo* is caused by its direct effect on DNA binding as observed *in vitro* (as opposed to an indirect effect through its interaction with another factor present in the cell).

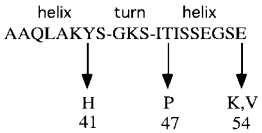
A possible molecular explanation for the behavior of EK54 Tnp is suggested by the properties of EV54 Tnp. Since the side-chain of Val is unable to form a hydrogen bond with DNA (although it may interact with DNA through van der Waals interactions with its hydrophobic side-chain), the fact that EV54 binds better to OE and IE than wt indicates that Glu54 in wt Tnp does not make a favorable interaction with OE or IE, but rather, might be slightly inhibitory due to its negative charges. In view of that, Lys54 may be even more unfavorable in contacting IE due to its bulky side-chain, but may be able to make a favorable interaction(s) with OE by being a hydrogen bond donor (the Glu side-chain can act only as a hydrogen bond acceptor). If this is true, thymine is a good candidate base that only the Lys side-chain can contact, since it can only be a hydrogen bond acceptor.

TP47 TnpΔ369 in crude extracts as well as in purified form significantly enhanced both OE and IE binding. It seems to differentiate OE *versus* IE in a manner similar to wt TnpΔ369 *in vitro*, although with higher affinities for both. Since the protein is added *in trans* to the DNA in these gel-retardation experiments, and since *in vivo*, TP47 Tnp differen-

tiates OE *versus* IE similar to wt Tnp *in trans*, the *in vitro* DNA binding data agree well with the *in vivo* transposition data. To test *in vitro* DNA-binding in *cis*, we would have to develop a system to transcribe and translate Tnp derivatives *in vitro* from a DNA that also contains the end sequences. Even if this were done, the total binding observed would be the combined result of *cis* and *trans* binding, and there is no rigorous way of measuring the *cis* binding alone.

Nevertheless, TP47 is an unusual mutant for two reasons: (a) it manifests its IE preference only *in cis*, and (b) it increases transposition and DNA-binding through a Pro substitution for Thr. One possible model that can be used to explain (a) is that the conformational change introduced by Pro47 that makes the Tnp optimal for interaction with IE may be unstable; perhaps only newly translated Tnp takes on such a conformation (thus the *cis* preference). Soon afterwards, a conformational change happens that renders the TP47 Tnp only moderately elevated in its ability to interact with both OE and IE, compared to wt Tnp. Note that such a mechanism is invoked here to explain the excessive *cis* preference of the TP47 Tnp in combination with IE, and does not replace the previously proposed dimerization inactivation model (Weinreich *et al*., 1994), which could explain the general *cis* preference of wt and other mutant Tnp proteins, including the *cis* preference of TP47 Tnp, when using two OEs.

Pro is generally thought to be a secondary structure breaker in proteins. The fact that TP47 enhances DNA-binding and transposition seems at first glance very peculiar. A.S. Silbergleit, V.A. Lanzov and N. Benuch (personal communication) have identified a possible helix-turn-helix motif in the Tnp N-terminal region, using a self-developed motif searching program that is similar to the neural network programs for protein secondary structure predictions (e.g. Rost & Sander, 1993). This motif is:



According to this model, Thr47 is the second residue from the N terminus of the second helix (the recognition helix in known HTH motifs). It is of interest that Richardson & Richardson (1988) reported that among 215 α-helices from 45 different globular proteins whose three-dimensional structures had been well defined, Pro was found to be preferred at (and only at) the N1 position (second residue) of an α-helix, and Thr was less preferred compared to Pro at this position. The relative frequency of Pro *versus* Thr at this position was 2.6:0.8, after being normalized by the relative abundance of the two amino acids. According to Richardson & Richardson (1988), Pro in the first turn of a helix fits well in terms of its own backbone conformation, and its rigid structure should have some stabilizing influence. In addition, Pro in the first turn should help block continuation of the helix in the NH_2-terminal direction (Richardson & Richardson, 1988). The TP47 substitution of Tnp may increase DNA-binding and transposition by stabilizing the α-helix in which it resides, and/or by slightly shifting the beginning position and/or orientation of the helix backbone.

Richardson & Richardson (1988) also confirmed previous studies (Ptitsyn, 1969; Shoemaker *et al*., 1987) that showed an asymmetrical preference for negatively charged side-chains in the first turn and positively charged side-chains in the last turn of an α-helix. This may be due to stabilizing interactions of these charged side-chains with the helix dipole, and/or the formation of side-chain to main-chain hydrogen bonding. According to their statistical data, Lys was more preferred than Glu at the last position (the C-cap position) of an α-helix, and His was more preferred than Tyr at the position just before the C-cap (the C1 position). Using these observations, we propose that YH41 and EK54 may increase DNA binding and transposition by stabilizing the respective α-helices in which they reside, and/or by slightly altering the ending positions of the respective helices.

It is formally possible that the increase observed in the DNA-binding activity in the gel retardation assays may be due to an increased proportion of the population of purified mutant Tnp molecules being generally active, rather than due specifically to an elevated DNA-binding activity. However, the fact that EK54 TnpΔ369 displays different effects in its binding activity to DNA containing OE and IE clearly demonstrates that for this mutant, it is the DNA-binding activity that is directly affected. The close proximity of the locations of the other three mutations to position 54 suggests that it is quite likely that they, too, directly affect the DNA-binding activity.

The fact that we were able to isolate hypertransposing Tnp mutants that bind DNA better than wt Tnp means that wt Tnp has a sub-optimal DNA-binding motif. This is likely to be the result of evolutionary selection, since optimal DNA-binding by Tnp (coupled with its nucleolytic activity causing DNA nicking and breaking) and subsequent high-level transposition causing frequent insertion mutations would be detrimental to the host cell, and therefore be disadvantageous to the transposable element itself. Indeed, many transposable elements seem to contain some form of down-regulation of their transposition frequency (Berg & Howe, 1989). Tn*5* transposition has been shown to be tightly controlled at multiple steps including *tnp* promoter activity, IE availability (both of which are regulated by host *dam* methylation), the abundance of the Inh protein, and a mechanism preventing translation of readthrough transcripts of *tnp* initiated from fortuitous promoters upstream (reviewed by Berg, 1989; Reznikoff, 1993). Our results suggest that Tn*5* transposition

is also down-regulated by an inefficient Tnp-end sequence binding reaction.

It appears that wt Tnp evolved to recognize OE and IE at a subtly balanced frequency to give both Tn*5* and IS*50* transposition. Transposition of Tn*5* propagates the antibiotic resistances that it carries, while transposition of IS*50* presumably can lead to the evolution of new Tn*5*-like elements. This could be another explanation of why Pro47 or Lys54 was not found in wt Tnp.

Finally, the approach utilized in this study may be useful in other systems in which the same protein binds to different sites, for identifying specific contacts between the protein and its targets. Such a genetic and biochemical approach can often complement a structural approach, and provide independent insight into the interaction between a protein and DNA.

Materials and Methods

Media and reagents

Bacterial strains were grown in LB for DNA cloning and mating-out assays. SOC medium (Sambrook *et al*., 1989) was used for all electroporations. Tryptone-phosphate medium (Moore *et al*., 1993) was used for overexpression of TnpΔ369 and its mutant derivatives, both for preparing crude extracts and for protein purification. Glucose minimal Miller medium (Miller, 1972) containing 0.3% (w/v) Casamino Acids, 40 μg/ml of 5-bromo-4-chloro-3-indolyl-β-D-galactoside, and 0.05% (w/v) phenyl-β-D-galactoside (referred to here as Trp^--XG-PG plates) was used for papillation assays. Antibiotics were purchased from Sigma and used in the following concentrations: ampicillin, 100 μg/ml; kanamycin, 40 μg/ml; tetracycline, 15 μg/ml; chloramphenicol, 20 μg/ml; nalidixic acid, 20 μg/ml; and gentamycin, 5 μg/ml. Restriction endonucleases were obtained from New England Biolabs and Promega Corp. Taq DNA polymerase (for PCR mutagenesis) and phage T4 DNA ligase were from Promega Corp. Sequenase 2.0 was from United States Biochemical Corp. AMV reverse transcriptase was from Molecular Genetic Resources. Calf intestinal alkaline phosphatase was from Boehringer Mannheim. All oligonucleotides were synthesized on an Applied Biosystems (Foster City, CA) model 391 DNA oligonucleotide synthesizer. Radioactive nucleotides were purchased from Amersham.

Bacterial strains and plasmids

All bacterial strains and plasmids used in this study are described in Table 1. All strains were derivatives of *Escherichia coli* K-12, except for the B-strain BL21(DE3) pLysS. The structures of some of the plasmids are shown in Figure 2. All cloning was done using strain DH5α. The results of all cloning and site-directed mutagenesis manipulations were confirmed by DNA sequence analysis.

Plasmid pRZ5412 that encodes full-length Tnp was constructed as follows: the small *Nhe*I-*Bam*HI fragment of pRZ3271-26G containing the part of the *tnp* gene encoding the C terminus of Tnp was substituted with the corresponding *Nhe*I-*Bam*HI fragment from pRZ7055, resulting in the insertion of a kanamycin resistance gene between the end of the *tnp* gene and IE. The resulting plasmid is called pRZ5400. An *Afl*II restriction site was introduced just upstream of the *tnp* gene but within the IS*50*R sequence in pRZ5400 by site-directed-mutagenesis (Kunkel *et al*., 1987), using the oligonucleotide sequence 5′ ACGTTACCATCTTAAGAGGTCACCT 3′. The strain CJ236 was used to prepare the single-stranded phagemid DNA for the mutagenesis procedure. An *Eco*RI-*Hpa*I fragment from the resulting plasmid containing the *Afl*II site was subcloned into the unmutagenized pRZ5400 to create pRZ5404. This ensured that no unintended mutation was introduced into the plasmid during the course of site-directed mutagenesis. Finally, pRZ5412 was constructed by substituting the small *Hpa*I-*Not*I fragment of pRZ5404 with the corresponding fragment from pRZ7016, introducing the MA56 mutation, which changes the start codon AUG of the *inh* gene into GCC, thereby eliminating Inh synthesis (Wiegand & Reznikoff, 1992). This unavoidably also changes the Met56 residue of Tnp into Ala56, which according to Wiegand & Reznikoff (1992), does not greatly alter Tnp function.

pFMA187(OE/OE), which was one of the plasmids used in *in vivo cis* transposition assays, was constructed by substituting the IE sequence in pFMA187(OE/IE) with an OE-containing sequence *via* site-directed mutagenesis (Kunkel *et al*., 1987). The oligonucleotide used in this mutagenesis was 5′ TATCATCGATAATACTTGTG-TATAAGAGTCAGGCGATTTGCAATTCCCCGGAT 3′. Each of the *tnp* gene single mutations that resulted in YH41, TP47, EK54 and EV54, was then subcloned from their original pRZ5412 vector into each of the three versions of pFMA187: (OE/IE), (IE/IE) and (OE/OE), between the corresponding *Hpa*I site and *Not*I site. Note that the MA56 mutation is present in all the original and derived plasmids in this series. pRZ5452, pRZ5453, and pRZ5454 were constructed by deleting the DNA between the two *Bsp*HI sites flanking the *tnp* gene in pFMA187(OE/IE), (IE/IE) and (OE/OE), respectively, resulting in the complete loss of the *tnp* sequence, while maintaining the OE and IE sequences.

pRZ9000 encodes the N-terminal 368 amino acid residues of Tnp followed by a glycine residue, under the control of the T7 promoter in a pET-21d(+) based vector. The *tnp* mutations YH41, TP47, EK54 and EV54 were each subcloned into pRZ9000 from their original pRZ5412 vector, between a *Hpa*I site and a *Bsm*I site in the *tnp* gene. Again, the MA56 mutation is present in all constructs.

PCR random mutagenesis

PCR reactions were carried out using two different reaction conditions: (a) similar to that described by Zhou *et al*. (1991), including, in a 100 μl reaction volume, 10 mM Tris-HCl (pH9 at 25°C), 50 mM KCl, 0.1% (v/v) Triton X-100, 1.5 mM $MgCl_2$, 50 μM of each dNTP, 2 to 3 fmol of template, 30 pmol of each primer, and five units of Taq DNA polymerase. (b) Same as (a) except containing 200 μM of each dNTP, and 250 μM $MnCl_2$. pRZ5412 (encoding the *tnp* gene) was used as template. The phage T7 primer (New England Biolabs, catalog no. 1248; which hybridizes upstream of the *tnp* gene), and the p801 primer (which corresponds to nucleotides 820 to 801 of the IS*50*R sequence) were used as the two primers to synthesize a PCR product of ~880 bp in length. An *Afl*II-*Not*I fragment of this PCR product was substituted for the corresponding fragment of pRZ5412. The ligation product was electroporated into strain MDW320, and screened for transposition mutants by a papillation assay (see below). After isolation and purifi-

cation of the mutants, the entire region between the *Afl*II and the *NotI* site in each of the mutants was sequenced.

Isolation of tnp mutants

The papillation assay (initially described by Krebs & Reznikoff, 1988) was used to screen for *tnp* mutants causing increased or decreased transposition frequency involving two OEs. In this assay, the frequency of transposition events is approximated by the accumulation of Lac^+ papillae in an otherwise Lac^- colony. The strain MDW320 contains a papillation factor pOXGen386, which is a modified F factor containing two OEs in inverted orientation flanking the *lacZYA′* genes and a tetracycline resistance gene. The *lacZ* gene lacks its transcriptional and translational initiation signals, thus the strain forms white colonies on Trp^--XG-PG plates. When pRZ5412 is electroporated into MDW320, the Tnp encoded by pRZ5412 causes occasional transposition of the sequence on pOXGen386 flanked by the two OEs. The events that fused the *lacZ* gene to an actively transcribed gene in the correct reading frame will manifest themselves as blue papillae on an otherwise white colony. The rate of appearance of these papillae is a measurement of the transposition frequency induced by the wt or mutant alleles of the *tnp* gene on pRZ5412.

We carried out the papillation assay similarly to that described by Krebs & Reznikoff (1988), except that our assay was performed *in trans*, the Trp^--XG-PG plates (with tetracycline and ampicillin) were used instead of lactose MacConkey plates for better sensitivity of papillae detection, and the plates were incubated at 32°C instead of 37°C, since we found that papillation is much more efficient at 32°C than at 37°C, 42°C or 28°C (data not shown).

Mating-out assay

The mating-out assay is performed essentially as described by Yin *et al*. (1988). It is more quantitative than the papillation assay. It measures the movement of a transposable element from a plasmid to the F factor in the cell. After transfer of the F factor into recipient cells through conjugation, the total number of exconjugants is compared with the number of exconjugants that have received the transposon. The ratio of the two is a measurement of the transposition frequency.

In this study, the *tnp* gene is present either *in cis* (as in pFMA187OE/OE, OE/IE and IE/IE; Table 1 and Figure 2), or *in trans* (as in pRZ5412), of the DNA that contains the end sequences flanking an antibiotic resistance marker (i.e. the transposable element whose movement is measured in this assay). In the *trans* situation, the ends are contained on pRZ5452, pRZ5453 or pRZ5454 (Table 1). The appropriate antibiotics were used to select for exconjugants that have or have not received the transposable element. The donor strain was JCM101/pOX38-Gen, which is a *dam* strain that contains the F factor derivative pOX38-Gen (Johnson & Reznikoff, 1984). The recipient (F^-) strain was 14R525, which is a nalidixic acid resistant prototroph.

Preparation of crude cell extracts and transposase purification

Crude extracts were prepared from BL21(DE3) pLysS strain (Studier *et al*., 1990) essentially as described (Weinreich *et al*., 1994) except that cells were grown in tryptone-phosphate medium (Moore *et al*., 1993) containing chloramphenicol and ampicillin instead of LB, 0.1 mM IPTG was used for induction, and the sonicate was centrifuged for 15 instead of five minutes at 4°C in a microfuge. The crude extracts were compared by SDS-PAGE. The Coomassie blue-stained gels were scanned by a densitometer. The wt and mutant TnpΔ369 band were compared, and an appropriate volume of each extract containing equivalent amount of the TnpΔ369 was used in the DNA-binding assay.

TnpΔ369 (wt and mutant) purification was carried out once according to the procedures described by de la Cruz *et al*. (1993) and once with a modified procedure. The previously described procedure includes induction with IPTG, harvesting, cell lysis with a French press, polyethylenimine (PEI) precipitation, 47% saturation $(NH)_2SO_4$ precipitation, dialysis to remove $(NH)_2SO_4$, and heparin-agarose column chromotography. In our modified protocol, precipitation with PEI was omitted because we observed that the majority of the overexpressed TnpΔ369 was lost at this step, rendering the protein we intended to purify a minor species in the supernatant after precipitation. We also did a 40% saturation $(NH)_2SO_4$ precipitation instead of 47%, since we found that 40% is more effective in purifying TnpΔ369 against other proteins, compared to higher concentrations. Finally, we used a high-affinity heparin-agarose matrix (Sigma catalog no. H0402, made by end-point attachment of heparin on 4% beaded agarose) that allowed TnpΔ369 to be recovered more efficiently. The equilibration and elution buffers used for the column contained 10% instead of 20% (v/v) glycerol described by de la Cruz *et al*. (1993).

The concentration of EK54 TnpΔ369 was calculated by using the Bradford (1976) assay to determine total protein concentration and densitometric scanning of a Coomassie-stained SDS-PAGE gel to determine the percentage of EK54 TnpΔ369 in the sample. The concentrations of other purified wt and mutant TnpΔ369 proteins were determined relative to EK54 TnpΔ369 in the same SDS gels. Each of the TnpΔ369 derivatives was ≈ 85% pure.

Gel shift assay

This was performed essentially as described (Wiegand & Reznikoff, 1992). For crude extract binding, about 2 ng (55 fmol) of a ^{32}P-labeled 53 bp DNA fragment containing the OE isolated from dam^+ cells, or a 56 bp DNA fragment containing the IE isolated from *dam* cells was incubated with crude cell extracts containing equivalent amounts of wt or mutant TnpΔ369 in a 24 μl reaction volume in a buffer containing 20 mM sodium phosphate (pH7.5), 100 mM potassium glutamate, 1 mM DTT, 0.5 mM EDTA, 0.1% Triton X-100, 200 μg/ml bovine serum albumin, 10 μg/ml unlabeled calf thymus DNA as non-specific competitor. After incubation at 30°C for 30 minutes, 6 μl of a loading solution containing 20% glycerol and 0.05% (w/v) xylene cyanol was added before loading 25 μl of each reaction onto a 1.2 mm thick, 20 cm long, 8% (39:1, w/w acrylamide to bis-acrylamide) polyacrylamide gel. Electrophoresis was carried out at 4°C and 17 V/cm for three hours in a running buffer that contained 45 mM Tris-borate, 1.25 mM EDTA. Gels were dried, quantified with a Betascope 603 blot analyzer (Betagen Corporation, Waltham, MA), and exposed to X-ray film overnight.

For purified TnpΔ369 binding, about 0.4 ng (11 fmol, final concentration 0.4 nM) of the same labeled DNA as above was incubated with various amounts of purified wt or mutant TnpΔ369 in a 27 μl reaction volume and same buffer composition as described above except with a tenfold higher concentration of unlabeled calf thymus DNA (100 μg/ml). Using 2 ng of the labeled DNA (as was done for the crude extract binding) yielded equivalent results (data not shown). After the incubation at 30°C for 30 minutes, 3 μl of a loading solution (50% glycerol, 0.02% xylene cyanol) was added, and subsequent steps were performed exactly as described above. Data were quantified with a Phosphorimager (Molecular Dynamics).

Acknowledgments

We thank Igor Y. Goryshin for helpful discussion, technical assistance and frequent encouragement, Dona York for help in protein purification, Michael D. Weinreich for technical advice, and the rest of the Reznikoff laboratory for suggestions. This work was supported by NIH grant GM50692 and a gift from the Promega Corporation.

References

Berg, D. E. (1989). Transposon Tn*5*. In *Mobile DNA* (Berg, D. E. & Howe, M. M., eds), pp. 185–210, American Society for Microbiology, Washington, DC.

Berg, D. E. & Howe, M. M. (1989). Editors of *Mobile DNA,* American Society for Microbiology, Washington, DC.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal Biochem.* **72**, 248–254.

Burt, D. W. & Brammar, W. J. (1982). The *cis*-specificity of the Q-gene product of bacteriophage lambda. *Mol. Gen. Genet.* **185**, 468–472.

de la Cruz, N. B., Weinreich, M. D., Wiegand, T. W., Krebs, M. P. & Reznikoff, W. S. (1993). Characterization of the Tn*5* transposase and inhibitor proteins: a model for the inhibition of transposition. *J. Bacteriol.* **175**, 6932–6938.

Delong, A. & Syvanen, M. (1991). *Trans*-acting tranposase mutant from Tn*5*. *Proc. Natl Acad. Sci. USA,* **88**, 6072–6076.

Derbyshire, K. M., Kramer, M. & Grindley, N. D. (1990). Role of instability in the *cis* action of the insertion sequence IS903 transposase. *Proc. Natl Acad. Sci. USA,* **87**, 4048–4052.

Dodson, K. W. & Berg, D. E. (1989). Saturation mutagenesis of the inside end of insertion sequence IS*50*. *Gene,* **85**, 75–81.

Ebright, R. H. (1991). Identification of amino acid-base pair contacts by genetic methods. *Methods Enzymol.* **208**, 620–640.

Echols, H., Court, D. & Green, L. (1976). On the nature of *cis*-acting regulatory proteins and genetic organization in bacteriophage: the example of gene Q of bacteriophage λ. *Genetics,* **83**, 5–10.

Fuller, R. S., Funnell, B. E. & Kornberg, A. (1984). The DnaA protein complex with the *E. coli* chromosomal replication origin (*ori*C) and other DNA sites. *Cell,* **38**, 889–900.

Goryshin, I. Y., Kil, Y. V. & Reznikoff, W. S. (1994). DNA length, bending, and twisting constraints on IS*50* transposition. *Proc. Natl Acad. Sci. USA,* **91**, 10834–10838.

Hanahan, D. (1985). Techniques for transformation of *E. coli*. In *DNA cloning* (Glover, D. M., ed.), pp. 109–135, IRL Press, Oxford, UK.

Isberg, R. R., Lazaar, A. L. & Syvanen, M. (1982). Regulation of Tn*5* by the right repeat proteins: control at the level of the transposition reactions? *Cell,* **30**, 883–892.

Jain, C. & Kleckner, N. (1993). Preferential cis action of IS*10* transposase depends upon its mode of synthesis. *Mol. Microbiol.* **9**, 249–260.

Jilk, R. A., Makris, J. C., Borchardt, L. & Reznikoff, W. S. (1993). Implications of Tn*5*-associated adjacent deletions. *J. Bacteriol.* **175**, 1264–1271.

Jilk, R. A., York, D. & Reznikoff, W. S. (1996). The organization of the outside end of transposon Tn*5*. *J. Bacteriol.* **178**, 1671–1679.

Johnson, R. C. & Reznikoff, W. S. (1983). DNA sequences at the end of transposon Tn*5* required for transposition. *Nature,* **204**, 280–282.

Johnson, R. C. & Reznikoff, W. S. (1984). Copy number control of Tn*5* transposition. *Genetics,* **107**, 9–18.

Johnson, R. C., Yin, J. C. & Reznikoff, W. S. (1982). Control of Tn*5* transposition in *Escherichia coli* is mediated by protein from the right repeat. *Cell,* **30**, 873–882.

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

Krebs, M. P. & Reznikoff, W. S. (1988). Use of a Tn*5* derivative that creates *lacZ* translational fusions to obtain a transposition mutant. *Gene,* **63**, 277–285.

Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987). Rapid and efficient site specific mutagensis without phenotypic selection. *Methods Enzymol.* **154**, 367–382.

Makris, J. C., Nordmann, P. L. & Reznikoff, W. S. (1988). Mutational analysis of insertion sequence 50 (IS*50*) and transposon 5 (Tn*5*) ends. *Proc. Natl Acad. Sci. USA,* **85**, 2224–2228.

Miller, J. (1972). *Experiments in Molecular Genetics,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Moore, J. T., Uppal, A., Maley, F. & Maley, G. F. (1993). Overcoming inclusion body formation in a high-level expression system. *Protein Expr. Purif.* **4**, 160–163.

Ptitsyn, O. B. (1969). Statistical analysis of the distribution of amino acid residues among helical and non-helical regions in globular proteins. *J. Mol. Biol.* **42**, 501–510.

Reznikoff, W. S. (1993). The Tn*5* transposon. *Annu. Rev. Microbiol.* **47**, 945–963.

Richardson, J. S. & Richardson, D. C. (1988). Amino acid preferences for specific locations at the ends of α helices. *Science,* **240**, 1648–1652.

Rost, B. & Sander, C. (1993). Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.* **232**, 584–599.

Sakai, J. & Kleckner, N. (1996). Two classes of Tn*10* transposase mutants that suppress mutations in the Tn*10* terminal inverted repeat. *Genetics,* **144**, 861–870.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual,* pp. A2, Cold

Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sasakawa, C., Carle, G. F. & Berg, D. E. (1983). Sequences essential for transposition at the termini of IS*50*. *Proc. Natl Acad. Sci. USA,* **80**, 7293–7297.
Schulz, V. P. & Reznikoff, W. S. (1991). Translation initiation of IS*50*R read-though transcripts. *J. Mol. Biol.* **221**, 65–80.
Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M. & Baldwin, R. L. (1987). Tests of the helix dipole model for stabilization of α-helices. *Nature,* **326**, 563–567.
Studier, F. 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.
Tomcsanyi, T. & Berg, D. E. (1989). Effects of adenine (Dam) methylation on transposition activity of O end mutants of IS*50*. *J. Mol. Biol.* **209**, 191–193.
Weinreich, M. D. & Reznikoff, W. S. (1992). Fis plays a role in Tn*5* and IS*50* transposition. *J. Bacteriol.* **174**, 4530–4537.
Weinreich, M. D., Mahnke-Braam, L. & Reznikoff, W. S. (1993). A functional analysis of the Tn*5* transposase, identification of domains required for DNA binding and multimerization. *J. Mol. Biol.* **241**, 166–177.
Weinreich, M. D., Gasch, A. & Reznikoff, W. S. (1994). Evidence that the *cis* preference of the Tn*5* transposase is caused by nonproductive multimerization. *Genes Devel.* **8**, 2363–2374.
Wiegand, T. W. (1993). Transposase mutants that increase the transposition frequency of Tn*5*, PhD thesis, University of Wisconsin-Madison.
Wiegand, T. W. & Reznikoff, W. S. (1992). Characterization of two hypertransposing Tn*5* mutants. *J. Bacteriol.* **174**, 1229–1239.
Wiegand, T. W. & Reznikoff, W. S. (1994). Interaction of Tn*5* transposase with the transposon termini. *J. Mol. Biol.* **235**, 486–495.
Yin, J. C. & Reznikoff, W. S. (1987). *dnaA*, an essential host gene, and Tn*5* transposition. *J. Bacteriol.* **169**, 4637–4645.
Yin, J. C., Krebs, M. P. & Reznikoff, W. S. (1988). Effect of *dam* methylation on Tn*5* transposition. *J. Mol. Biol.* **199**, 35–45.
York, D. & Reznikoff, W. S. (1996). Purification and biochemical analyses of a monomeric form of Tn*5* transposase. *Nucl. Acids Res.* **24**, 3790–3796.
York, D. & Reznikoff, W. S. (1997). DNA-binding and phasing analyses of Tn*5* transposase and a monomeric variant. *Nucl. Acids Res.* **25**, 2153–2160.
Zhou, H. M. (1997). Interactions between Tn*5* transposase and Tn*5* end sequences: study of the DNA binding step. Phd thesis, University of Wisconsin-Madison.
Zhou, Y., Zhang, X. & Ebright, R. H. (1991). Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase. *Nucl. Acids Res.* **19**, 6052.

Edited by G. Smith

(Received 11 February 1997; received in revised form 4 June 1997; accepted 4 June 1997)