Published in final edited form as:
J Mol Biol. 2010 August 13; 401(2): 167–181. doi:10.1016/j.jmb.2010.06.003.

Architecture of the *Tn7* Post-Transposition Complex: an Elaborate Nucleoprotein Structure

Jason W. Holder[1] and **Nancy L. Craig**[*]
Howard Hughes Medical Institute, Department of Molecular Biology & Genetics, Johns Hopkins University School of Medicine, 725 N. Wolfe Street, 502 PCTB, Baltimore MD 21205

Abstract

Four transposition proteins encoded by the bacterial transposon *Tn7*, TnsA, TnsB, TnsC, and TnsD, mediate its site- and orientation-specific insertion into the chromosomal site *attTn7*. To establish which Tns proteins are actually present in the transpososome that executes DNA breakage and joining, we have determined the proteins present in the nucleoprotein product of transposition, the Post-Transposition Complex (PTC) using fluorescently labeled Tns proteins. All four required Tns proteins are present in the PTC in which we also find that the *Tn7* ends are paired by protein-protein contacts between Tns proteins bound to the ends. Quantification of the relative amounts of the fluorescent Tns proteins in the PTC indicates that oligomers of TnsA, TnsB, and TnsC mediate *Tn7* transposition. High-resolution DNA footprinting of the DNA product of transposition *attTn7::Tn7* revealed that about 350 bp of DNA on the transposon ends and on *attTn7* contact the Tns proteins. All seven binding sites for TnsB, the component of the transposase that specifically binds the ends and mediates 3' end breakage and joining, are occupied in the PTC. However, the protection pattern of the sites closest to the *Tn7* ends in the PTC are different from that observed with TnsB alone, likely reflecting the pairing of the ends and their interaction with the target nucleoprotein complex necessary for activation of the breakage and joining steps. We also observe extensive protection of the *attTn7* sequences in the PTC and that alternative DNA structures in substrate *attTn7* that are imposed by TnsD are maintained in the PTC.

Keywords

Transposition; Transposase; Nucleoprotein complex; Expressed Protein Ligation; DNA-footprinting

[*]corresponding author Howard Hughes Medical Institute, Department of Molecular Biology & Genetics, Johns Hopkins University School of Medicine, 502 PCTB, 725 N. Wolfe Street, Baltimore, MD 21205, Phone: 410 955 3933, Fax: 443 287 7798, ncraig@jhmi.edu.
[1]present address Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Disclosure Statement:
All authors declare no actual or potential conflict of interest.

INTRODUCTION

Transposable elements are discrete DNA segments that can move between non-homologous positions within genomes. They are present in virtually every genome that has been examined and in some instances form a substantial fraction of the genome; for example, at least 45% of the human genome is derived from transposable elements[1]. Transposable elements also have substantial impact on bacterial genomes[2].

Element-encoded transposases bind specifically to transposon ends and mediate the catalytic steps in transposition, cleaving the phosphodiester bonds that link the transposon to the donor site and joining the transposon ends to the insertion site[3; 4; 5]. The transposase often acts in concert with other element- and/or host-encoded proteins to assemble the nucleoprotein complexes called transpososomes that mediate key steps in transposition such as pairing of the transposon ends and selection of a target site[6; 7; 8]. The proper nucleoprotein architecture of these complexes is central to the activation and coordination of the DNA breakage and joining events that underlie transposition.

Tn7 transposition occurs by a “cut & paste” mechanism in which the transposon is excised from the donor site and integrated into the target site[9; 10]. This process is tightly regulated by the recognition of the target site and assembly of a TnsABC+D transpososome: no excision of *Tn7* from the donor site occurs unless *attTn7* and all the Tns proteins are present[11]. Using DNA crosslinking, we have previously identified nucleoprotein complexes containing both a donor DNA from which *Tn7* has not yet been excised and a target DNA whose formation requires all the Tns proteins[12]. The architecture of the transpososome that mediates DNA breakage and joining must result in the juxtaposition of three DNA sites, the Left (*Tn7*.L) and Right (*Tn7*.R) ends of the transposon and a target DNA. We show here that all the Tns proteins required for insertion into *attTn7* are still present in the nucleoprotein complex in which breakage and joining have occurred.

Tn7’s transposition machinery is especially elaborate[9; 10]. Unlike simpler systems that utilize one or two transposition proteins[13], four *Tn7*-encoded proteins, TnsA, TnsB, TnsC, and TnsD, are required to insert *Tn7* into a specific site in bacterial chromosomes called *attTn7*[11]. *Tn7*.L and *Tn7*.R both contain multiple specific binding sites for the transposase subunit TnsB but these sites are differently positioned on each end, making the ends structurally asymmetric[14; 15]. The ends of *Tn7* are also functionally asymmetric[15; 16] elements with two *Tn7*.Rs transpose whereas those with two *Tn7*.Ls do not. Furthermore, *Tn7* insertion into *attTn7* downstream of the highly conserved *glmS* gene is orientation-specific as well as site-specific[17].

attTn7 is chosen as a specific site for *Tn7* insertion by the binding of TnsD to a particular sequence at the end of the highly conserved *glmS* gene[11; 18; 19]. The exceptional conservation of the TnsD binding site within the region of the *glmS* encoding the GlmS active site provides a specific insertion site for *Tn7* in virtually all prokaryotic genomes that have been examined[20]. Similar sequences in the human *glmS* homologs provide specific sites for the high frequency insertion of *Tn7* insertion *in vitro*[19; 21].

TnsA and TnsB together form the transposase that carries out *Tn7* DNA breakage and joining[22; 23; 24]. The TnsAB transposase is not, however, constitutively active: transposase activity is controlled by interaction of TnsAB with TnsCD bound to *attTn7*[11]. Selection of *attTn7* as the target DNA is initiated by TnsD binding to its 30 bp binding site, which generates distortions in *attTn7*[11; 18]; TnsC is recruited to this distorted DNA[25]. The interaction of TnsC with TnsD-*attTn7* is ATP-dependent[11] and leads to the regulation of transposase activity via TnsC interactions with both TnsA and TnsB[12; 26; 27; 28].

At what stages of transposition are the various Tns proteins required? For example, does TnsD simply serve as an "assembly" factor that dissociates from the rest of the transposition machinery once TnsC is loaded onto the *attTn7* target DNA as is seen for the targeting protein UvrA protein within the Nucleotide Excision Repair pathway[29]. We report here our analysis of the architecture of the *Tn7* transpososome in which transposition has occurred, the Post-Transposition Complex (PTC). We demonstrate that the PTC is an elaborate nucleoprotein complex in which the *Tn7* ends are paired and contains all four Tns proteins that are required to execute transposition, TnsABCD, in addition to the DNA product of transposition, the Simple Insertion DNA *attTn7*::*Tn7*. We also characterize the extensive Tns protein-DNA contacts within the PTC.

After Tns-mediated DNA breakage and joining, conversion of the *attTn7*::*Tn7* transposition product to intact duplex DNA requires repair by host proteins of the single-strand gaps that flank the 5' ends of the newly inserted transposon. These single-strand gaps result from the attack of the transposon's 3' ends on staggered positions in the target DNA. We find that such repair cannot occur on the PTC *in vitro*. Thus host proteins are also likely involved in disassembly of the PTC to remove the Tns proteins such that host DNA repair machinery can access the newly inserted transposon. Host proteins have been shown to remodel the Mu transpososome after transposition to allow DNA replication and repair[30; 31; 32].

RESULTS

Tn7 PTCs can be visualized by Atomic Force Microscopy

We carried out transposition reactions using a supercoiled *Tn7* donor plasmid and a small linear *attTn7* target fragment as DNA substrates and then used multiple cycles of size filtration and washing to isolate the stable PTCs from unincorporated Tns proteins (Fig 1a). In contrast to pre-transposition complexes that are detected only upon crosslinking[12], the PTC is stable in the absence of crosslinking, indicating that an increase in stability of Tns-DNA complexes accompanies *Tn7* transposition. We used Atomic Force Microscopy (AFM) to directly examine the architecture of PTCs (Fig 1b). Two types of PTCs were present that contained circular DNAs with measured lengths appropriate to be the *Tn7* element with a single Tns protein complex holding the *attTn7*::*Tn7*.L and *Tn7*.R::*attTn7* ends together. One type of PTC had open circular, i.e. relaxed, DNA, and the other had twisted, i.e. supercoiled, DNA. The flanking donor DNA was not contained in the PTCs.

The *Tn7* PTC contains TnsABCD

We also visualized PTCs isolated by size filtration by gel electrophoresis and directly characterized the PTC Tns and DNA components. The excision of *Tn7* from the donor site and its insertion into *attTn7* requires TnsABCD[9; 10; 33]. To identify the protein components of the PTC, we generated fluorescently labeled derivatives of all four Tns proteins by Expressed Protein Ligation[34] that added a short peptide containing fluorescein (green) to the carboxylterminus of each protein (Tns_f); the specific activities of the labeled proteins in transposition were equivalent to those of unlabeled proteins (data not shown). We carried out transposition reactions using the Tns_f proteins, a supercoiled donor plasmid containing a *Tn7* element and an *attTn7* fragment, either unlabeled or end-labeled with a different fluorophore, Alexa Fluor 555 (red). After transposition, reactions were digested with *SmaI* that cuts the *Tn7* element between the *Tn7*.L and *Tn7*.R ends, the PTCs purified by filtration, and then displayed on non-denaturing gels. Schematic diagrams of the substrates and products of transposition including the χ–PTC that results from *SmaI* digestion are shown in Fig 2a.

The four panels in Fig 2b display reactions in which one of the Tns proteins is fluorescently labeled. In each panel, lanes 1–3 display reactions with unlabeled *attTn7* and lanes 4–7 display reactions with $attTn7_f$; the reactions in lanes 2 and 5 were deproteinized while those in lanes 1 and 4 were not as indicated; Tns proteins were omitted in lane 3 and 9. Fluorescence was used to track the Tns_f proteins (green) and $attTn7_f$ (red); merging of coincident Tns_f and $attTn7_f$ signals results in a yellow signal. *Tn7* end-containing fragments were identified by hybridization to ^{32}P-labeled *Tn7*.L (lanes 4–6) or *Tn7*.R (lanes 7–9) probes following southern blotting of the gels (Fig 3).

The χ–PTC isolated by size filtration and gel electrophoresis (Fig 2,3), is observed with each Tns_f protein and thus contains all four Tns proteins, TnsABCD. The χ–PTC also contains *attTn7*::*Tn7* as indicated by the merged (yellow) signal of the Tns_f proteins (green) and $attTn7_f$ (red) (Fig 2 lane 4). The χ–PTC seen without deproteinization also contains *Tn7*.L (Fig 3 lane 5) and *Tn7*.R (lane 8), although the ends are no longer covalently linked because of cleavage of the *Tn7* element by *SmaI*. Thus, consistent with the AFM analysis shown above, the ends of *Tn7* are paired within the χ-PTC. In the absence of *SmaI* digestion, relaxed and supercoiled PTCs were also observed by gel electrophoresis (data not shown). Equivalent amounts of recombinant *attTn7*::*Tn7*.L and *Tn7*.R::*attTn7* products are present in the PTC when reactions are deproteinized. (Fig 2 lane 5 and Fig 3). Several conditions contribute to the smears below the χ-PTC. In reactions with $TnsA_f$ and $TnsC_f$ (Fig 2,3), a TnsACD-att*Tn7* complex [26; 35] is detectable. Also, some dissociation of the *Tn7* ends does occur as evidenced by the *Tn7*.L (Fig 2, all panels, lane 4) and *Tn7*.R (lane 5) smears and these ends are bound by $TnsB_f$ (Fig 2, TnsBf panel, lane 1).

Determination of the relative amounts of the Tns proteins in the PTC

We estimated the relative amounts of each of the Tns proteins in the PTC by comparing Tns_f fluorescence to the amount of end-labeled ^{32}P *attTn7*::*Tn7*. After *in vitro* transposition, we filter-purified the PTCs, crosslinked them to decrease protein dissociation during electrophoresis and displayed them by native agarose electrophoresis. The amount of fluorescence and ^{32}P was quantified for PTCs formed with each of the Tns_f proteins and ratios of the Tns_f/ *attTn7* measurements compared (Table 1).

We believe that the observed $TnsD_f$/^{32}P *attTn7* ratio of 0.12 reflects the presence of 1TnsD in each PTC. TnsD binds specifically to a site in *attTn7* that lies only to one side of the point of *Tn7* insertion[18] and DNA footprinting shows that this site is occupied in the PTC (see below). Moreover, there are no symmetrical features within the TnsD binding site. These observations are consistent with the suggestion that one TnsD protein binds to *attTn7*.

We then compared the $TnsD_f$/ ^{32}P *attTn7* value to the Tns_f/ ^{32}P *attTn7* values for all the other Tns_f proteins. The value of the $TnsD_f$/ ^{32}P *attTn7* was consistently much lower than that for any other Tns protein in the PTC, suggesting that there are multiple copies of the other Tns proteins in the PTC. There are 7 TnsB binding sites on the ends of *Tn7*. The PTC $TnsB_f$/ ^{32}P *attTn7* value was 6.3 fold higher than that for TnsDf/ ^{32}P *attTn7*, a value consistent with the presence of 7 TnsBs in each PTC. All 7 TnsB binding sites in the PTC are protected in DNA footprinting experiments (see below).

The ratios of $TnsA_f$ and $TnsC_f$ to ^{32}P *attTn7* in the PTC were much higher than for TnsD; the $TnsA_f$/ ^{32}P *attTn7* value was 18 fold-higher than $TnsD_f$/ ^{32}P *attTn7* and the $TnsC_f$/ ^{32}P *attTn7* value was 25 fold-higher than TnsDf/ ^{32}P *attTn7*. Thus the PTC contains more TnsA and TnsC than the 1 TnsD and 7 TnsBs suggested above. Previous work has shown that TnsA and TnsC form a $TnsA_2C_2$ complex[36]. DNA footprinting data is consistent with the suggestion that multiple $TnsA_2C_2$ complexes are present in the PTC (see below). The lower

value for Tns_f/ ^{32}P *attTn7* DNA for TnsA (18) compared to TnsC (25) likely reflects preferential loss of TnsA from the PTC during filter purification[35].

Extensive Regions of the *Tn7* End Sequences in the PTC Interact with the Tns Proteins

Where do the multiple Tns proteins in the PTC contact the Simple Insertion product *attTn7*::*Tn7*? To examine these protein-DNA contacts at base pair resolution, we carried out DNA-footprinting with line graph analysis using hydroxyl radicals, $KMnO_4$, and DMS as attack reagents. PTC footprinting was compared to that of naked Simple Insertion DNA of identical sequence but lacking the single-strand gaps at the transposon ends that are present in the PTC. By using *attTn7* substrates radiolabeled at either 3' end of *attTn7* DNA, we evaluated protection of the *Tn7* L and R 3' ends and *attTn7* strands that are covalently joined in the Simple Insertion DNA. Fig 4 shows hydroxyl radical footprinting gels and line traces and Fig 5 shows $KMnO_4$ and DMS footprinting gels for analysis of protection the *Tn7* ends. The results of these footprinting experiments are summarized at the nucleotide level in Fig 6.

Large segments of *Tn7*.L (Fig 4a) and *Tn7*.R (Fig 4b) end sequences within the PTC are protected from hydroxyl radical attack, indicating extensive contacts between the *Tn7* ends and the Tns proteins. Most of the end protection in the PTC reflects the binding of the transposase subunit TnsB to its specific binding sites in the transposon ends. When examined alone with individual *Tn7* ends, TnsB (purple circles) binds to seven sites, three in *Tn7*.L including the terminal α site, and the internal β and γ sites, and four in *Tn7*.R including the internal sites φ, χ, ψ, and the terminal ω site[14; 15]. The presence of all these sites in a *Tn7* element is required for wild-type transposon activity[16]. In the PTC, the *attTn7*::*Tn7* DNA is protected at all seven TnsB binding sites, indicating that transposition does not involve the loss of TnsB from these sites.

There are, however, several striking differences between the hydroxyl radical protection patterns observed in the terminal TnsB binding sites of the PTC and the previously observed TnsB-only protection of the individual *Tn7* ends that likely reflect changes in the transpososome when it is activated for breakage and joining. One PTC-dependent difference is that the terminal *Tn7*.L α and the terminal *Tn7*.R ψ and ω sites are more completely protected in the PTC (Figs. 4, dark blue brackets, yellow highlights). On the transferred strands in which the 3' ends of *Tn7* are covalently linked to *attTn7*, TnsB-only binding on individual *Tn7* ends results in characteristic regions of protection designated I, III and V within the TnsB binding sites, separated by regions of susceptibility to hydroxyl radical attack. In the PTC, however, in the terminal *Tn7*.L α and the terminal *Tn7*.R ψ and ω sites, the susceptible regions that are between the TnsB domains I, III and V are not cleaved by hydroxyl radicals, indicating a more extensive nucleoprotein complex at the termini. Another PTC-dependent difference is that the -ACAs at the 3' ends of *Tn7* that join to the target DNA are more completely protected, likely reflecting the positioning of the active transposase at the *Tn7* ends.

In addition to increased protection at *Tn7* termini within the PTC, we also find regions of strong protection immediately internal to the terminal *Tn7*L TnsB α site on the internal edge of the site between the α and β sites (Fig. 4, dark blue bracket and yellow highlight) and also weaker protections between the internal *Tn7*L β and γ sites (Fig. 4, light blue brackets and yellow highlight).

We also used $KMnO_4$ to probe the structure of *Tn7*.R within the PTC (Fig. 5a and summarized at the nucleotide level in Fig. 6). Hyper-reactivity to $KMnO_4$ results from distortions in DNA structure that lead to an increased accessibility to $KMnO_4$[18]. Several positions of $KMnO_4$ hyper-reactivity are observed in the internal *Tn7*.R TnsB sites φ and χ

at thymidine nucleotides flanking binding region III of these sites at R51, R58, and R73 (Fig 5a, red triangles). Despite similarly positioned thymidine nucleotides in the *Tn7*.R ψ and ω sites, no such $KMnO_4$ hyper-reactivity is observed at these sites, thus providing additional support for the view that the protein environment of the terminus of *Tn7*.R is distinct from that of the interior regions.

Tn7 insertion occurs by the attack of the 3'OH ends of *Tn7* on staggered positions on the top and bottom strands in *attTn7*, resulting in single strand gaps at the ends of *Tn7*. It is notable that in the PTC, the single-stranded gap that flanks the 5' end of *Tn7* at the *Tn7*.R::*attTn7* junction is also protected from hypersensitivity to $KMnO_4$ whereas in the deproteinized *Tn7*.R::*attTn7* transposition product, the single-strand gap was hyper-cleaved in the presence of $KMnO_4$ (Fig 5a, blue triangles).

Extensive Regions of the *attTn7* sequences in the PTC Interact with the Tns proteins

The middle bp of the 5bp *attTn7* target sequence duplicated upon *Tn7* insertion is designated "0", sequences that lie rightward towards *glmS* are "+" and those leftward are "−"; thus the *Tn7* target site duplication is *attTn7* −2 to +2. The target information required to promote *Tn7* insertion at *attTn7* −2 to +2 is contained within the TnsD binding site, which is located at about *attTn7* +25 to +55. There is no evidence that any particular DNA sequence is required to the left of the TnsD binding site in *attTn7* including that at the actual point of insertion[11; 19; 37].

DNA-footprinting analysis using hydroxyl radicals of the PTC revealed strong and extensive protection of the *attTn7* sequences from *attTn7* −58 to *Tn7*.L (Fig 7) and from *Tn7*.R to *attTn7* +55 (Fig 5b), including the single-strand gaps at the transposon ends. Thus the proteins within the PTC strongly interact with 114 bp of *attTn7* DNA sequences. As no particular sequence leftwards of the TnsD binding site at *attTn7* +25 to +55 is required for transposition[11; 18; 19; 37], it is unclear why the left border of protection in the PTC stops sharply at *attTn7* −58,

Our DNA-footprinting analysis with DNA distortion-sensitive probes (Fig 5) also supports the view that TnsD remains bound to *attTn7* in the PTC. TnsD binding to the pre-transposition substrate *attTn7* results in DNA distortions at *attTn7* +27 as evidenced by $KMnO_4$ hyper-reactivity and DMS hyper-reactivity at *attTn7* +45[18]. These distortions are still present in the PTC, confirming the presence of TnsD and the TnsD-induced *attTn7* distortions.

The Tns proteins in the PTC block repair of the terminal gaps in *attTn7*::*Tn7*

The 5' ends of *Tn7* in the Simple Insertion product *attTn7*::*Tn7* are flanked by single-strand gaps that require repair by host factors to generate intact duplex DNA. We find that when the PTC is deproteinized, repair of these gaps is readily observed upon the addition of DNA polymerase I and DNA ligase to *attTn7*::*Tn7* DNA as evaluated by denaturing gel electrophoresis showing that the top 5' strand of *Tn7*.L becomes covalently linked to the flanking *attTn7* target (Fig 8, lanes 1). When *Tn7* inserts into *attTn7* to form *attTn7*::*Tn7*, the gap sequence adjacent to the top 5' strand of *Tn7*.L which must be synthesized is 5'-GGGCG. If DNA synthesis proceeds into the 5' end of *Tn7*.L, the newly synthesized DNA extending from the gap into the end would include 5' TGT. However, efficient repair is still observed when dA or dT is omitted from the repair reaction (lanes 2 and 3), indicating that repair synthesis need not proceed into the end of *Tn7* to repair the transposition-generated gap.

Repair of the gap flanking *Tn7*.L of the Simple Insertion DNA in the PTC is not observed without deproteinization (lanes 4–6), revealing that the Tns proteins block access of the

repair enzymes to the flanking gap. The fact that this gap remains refractory to repair even in PTCs that have been incubated for several days at 4°C indicates the stability of the PTC (data not shown). Thus a key step in *Tn7* transposition must be the disassembly of the PTC to allow repair. Incubation of the *Tn7* PTC in *E. coli* crude extracts can remove the block to DNA repair observed with PTCs, suggesting that host factors can actively promote disassembly of the PTC to allow repair (J. Holder, C. Johnson, and NLC unpubl. observations).

Discussion

Tn7 Transpososomes are Elaborate and Stable Nucleoprotein Complexes

Successful transposition requires specific recognition and juxtaposition of the transposon ends and the target DNA, and careful coordination of the DNA breakage and joining events, avoiding non-productive DNA breaks at the transposon ends and promoting equivalent joining of the transposon ends to the target DNA. These reactions occur in nucleoprotein complexes called transpososomes[6; 8]. We have defined the components and architecture of the transpososome product of *Tn7* transposition, the Post-Transposition Complex (PTC), which mediated the site-and orientation-specific insertion of *Tn7* into the *attTn7* target DNA to form the Simple Insertion product *attTn7*::*Tn7*.

We have found that the *Tn7* ends are stably paired within the PTC such that supercoiling of the *Tn7* substrate can be maintained within the PTC. Such stable end pairing must be key to the coordination of Tns protein activity on both transposon ends and has also been observed in other transposition systems[38; 39] The fact that PTCs including those in which the *Tn7* segment has been cut by a restriction enzyme (χ-PTCs) can be observed after filter purification and gel electrophoresis in the absence of stabilization by crosslinking attests to their stability. By contrast, pre-transposition *Tn7* donor/ target complexes can be observed only in the presence of crosslinking[12]. Increasing transpososome stability provides a mechanism for guiding transposition to convert the *Tn7* donor and *attTn7* target substrate DNAs to the *attTn7*::*Tn7* product as has been suggested for other transposition systems[8; 40].

In addition to promoting breakage at the transposon ends and joining to the target DNA, another key role of transpososomes is to prevent side reactions that would convert the transposition product to a form that cannot be repaired to form an simple insertion containing intact duplex DNA[41; 42; 43] .

The production of a highly stable PTC presents a challenge to the cell. We have found that the PTC is refractive to the repair reactions necessary to convert the single strand gaps at the *Tn7* ends in the actual transposition product to duplex DNA, suggesting that active disassembly of the PTC occurs *in vivo*. Active disassembly of the highly stable MuA tetramer that is the core of the Mu transposition machinery has been demonstrated to play a key role in the completion of Mu transposition[44].

Multiple Tns proteins and extensive *attTn7*: :*Tn7* sequences interact in the PTC

A key finding of this work is that the *Tn7* PTC contains the four Tns proteins that are essential for transposition into *attTn7*, TnsABCD, and that multiple copies of TnsA, TnsB and TnsC are present in the PTC. We have also defined the extensive, nearly 360 bp, interactions between the Tns proteins and the DNA product of transposition *attTn7*::*Tn7*. The extensive protein-DNA interactions both at the ends of *Tn7* and on the *attTn7* sequences reflect the complexity of the *Tn7* DNA substrates and their interaction with many Tns proteins required to promote high-frequency site- and orientation-specific *Tn7* insertion into *attTn7*.

Multiple copies of the TnsB transposase subunit are present on the ends of *Tn7* in the PTC

TnsB is the transposase subunit that mediates DNA breakage and joining at the 3' ends of *Tn7*[23] and transposition requires multiple TnsB binding sites in both ends of *Tn7*[14; 15; 16]. Using DNA footprinting, we have found that all 7 of these TnsB binding sites are occupied in the PTC. The stoichiometry of TnsB in the PTC that we have determined, 6.3 TnsBs/ *attTn7*, is consistent with these protection results.

There are, however, significant differences in the Tns protein interactions with the ends of *Tn7* in the PTC in which both pairing of the *Tn7* ends and their juxtaposition to *attTn7* has occurred as compared to the interaction of TnsB alone: notably, protection is increased within the terminal *Tn7*L TnsB site α and on the *Tn7*.R TnsB sites ψ and ω. Moreover, in the PTC, DNA protection extends to the 3' *Tn7* ends where TnsB-mediated breakage and joining occurs. These changes in the interaction of the Tns proteins with the ends of *Tn7* must reflect the TnsB-mediated pairing of the *Tn7* ends[12], the positioning of the TnsBs at the 3' ends for cleavage, and likely additional end interaction by TnsA, the other subunit of the transposase that promotes cleavage at the 5' termini[23; 24].

TnsB is a member of the Retroviral Integrase superfamily of transposases which includes the well-characterized *Tn5/ Tn10* transposases whose catalytic centers are RNaseH-like domains[4; 5]. In the *Tn5* system, one transposase protomer and hence one active site is positioned at each transposon end and mediates all the DNA breakage and joining events at that end[45; 46; 47]. An attractive hypothesis is that, similarly, one TnsB bound at each 3' end mediates the DNA breakage and joining events at that end and that one TnsA mediates the breakage events at each 5' end. In this view, TnsBs bound at the internal TnsB binding sites are thus likely involved in positioning the catalytic machinery at the point of insertion and imposing the orientation-specific insertion of *Tn7*, presumably by interaction with an asymmetric component of a target complex (see below).

TnsD is present in the PTC

Tn7 is recruited specifically to *attTn7* by the specific binding of TnsD to a sequence within *attTn7*[11; 18]. A reasonable question is whether TnsD remains in the pre-transposition complex once the TnsABC machinery has been positioned on *attTn7* or does it simply serve as an "assembly" factor. We have found that TnsD remains bound to its recognition site on *attTn7* in the PTC. The continued presence of TnsD and the distortions in *attTn7* DNA structure it provokes may be necessary for asymmetric activation of TnsABC to promote orientation-specific insertion of *Tn7* into a unique position in *attTn7* (below).

TnsC is positioned on the insertion site by interaction with DNA distortions imposed on *attTn7* by TnsD18 and perhaps through protein-protein interactions between TnsC and TnsD (R. Mitra and N.L. Craig unpubl.). TnsC also interacts with both transposase subunits, TnsA and TnsB,[12; 26; 27] (see below).

Extensive *attTn7* sequences beyond the TnsD binding site are bound by protein in the PTC

A striking feature of the *Tn7* PTC is the extensive protection of *attTn7* sequences, extending from its leftward border at *attTn7* −58 to its rightward border at *attTn7* +55. TnsD binding to *attTn7* +25 to +55 accounts for some of this protection. We suggest that the extended region of protection leftward of the TnsD binding site in the PTC reflects the interaction of TnsA and TnsC, likely as multiple $TnsA_2C_2$ complexes which have been previously identified[36]. This would account for the much higher amounts of TnsA and TnsC in the PTC compared to TnsD (1 TnsD/ PTC) and TnsB (7 TnsB/ PTC), being (at least) 18 TnsA/PTC and 25 TnsC/ PTC. It remains to be determined why there is a sharp border of protection at

attTn7 −58 as no specific sequences between *attTn7* −58 and *attTn7* +25 are required for transposition.

The interaction of pre-transposition TnsACD-*attTn7* target and TnsB-*Tn7* donor complexes positions and activates the TnsAB transposase

We have previously observed the formation of distinct TnsD-*attTn7*[11]. TnsCD-*attTn7*[18] and TnsACD-*attTn7* complexes[26; 35] and have argued that TnsC interacts with both the target DNA[11; 25] and the TnsAB transposase[48]. We suggest here that a $TnsA_2C_2$ complex is positioned at the point of *Tn7* insertion in *attTn7* by interaction with TnsD-*attTn7* to form a target complex that recruits TnsB bound to the *Tn7* ends. There are multiple protein-protein interactions between TnsA, TnsB and TnsC: TnsA interacts with TnsC[36; 49]. TnsB interacts with TnsC[28; 49] and TnsA and TnsB likely interact because TnsA is required to provoke TnsB-mediated 3' end breakage and joining in a minimal TnsAB-only recombination system[22]. TnsA and TnsB form the transposase, TnsA mediating DNA breakage at the transposon 5' ends and TnsB mediated breakage and joining at the 3' transposon ends[23; 24]. Notably, breakage and joining is seen only in the presence of both TnsA and TnsB[11]. Thus only when a TnsACD-*attTn7* target complex interacts with a TnsB-end complex is the TnsAB transposase formed and activated.

The molecular mechanism by which apparently only the $TnsA_2C_2$ complex at the insertion site can capture the TnsB-bound ends for insertion in highly site- and orientation-specific insertion remains to be determined but it seems likely that its asymmetric interaction with TnsD-*attTn7* including the distortions introduced by TnsD into *attTn7* will be key in this process[18; 25]. Perhaps the interaction of TnsD with one of the C subunits of a $TnsA_2C_2$ complex at the point of insertion defines an asymmetric target site.

A striking observation from this work is that almost 60 bp of target DNA leftward to *Tn7*L in *attTn7::Tn7* are protected within the PTC. This is surprising because we have demonstrated that no specific DNA sequences are required in this region of a target DNA[18]. We suggest that this region is also bound by role of $TnsA_2C_2$ complexes beyond the point of insertion but their role remains to be determined. Perhaps they mediate interactions with TnsBs that determine including orientation- and site-specific insertion. If multiple $TnsA_2C_2$s are present in pre-transposition complexes, they may provide for a multi-valent landing site for the TnsB-*Tn7* ends that can channel the ends to the point of insertion.

The *Tn7* transposition machinery is complex

The simplest model that accounts for the chemical steps of *Tn7* transposition is that one TnsB bound at each end of *Tn7* mediates 3' end breakage and joining and one TnsA bound at each end mediates 5' end breakage and that this TnsAB transposase is part of a $TnsA_2B_2C_2$ complex that can be positioned at a specific site by interaction with TnsD. However, our studies of the *Tn7* PTC have revealed a far more complex and elaborate transposition machine. It seems likely the additional TnsBs and $TnsA_2C_2$ complexes that we have established are in the PTC play important roles in *Tn7* end identification, target site selection and controlling and directing the highly site- and orientation-specific insertion of *Tn7* insertion into *attTn7*. The elaborate architecture of the PTC we have described here has deepened our understanding of the control of *Tn7* transposition and raised new questions for future investigations.

METHODS AND MATERIALS

Tns proteins

The *tnsA, tnsB, tnsC and tnsD* genes were cloned into pCYB1 (New England Biolabs = NEB) to express Tns-Intein-Chitin Binding Domain fusion proteins for affinity purification on chitin beads. 1 liter cultures of CAG456[23] containing Tns-Intein-chitin binding domain fusion proteins in the pCYB1 vector were grown at 30°C to an OD of 0.5, and cooled on ice for 20 minutes prior to induction with 400 μM IPTG for 18 hours at 16°C. Cells were harvested by centrifugation and resuspended in 10 ml Lysis $Buffer_{TnsABD}$ = 25 mM HEPES pH 8.0, 10 % glycerol, 0.1 mM EDTA, 500 mM NaCl, 0.1% Tween 20 or Lysis $Buffer_{TnsC}$ = 25 mM HEPES pH 8.0, 10 % glycerol, 1 mM EDTA, 1 M NaCl, 1 mM ATP. Cells were lysed in a French Press with pressure in excess of 1000 psi and then cell extracts were spun at 100,000 × G for 20 minutes. The resulting supernatant was filtered through a 0.45 μM filter (Millipore) prior to mixing with 1 ml Lysis Buffer-equilibrated chitin beads (NEB) in a capped and sealed 20 ml column (BioRad), followed by gentle shaking at 4°C for 20 minutes. The chitin beads were allowed to settle, the column allowed to drain and washed with 4 × 20 ml Lysis Buffer. For TnsA, TnsB, and TnsD, the beads were resuspended with shaking for 60 minutes at room temperature with 10 ml Release Buffer = 20 mM HEPES pH 8.0, 0.1% Tween 20, 10 % glycerol, 500 mM NaCl, 10 mM $MgCl_2$, 10 mM ATP to remove *E. coli* chaperone proteins. The chitin beads were then allowed to settle, the column allowed to drain and washed with 10 ml Lysis Buffer.

To generate fluorescently labeled Tns proteins, a fluorescent peptide CKH_6 labeled on the ε amino group of the lysine with fluorescein and HPLC purified (Sigma-Genosys) was ligated to each Tns protein by Expressed Protein Ligation[34]. Intein-mediated cleavage of the Tns fusion protein is coupled to peptide incorporation at the C-terminus of the Tns proteins. A 1 ml bed of chitin beads bound by Tns fusion protein was incubated with 500 μM peptide and 200 mM sodium 2-mercaptoethanesulfonate (MESNA) pH 8.0 (Sigma-Aldrich) in Lysis Buffer overnight at 4°C in the dark. The beads were then allowed to settle, the column drained and washed with 3 ml Lysis Buffer to yield a final eluate of 4 ml containing Tns_f and unincorporated peptide. Heparin chromatography was used to separate $TnsA_f$, $TnsB_f$, and $TnsD_f$ from the peptide. Peptide was removed from $TnsC_f$ by ammonium sulfate precipitation of TnsC (30 % saturation = 178 mg/ml) and washes in TnsC Low Salt buffer as previously published[50]. The Tns_f protein was at least 70% of each preparation and had the same specific activity as Tns protein (data not shown).

$TnsA_f$ is stored in 25 mM HEPES pH 8.0, 0.1% Tween 20, 10 % glycerol, 150 mM NaCl, 1 mM DTT. $TnsB_f$ is stored in 25 mM HEPES pH 8.0, 0.1% Tween 20, 25 % glycerol, 500 mM NaCl, 1.0 mM EDTA, 1 mM DTT. $TnsC_f$ is stored in 25mM HEPES pH 8.0, 0.1 mM EDTA, 1.0 M NaCl, 2.5 mM DTT, 10 % glycerol, 10 mM CHAPs, 10 mM $MgCl_2$, 1 mM ATP. $TnsD_f$ is stored in 25 mM HEPES pH 8.0, 500 mM NaCl, 1 mM EDTA, 2 mM DTT, 25 % glycerol.

TnsA, TnsB, TnsC, and TnsD proteins were purified as previously described[19] or TnsC, and TnsD were derived from purified Intein fusion proteins as for Tns_f proteins except that 40 mM DTT was used for cleavage.

Transposition substrates

Tn7 donor plasmids used as indicated in the Figure legends were pEMΔ (5.9 kB) that contains a 1,625 bp *Tn7* element with 166 bp *Tn7*.L and 199 bp *Tn7*.R in a donor backbone of 4301 bp and pPK21 (3.2 kB) that contains a 373 bp *Tn7* element with 166 bp *Tn7*.L and 199 bp *Tn7*.R in a donor backbone of 2785 bp[19]. *attTn7* target fragments used as indicated

in the Figure legends were generated as described below. The target *attTn7* plasmid used in Fig 8 was pRM2.

Construction and Labeling of *attTn7* Target Fragments

The 244 bp *attTn7* −123 to +120 target fragment used in the AFM analysis of Figure 1 was generated by PCR using pRM2 as a template with oligonucleotides NLC1193 5'-agccggaattcgacagaaaattttcattctg and NLC1165 5'-ctttgatcagcgcgacatggtaaggcggccgcattcttat, followed by gel purification.

The 107 bp *attTn7* −38 to +68 used in the experiments of Figures 2 and 3 was generated by PCR using pRM2 as a template with oligonucleotides NLC 1456 5'-tccaaagccggaattcttgattaaa-aacataacaggaag and NLC 1667 5'-gccaggttacgcggctggtgaat. The PCR fragment was then digested with *EcoRI* at the two sites that flank the *attTn7* sequences and the resulting 3' ends filled in with dA and dU-Alexa Fluor 555 using the ARES DNA labeling Kit (Molecular Probes) followed by gel purification.

The target used in the experiment of Table 1 was the 107 bp *attTn7* −39 to +68 fragment generated by PCR with primers 1456, and 1667 using pRM2 as a substrate and was labeled by a ^{32}P dATP incorporation following *EcoRI* digestion.

In the footprinting experiments of Figures 4, 5 and 7, the 204 bp *attTn7* −83 to +120 fragment bounded by *HindIII* and *XbaI* sites in pPK13[18] was used as the target DNA for the generation of PTCs. For analysis of *attTn7*::*Tn7*.R, the fragment was labeled at the 3' end of the *attTn7*.R sequences (*attTn7*.R +120, *XbaI*) and for analysis of *attTn7*::*Tn7*.L, the fragment was labeled at the 3' end of the *attTn7*.L sequences (*attTn7*.L −81; *HindIII*). End-specific labeling was achieved by digestion of the plasmid at the end to be labeled, filling in with α-^{32}P dATP using Klenow exo$^-$ DNA polymerase (NEB), the plasmid digested at the other end and then gel purified. In each case, the 3' labeled *attTn7* strand becomes covalently linked to the 3' end of *Tn7*.

In the footprinting experiment of Figure 7 where the left boundary of PTC protection was determined, the 424 bp *attTn7* −300 to +120 fragment bounded by *NdeI* and *XbaI* sites in pPK13[18] was used as the target DNA for the generation of PTCs. The 424 bp fragment was labeled at the 3' end of the *attTn7*.L sequences by digesting with *NdeI* and filling in with dTTP and ^{32}P α-dATP using Klenow exo$^-$ DNA polymerase (NEB), the plasmid digested with *XbaI* and the labeled *attTn7* fragment purified using a PCR clean up kit (Qiagen).

Generation of the repaired simple insertion *attTn7*::*Tn7*

A repaired Simple Insertion DNA containing the same sequences as the gapped simple insertion DNA product in the PTC that was used as the control in the DNA footprinting experiments was generated by *in vitro* transposition and cloning. The 204 bp fragment containing *attTn7* −83 to +120 was isolated by digestion of pPK13 with *HindIII* and *XbaI* and used as a target in an *in vitro* transposition reaction using as a donor plasmid pPK21 which contains the 373 bp *Tn7* element. The Simple Insertion product *attTn7 HindIII-XbaI*::*Tn7*-373 bp was isolated from a gel, purified by Qiagen treatment, ligated to pPK13 digested with *HindIII* and *XbaI* and the *attTn7*::*Tn7* products recovered by transformation creating plasmid pJH375. Digestion of the resulting plasmid with *HindIII* and *XbaI* yields a double-stranded, i.e. repaired, *attTn7*::*Tn7* fragment that contains the same sequences as are contained in the gapped Simple insertion DNA in the PTC as in Figure 4, 5 and 7.

Transposition reactions

Transposition reactions were carried out as previously described[19] except where indicated. Reactions were performed in 24 μl with 22 mM HEPES pH 7.6, 3.3 mM DTT, 375 ng herring sperm DNA (Roche), 1.7 mM ATP, 1.6 μl TnsA, 1 μl TnsB, 1 μl TnsC, and 3 μl TnsD. In Figure 2, the donor was 0.8 nM pEMΔ, the target was 11 nM Alexa Fluor 555-labeled *attTn7* −38 to +68 and Tns proteins including Tns_f were 35 nM TnsA, 7 nM TnsB, 25 nM TnsC, and 3 nM TnsD. In Figures 4 and 5, the donor was 1.5 nM pPK21 and the 1.7 nM ^{32}P-labeled targets were either *attTn7* −83 to +120 for analysis of *attTn7*::*Tn7*.L (Fig 4a) and *Tn7*.R::*attTn7* (Figs. 4b,5) and *attTn7* −300 to +120 for analysis *attTn7*::*Tn7*.L (Fig 7) and the Tns proteins were 52 nM TnsA, 7 nM TnsB, 50 nM TnsC, and 28 nM TnsD.

Purification of PTCs by filtration

Reaction mixtures were added after transposition to 400 μl Wash Buffer (25 mM HEPES 8.0, 150 mM NaCl, 1mM DTT, 1 mM ATP pH 8.0) in an equilibrated 100 kDa cut off microcon filter (Millipore). Samples were concentrated by spinning as directed by manufacturer, 400 μl of buffer added with mild agitation and respun. This cycle was repeated 5 times to a 20–30 μl final retentate that contains PTCs, donor, and flanking donor DNA that were separated during gel electrophoresis or on a mica grid for AFM.

AFM of PTCs

Filter purified PTCs were resuspended in AFM buffer (25 mM HEPES pH 7.6, 25 mM KOAc pH 7.2, 5 mM DTT, 1 mM ATP, 2 mM MgOAc). 10–20 μl of sample was adsorbed onto a 12 mm disk of freshly cleaved mica for 5–10 minutes and rinsed with water. The liquid was blown off the mica with compressed gas and the sample allowed to dry. AFM imaging was performed in ambient Tapping Mode using a Nanoscope IIIa controller with a multimode AFM (Digital Instruments, Santa Barbara, CA). Conventional silicon cantilevers type TESP (Veeco Instruments, Santa Barbara, CA) were used.

Electrophoresis and visualization of the recombination products

In the experiments in Figure 2 and 3, filter-purified χ-PTCs were loaded in 10% glycerol onto a 19 cm long 1.0% 0.5 × TBE agarose gel (Life Tech) and run for 714 V-Hr. Gels to be blotted were transferred to Gene Screen (NEN) and southern blotted with oligonucleotide probes specific for the *Tn7* ends, NLC 272 5'-attttcgtattagcttacgacgctacaccc for *Tn7*.L and NLC 925 5'-gttcagtttaagactttattgtc for *Tn7*.R, that had been phosphorylated at their 5'ends by T4 polynucleotide kinase (NEB) and γ ^{32}P ATP.

DNA footprinting

Hydroxyl radical footprinting reactions were done by mixing 3 μl 3 mM $Fe(NH_4)_2(SO_4)_2$, 3 μl 6 mM EDTA, and 3 μl 30 mM Na-Ascorbate with 20 μl retentate containing filter-purified PTCs and allowed to react at room temperature. The reactions were quenched after 1 minute by addition of 1 μl βME; protein was stripped off the DNA by addition of EDTA and SDS to 10 mM EDTA/ 0.08% SDS and the samples incubated at 65°C for 10 minutes. Simple Insertion DNA was then purified by agarose gel electrophoresis and gel extraction (Qiagen). For $KMnO_4$ footprinting, filter-purified PTCs, repaired SI DNA and gapped SI DNA were treated with 2.5 mM $KMnO_4$ for 1 minute at 24°C in 20 μl total volume[18]; the DNAs were then stripped of protein, purified as described above and then treated with 10% piperidine for 20 minutes at 90°C. For DMS footprinting, filtered-purified PTCs and SI DNA in 20 μl reactions were treated with 0.2% dimethyl sulfate (DMS) for 30 seconds and then stopped by the addition of 50 μl 25 mM HEPES pH 7.6, 125 mM βME, 1 mM EDTA[18]. Proteins were stripped from the DNA, the DNA purified as described above and described and then treated with 0.1 M NaOH for 20 minutes at 90°C.

In the footprinting experiments of Figure 4, 6, and 7b, the DNA was displayed on an 8M urea 7% 29:1 acrylamide: bis-acrylamide gel (Explorer, JT Baker) run for 9375 V-Hr. In Figures 7a and supplementary figures 3 and 5 the DNAs were displayed on 8M urea 8% 29:1 acrylamide: bis-acrylamide gels run for 6868V-Hr. Footprinting gels were dried down on 3mM paper, exposed to a phosphor screen and visualized by phospho-imaging.

DNA and fluorescent protein imaging

Gels were imaged on a Typhoon 9410 (Molecular Dynamics). ^{32}P was evaluated by phosphorimaging. Quantification of DNA fragments in denaturing gels was done using line graph analysis performed with Image Quant 5.0. The intensity values for each lane of the gel were transferred into Microsoft Excel for graphing DNA-cleavage intensity as a function of distance from the well. Scans for *attTn7*$_f$ DNA labeled with Alexa Fluor 555 were done using a 532 nm laser for excitation and a 580BP30 emission filter; the emitted light passed through a photomultiplier tube PMT at 800V. The fluorescein dye contained on the Tns$_f$ proteins using a 488 nm laser and a 520BP40 emission filter.

Determination of Tns protein ratios in the PTC

For the experiment in Table 1, after transposition incubation, PTCs were filter purified and then crosslinked with 0.01 % gluteraldehyde for 10 minutes and then quenched. They were then loaded onto a 1 % agarose gel and run for 714V-Hr at room temperature. TnsA, TnsB and TnsC, the amount of fluorescence was determined for supercoiled PTC only. The amount of fluorescence in both the relaxed and supercoiled PTCs was determined for TnsD. The DNA gel was imaged on a Typhoon 9410 for the fluorescein dye contained on the Tns$_f$ proteins using a 488 nm laser and a 520BP40 emission filter; the amount of fluorescence for each protein was determined by subtraction of background fluorescence. After fluorescent imaging was performed, the agarose gel was dried down onto DEAE paper and exposed for phophorimaging. It should be noted that some TnsA can be lost from the PTC during purification of PTCs[35] and thus the relative amount of TnsA is likely closer to the TnsC value. This measurement does not give absolute stoichiometry because of local quenching effects on the fluorescent emission that may vary for each fluorophores location within the PTC.

Gap Repair

Tn7 donor plasmid was pEMΔ which contains a 1.6 kb *Tn7* element with 166 bp *Tn7*.L and 199 bp *Tn7*.R[51]. The *attTn7* target plasmid pRM2 contains *attTn7* −342 to +165[51]. Gap repair was performed by addition of 2.5 units of DNA Polymerase I (NEB), 100 units of T4 DNA ligase (NEB), 200 μM dNTPs, and 10 mM MgOAc to 25 μl of Wash Buffer containing filter purified PTCs or deproteinized transposition reactions. Following repair, proteins were removed (Qiagen PCR clean-up kit) and DNA was then digested *PstI* (NEB). Digested repair reactions were displayed on a denaturing gel that was then Southern blotted with a *Tn7*.L-specific oligonucleotide.

Acknowledgments

We thank Robert Sarnovsky cloning of TnsC into pCYB1, Prasad Kuduvalli, Gregory McKenzie, Chuck Merryman, and Rupak Mitra for helpful discussions, Jan Hoh for assistance with AFM, Iva Ivanovska, Helen McComas and Patti Kodeck for assistance with manuscript preparation. This work was supported by NIH grants P01 CA16519 and RO1 GM076425 to NLC. NLC is an Investigator of the Howard Hughes Medical Institute.

Role of the Funding Source:

JWH and NLC conceived of and designed the experiments, JH performed the experiments and JWH and NLC wrote the paper. This work was supported by NIH grants P01 CA16519 and RO1 GM076425 to NLC and NIH

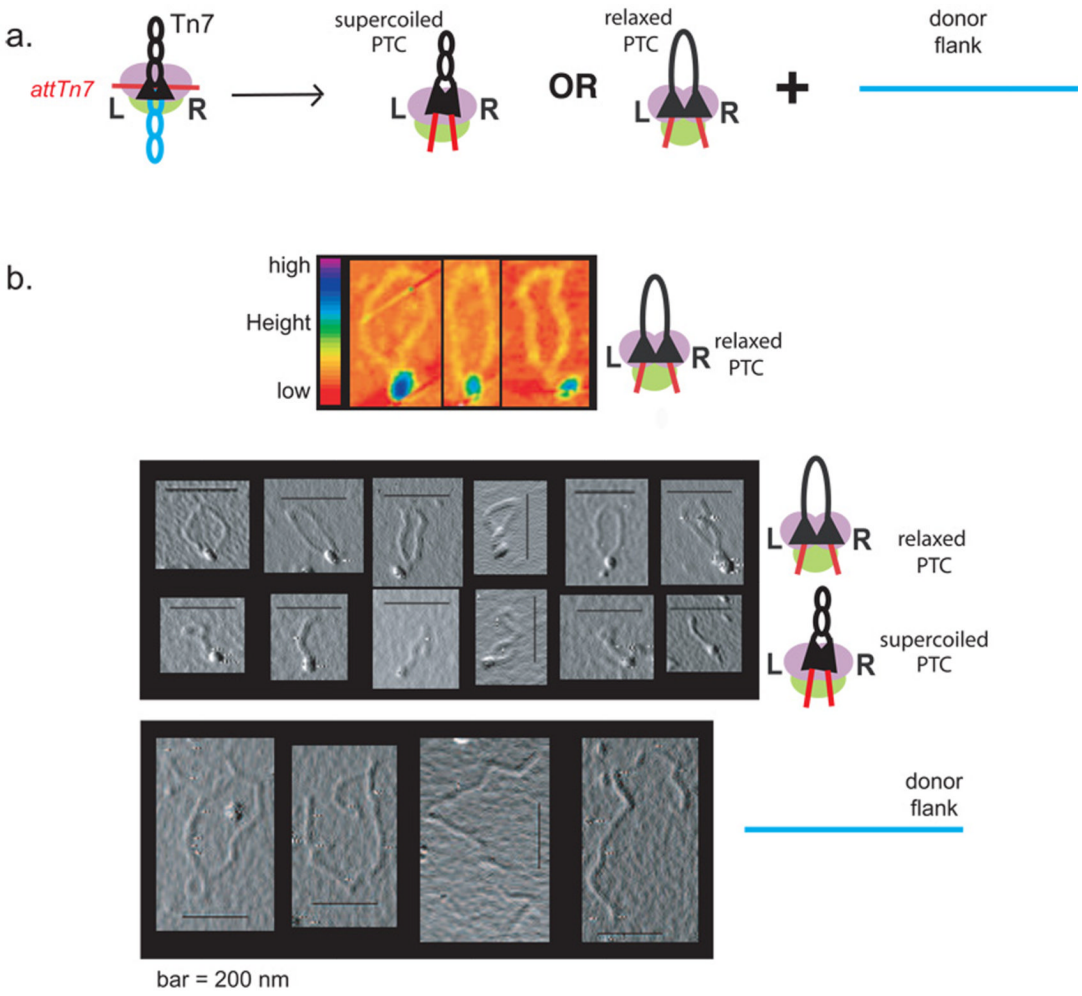
Training Grant T32-GM07445 to the Biological Chemistry Molecular Biology Program at Johns Hopkins University. NLC is an Investigator of the Howard Hughes Medical Institute.

REFERENCES

1. Lander E, Linton L, Birren B, Nusbaum C, Zody M, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczk yJ, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov J, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin J, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston R, Wilson R, Hillier L, McPherson J, Marra M, Mardis E, Fulton L, Chinwalla A, Pepin K, Gish W, Chissoe S, Wendl M, Delehaunty K, Miner T, Delehaunty A, Kramer J, Cook L, Fulton R, Johnson D, Minx P, Clifton S, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng J, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921. [PubMed: 11237011]
2. De Palmenaer D, Siguier P, Mahillon J. IS4 family goes genomic. BMC Evol Biol 2008;8:18. [PubMed: 18215304]
3. Mizuuchi, K.; Baker, T. Chemical mechanisms for mobilizing DNA. In: Craig, N.; Craigie, R.; Gellert, M.; Lambowitz, A., editors. Mobile DNA II. Washington, DC: ASM Press; 2002. p. 2-23.
4. Haren L, Ton-Hoang B, Chandler M. Integrating DNA: transposases and retroviral integrases. Annu Rev Microbiol 1999;53:245–281. [PubMed: 10547692]
5. Hickman A, Chandler M, Dyda F. Integrating prokaryotes and eukaryotes: DNA transposases in light of structure. Crit Rev Biochem Mol Biol. 2010 Epub ahead of print.
6. Craigie R. Quality control in Mu DNA transposition. Cell 1996;85:137–140. [PubMed: 8612265]
7. Nagy Z, Chandler M. Regulation of transposition in bacteria. Res Microbiol 2004;155:387–398. [PubMed: 15207871]
8. Gueguen E, Rousseau P, Duval-Valentin G, Chandler M. The transpososome: control of transposition at the level of catalysis. Trends Microbiol 2005;11:543–549. [PubMed: 16181782]
9. Craig, N. Tn7. In: Craig, N.; Craigie, R.; Gellert, M.; Lambowitz, A., editors. Mobile DNA II. Washington, D.C: ASM Press; 2002. p. 423-456.
10. Peters J, Craig N. Tn7: smarter than we thought. Nat Rev Mol Cell Biol 2001;2:806–814. [PubMed: 11715047]
11. Bainton R, Kubo K, Feng J-N, Craig N. Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. Cell 1993;72:931–943. [PubMed: 8384534]
12. Skelding Z, Sarnovsky R, Craig N. Formation of a nucleoprotein complex containing Tn7 and its target DNA regulates transposition initiation. EMBO J 2002;21:3494–3504. [PubMed: 12093750]
13. Craig, N.; Craigie, R.; Gellert, M.; Lambowitz, A. Mobile DNA II. Washington, D.C: ASM Press; 2002.
14. Arciszewska LK, McKown RL, Craig NL. Purification of TnsB, a transposition protein that binds to the ends of Tn7. J Biol Chem 1991;266:21736–21744. [PubMed: 1657979]
15. Arciszewska L, Craig N. Interaction of the Tn7-encoded transposition protein TnsB with the ends of the transposon. Nucleic Acids Res 1991;19:5021–5029. [PubMed: 1656385]
16. Arciszewska L, Drake D, Craig N. Transposon Tn7. cis-acting sequences in transposition and transposition immunity. J Mol Biol 1989;207:35–52. [PubMed: 2544738]
17. Lichtenstein C, Brenner S. Unique insertion site of Tn7 in E. coli chromosome. Nature 1982;297:601–603. [PubMed: 6283361]
18. Kuduvalli P, Rao J, Craig N. Target DNA structure plays a critical role in Tn7 transposition. EMBO J 2001;20:924–932. [PubMed: 11179236]
19. Kuduvalli P, Mitra R, Craig N. Site-specific Tn7 transposition into the human genome. Nucleic Acids Res 2005;33:857–863. [PubMed: 15701757]

20. Parks A, Peters J. Transposon Tn7 is widespread in diverse bacteria and forms genomic islands. J Bacteriol 2007;189:2170–2173. [PubMed: 17194796]
21. Cleaver S, Wickstrom E. Transposon Tn7 gene insertion into an evolutionarily conserved human homolog of Escherichia coli attTn7. Gene 2000;254:37–44. [PubMed: 10974534]
22. Biery M, Lopata M, Craig N. A minimal system for Tn7 transposition: the transposon-encoded proteins TnsA and TnsB can execute DNA breakage and joining reactions that generate circularized Tn7 species. J Mol Biol 2000;297:25–37. [PubMed: 10704304]
23. Sarnovsky R, May E, Craig N. The Tn7 transposase is a heteromeric complex in which DNA breakage and joining activities are distributed between different gene products. EMBO J 1996;15:6348–6361. [PubMed: 8947057]
24. May E, Craig N. Switching from cut-and-past to replicative Tn7 transposition. Science 1996;272:401–404. [PubMed: 8602527]
25. Rao J, Miller P, Craig NL. Recognition of triple-helical DNA structures by transposon Tn7. Proc Natl Acad Sci U S A 2000;97:3936–3941. [PubMed: 10737770]
26. Lu F, Craig N. Isolation and characterization of Tn7 transposase gain-of-function mutants: a model for transposase activation. EMBO J 2000;19:3446–3457. [PubMed: 10880457]
27. Stellwagen AE. Analysis of gain of function mutants of an ATP-dependent regulator of Tn7 transposition. J Mol Biol 2001;305:633–642. [PubMed: 11152618]
28. Skelding Z, Queen-Baker J, Craig N. Alternative interactions between the Tn7 transposase and the Tn7 target DNA binding protein regulate target immunity and transposition. EMBO J 2003;22:5904–5917. [PubMed: 14592987]
29. Thiagalingam S, Grossman L. The multiple roles for ATP in the Escherichia coli UvrABC endonuclease-catalyzed incision reaction. J Biol Chem 1993;268:18382–18389. [PubMed: 8349713]
30. Levchenko I, Luo L, Baker T. Disassembly of the Mu transposase tetramer by the ClpX chaperone. Genes Dev 1995;9:2399–2408. [PubMed: 7557391]
31. Nakai H, Doseeva V, Jones JM. Handoff from recombinase to replisome: insights from transposition. Proc Natl Acad Sci U S A 2001;98:8247–8254. [PubMed: 11459960]
32. Nakai H, Kruklitis R. Disassembly of the bacteriophage Mu transposase for the initiation of Mu DNA replication. J Biol Chem 1995;270:19591–19598. [PubMed: 7642646]
33. Bainton, R. Ph.D. Thesis. San Francisco: University of California; 1992. Tn7 transposition in vitro.
34. Muir T, Sondhi D, Cole P. Expressed protein ligation: a general method for protein engineering. Proc Natl Acad Sci U S A 1998;95:6705–6710. [PubMed: 9618476]
35. Holder, J. Ph.D. Thesis. The Johns Hopkins University; 2006. Dissertation: Assembly and Architecture of Tn7 Transposomes.
36. Ronning D, Li Y, Perez Z, Ross P, Hickman A, Craig N, Dyda F. The carboxy-terminal portion of TnsC activates the Tn7 transposase through a specific interaction with TnsA. EMBO J 2004;23:2972–2981. [PubMed: 15257292]
37. Gringauz E, Orle K, Orle A, Waddell C, Craig N. Recognition of *Escherichia coliattTn7* by transposon Tn7: lack of specific sequence requirements at the point of Tn7 insertion. J Bacteriol 1988;170:2832–2840. [PubMed: 2836374]
38. Surette M, Buch S, Chaconas G. Transpososomes: stable protein-DNA complexes involved in the in vitro transposition of bacteriophage Mu DNA. Cell 1987;49:254–262.
39. Sakai J, Chalmers R, Kleckner N. Identification and characterization of a pre-cleavage synaptic complex that is an early intermediate in Tn10 transposition. EMBO J 1995;14:4374–4383. [PubMed: 7556079]
40. Yanagihara K, Mizuuchi K. Progressive structural transitions within Mu transpositional complexes. Mol Cell 2003;11:215–224. [PubMed: 12535534]
41. Stewart BJ, Wardle SJ, Haniford DB. IHF-independent assembly of the Tn10 strand transfer transpososome: implications for inhibition of disintegration. EMBO J 2002;21:4380–4390. [PubMed: 12169640]
42. Au TK, Pathania S, Harshey RM. True reversal of Mu integration. EMBO J 2004;23:3408–3420. [PubMed: 15282550]

43. Lemberg KM, Schweidenback CT, Baker TA. The dynamic Mu transpososome: MuB activation prevents disintegration. J Mol Biol 2007;374:1158–1171. [PubMed: 17988683]
44. Chaconas G, Harshey R. Transposition of Phage Mu DNA. Mobile DNA II 2002:384–402.
45. Davies D, Goryshin I, Reznikoff W, Rayment I. Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. Science 2000;289:77–85. [PubMed: 10884228]
46. Kennedy A, Haniford D, Mizuuchi K. Single active site catalysis of the successive phosphoryl transfer steps by DNA transposases: insights from phosphorothioate stereoselectivity. Cell 2000;101:295–305. [PubMed: 10847684]
47. Bolland S, Kleckner N. The three chemical steps of Tn10/IS10 transposition involve repeated utilization of a single active site. Cell 1996;84:223–233. [PubMed: 8565068]
48. Stellwagen A, Craig N. Mobile DNA elements: controlling transposition with ATP-dependent molecular switches. Trends Biochem Sci 1998;23:486–490. [PubMed: 9868372]
49. Stellwagen A, Craig N. Avoiding Self: Two Tn7-encoded proteins mediate target immunity in Tn7 transposition. EMBO J 1997;16:6823–6834. [PubMed: 9362496]
50. Gamas P, Craig N. Purification and characterization of TnsC, a Tn7 transposition protein that binds ATP and DNA. Nucleic Acids Res 1992;20:2525–2532. [PubMed: 1317955]
51. Gary P, Biery M, Bainton R, Craig N. Multiple DNA processing reactions underlie Tn7 transposition. J Mol Biol 1996;257:301–316. [PubMed: 8609625]



c.

DNA	bp	expected length nm	molecules measured	observed average length nm	standard deviation nm	observed/ expected length	observed average length nm/bp
donor flank	4301	1290	9	1290	148	100%	0.30
PTC Tn7::*attTn7*	1869	560	47	510	65	91%	0.27

Figure 1. AFM images of *Tn7* PTCs

a. Schematic of *Tn7* Transposition

Insertion of *Tn7* (black) into a linear att*Tn7* fragment (red) mediated by proteins that act on the *Tn7* ends (purple) and *attTn7* target (green) results in the excision of *Tn7* from the flanking donor DNA (blue) in formation of the PTC in which the *Tn7* ends are paired.

b. AFM of PTCs

Transposition reactions using a 5926 bp pEMΔ donor plasmid and a 244 bp att*Tn7* −120 to +123 as a target were performed, filter purified, and imaged with an Atomic Force Microscope. Differences in the measured height of PTCs are represented in a 32-color image spectrum where blue is the tallest point in the image and red is the lowest (top row). PTCs

with relaxed *Tn7* DNA (2nd row), PTCs with supercoiled *Tn7* DNA (3rd row) and broken donor DNAs from which *Tn7* has excised (4th row) are shown. Bar = 200 nm.

c. DNA Length Measurements

The length of the linear donor flank from which *Tn7* was measured as control, yielding an observed value of 0.30 nm/ bp, consistent with literature values. The underestimate between the 560 nm expected length of the DNA in the relaxed PTC (1869 bp) and the 510 nm measured length of the DNA in the relaxed PTC must result from the binding of the Tns proteins to the Simple Insertion DNA.

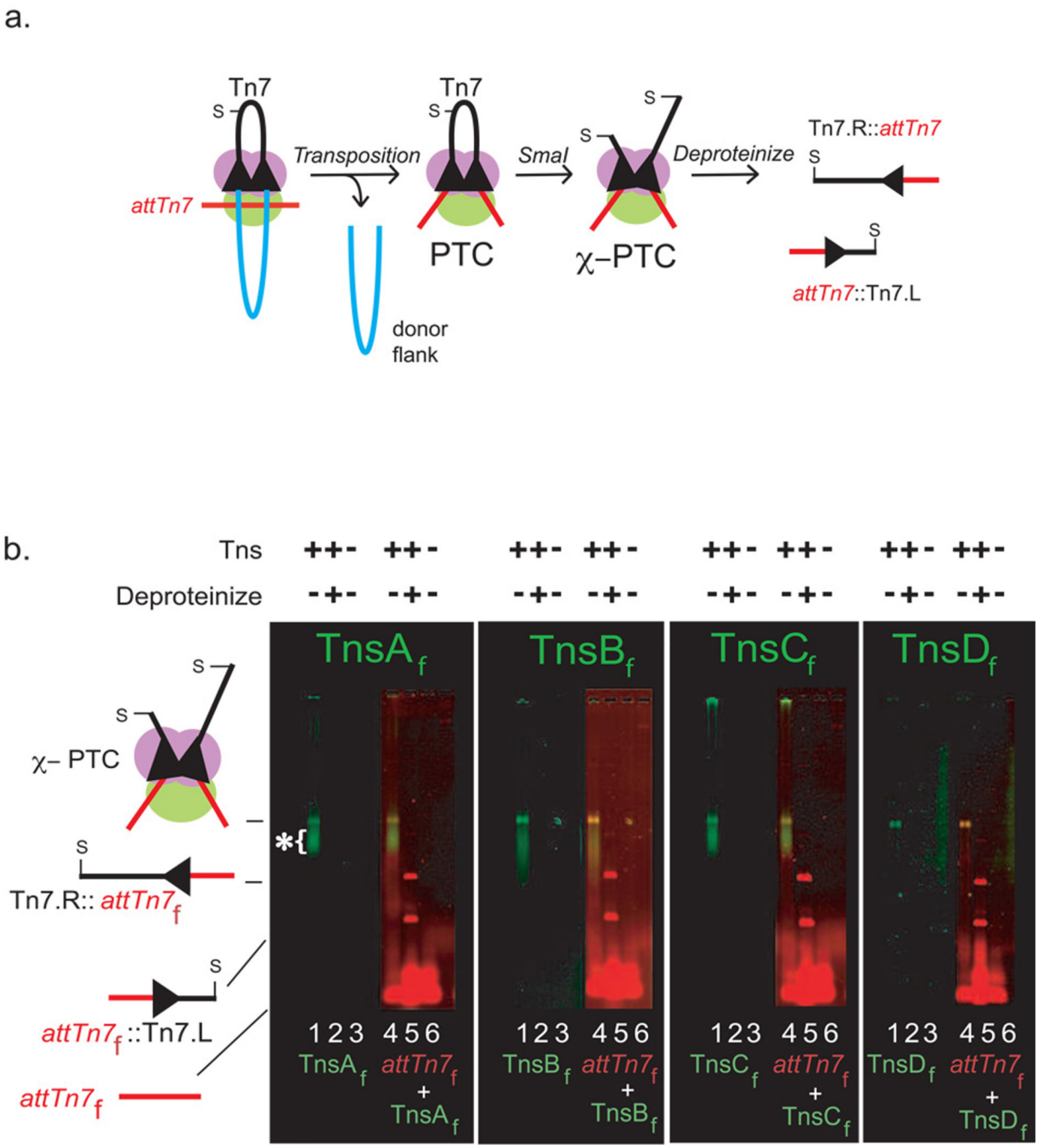


Figure 2. Identification of the Tns proteins in the *Tn7* PTC

a. Schematic of Formation of the *Tn7* PTC and χ-PTC.

Insertion of *Tn7* into a linear *attTn7* fragment results in formation of the PTC in which the *Tn7* ends are paired. Digestion with *SmaI* generates the χ-PTC. Protein removal releases *attTn7*::*Tn7*.L and *Tn7*.R::*attTn7*. The colors used are the same as in Fig 1.

b. Identification of PTC proteins using Fluorescent Tns Proteins

Four different transposition reactions were performed using the indicated Tns_f protein (green), the *Tn7* donor plasmid pEMΔ and *attTn7* $_f$ −38 to +68, digested with *SmaI*, purified by filtration, deproteinized or not as indicated and displayed on a non-denaturing agarose gel. Lanes 1–3 were visualized by fluorescent imaging specific for the Tns_f proteins (green);

lanes 4–6 were visualized by merging the fluorescent imaging for Tns_f proteins (green) and *attTn7*$_f$ (red) such that nucleoprotein complexes are evident as yellow. Reaction products are χ-PTCs containing both *attTn7*::*Tn7*.L and *Tn7*.R::*attTn7* as shown in Fig 3 by hybridization to these gels with *Tn7*L and *Tn7*R probes. The white bracket indicates the target TnsACD-*attTn7* complex (below) and *attTn7*::*Tn7* end complexes that have dissociated from χ-PTCs but are still bound by Tns proteins. Lane 1,4 = non-deproteinized, lane 2,5 = deproteinized, lane 3,6 = no Tns proteins.

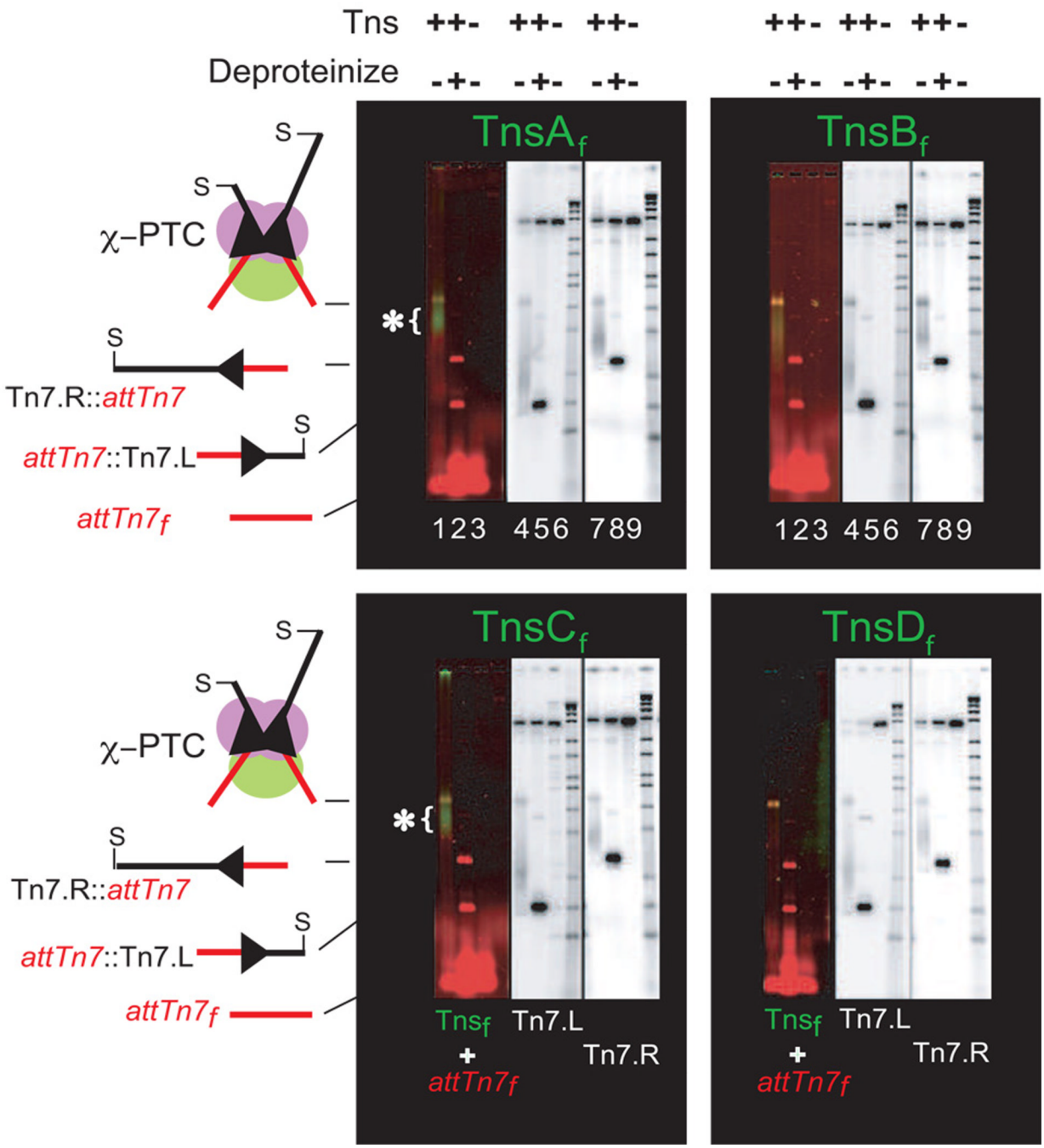


Figure 3. Identification of the *Tn7* ends in the PTC

As in Figure 2a, transposition reactions were performed using Tns proteins labeled with fluorescein (Tns$_f$) (green), a plasmid *Tn7* donor and *attTn7*$_f$(red). The reaction mixtures were digested with *SmaI* to generate χ-PTCs, purified by filtration, deproteinized or not as indicated to yield *attTn7*::*Tn7*.L and *Tn7*.R::*attTn7* DNA fragments, displayed on a non-denaturing agarose gel and visualized by: lanes 1–3 = fluorescent imaging for both the Tns$_f$ proteins and *attTn7*$_f$ target fragment such that nucleoprotein complexes are yellow or phospho-imaging following southern blotting with ^{32}P-labeled oligonucleotide probes specific for lanes 4–6 = *Tn7*.L or lanes 7–9 = *Tn7*.R. The region marked by the white

bracket contains TnsACD-*attTn7* target complexes and *attTn7*::*Tn7* ends that have dissociated from the χ–PTCs but are still bound by Tns proteins.

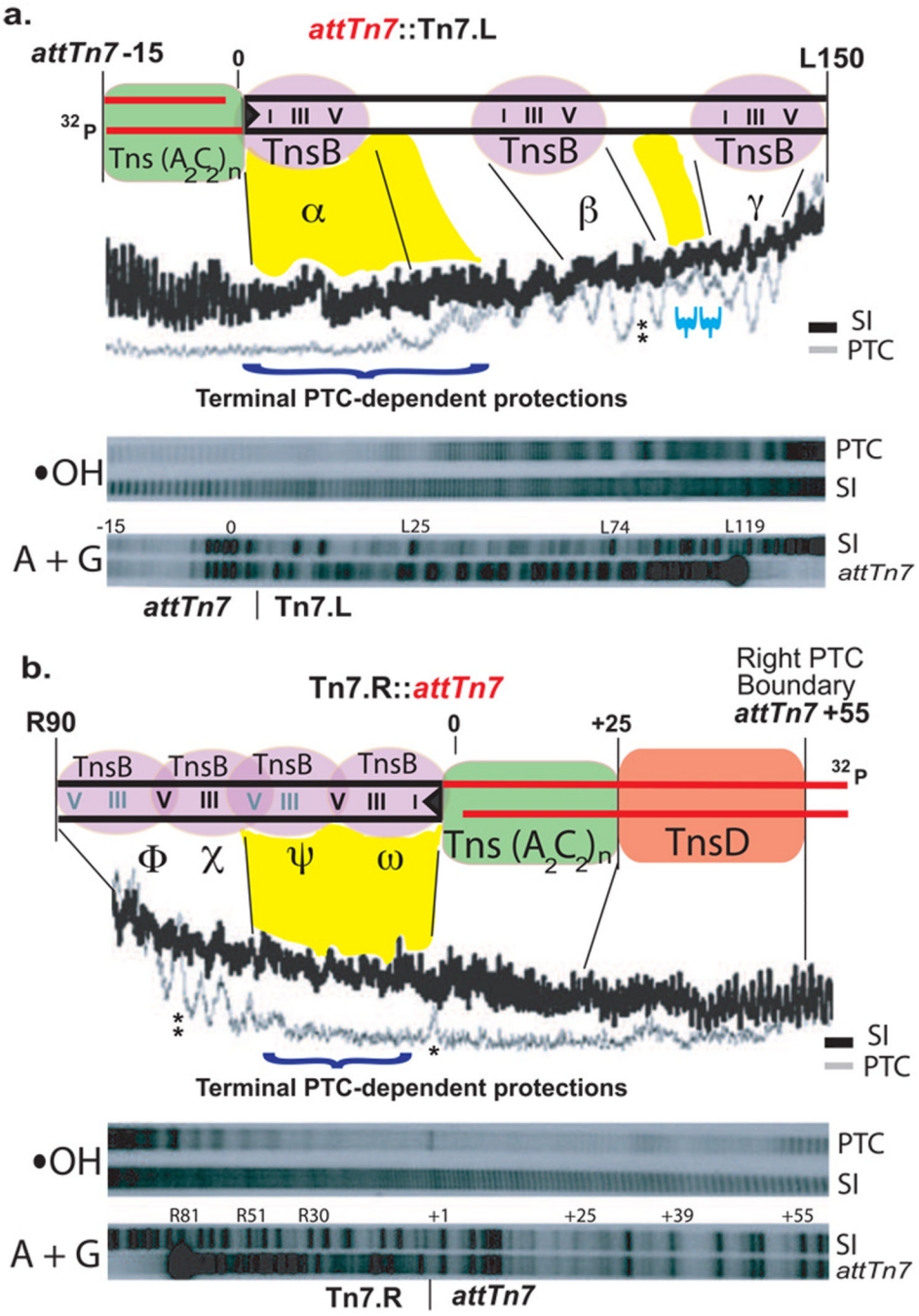


Figure 4. Hydroxyl radical footprinting of the *Tn7* ends in *attTn7*::*Tn7* in the PTC

a. The gel images and line graphs of the hydroxyl radical footprinting comparing *Tn7*L in *attTn7*::*Tn7*.L in the PTC and naked Simple Insertion DNA.

In the PTC, the bottom strand of *attTn7* −83 to +120 is covalently linked to the 3' bottom strand of *Tn7*.L.; the 5' ends of *Tn7* are flanked by gaps. DNA strands in PTCs that contain gapped *attTn7*::*Tn7* and naked non-gapped Simple Insertion *attTn7*::*Tn7* labeled on the indicated strands (^{32}P) were treated with hydroxyl radicals and then displayed on a denaturing sequencing gel. The image of the gel as visualized by phospho-imaging is shown; the positions of hydroxyl radical cleavage were determined by comparison to the products of Maxam-Gilbert (A + G) sequencing reactions. The line graph analysis evaluates

hydroxyl radical cleavage intensity at each nucleotide position; PTC = gray line; naked Simple Insertion (SI) DNA = black line. Binding to all three TnsB binding sites (purple) on *Tn7*.L, the terminal α site and the internal β and γ sites, is observed in the PTC. With *Tn7*.L and TnsB alone, each TnsB site displays a characteristic pattern of protection in regions I, III, and V that are separated by regions of hydroxyl radical attack. In the PTC, the terminal *Tn7*.L α site is more completely protected, as highlighted with dark blue bracket and yellow; weaker PTC-specific protections are also seen between the β and γ TnsB binding sites (light blue brackets and yellow). In the PTC, the *Tn7*.L terminus is flanked by target DNA bound by $Tns(A_2C_2)_n$ complex (green) protecting *attTn7* DNA from hydroxyl radical attack. We suggest the *attTn7* protections outside of the TnsD binding site reflect the binding of an oligomer of $TnsA_2C_2$ subunits (see text).

b. The gel images and line graphs of the hydroxyl radical footprinting comparing *Tn7*R in *Tn7*.R⁖*attTn7* in the PTC and naked Simple Insertion (SI) DNA. The 3' top strand of *Tn7*.R is covalently linked to the top strand of *attTn7* –83 to +120. The TnsD binding site (orange) and multiple TnsB binding sites ω, ψ, χ and ϕ (purple) on *Tn7*.R are indicated. In the PTC, more complete protection at the terminal TnsB sites ω, ψ in which the regions between the characteristic III and V are now impervious to hydroxyl radical attack as highlighted with dark blue bracket and yellow. *attTn7* DNA is protected in the region between the TnsD binding site and the *Tn7*.R terminus by $Tns(A_2C_2)_n$ complex (green). (**) = position of *attTn7* DNA that contaminates the PTC preparation; (*) = position of a substrate DNA contaminant, i.e. is not related to recombination.

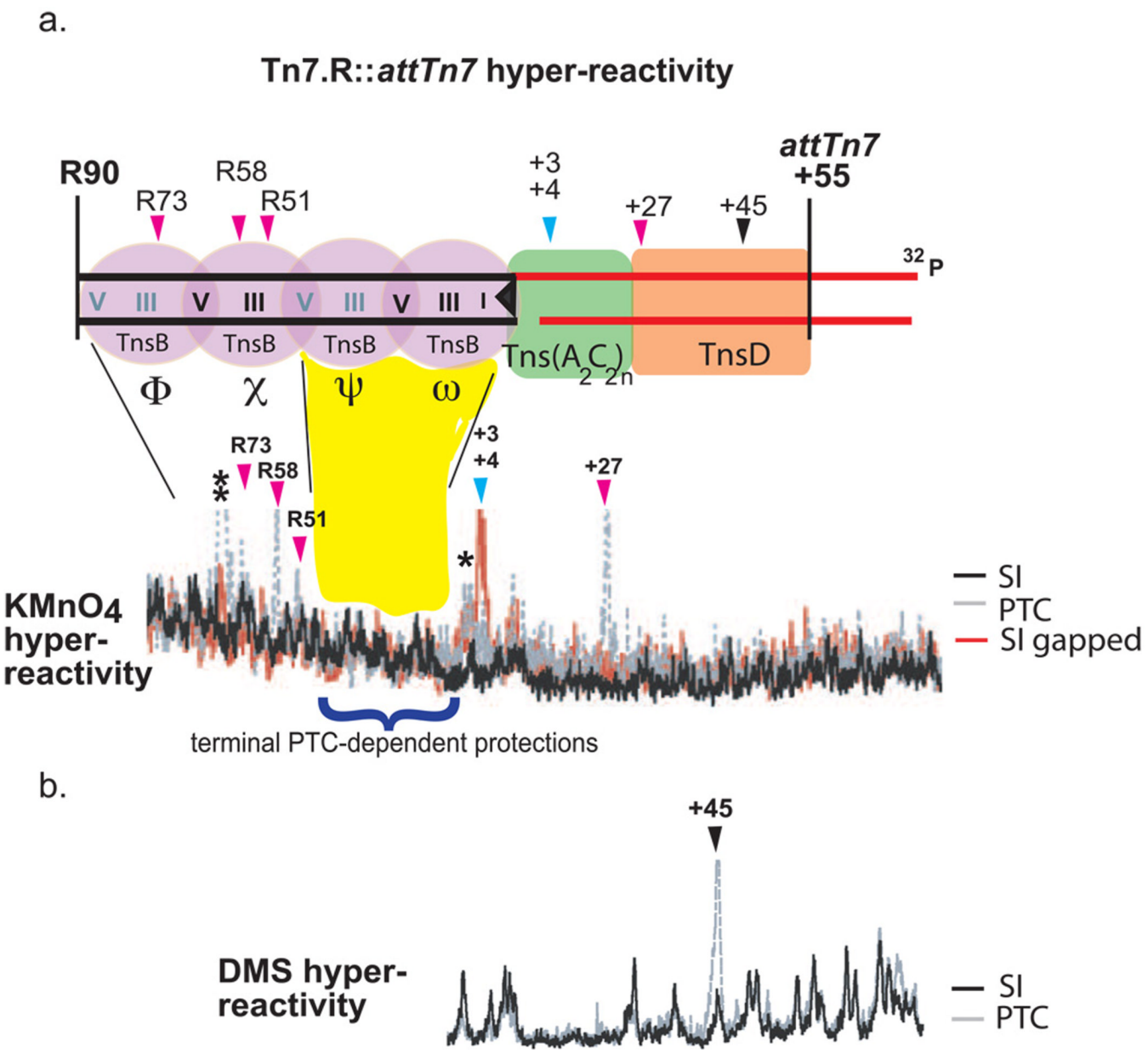


Figure 5. Distorted DNA structures are present in the PTC

In both a and b, *attTn7* −83 to +120 labeled at the 3' end of the top strand (^{32}P) was used as a target for *in vitro* transposition using a donor plasmid containing the 373 bp *Tn7* element, the product PTCs purified by filtration, footprinted as indicated, the products displayed on a denaturing gel and subjected to line graph analysis.

a. Footprinting of *Tn7*R::att*Tn7* with $KMnO_4$

$KMnO_4$ reacts preferentially with distorted thymidine nucleotides. *Tn7*R::*attTn7* in the PTC (gray line graph), as naked duplex DNA (black line graph) and gapped, DNA (red line graph) were footprinted. Four positions of $KMnO_4$ hyper-reactivity (magenta triangles) are observed in the PTC: in *Tn7*.R at R73, R58, and R51 and at *attTn7* +27. The blue triangle indicates positions of hypersensitivity in the att*Tn7* gap flanking *Tn7*R in naked, gapped *Tn7*R::*attTn7* DNA. (**) indicates contaminating *attTn7* DNA in the PTC preparation.

b. Footprinting of *Tn7*R::*attTn7* with DMS

DMS reacts preferentially with distorted G and A nucleotides. One position of DMS hyper-reactivity (black triangle) is observed at *attTn7* +45.

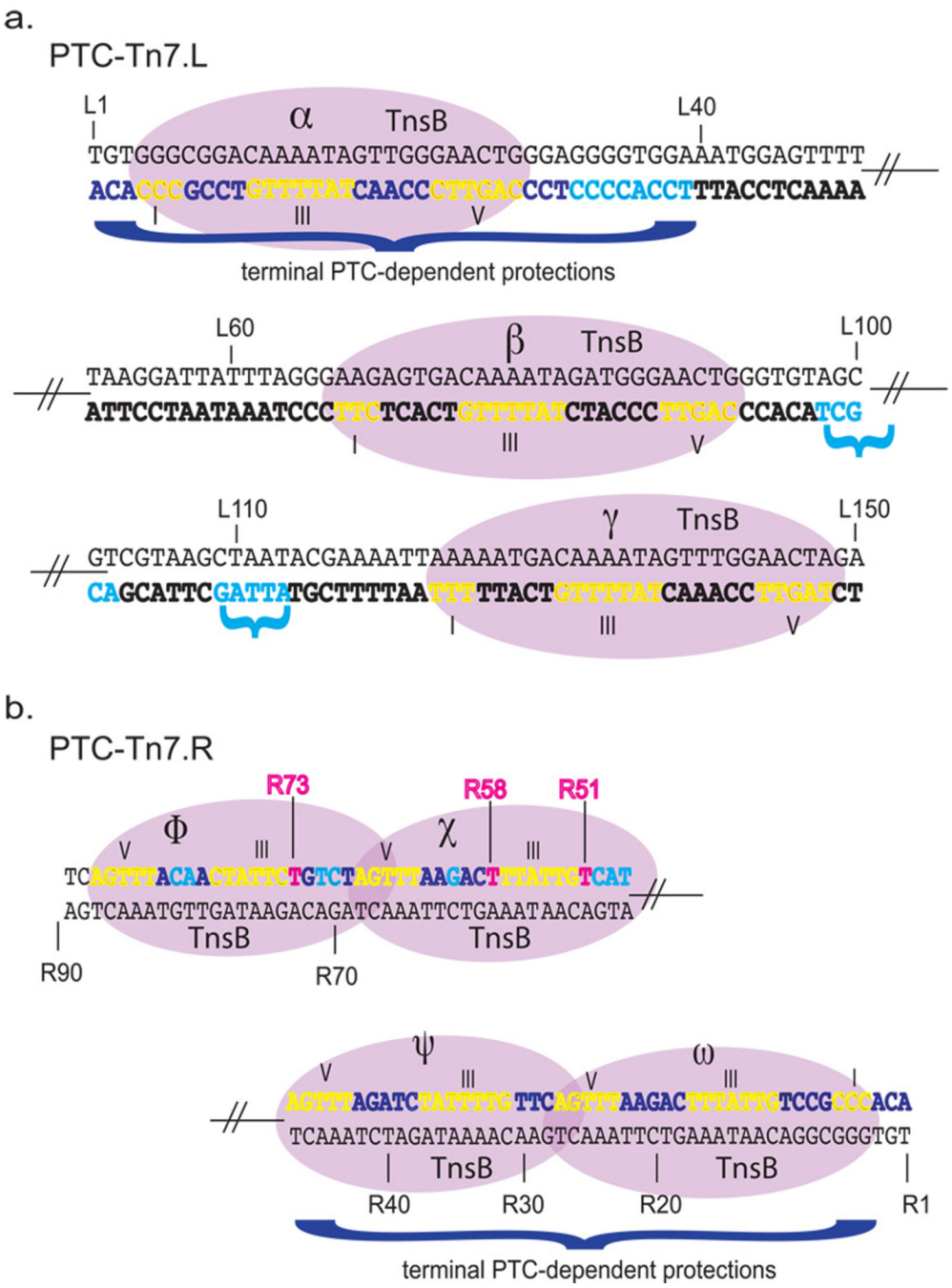


Figure 6. Summary of footprinting of the *Tn7* ends in *attTn7*::*Tn7* the PTC

a. DNA sequence summary of hydroxyl radical footprinting of *Tn7*.L.

The bottom strand of *Tn7*.L that is covalently linked to *attTn7* at (L1) in the PTC is represented with colors indicating degree of protection. Purple ovals = TnsB binding sites; black = unprotected; yellow = protection observed with in the PTC and with just TnsB[14]; 15 dark blue = strong PTC-specific protection and light blue = weaker PTC-specific protection. Dark blue bracket indicate regions of strong terminal PTC protections while light blue brackets indicate regions of weaker protection.

b. DNA sequence presents a summary of hydroxyl radical footprinting of *Tn7*.R and hyper-reactivity to $KMnO_4$.

The *Tn7*.R protections of the top strand that is covalently linked to *attTn7* at (R1) are represented as in (A). $KMnO_4$ hyper-cleavages are magenta triangles at positions R51, R58, and R73.

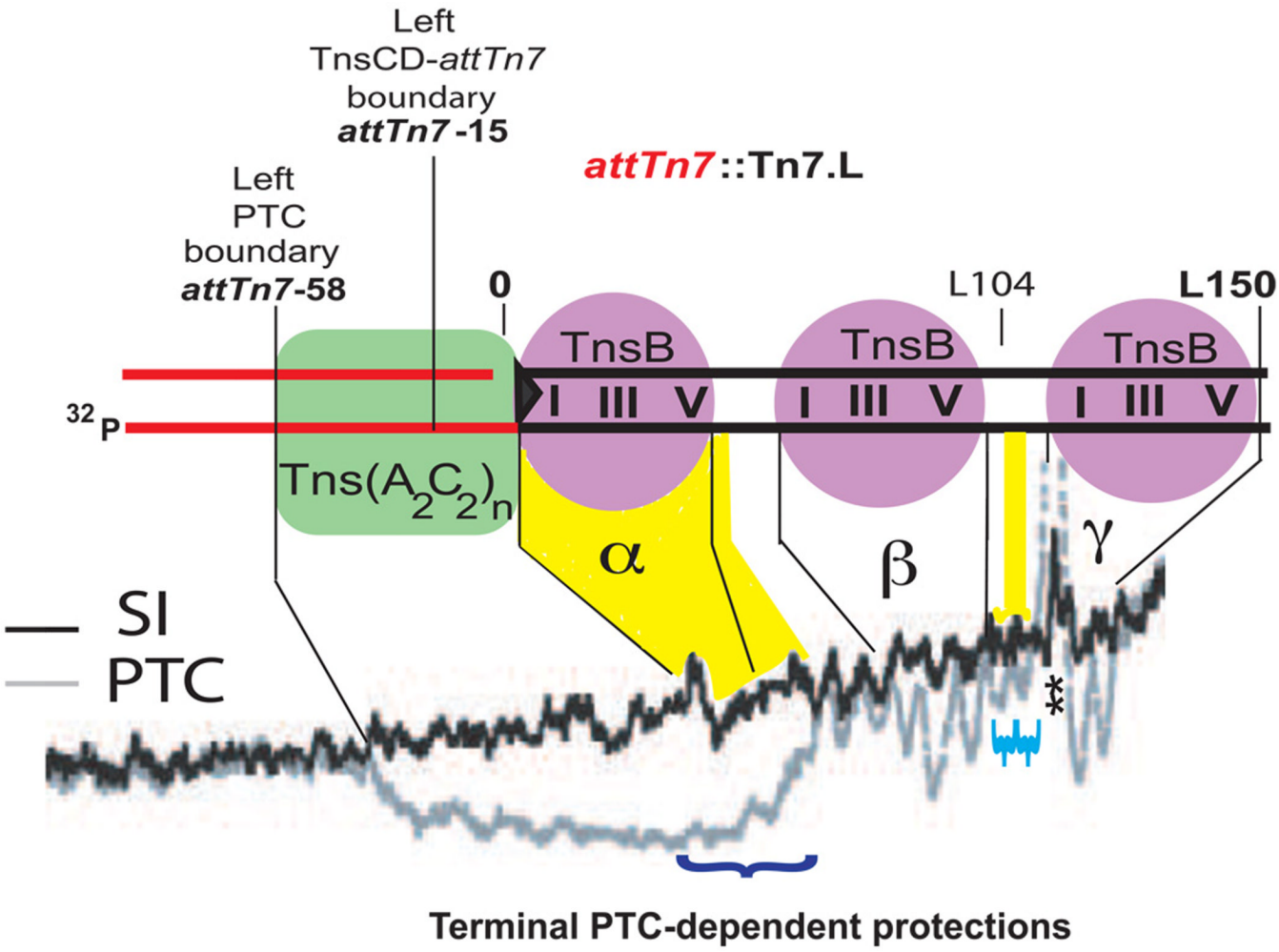


Figure 7. Hydroxyl radical footprinting of *attTn7::Tn7L* in the PTC

Transposition using a donor plasmid containing a 373 bp *Tn7* element into bottom-strand 3' end-labeled (^{32}P) *attTn7* −300 to +120 results in labeled PTCs. Labeled target DNA (red) is joined to 3'OH *Tn7* ends (black). Line graph analysis of labeled PTCs evaluates hydroxyl radical cleavage intensity at each nucleotide position: PTC = gray dotted line, naked Simple Insertion DNA = black line. Protections are observed at previously identified TnsB binding sites (purple) at sites α, β, and γ on *Tn7*.L with conserved domains of protection I, III, and V and flanked by *attTn7* DNA bound by Tns$(A_2C_2)_n$ complex (green) protecting *attTn7* to a left boundary ~ 60 nucleotides from the *Tn7* point of insertion. Strong terminal PTC-dependent protections (dark blue bracket, yellow highlight) and weaker protections are observed between the β and γ sites (light blue brackets, yellow highlight). (**) indicates contaminating *attTn7* DNA in the PTC preparation.

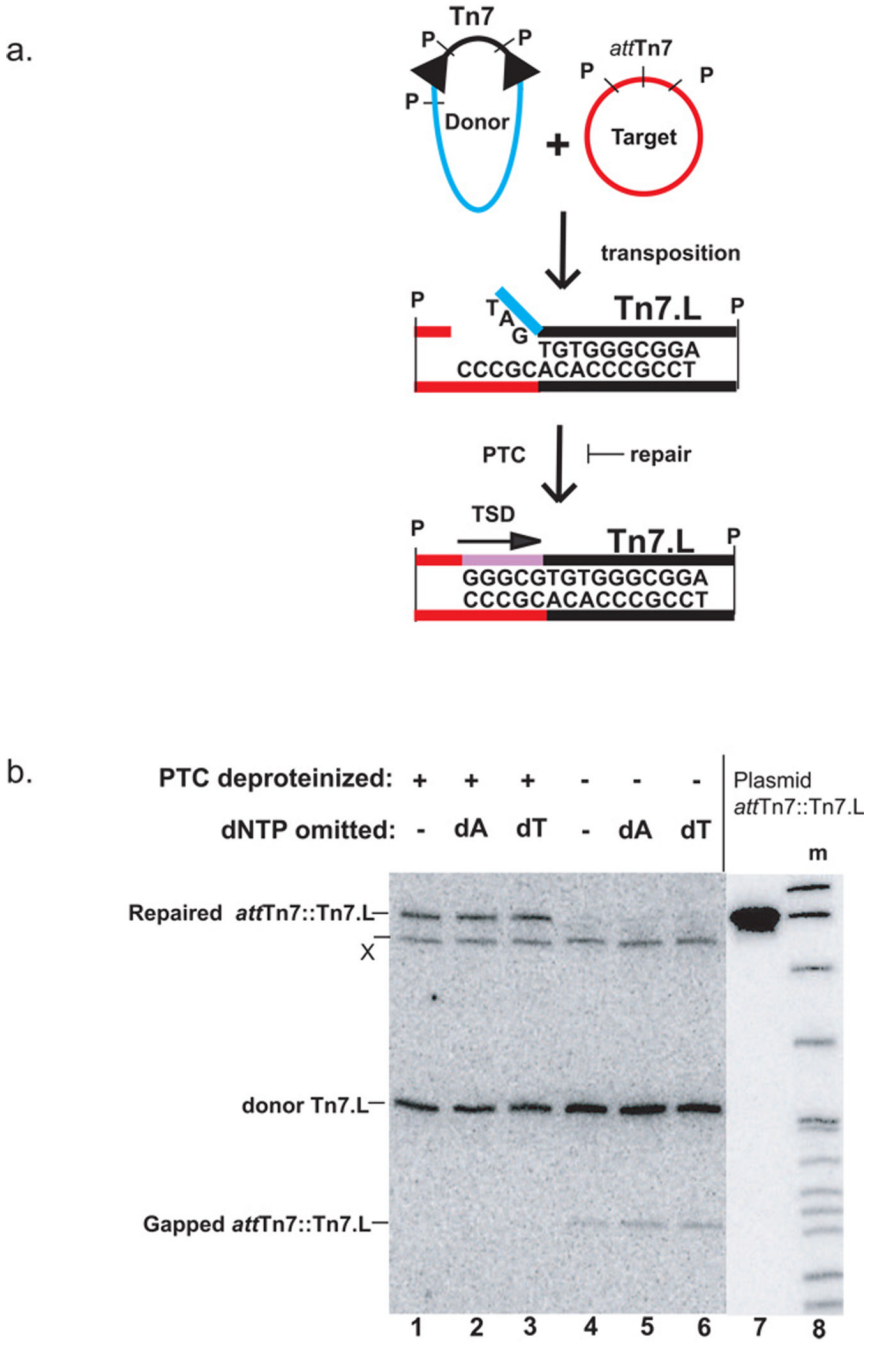


Figure 8. The *Tn7* PTC blocks gap repair at the 5' ends of *Tn7*

a. Schematic of gap repair required at the 5' end of *Tn7* following transposition. *Tn7* insertion into *attTn7* (red) results in covalent joining of the 3' end of *Tn7* to *attTn7* and a gap at the 5' end of *Tn7* which is flanked by 3 nucleotides of donor DNA (blue). Gap repair using the bottom strand of *attTn7* as a template results in the 5 bp target site duplication characteristic of *Tn7* insertion.

in vitro repair of *Tn7* transposition products with DNA pol I and DNA ligase in the presence and absence of deproteinization

b. The gapped transposition products that were deproteinized as indicated were subjected to in vitro repair with DNA pol I and DNA ligase. Following repair, reactions were digested by

PstI (P) to release *attTn7*::*Tn7*.L. Denaturing PAGE and southern blotting with the *Tn7*.L end-specific probe NLC 272 reveals that deproteinized transposition products are repaired as evidenced by the joining of *Tn7* 5' end to *attTn7* DNA (lanes 1–3), repair occurs in the absence of nucleotides necessary for synthesis inside the *Tn7* end (lanes 2,3). Repair does not occur in the PTC (lanes 4–6), control repaired Simple Insertion (lane 7) and DNA ladder (pBR322 + *MspI*) (lane 8). X = non-specific crosshybridization of the probe to target plasmid substrate.

Table 1

Determination of Tns Stoichiometry in the PTC

Fluorescent Tns protein	Fluorescein / ^{32}P *attTn7*	ratio relative to $TnsD_f$ / ^{32}P *attTn7*
$TnsA_f$	2.1	18
$TnsB_f$	0.76	6.3
$TnsC_f$	3.0	25
$TnsD_f$	0.12	1.0

Transposition reactions were performed with each of the 4 Tns_f proteins (column 1) in separate reactions using 1.6 kB *Tn7* and 107 bp ^{32}P-labeled *attTn7*, the PTCs filter-purified, crosslinked with 0.01% gluteraldehyde for 10 min and then quenched with 10 mM Tris/ 10 mM lysine. The reaction products were displayed on a 1% agarose gel. Supercoiled PTCs were used for $TnsA_f$, $TnsB_f$, and $TnsC_f$ protein and DNA quantification while both relaxed and supercoiled PTCs were used for $TnsD_f$. Complexes containing $TnsA_f$, $TnsB_f$, $TnsC_f$ and $TnsD_f$ were quantified by fluorescence. The gel was then dried down onto DEAE paper and the amount of ^{32}P *attTn7* in each complex determined by phorphorimaging. The amount of fluoroscein to ^{32}P in each PTC reaction is shown in arbitrary units (column 2). The ratio of $TnsD_f$/ ^{32}P *attTn7* is then set equal to 1 and the amount of the other complexes expressed relative to this value (column 3).