Purification and biochemical analyses of a monomeric form of Tn5 transposase

Dona York and William S. Reznikoff*

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

Received May 31, 1996; Revised and Accepted August 22, 1996

ABSTRACT

The binding of transposase (Tnp) to the specific Tn5 end sequences is the first dedicated reaction during transposition. In this study, comparative DNA-binding analyses were performed using purified full-length Tnp and a C-terminal deletion variant (\triangle 369) that lacks the putative dimerization domain. The shape of the binding curve of full-length Tnp is sigmoidal in contrast to the hyperbolic-shaped binding curve of \triangle 369. This observation is consistent with previous observations as well as a rate of binding study presented here, which suggest that the full-length Tnp-end interaction, unlike that of the truncated protein, is a complex timedependent reaction possibly involving a subunit exchange. Circular permutation assay results indicate that both proteins are capable of distorting the Tn5 end sequences upon binding. Molecular weight determinations based on the migratory patterns of complexed DNA in polyacrylamide gels has shown that \triangle 369 specifically binds the Tn5 end sequences as a monomer while full-length Tnp in complex represents a heterodimer.

INTRODUCTION

Tn5 is a composite transposon composed of two insertion sequences: IS50R and IS50L that flank antibiotic resistance genes. The functional Tn5 transposase (Tnp) is expressed from IS50R (Fig. 1). The inhibitor protein (Inh), translated from the same reading frame, lacks the N-terminal 55 amino acids. Tnp is a *cis*-active transposase that specifically recognizes two unique, 19 bp end sequences (OE and IE) positioned at the termini of each IS50. With the exception of position 4, the first 9 bp of each site are identical. The non-identical sections of each site contain binding sites for host proteins (for a review see ref. 1).

The critical initial step of Tn5 transposition requires the specific recognition of end sequences by the Tnp protein. Two distinct Tnp–OE complexes have been observed in gel retardation assays: Complex I and the faster migrating Complex II (2). Depending on the Tnp protein preparations used, wt Tnp or Tnp MA56 (which eliminates production of the Inh protein), the proteins present in Complex I include either Tnp, Inh and a naturally-occurring N-terminal deletion product, Tnp α , or Tnp

and Tnp α , respectively (Fig. 1). Since neither Inh nor Tnp α can themselves form specific nucleoprotein complexes with the OE, their presence in Complex I indicates a protein-protein interaction with full-length Tnp. The migratory pattern of Complex II, shown to represent binding of C-terminal deletion products, Tnp γ and Tnp δ (Fig. 1), is indicative of a change in the oligometric states of these proteins when complexed with the OE (2). The protein composition of Complex I as well as a number of other observations have suggested that the full-length Tnp-OE interaction is a complicated reaction. Addition of Inh protein stimulates the binding activity of Tnp to the OE presumably through a protein–protein interaction (3). In addition, a prolonged incubation time increases the amount of Tnp-OE complex formed (2). This may be related to a disaggregation and subunit exchange between full-length Tnp and Inh resulting in heterooligomer formation.

As has been demonstrated for a number of transposases from other systems (4–6), Tnp also distorts the OE DNA upon binding as determined by a circular permutation assay (7). This distortion is the result of bending as determined by phasing analysis (York and Reznikoff, in preparation). The apparent bending angle is >100° and centers near the first to third nucleotide of the 19 bp OE fragment (7).

Preliminary domain mapping of the Tn5 Tnp protein was accomplished by deletion analysis. Various restriction fragments of the Tnp gene were used in an *in vitro* transcription/translation experiment to generate a family of C-terminal variants of Tnp protein (8). DNA-binding analysis with these Tnp products have suggested that deletion of the C-terminal 107 amino acids restricts Tnp oligomerization. This Tnp variant, $\Delta 368$, also demonstrated an apparent increased binding affinity. This result implies that the C-terminus of the protein may partially block the DNA-binding domain of the protein, therefore reducing the overall DNA binding affinity for the OE.

The goal of this paper is to further characterize the truncated variant of transposase *in vitro*. A T7 expression vector was made to facilitate the overproduction of a transposase variant lacking 107 amino acids from the C-terminus. The protein was purified using a modified procedure previously reported. Purified preparations of this protein, designated Δ 369, and full-length Tnp were tested in gel retardation assays for binding affinity and ability to bend the OE. Molecular weight determinations based on the migration patterns of complexed DNA in polyacrylamide gels using both the truncated and full-length protein were also

^{*} To whom correspondence should be addressed

MATERIALS AND METHODS

Strains and plasmids

Escherichia coli strains DH5 α and BL21(DE3) pLysS were used for plasmid isolation and transposase overexpression, respectively. Plasmids pRZ7074 MA56, pRZ7067 (2), pRZ4826 (8) and pRZ9012 (7) were described previously. pRZ9000 (this study) is described below.

Plasmid construction

Plasmid pRZ9000 used for overexpression of the truncated Tnp, was constructed from pRZ7074 (2) and pRZ4826 (8). pRZ7074 (2), is a pET21d derivative which contains the entire Tnp gene (with MA56 mutation that eliminates Inh production) under control of the IPTG-inducible T7 promoter. Digestion with HindIII results in two fragments: a 6448 bp fragment containing the T7 promoter region and ampicillin resistance gene and the amino-half of the Tnp gene (encoding the first 368 amino acids) and a 1900 bp fragment containing the carboxy-half of the Tnp gene and the rest of the vector. These fragments were treated with Mung Bean nuclease (New England Biolabs) according to manufacturer's instructions and the large fragment was isolated and ligated to a 593 bp SmaI-NruI DNA fragment, isolated from plasmid pRZ4826 (8), which includes stop codons in all three reading frames at the 5' end. The resulting construct, pRZ9000, was confirmed by sequence analysis and used for the overexpression of the truncated form of Tnp [368 amino acids from Tnp and an additional codon (gly) from the stop cassette].

Purification of truncated ($\Delta 369)$ and full-length Tnp protein

An overnight culture of BL21 (DE3) pLysS containing pRZ9000 MA56 (eliminates production of Inh protein) was used to inoculate 2 l of tryptone-phosphate broth (2% bacto-tryptone, 0.2% Na₂HPO₄, 0.1% KH₂PO₄, 0.8% NaCl and 2% yeast extract) (9). Cells were grown at 37°C to an OD₆₀₀ of 0.5. Protein overexpression was induced with IPTG at a final concentration of 0.1 mM. After an additional 1.5 h incubation, cells were harvested and the protein purified as previously described (10). The homogeneity of Δ 369 was analyzed on a denaturing 10% SDS–PAGE gel followed by Coomassie staining. Full-length Tnp protein, overexpressed and purified from BL21 (DE3) pLysS cells containing pRZ7074 MA56, was a gift from Maggie Zhou. This full-length Tnp preparation, as analyzed on a Coomassie-stained SDS–PAGE, was determined to be ~90% pure (Fig. 2).

Gel retardation assay

The binding affinity assay was performed using an α -³²P-labeled 52 bp *Eco*RI–*Hin*dIII fragment from pRZ7074 MA56 (2). Reaction mixtures of 15 µl [100 mM potassium glutamate, 25 mM Tris-acetate (pH 7.5), 10 mM magnesium acetate, 0.5 mM β -mercaptoethanol, 400 mg/ml bovine serum albumin (BSA), 600 ng tRNA and 20% glycerol] containing 1 nM (0.5 ng) of labeled probe were incubated for 30 min at 30°C with increasing amounts of purified full-length Tnp MA56 and Δ 369 MA56 (0, 62.5, 125, 625 and 1250 nM). 3.5 µl of 20% glycerol/0.1% bromophenol blue/0.1% xylene cyanol FF was added to each sample and loaded onto an 8% non denaturing 0.5× TBE polyacrylamide (29:1) gel. The gel was run at 300 V, 4°C for 2 h, dried and exposed to film overnight. The percentage of free and bound DNA was determined by a Molecular Dynamics Phosphorimager, ImageQuant software.

For the rate of binding studies, 625 nM amounts of full-length Tnp and Δ 369 were incubated for various times (0, 15, 30, 60 and 120 min) with 1 nM (0.5 ng) of the 52 bp OE probe in 15 µl of binding buffer (see above). Samples were processed as described above.

Molecular weight determination

Molecular weight determination was performed as described by Orchard and May (11). The 52 bp OE fragment was incubated with 500 nM of full-length Tnp and Δ 369, as described above, and run on a series of $0.5 \times \text{TBE}$ polyacrylamide gels (5–10%). Five micrograms of protein standards: lactalbumin (MW 14 200), carbonic anhydrase (MW 29 000), chicken egg albumin (MW 45 000), and BSA, which runs as a monomer (MW 66 000) and a dimer (MW 132 000) were also run simultaneously on each set of gels. For each gel the position of the bromophenol blue dye was marked. The half of each gel containing the binding reactions was dried and exposed to film overnight while the protein standard half of each gel was stained with Coomassie. The migration distances of each complex as well as each protein standard was measured and divided by the distance the bromophenol blue migrated in the same sample (relative mobility = $R_{\rm f}$). The calibration curves from the known protein standards as well as those of the full-length Tnp–OE and Δ 369–OE complexes were derived from a plot of $100[\log(R_f \times 100)]$ as a function of gel concentration (%). The Ferguson plot is derived from the negative slopes of the calibration curves of each of the protein standards as a function of the known molecular weights and used to determine the molecular weight of Tnp protein complexes.

Circular permutation assay

Gel retardation assays were performed with four circularly permutated 182 bp fragments from pRZ9012 containing the 19 bp OE consensus sequence as described previously (7). The apparent bend angle was computed by using the formula of Zhou *et al.* (13), a derivative of the Thompson and Landy bending equation (14). This equation is as follows;

$$\frac{\mu_i}{\mu_j} = \frac{[1-2(\chi_i./L)(1-\cos\alpha) + 2(\chi_i./L)^2(1-\cos\alpha)]^{0.5}}{1-2(\chi_j./L)(1-\cos\alpha) + 2(\chi_j./L)^2(1-\cos\alpha)]^{0.5}}$$

where μ_i and μ_j are the relative mobilities of the complexes with Tnp bound to a centrally positioned OE and a peripherally positioned OE respectively, χ_i/L and χ_j/L are the fractional distances from the center of bending (+3) to the left end of the bending probes of the same OEs, and α is the angle of the bend.

RESULTS

Purification of a truncated form of Tn5 transposase

The binding properties of the Tnp variant lacking the C-terminal 107 amino acids suggested that the protein alters both DNA-binding



Figure 1. The structure of IS50R. The top shows the organization of the 1534 bp IS50R which encodes two proteins: the 476 amino acid multifunctional transposase (Tnp) protein and the inhibitor (Inh). The Inh, read from the same reading frame as Tnp, lacks the first 55 amino acids. Three proteolytic products of Tnp that are a result of overexpression and purification are indicated above IS50R. Tnp α lacks 20–25 amino acids from the N-terminus while Tnp δ and Tnp γ are a result of cleavage in the C-terminus. The bottom of the figure shows the 19 bp sequence of the OE and IE. Consensus sequences of host proteins within each sequence have been indicated by lines or arrows.

activity and oligomerization. Since DNA-binding and Tnp oligomerization are essential steps in transposition, this study focused on more detailed biochemical analyses of this truncated Tnp variant. A T7 overexpression vector containing the portion of the Tnp MA56 gene (which eliminates Inh production) encoding the first 368 amino acids was constructed (pRZ9000) as described in Materials and Methods. This truncated protein, designated Δ 369, when overexpressed under conditions previously outlined for full-length Tnp (10), formed insoluble inclusion bodies (data not shown). Denaturation and renaturation of these insoluble aggregates were not successful therefore attempts were made to increase the solubility of the protein by modifying overexpression conditions. Using a procedure outlined by Moore *et al.* (9), Δ 369 was overexpressed in BL21 (DE3) pLysS cells grown in a tryptone-phosphate broth as described in Materials and Methods. The soluble fraction of the $\Delta 369$ protein, ~10–25% of the total Δ 369 protein (Fig. 2), was purified by a previously reported protocol (10). The homogeneity of the peak fraction was determined by a denaturing SDS-PAGE gel to be ~85% (Fig. 2).

In vitro DNA binding analysis of full-length Tnp and $\triangle 369$

Purified preparations of full-length Tnp MA56 (lacking Inh protein) and Δ 369 MA56 were tested for DNA-binding activity by a gel retardation assay. Increasing amounts of full-length Tnp and Δ 369 protein preparations were incubated with a 52 bp fragment containing the 19 bp OE consensus sequence as described in Materials and Methods. As shown in Figure 3A, full-length protein forms the characteristic Complex I and the faster migrating Complex II (2), while Δ 369 forms a single faster migrating complex. This result mimics that reported for the *in vitro* transcribed/translated 368 amino acid residue long C-terminal deletion variant (8).

The DNA-binding activity of each protein was compared by generating a binding curve from a plot of the percent disappearance of OE DNA as a function of protein concentration (Fig. 3B).



Figure 2. Coomassie-stained 10% SDS gel of various stages of Δ 369 purification. Lane 1, the molecular weights in kDa of the proteins used as standards (std.) are denoted on the left; lane 2, 7.5µg of purified wt (full-length) Tnp transposase (MW 53 311); lane 3, uninduced cells containing pRZ9000; lane 4, IPTG-induced cells containing pRZ9000; lane 5, insoluble protein fraction; lane 6, soluble protein fraction; lane 7, 7.5µg of Δ 369 from the peak fraction off the heparin–agarose column (MW 41 400).

The relative binding affinities, $K_{observed}$ (K_{obs}), of each protein preparation was determined from the protein concentration at which 50% of the OE DNA was bound. Surprisingly, the K_{obs} of the $\Delta 369$ protein preparation is ~2-fold lower than that of full-length Tnp protein. This result was unexpected based on the previous binding analysis of *in vitro* transcription/translation products in which a nearly identical truncated protein demonstrated an apparent higher binding activity than that of full-length Tnp (8). This discrepancy will be addressed in the Discussion.

The shape of the two curves in Figure 3B generated by full-length Tnp and $\Delta 369$ can be an indication of the type of binding reaction for each protein. The hyperbolic binding curve of $\Delta 369$ is indicative of a simple binding reaction in which a preformed protein component specifically binds DNA in a one step reaction. However, the sigmoidal nature of the binding curve of full-length Tnp, suggests a more complex reaction in which protein-protein interactions play an important role. This observation is consistent with the results of previous studies which demonstrated the presence of more than one species of Tnp protein in OE complexes. A Western blot analysis of the two retarded complexes formed in the presence of full-length Tnp demonstrated that three proteins are present on an immunoblot of Complex I (2,3). These three proteins are full-length Tnp, Tnp α and the inhibitor (Inh) protein. The protein components of Complex I formed in the presence of full-length Tnp MA56, which lacks Inh protein (as used in this study), contain only Tnp and Tnp α . Tnp α is a naturally-occurring proteolytic product of Tnp lacking ~25 amino acids from the N-terminus (3) and like the Inh protein (15) has been shown to lack any specific DNA-binding activity (3). The presence of Tnp α in Complex I strongly suggests that it forms heteromultimers with full-length Tnp. Protein composition of the faster migrating Complex II was shown by Western analysis to be two other proteolytic products of Tnp: Tnp δ or γ . Both of these proteins are C-terminal deletion products of full-length Tnp. Based on the migratory position of



Figure 3. Gel retardation assay (A) and binding analysis (B) of wt (full-length) Tnp and Δ 369. (A) The wedge above the lanes for Wt Tnp and Δ 369 represents increasing amounts of protein. Complex I and Complex II formed by wt Tnp with the OE are shown by arrows. (B) Percent disappearance of OE DNA at each protein level was determined using a phosphorimager and plotted as a function of concentration.



Figure 4. Extent of binding as a function of incubation time. Labeled OE DNA was incubated at 30°C with a constant amount of Wt (full-length) Tnp or Δ 369 protein (625 nM) for the times indicated and used in a gel retardation assay as described in Materials and Methods. The extent of complex formation, as determined from a phosphorimage, was plotted as a function of time. This plot is the average of three independent experiments.

Complex II, it has been suggested that these proteins bind OE DNA in monomeric form (2).

DNA binding activity as a function of incubation time

Previous work had suggested that formation of Complex I increases with prolonged incubation (2). This observation is consistent with a subunit exchange between full-length Tnp aggregates and Tnp α or Inh. We decided to further explore this possibility by incubating full-length Tnp MA56 and Δ 369 with OE DNA. Binding reactions were carried out as outlined in Materials and Methods, however, incubation times were varied (0, 15, 30, 60 and 120 min). As shown in the gel retardation analysis in Figure 4, the formation of full-length Tnp-OE complexes increases with longer incubation times. In contrast, >50% of the $\Delta 369$ –OE complexes form almost instantaneously with the full extent of binding occurring within 15 min. It is of interest that the quantity of specific $\Delta 369$ -OE complex formation decreases with longer incubation times which may be a reflection of the instability of Δ 369–OE complex based in part on the inherent ability of the protein to aggregate.

Molecular weight determination

The oligomeric state of full-length Tnp MA56 (lacking Inh protein) and Δ 369 MA56 in complex with the OE was determined using a modified Ferguson analysis reported by Orchard and May (11,12). This method allows the molecular weight of protein–DNA complexes to be determined indirectly by running them on a series of non-denaturing polyacrylamide gels of differing acrylamide concentrations alongside a number of protein standards of known molecular weights. Mobility is therefore related to the sieving effects of the gel and consequently to the size and shape of the protein in complex. The molecular weight of an unknown protein in a DNA complex can be determined graphically from mobility differences.

The 52 bp fragment containing the 19 bp OE consensus sequence was incubated with a single concentration of full-length Tnp and Δ 369 under the conditions described in Materials and Methods. Reactions were run on a series of polyacrylamide gels (5-10%) alongside a set of protein standards and processed as described under Materials and Methods. The relative mobility $(R_{\rm f})$ of Complex I, Complex II and $\Delta 369$ –OE complex as well as each protein standard was determined for each gel concentration. A calibration curve was derived for each of the protein standards and each of the Tnp-OE complexes (Fig. 5A). The gradient of the calibration curve of the protein standards was plotted as a function of their known molecular weights on a log scale (Ferguson plot). From this linear plot (Fig. 5B) the approximate molecular weight of the protein-DNA complexes was determined and after subtracting the contribution of the OE DNA (MW 36 075), the molecular weight of the protein in complex was approximated. The protein component of Complex I (MW 110 000) represents a homodimer of full-length Tnp (MW 53 311) and/or a heterodimer of full-length and Tnp α (MW ~49 000/monomer). The protein component of Complex II (MW 44 000) most likely represents a monomeric form of Tnp δ and/or γ . The molecular weight of $\Delta 369$ (MW 41 473 based on amino acid content) in complex is ~41 000 which coincides with a monomeric form of the protein.

Circular permutation analysis

Recently it was shown that full-length Tnp bends the OE upon binding at an apparent bend angle >100°, centered at position 3 (7). Comparison analysis based on a circular permutation assay was performed using Δ 369 protein as well as full-length Tnp.



Figure 5. Molecular weight determination. (A) Calibration curve of molecular weight standards and Tnp–OE complexes. The logarithm of the relative mobility (R_f) is plotted against the percent gel concentration. A curve is generated for each molecular weight standard, Complex I (full-length Tnp), Complex II (full-length Tnp) and $\Delta 369$ (indicated in the graph legend). (B) Ferguson plot. The negative slope of each of the standard curves in (A) was plotted against the known molecular weights of the standards (open circles). Using this curve, the negative slope of each of the calibration curves from the Tnp-complexes (A) was applied to the standard curve (positions indicated by arrows) and the molecular weights of Complex I, Complex II and $\Delta 369$ were determined.

Four circularly permuted 182 bp fragments containing the 19 bp OE segment at various positions (Fig. 6A) were used in a gel retardation analysis with full-length and Δ 369 protein. As seen in Figure 6B, mobility of complexes is dependent on the position of the OE sequence in the fragment, an indication that both proteins bend the OE. The *R*_f of Complex I (full-length Tnp MA56) and the Δ 369–OE complex was measured and plotted as a function of the position of the center of each of the bending probes (Fig. 6A). These curves (Fig. 6C) allowed us to map the center of the full-length Tnp-induced and Δ 369-induced bend, located at the lowest point of the curve. For both these proteins, the lowest point is at position 135 which maps approximately to the third nucleotide in the 19 bp OE sequence.

A derivation of the Thompson and Landy (14) bending equation as formulated by Zhou *et al.* (13) was used to determine the angle of the bend (see Materials and Methods). Having determined the R_f of the center-most (*Eco*RV, fractional distance = 0.47) and end-most (*MluI*, fractional distance = 0.74) digest to be 0.358 and 0.462 for full-length Tnp and 0.551 and 0.695 for Δ 369, respectively, the angle of the bend was then calculated to be ~116° for both proteins. The same angle is found using the fractional distances and relative mobilities of the *NheI* [μ = 0.68, χ/L = 0.41 (full-length Tnp); 0.631 (Δ 369)] and *NruI* [μ = 0.34, χ/L = 0.40 (full-length Tnp); 0.61 (Δ 369)] digests. These results indicate that the monomeric form of Tnp bends the OE DNA in a similar manner as the full-length protein.

DISCUSSION

Tn5 transposase is a multifunctional protein that is an essential component in all the steps involved in transposition: site-specific binding to end sequences, synaptic complex formation through oligomerization, double-stranded cleavage of the DNA immediately adjacent to the Tn5 ends and subsequent cleavage of the

target DNA, and strand exchange of the free 3' ends to the 5' ends of the target. Preliminary domain mapping has suggested that the N-terminus contains the determinants for specific end recognition while the oligomerization domain lies in the C-terminus of the protein (8; Zhou and Reznikoff, in preparation). Amino acid homologies with other transposase and integrase proteins (16) suggest that the catalytic domain of Tnp is located near the central portion of the protein.

Previous binding analyses of a family of C-terminal deletion variants of Tnp has shown that deletion of the last 107 amino acids changes the oligomeric state of Tnp in complex with the OE. In addition, this deletion variant of Tnp demonstrated a higher apparent binding affinity for the OE suggesting that the C-terminus of the protein partially inhibits Tnp binding (8). This study focuses on further in vitro characterization of this protein variant. This truncated form of Tnp, Δ 369, was purified and analyzed *in* vitro in a comparison study with full-length Tnp protein. DNA-binding ability of each protein, tested in a gel retardation assay (Fig. 3), showed that there is only a 2-fold difference in the relative binding affinity of each protein preparation. This result is in sharp contrast with that reported in the deletion studies (8). However, this can be explained three different ways: (i) the protein concentrations used in the previous study may fall in the early part of the curve (Fig. 3B) where it appears that Δ 369 has a four to five times higher binding affinity for the OE in comparison with full-length Tnp protein, (ii) since full-length Tnp binding is time dependent (Fig. 4), the incubation time used in the previous study may not have been sufficient to observe the full extent of binding, and (iii) the availability of N-terminal proteolytic products, which seem to have a direct effect on full-length Tnp-DNA complex formation, may be significantly lower in the Tnp protein preparation produced in the eukaryotic in vitro transcription/translation system.



Figure 6. Bending of OE DNA by Tnp. (**A**) DNA probes used in OE bending experiments. The top line shows the 256 bp *MluI–NruI* region of the pBend 3 derivative, pRZ9012. The nucleotide positions of the 19 bp OE sequence (hatched box) are indicated. Four circularly permuted probes of equal size (182 bp), each containing the OE sequence, were generated by digestion with *MluI*, *NheI*, *Eco*RV and *NruI*, respectively. The arrowheads mark the center of each probe. The nucleotide position of each center is also shown. (**B**) Gel retardation assay of Tnp binding to each of the four bending probes: *MluI/MluI*, *NheI*, *Eco*RV and *NruI*, Each probe was incubated with 600 ng of full-length Tnp or Δ 369 protein as indicated. Free and complexed DNA were separated on a 5% polyacrylamide gel. (**C**) Mapping of the bending center of the OE fragment. The relative mobility, *R*_f (distance of migration of DNA–Tnp complex/distance of mobility of free DNA), of each bending probe was plotted as a function of the center in each probe as shown in (A). A third order polynomial curved fit of the data points allowed us to map the center of full-length Tnp-induced and Δ 369-induced bending to position 3 in the OE sequence).

In addition to similar relative binding affinities, both proteins show the same severe distortion of OE DNA upon binding as demonstrated by the circular permutation assay. This distortion has recently been shown by a phasing analysis to be a direct result of protein-induced bending (York and Reznikoff, in preparation). Bending of the OE by either protein is ~116° with the bend center located at the third position in the 19 bp sequence. This result is in agreement with that reported previously with full-length protein (7). An asymmetrical bend center relative to a protein binding site has also been demonstrated for the TnsB protein of Tn7 (4). The proximity of the Tnp-induced bend center near the first nucleotide in the OE strongly suggests that additional nucleotide contacts are needed upstream for efficient binding and bending by Tnp. Experiments to determine the minimal binding site needed for optimal protein binding to the OE are currently in progress.

While both full-length Tnp and $\Delta 369$ exhibit strong similarities in a number of the *in vitro* assays presented here, two significant differences were observed. One difference involves the shape of the binding curve (Fig. 3B). The hyperbolic shape of the binding curve of $\Delta 369$ is generally representative of a simple one step binding reaction in which a single preformed protein component binds to DNA. However, the sigmoidal curve binding of full-length Tnp suggests a more complex reaction. This hypothesis is supported by a number of previous observations. Protein composition studies of Complex I have shown the presence of

full-length Tnp, Inh and Tnp α , a naturally-occurring proteolytic product lacking ~25 amino acids from the N-terminus, when using wild-type Tnp preparations, or full-length Tnp and Tnp α when using Tnp MA56 (as in this study). Full-length Tnp protein appears in an equimolar ratio with the other polypeptides (2,3). Since Tnp α does not possess sequence-specific binding abilities (3), its presence in Complex I indicates that full-length Tnp forms complexes with Tnp α through a protein–protein interaction. Based on the molecular weight determination presented in this study with Tnp MA56 (lacking Inh protein), this association is in the form of heterodimers: full-length Tnp/Tnp α . The molecular weight determination does not rule out the possibility that full-length Tnp homodimers also bind to OE DNA. Experiments are currently underway to determine what different dimeric species of Tnp make up Complex I and the relative amounts of each that are present. This information would give us an indication of the relative affinity of each Tnp dimeric species for OE DNA. The second important difference between full-length Tnp and $\Delta 369$ involves the time required for each protein to form complexes with the OE. Full-length Tnp complex formation steadily increases over time. In contrast, Δ 369–OE complex formation maximizes very rapidly. Again this result suggests that a subunit exchange occurs between full-length Tnp aggregates and Tnp α .

We believe that the formation of Tnp heterodimers through a subunit exchange is the key to the complexity of full-length Tnp-OE binding reflected in part by the sigmoidal shape of the binding curve and in part by the slow formation of full-length Tnp complexes. Equilibrium-binding studies with the EcoRI restriction endonuclease in the presence of DNA lacking EcoRI recognition sites displayed a sigmoidal-shaped curve. This sigmoidicity was shown to reflect the addition of another enzyme subunit (17). A model involving Tnp subunit exchange is supported by a number of observations. Previous experiments have shown that titration of reactions containing constant amounts of Tnp MA56 protein preparations and OE DNA with Inh protein results in a steady increase in Complex I formation. This is presumably due to the increased availability of Inh to form heterodimers with full-length Tnp. In addition, as shown in this study, longer incubation times also increases Complex I formation (Fig. 4). This implies that heterodimer formation and subsequent OE binding are dynamic processes limited by the availability of Inh and/or Tnp α . This hypothesis also implies that the heterodimer is more active in binding than homodimers of full-length Tnp or that homodimers of full-length Tnp are trapped in aggregates under the conditions used in the assays. The possible role of heterodimers in Tn5 transposition is discussed below.

Heterodimer formation between a number of eukaryotic transcriptional regulatory proteins that share dimerization motifs have been shown to be involved in the regulation of their activity by modifying their affinity for different DNA targets (for a review see ref. 18). In addition, heterodimer formation has also been shown to act as an inhibition mechanism. Several heterodimer forms of the mammalian transcriptional activator protein, CREB, act as repressors of cAMP-induced transcription while the monomeric form acts as a strong constitutive activator (19). This model is not unlike that proposed for the regulation of Tn5 transposition (2) as described below.

The mechanism of transposition is a highly regulated process. In Tn5, the Tnp protein has two opposing activities: (i) in cis the protein is able to catalyze transposition (20) and (ii) Tnp acts primarily as an inhibitor in trans. The Inh protein acts exclusively as an inhibitor of transposition (10). It is proposed that immediately after translation, Tnp exists briefly as the active monomeric form which is able to bind an end and mediate synapse formation through dimerization of a second end-bound monomer. This intermediate is competent for subsequent steps of transposition. The Tnp protein soon after translation is subject to dimerization and possibly proteolysis in vivo, both of which may contribute to inhibition of transposition. All the proteolytic products of Tnp as well as the Inh protein have been found to be inactive in promoting transposition in vivo (2,3). However, all may contribute to inhibition. This mechanism of inhibition can occur by two not mutually exclusive steps. The active Tnp

monomers form inactive homodimers and/or heterodimers with Inh or a proteolytic product. This dimer formation reduces the active population of Tnp. In addition, these dimeric isoforms can specifically bind to end sequences blocking access of the active monomeric form of Tnp.

The mechanism of Tnp heterodimer formation is currently under investigation. Preliminary observations suggest that the subunit exchange indicated by the sigmoidal binding curve as well as the rate of binding study occurs primarily in the presence of OE DNA (York and Reznikoff, data not shown). This is not unlike that seen with CAP protein (21). This line of investigation may give us a better insight into the mechanism of transposition inhibition.

ACKNOWLEDGEMENTS

This work was supported by NIH grant GM50692 and a gift from the Promega Corporation. D.Y. was supported by NCI training grant T32 CA09075. W.S.R. is the Evelyn Mercer Professor of Biochemistry and Molecular Biology.

REFERENCES

- 1 Reznikoff, W. S. (1993) Annu. Rev. Microbiol. 47, 945-963.
- Wiegand, T. W. and Reznikoff, W. S. (1994) *J. Mol. Biol.* 235, 486–495.
 de la Cruz, N. B., Weinreich, M. D., Wiegand, T. W., Krebs, M. P. and
- Reznikoff, W. S. (1993) *J. Bacteriol.* 175, 6932–6938.
 Arciszewska, L. K. and Craig, N. L. (1991) *Nucleic Acids Res.* 19, 1000 (1991)
- 5021–5029. 5 Derbyshire, K. M. and Grindley, N. D. (1992) *EMBO J.* **11**, 3449–3455.
- 6 Kuo, C. F., Zou, A. H., Jayaram, M., Getzoff, E. and Harshey, R. (1991)
- EMBO J. 10, 1585–1591.
 7 Jilk, R. A., York, D. and Reznikoff, W. S. (1996) J. Bacteriol. 178, 1671–1679.
- 8 Weinreich, M. D., Mahnke, B. L. and Reznikoff, W. S. (1994) J. Mol. Biol. 241, 166–177.
- 9 Moore, J. T., Uppal, A., Maley, F. and Maley, G. F. (1993) Protein Expr. Purif. 4, 160–163.
- 10 Wiegand, T. W. and Reznikoff, W. S. (1992) J. Bacteriol. 174, 1229-1239.
- 1 Orchard, K. and May, G. E. (1993) Nucleic Acids Res. 21, 3335–3336.
- 12 Cui, Y., Midkiff, M. A., Wang, Q. and Calvo, J. M. (1996) J. Biol. Chem. 271, 6611–6617
- Zhou, Y., Zhang, X. and Ebright, R. H. (1993) *Proc. Natl. Acad. Sci. USA* 90, 6081–6085.
- 14 Thompson, J. F. and Landy, A. (1988) Nucleic Acids Res. 16, 9687–9705.
- 15 Krebs, M. P. and Reznikoff, W. S. (1986) J. Mol. Biol. 192, 781–791.
- 16 Rezsohazy, R., Hallet, B., Delcour, J. and Mahillon, J. (1993)
- Mol. Microbiol. 9, 1283–1295.
 Halford, S. E. and Johnson, N. P. (1980) Biochem. J. 191, 593–604.
 - 17 Frankright, S. E. and Joinison, N. F. (1980) *Biochem. J.* 191, 393–604.
 18 Lamb, P. and McKnight, S. L. (1991) *Trends Biochem. Sci.* 16, 417–422.
 - 18 Land, F. and McKinght, S. L. (1991) Trends Biochem. Sci. 10, 417–42.
 19 Krajewski, W. and Lee, K. A. (1994) Mol. Cell. Biol. 14, 7204–7210.
 - Krajewski, W. and Lee, K. A. (1994) Mol. Cell. Biol. 14, 7204–7210.
 Johnson, R. C., Yin, J. C. and Reznikoff, W. S. (1982) Cell 30, 873–882.
 - Brown, A. M. and Crothers, D. M. (1989) Proc. Natl. Acad. Sci. USA 86, 7387–7391.