Interaction of the Tn7-encoded transposition protein TnsB with the ends of the transposon

Lidia K.Arciszewska[1,3+] and Nancy L.Craig[1,2,3]*

[1]Department of Microbiology and Immunology, [2]Department of Biochemistry and Biophysics and [3]The George W.Hooper Foundation, University of California, San Francisco, CA 94143, USA

ABSTRACT

We have used several high resolution methods to examine the interaction of TnsB, a transposition protein encoded by the bacterial transposon Tn7, with its binding sites at the ends of the transposon. These binding sites lie within the DNA segments that are directly involved in transposition. We show that the binding of TnsB to DNA can promote DNA bending, suggesting that the interaction of TnsB with the ends may result in formation of a highly organized protein-DNA complex. We also identify likely positions of close contact between TnsB and its binding sites. Analysis of the interaction of TnsB with intact Tn7 ends reveals TnsB occupies its binding sites in a particular order, the sites immediately adjacent to the transposon termini being occupied only after other inner sites are bound. Such ordered occupancy suggests that the various binding sites have differing apparent affinities for TnsB.

INTRODUCTION

Transposable elements are discrete DNA segments that can move from one genomic location to another. A key step in transposition is the recognition of DNA sequences at the ends of the element that are directly involved in transposition by recombination proteins. Specific recognition of the ends is the initial step in a series of events that likely involves formation of higher-order protein-DNA complexes that contain both ends of the element, breakage of DNA strands at the transposon termini and joining of the transposon ends to target DNA. Both transposon- and host-encoded proteins may participate in end recognition (1).

We are interested in understanding the transposition mechanism of the bacterial transposon Tn7 (2,3). We have previously established that TnsB, a Tn7-encoded transposition protein, binds specifically to both ends of Tn7 (4,5). *tnsB* is one of an elaborate array of Tn7-encoded transposition genes, *tnsABCDE*, that mediate two distinct but overlapping transposition pathways differing in target site-selectivity (6—8). *tnsABC + tnsD* promote high-frequency insertion into a specific site in the *Escherichia coli* chromosome called *attTn7* whereas *tnsABC + tnsE* promote

low-frequency insertion in apparently random target sites. The interaction of TnsB with the transposon ends accounts for the requirement for *tnsB* in all Tn7 transposition reactions.

Previous DNase I protection studies (5) revealed that TnsB binds to several sites which are highly related in nucleotide sequence that lie within the Tn7 end segments that participate directly in recombination (9). We report here the results of high resolution studies of the interaction of TnsB with DNA using various chemical footprinting and binding interference techniques. Analysis of the interaction of TnsB with a single binding site has identified likely positions of close contact between TnsB and DNA and has also shown that TnsB binding can introduce a bend into DNA. We have also explored in more detail the interaction of TnsB with intact Tn7 ends. We have observed that the TnsB binding sites are occupied in a particular order, suggesting the various binding sites differ in their relative apparent affinities for TnsB. Our analyses support the view that Tn7 recombination involves the ordered assembly of specialized protein-DNA complexes containing the ends of Tn7.

MATERIALS AND METHODS

Strains and Plasmids

Fragments containing intact Tn7R (R1—199) and Tn7L (L1—166) were isolated from pKS⁺Ra and pKS⁺La (5), respectively. Fragments containing L109—166 and L89—166 were isolated from pLA77 (5) and pLA76. pLA76 was constructed by inserting by blunt-end ligation a *HgaI-HincII* fragment, obtained by digestion of the *EcoRI-HincII* Tn7L fragment from pLA26 (9), into the *EcoRV* site of Bluescript-KS (Stratagene); in pLA76, position L166 is adjacent to the vector *EcoRI* site. To construct pLA78, a fragment containing the L119—155 segment, flanked on both ends by *XbaI* staggered ends, was obtained by annealing two 42 bp synthetic oligonucleotides,

5′-CTAGAATTAAAAATGACAAAATAGTTTGGAACTAGATTTCAT and
5′-CTAGATGAAATCTAGTTCCAAACTATTTTGTCATTTTTAATT.

and then inserted into the *XbaI* site of pBend2 (10) with L155 adjacent to the vector *EcoRI* site.

Preparation and labeling of DNA fragments

Plasmid DNA was digested with appropriate restriction enzymes and electrophoresed through 6% polyacrylamide gels. Slices containing relevant fragments were excised, the DNA electrophoretically transferred onto a DEAE membrane (Schleicher and Schuell) and recovered as suggested by the manufacturer. Purified DNAs were labeled at their 3' ends by incubation with DNA polymerase I Klenow fragment and appropriate α-^{32}P-NTPs (11). Fragments from pKS$^+$Ra were used for analysis of R1 − 199; the 306 bp SmaI-HindIII fragment labeled at HindIII was used for bottom strand analysis and the 294 bp EcoRI-BamHI fragment labeled at EcoRI was used for top strand analysis (top and bottom strands are as shown in Figure 5). Fragments from pKS$^+$La were used for analysis of L1 − 166; the 270 bp SmaI-HindIII fragment labeled at HindIII was used for top strand analysis and the 258 bp EcoRI-BamHI fragment labeled at EcoRI was used for bottom strand analysis. For analysis of L109 − 166, fragments from pLA77 were used; the 85bp SmaI-HindIII fragment labeled at HindIII was used for top strand analysis and the 91 bp EcoRI-HincII fragment labeled at EcoRI was used for bottom strand analysis. For analysis of the bottom strand of L89 − 166, the 104 bp HindIII-SmaI fragment labeled at HindIII from pLA76 was used.

DNA Binding and Band Shift Assays

Unless otherwise indicated, binding reactions (100 μl contained 13.1 mM Tris·HCl (pH 8.0), 1.03 mM DTT, 0.03 mM EDTA, 110 mM NaCl, 5 mM MgCl$_2$, 1 mM CaCl$_2$, 10.3% (v/v) glycerol, 2 ng/ml poly(dI-dC)·poly (dI-dC), 275 μg/ml BSA, approximately 0.15 pmol of 3' end-labeled DNA fragment and TnsB as indicated. After a 20 min incubation at room temperature, protein-DNA complexes were attacked with hydroxyl radicals or separated by electrophoresis (5) and analyzed as described below.

Hydroxyl radical footprinting

Reactions (150 μl) contained 13.3 mM Tris·HCl (pH 8.0), 0.003 mM DTT, 0.13 mM EDTA, 115 mM NaCl, 1 mM CaCl$_2$, 0.15% (v/v) glycerol, 2 μg/ml poly(dI-dC)·poly(dI-dC), 250 μl BSA, approximately 0.23 pmol 3' end-labeled DNA fragment and TnsB as indicated. After incubation with TnsB, hydroxyl radical cleavage of DNA was performed as described (12) except that the reaction was carried out for 6 min. DNA was recovered by ethanol precipitation and analyzed on 8% sequencing gels. To decrease agents that interfere with cleavage, TnsB was dialyzed against 25 mM Tris·HCl (pH 8.0), 0.1 mM DTT, 1 mM EDTA, 500 mM NaCl and 5% (v/v) glycerol.

Missing nucleoside binding interference

Experiments were performed as described (13). 3' end-labeled DNA fragments were modified by hydroxyl radical attack as described for footprinting in 12.5 mM Tris.HCl (pH 8.0), 0.1 mM EDTA, 100 mM NaCl and 1 mM CaCl$_2$. Protein-DNA complexes were formed in reactions as described and then free and protein-bound DNAs were separated by electrophoresis. Slices containing the complexes of interest were excised, the DNA electrophoretically transferred to a DEAE membrane (Schleicher and Schuell), eluted, recovered by ethanol precipitation and analyzed on a DNA sequencing gel.

Methylation binding interference

Alkylation of 3' end-labeled DNA fragments was performed as described (14). Protein-DNA complexes were formed in 100 μl reactions containing 13.7 mM Tris.HCl (pH 8.0), 1.1 mM DTT, 0.1 mM EDTA, 10.5% (v/v) glycerol, 130 mM NaCl, 300 μg/ml BSA, approximately 0.15 pmol modified DNA fragment and TnsB as indicated. After incubation for 20 minutes at room temperature, free and protein-bound DNAs were resolved and recovered as described above, cleaved with piperidine and analyzed on a sequencing gel.

Copper-phenanthroline footprinting

DNA binding reactions were performed as above except that the reactions contained approximately 0.28 pmol 3' end-labeled DNA fragment. Protein-DNA complexes were resolved by electrophoresis in a 4% polyacrylamide gel. The gel was subjected to chemical attack with 1,10-phenanthroline copper ion as described (15). Slices containing the DNA of interest were excised and DNA electrophoretically transferred into a dialysis bag (16). DNA was recovered by ethanol precipitation and analyzed on sequencing gel.

Analysis of DNA bending by circular permutation

pLA78 DNA was digested with appropriate restriction enzymes and fragments recovered as described above. Approximately 1 pmol of DNA was incubated in 20 μl reaction mixtures containing 13.7 mM Tris·HCl (pH 8.0), 1.05 mM EDTA, 2.05 mM DTT, 125 mM NaCl, 10.5% (v/v) glycerol, 2 ng/ml poly(dI-dC)·poly(dI-dC), 300 μg/ml BSA and 1.3 μg/ml TnsB for 7 min at room temperature and then electrophoresed through a 6% polyacrylamide gel.

Other methods

Preparation of purified TnsB protein was as described (5). DNA sequencing was performed using the chemical degradation method (14).

RESULTS

The interaction of TnsB with a single binding site

DNase I protection experiments demonstrated that TnsB binds to multiple sites in the ends of Tn7 (5). We have now performed hydroyxl radical protection and binding interference studies using a fragment containing a single TnsB binding site, designated γ, which is the innermost TnsB binding site in Tn7L. Hydroxyl radical footprinting provides a high resolution picture of protein-DNA interactions because of the small size of the probing reagent which attacks deoxyribose (12). Experiments showing TnsB protection of the γ binding site against hydroxyl radical attack are shown in Figure 1 and summarized in Figure 2. TnsB protects several distinct clusters of nucleotides, designated I, II, III, IV, and V, that lie on both strands within the region protected by TnsB against DNase I attack.

Three protected clusters, III, IV and V, fall at positions that are highly conserved in sequence identity among the seven proposed TnsB binding sites (5; Figure 5C). A notable feature of these conserved sequences is a four bp inverted repeat located asymmetrically, i.e. towards the right edge, within the TnsB binding site (Figure 2). (In the γ site, these repeats are five bp.) Two clusters, IV and V, lie symmetrically on the inverted repeat. The 'spacer' region between the repeats is hypersensitive to
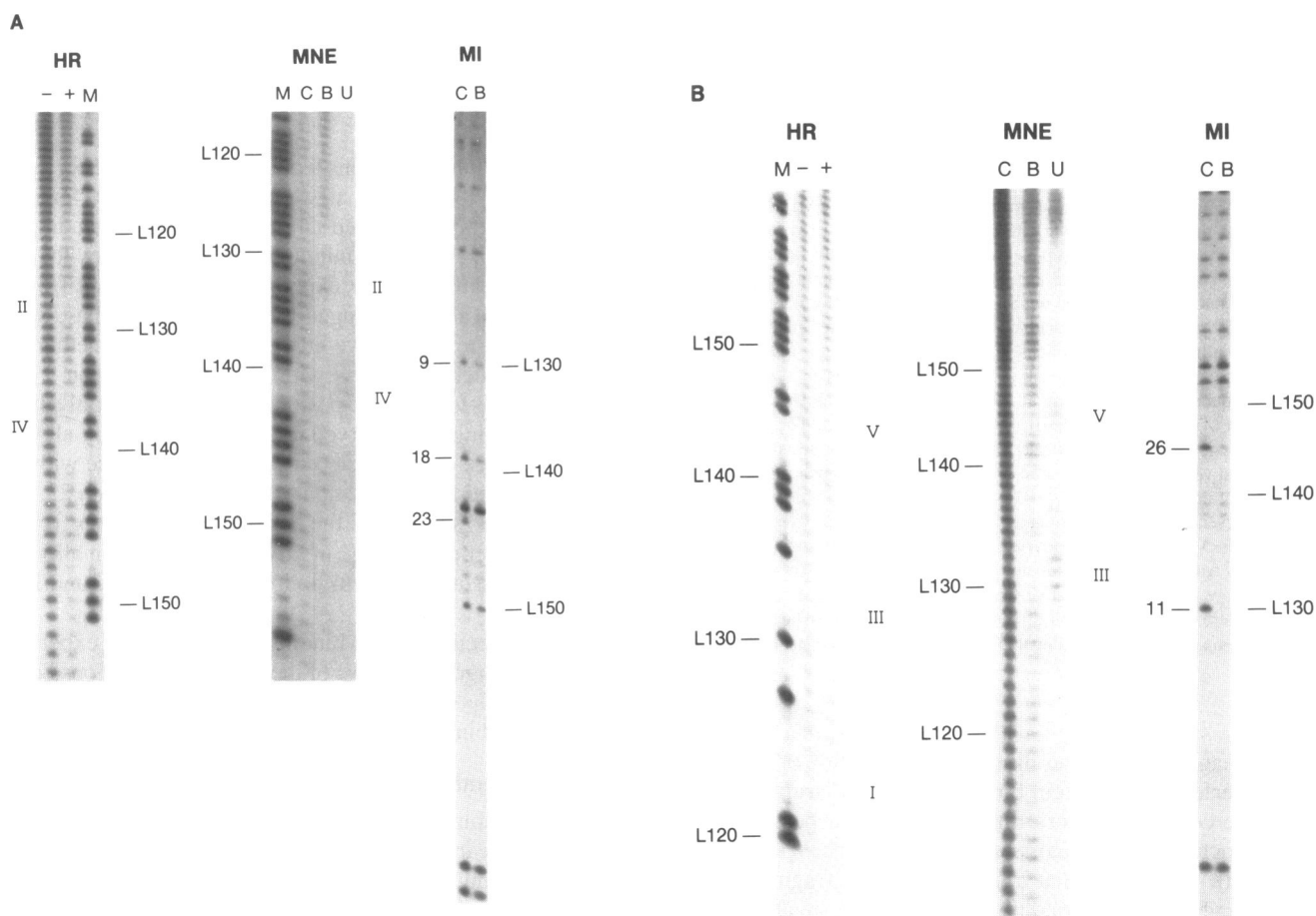
hydroxyl radical attack and DNase I attack in the presence of TnsB. Cluster III is located on one side (leftwards) of the inverted repeats, approximately in the middle of the binding site. Cluster II and the much weaker cluster I lie in a region of considerably less sequence similarity among the sites; this region is, however, also protected by TnsB from DNAse I attack.

We also performed binding interference studies with γ site DNA that was previously modified by hydroxyl radical attack to introduce single, random nucleoside gaps (13,17). The gapped DNA was incubated with TnsB under conditions where more than 95% of the DNA bound to TnsB, protein-bound and free DNA were separated by gel electrophoresis, extracted and analyzed on a DNA sequencing gel (Figure 1). Fragments lacking a nucleoside whose presence is important for binding will be overrepresented in the DNA that did not bind to TnsB (unbound) compared to the input (control) DNA; these fragments will also be underrepresented in bound DNA compared to the input control. Fragments lacking nucleosides that are not involved in binding will be present in the same proportion in the bound,
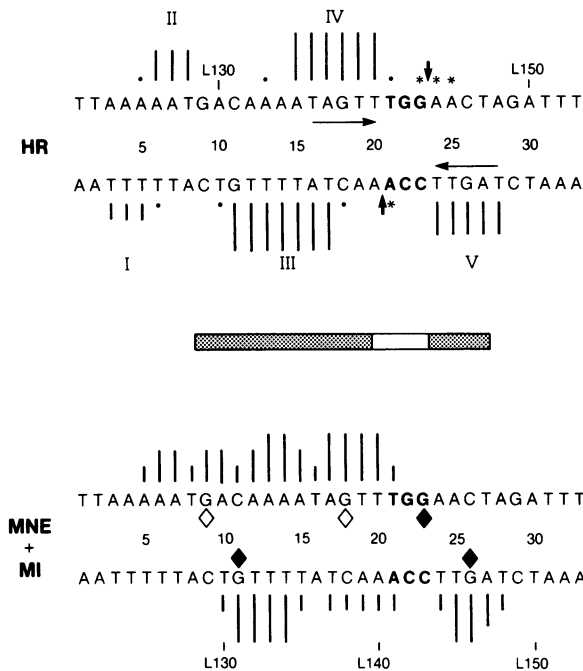
unbound and input DNAs. These experiments are summarized in Figure 2.

In general, binding of γ site DNA to TnsB is most impaired by the absence of nucleosides at the same positions where the strongest protection from hydroxyl radical attack was observed, i.e. within protected clusters II, III, IV and V. This finding suggests that these clusters reflect direct protein-DNA contacts. We also observe a modest effect on binding when nucleosides are missing at several positions not obviously protected from hydroxyl radical attack, i.e. at positions 9−14 on the top strand and 18−21 on the bottom strand. These modest effects may reflect the absence of specific nucleoside contacts or by local structural changes in the DNA backbone upon removal of the nucleosides at these positions (18). Many of the positions where nucleoside removal has only a modest effect on TnsB binding are base-paired to nucleosides that show strong protection and binding interference effects.

We also probed the interaction of TnsB with alkylated DNA. Dimethylsulphate (DMS) alkylation occurs in the major groove
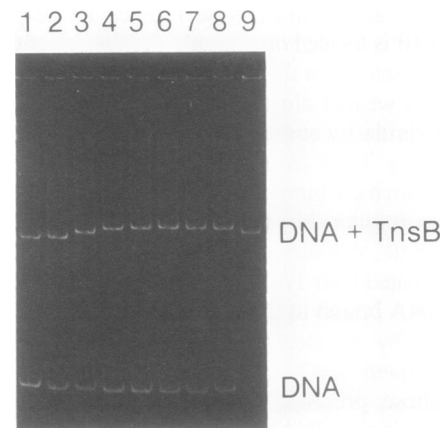

Figure 1. Interaction of TnsB with a single binding site evaluated by protection and binding interference analyses. Interaction of TnsB with the Tn7L γ binding site was evaluated by protection from hydroxyl radical attack and by binding interference with DNA modified by either hydroxyl radical attack or alkylation. Panels A and B present analyses of the top and bottom strands, respectively. Hydroxyl radical attack (HR) was performed in the absence (−) or presence (+) of 2 μg/ml TnsB. Clusters of protection (I, II, III, IV and V) are indicated. In the binding interference experiments which used DNA modified by hydroxyl radical attack, i.e. missing nucleosides (MNE), the modified DNA was incubated with 2.5 μg/ml TnsB. Protein-bound DNA (B) was separated from unbound DNA (U) by native gel electrophoresis and compared to control DNA (C). Clusters of nucleosides whose absence interferes with TnsB binding (II, III, IV and V) are indicated. A + G chemical sequencing reactions were used as markers (M) in the HR and MNE analyses. Nucleotides within the left end of Tn7 are indicated, L1 being the terminal nucleotide. Binding reactions for the methylation interference (MI) analysis, which used DNA modified by alkylation, were performed with 0.1 μg/ml TnsB. Protein-bound DNA (B) was separated from free DNA by native gel electrophoresis and compared to control DNA (C). Bases whose methylation affects TnsB binding are indicated by numbers which refer to their position within the γ binding site.

Figure 3. Circular permutation analysis of TnsB-induced DNA bending at the γ binding site. The mobility of TnsB-DNA complexes in which the γ binding site was located at different positions in the DNA fragment was examined. Binding reactions were performed with 1.3 μg/ml TnsB. The mobility of free and protein-bound DNA was examined on a native polyacrylamide gel and DNA visualized with ethidium bromide. The DNA fragments used were obtained from pLA78 by digestion with the indicated restriction enzymes: lane 1, *MluI*; lane 2, *NheI*; lane 3, *SpeI*; lane 4, *EcoRV*; lane 5, *PvuII*; lane 6, *StuI*; lane 7, *NruI*; lane 8, *KpnI*; lane 9, *BamHI*.

Figure 2. Summary of protection and interference analyses of a single TnsB binding site. The top panel summarizes the results of the hydroxyl radical protection (HR) experiments and the bottom panel summarizes the results of the binding interference studies that used DNA modified by hydroxyl radical attack, i.e. missing nucleosides (MNE) or by alkylation, i.e. methylation interference (MI) (Figure 1). The sequence of the γ TnsB binding site as defined by DNase I protection is shown (5). Numbers above the top panel and below the bottom panel refer to positions within Tn7L, L1 being the terminal nucleotide; numbers between the strands refer to nucleotide positions within a TnsB binding site (5). Horizontal arrows mark a 5 bp inverted repeat within the γ site; this inverted repeat is 4 bp in most TnsB binding sites. Nucleotides in this spacer region are marked with bold letters. The grey box indicates the nucleotide positions that are highly conserved in sequence identity among all identified TnsB binding sites. HR (top panel): Clusters of protection (I, II, III, IV and V) are marked. The degree of protection (strong, moderate and weak) observed in the presence of TnsB is indicated (longest vertical bar = strongest protection); filled dots indicate positions of very weak protection. Stars mark positions hypersensitive to hydroxyl radical attack in the presence of TnsB; vertical arrows indicate positions hypersensitive to DNaseI attack. MNE + MI (bottom panel): The effects of missing a particular nucleoside on TnsB binding as evaluated by examination of both bound and unbound DNA are indicated (longest vertical bar = strongest block to binding). Nucleotides where base methylation affects TnsB binding are indicated (filled diamonds = strong inhibition of binding and open diamonds = modest inhibition).

at N-7 of guanine and in the minor groove at N-3 of adenine and may interfere with protein binding through disruption of a specific base contact, by steric inhibition or by a perturbation of DNA structure (19). Fragments containing the γ site were modified with DMS such that a single base per fragment was alkylated, incubated under conditions where only a small fraction of the DNA bound to TnsB, protein-bound and free DNA separated by electrophoresis on a native polyacrylamide gel, extracted and analyzed on a DNA sequencing gel (Figure 1). Nucleotide positions at which alkylation interferes with the formation of a TnsB-DNA complex will be underrepresented in the bound DNA compared to the input control DNA. The results of this analysis are summarized in Figure 2.

Guanine alkylation on the bottom strand at positions 11 and 26 strongly interferes with TnsB binding. In the top strand, guanine alkylation at position 23 noticeably hindered binding and

guanine alkylation at positions 9 and 18 had a weak effect. Three positions (11, 18 and 26) at which alkylation interferes with binding are highly conserved among all the TnsB binding sites and are likely in close proximity to TnsB, as shown by the above protection and interference experiments. Inhibition of binding through guanine alkylation suggests that TnsB makes base-specific contacts in the major groove at these positions. Whether TnsB makes base-specific contacts with guanine at position 9 which is also a highly conserved position is less clear; only modest effects on binding are observed upon modification by alkylation or in the absence of this nucleoside. The interference with TnsB binding observed upon alkylation of guanine at position 23, which lies within the 'spacer' of the inverted repeat, may reflect a secondary effect of a change in DNA structure upon alkylation, as this region is accessible to both hydroxyl radical and DNase I attack.

These protection and binding interference experiments consistently suggest that most of the specific interactions between TnsB and the γ site occur within the highly conserved 'core' region of the binding site.

TnsB Can Bend DNA

Several recombination proteins can bend DNA (20−25). We asked if TnsB can bend DNA using a circular permutation assay (26). We introduced the L119−155 segment, which contains the γ binding site, into a tester plasmid (10) designed such that digestion with different restriction enzymes generates a series of identically sized fragments in which the position of the γ binding site is permuted with respect to the fragment ends. We examined the mobility of a variety of free and TnsB-bound DNA fragments by polyacrylamide gel electrophoresis (Figure 3). The free DNA fragments migrate with generally similar mobilities, indicating there is no considerable intrinsic bend within the γ site DNA. TnsB appears to induce a bend in DNA as evidenced by the different mobilities of complexes that contain the γ binding site in different relative positions. We estimate the TnsB-induced
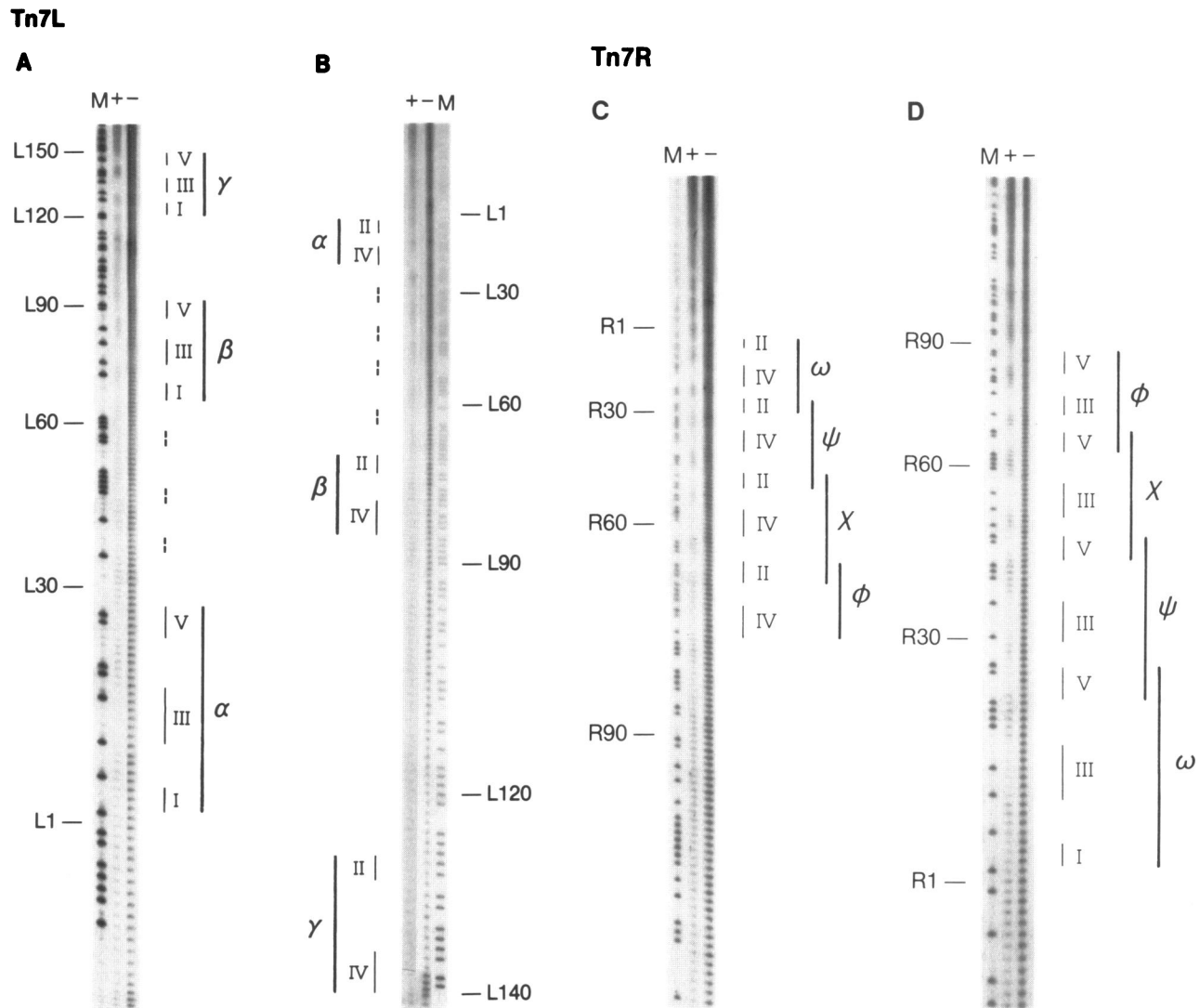
Figure 4. Interaction of TnsB with intact Tn7 ends evaluated by hydroxyl radical footprinting. Hydroxyl radical attack on Tn7 end fragments, L1−166 and R1−199, was performed in the presence (+) or absence (−) of 1.5 µg/ml TnsB. Panel A, bottom strand of Tn7L; panel B, top strand of Tn7L; panel C, bottom strand of Tn7R and panel D, top strand of Tn7R (the top and bottom strands are as shown in Figure 5). A + G chemical sequencing reactions were used as markers (M). The nucleotide positions within the Tn7 ends are indicated, L1 and R1 being the terminal nucleotides of Tn7L and Tn7R, respectively. Protection clusters (I, II,III, IV and V) are indicated with thin lines. The TnsB binding sites (α, β and γ in Tn7L and φ, χ, ψ and ω in Tn7R) are indicated with thick lines. The broken line indicates a region of weak protection between the α and β sites in Tn7L.
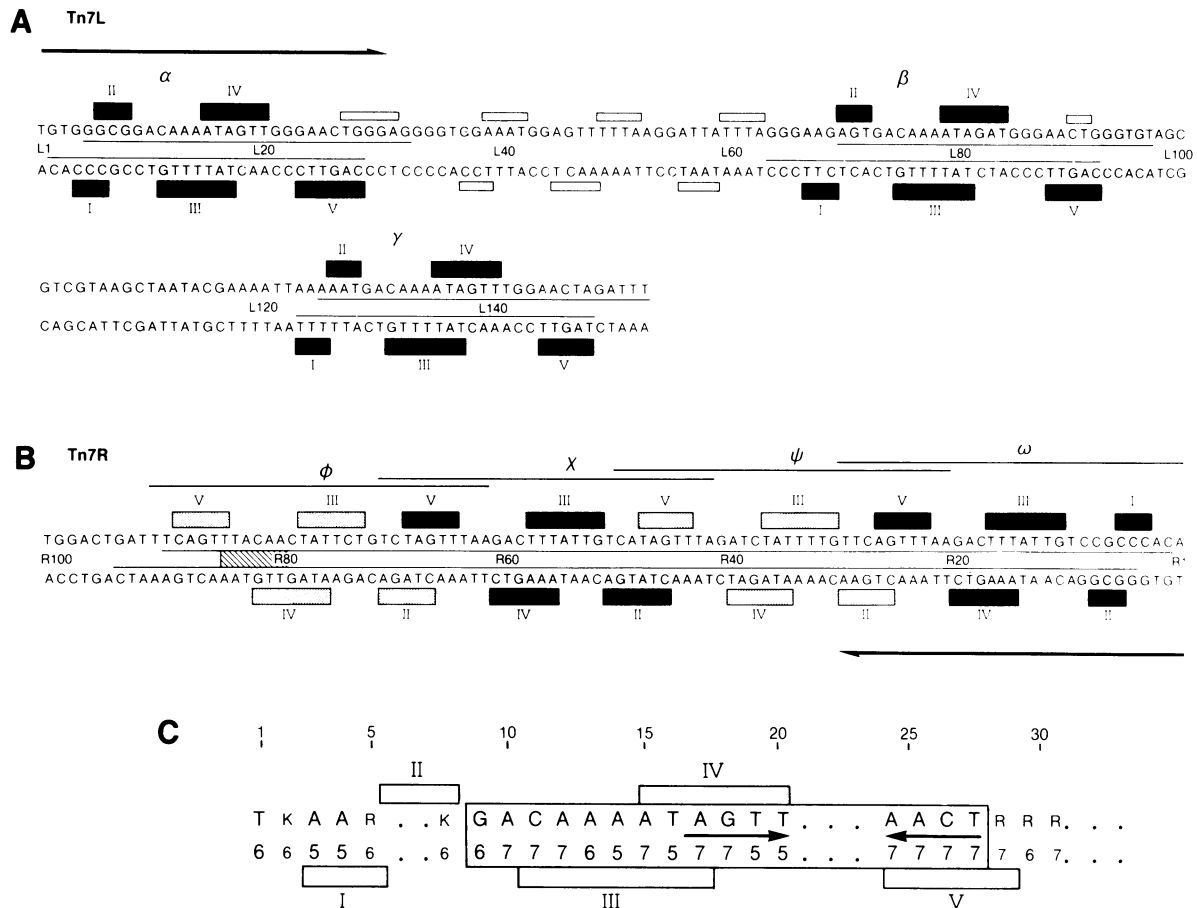
DNA bend angle to be near 50°, based on calculations utilizing Rf values for the lowest and highest mobility complexes (24). Similar analysis of a number of different fragments suggests that the TnsB-induced bend lies between positions 1 and 10 of the γ binding site (data not shown). Thus, the bending center is located asymmetrically within the TnsB binding site, within the region of relatively weak TnsB-DNA interaction.

The interaction of TnsB with intact Tn7 ends

In DNase I footprinting studies (5), three separate 30 bp regions of protection, designated α, β and γ, were evident in Tn7L; in Tn7R, a single 90 bp region of protection was observed which we proposed contains four overlapping TnsB binding sites, designated α, χ, ψ, and ω. Hydroxyl radical footprinting experiments using fragments containing intact Tn7 ends are presented in Figure 4 and summarized in Figure 5.

In Tn7L, three groups of five clusters of protection are observed. These clusters are similar in pattern to those observed at the single γ binding site. The positions of these groups correspond to the α, β and γ binding sites identified by DNase I footprinting. Several clusters of decreased sensitivity to hydroxyl radical attack are also evident between the α and β binding sites; this region is also weakly protected from DNase I attack at high TnsB concentrations (data not shown). We suspect that this weak protection represents non-specific TnsB binding because there is no obvious sequence similarity of the region between α and β to the actual TnsB binding sites. The α and β sites in Tn7L occupy the same relative positions as do φ and ω in Tn7R, so that the non-specific binding positions correspond spatially to the χ and ψ sites.

In Tn7R, four groups of protected clusters are evident, each group corresponding to clusters II, III, IV and V observed at

A Tn7L

TGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTCGAAATGGAGTTTTTAAGGATTATTTAGGGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGC
ACACCCGCCTGTTTTATCAACCCTTGACCCTCCCCACCTTTACCTCAAAAATTCCTAATAAATCCCTTCTCACTGTTTTATCTACCCTTGACCCACATCG

GTCGTAAGCTAATACGAAAATTAAAAATGACAAAATAGTTTGGAACTAGATTT
CAGCATTCGATTATGCTTTTAATTTTTACTGTTTTATCAAACCTTGATCTAAA

B Tn7R

TGGACTGATTTCAGTTTACAACTATTCTGTCTAGTTTAAGACTTTATTGTCATAGTTTAGATCTATTTTGTTCAGTTTAAGACTTTATTGTCCGCCCACA
ACCTGACTAAAGTCAAATGTTGATAAGACAGATCAAATTCTGAAATAACAGTATCAAATCTAGATAAAACAAGTCAAATTCTGAAATAACAGGCGGGTGT

C

|		II				IV		
T	K	A A R . . K	G A C A A A A T A G T T . . . A A C T	R R R . . .
6	6	5 5 6 . . 6	6 7 7 7 6 5 7 5 7 7 5 5 . . . 7 7 7 7	7 6 7 . . .
		I				III		V

Figure 5. Summary of hydroxyl radical protection analysis of intact Tn7 ends. The nucleotide sequences of L1 − 166 (panel A) and R1 − 100 (panel B) are shown. L1 and R1 are the terminal nucleotides of each end; the bold arrows mark the 30 bp terminal inverted repeats. Regions protected against DNase I attack (5) are indicated by thin lines between the strands; the TnsB binding sites (α, β and γ in Tn7L and ϕ, χ, ψ and ω in Tn7R) are indicated. Clusters of nucleotides protected by TnsB against hydroxyl radical attack (I, II, III, IV and V) are marked with boxes (black in Tn7L and black and gray in Tn7R). The protection clusters within the intact ends correspond to those observed at a single TnsB binding site (Figures 1 and 2) except that in Tn7R, a single cluster of protection (V) is observed in sites ϕ, χ and ψ at the positions corresponding to clusters I and V of the isolated γ site because of the overlap between these sites. The open boxes indicate regions of weak protection. The −35 region of the putative *tns*AB promoter in Tn7R is highlighted by a hatched box. Panel C shows the top strand of the consensus sequence for TnsB binding (with T or G indicated by K and A or G indicated by R (5). The numbers below each nucleotide indicate the degree of conservation among the 7 characterized TnsB binding sites. The region of sequence identity that is most highly conserved is boxed. The horizontal arrows indicate a conserved 4 bp inverted repeat. The numbers above the sequence refer to nucleotide positions in a single TnsB binding site. Protection clusters (I, II, III, IV and V) are marked by open boxes.

the isolated γ site. These results support the view that Tn7R contains four TnsB binding sites. There is considerable overlap among the sequences of the Tn7R TnsB binding sites. Consistent with such overlap, protected cluster I is not observed except at the terminal ω site; at the other three sites, cluster V from the adjoining site overlaps these positions.

Sequential occupancy of the TnsB binding sites

What roles do the multiple TnsB binding sites in the ends of Tn7 have in recombination? Does TnsB have a similar affinity for all sites or do these sites differ in their apparent affinities so that the sites are filled in a particular order? TnsB forms several distinct protein-DNA complexes with fragments containing either Tn7L or Tn7R in band shift assays (4,5). We show here (Figure 6) that three TnsB-DNA complexes, designated LI, LII and LIII, are observed with Tn7L and four complexes, designated RI, RII, RIII and RIV, are observed with Tn7R. The relative ratios of the complexes is dependent on TnsB concentration. At low concentrations, faster migrating species (such as LI and RI) predominate, whereas at higher TnsB concentrations, slower migrating complexes (such as LIII and RIV) predominate. Because there is a similar difference between the apparent molecular weights of successive complexes, as judged by comparison to DNA markers (data not shown), we suspect that these different complexes are DNA molecules bound by different numbers of TnsB molecules distributed among its binding sites. The distinctive pattern of complexes suggests that TnsB may bind to each end in an ordered fashion determined by differing apparent binding affinities of TnsB for the various binding sites. Differing apparent affinities for each site would be determnined by the intrinsic binding affinity of TnsB for each site and potentially by structural deformations in DNA induced by TnsB binding to adjacent sites and by protein-protein interactions between TnsB molecules bound at different sites. The gradual transition between the various TnsB-Tn7 end complexes (Figure 6) is notable and suggests there is little cooperativity in TnsB binding under these conditions.
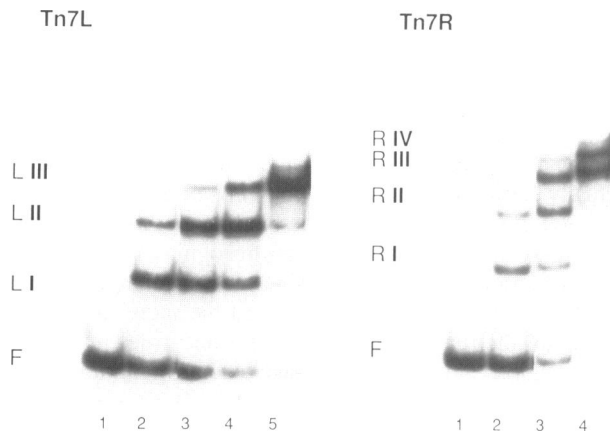
Figure 6. Protein-DNA complexes formed between TnsB and DNA fragments containing Tn7L and Tn7R. Protein-DNA complexes were formed in reactions containing the indicated amounts of TnsB and a Tn7L fragment (left panel) and a Tn7R fragment (right panel) and analyzed by native polyacrylamide gel electrophoresis. Three protein-DNA complexes (LI, LII and LIII) were observed with the Tn7L fragment and four complexes (RI, RII, RIII and RIV) were observed with the Tn7R fragment. Lane 1, no addition; lane 2, 0.31 μg/ml TnsB; lane 3, 0.62 μg/ml; lane 4, 1.25 μg/ml and lane 5, 2.5 μg/ml.

To evaluate the relative apparent affinity of TnsB for its binding sites, we directly characterized their occupancy in the various TnsB-DNA complexes by *in situ* footprinting. We used copper-phenanthroline as an attack reagent on gels containing the complexes (15), followed by extraction and analysis of the DNA (Figure 7). In the fastest migrating Tn7L complex LI, a single protected region corresponding to the γ site is observed. In complex LII, two regions of protection corresponding to the γ and β sites are observed. The slowest migrating species, complex LIII, displayed three regions of protection corresponding to the γ, β and α sites. The same ordered occupancy is observed in solution DNase I footprinting experiments (data not shown). These results show that TnsB occupies its Tn7L binding sites in sequential fashion, binding first to the innermost γ site, then to β and, finally, to the terminal α site. These results also indicate that various Tn7L sites each have a distinct apparent affinity for TnsB. We have no information about the form of TnsB that occupies each binding site.

Similar analysis of the Tn7 complexes reveal that the fastest migrating complex RI, represents TnsB binding to the χ binding site. In both complex RII and complex RIII, in addition to the χ site, considerable protection of the ψ and terminal ω sites is observed. In complex RII, the ω site appears to be slightly more protected than does the ψ site whereas in complex RIII, comparable protection of both the ψ and ω sites is evident. We suggest that 'complex RII' is heterogenous and is actually a mixture of two distinct but similar migrating complexes, each containing two TnsB molecules. We imagine that one complex contains TnsB at the χ and ψ sites and the other contains TnsB at the χ and ω sites. In complex RIII, three molecules of TnsB appear to be present, occupying the χ, ψ and ω sites. In the slowest migrating complex RIV, all four TnsB binding sites are protected, indicating that TnsB occupies the ϕ, χ, ψ and ω sites. Our studies suggest that the ϕ site, which overlaps a likely *tns* promoter (27), fills only after the other Tn7R sites are occupied and hence is the Tn7R site with the lowest apparent affinity for TnsB.
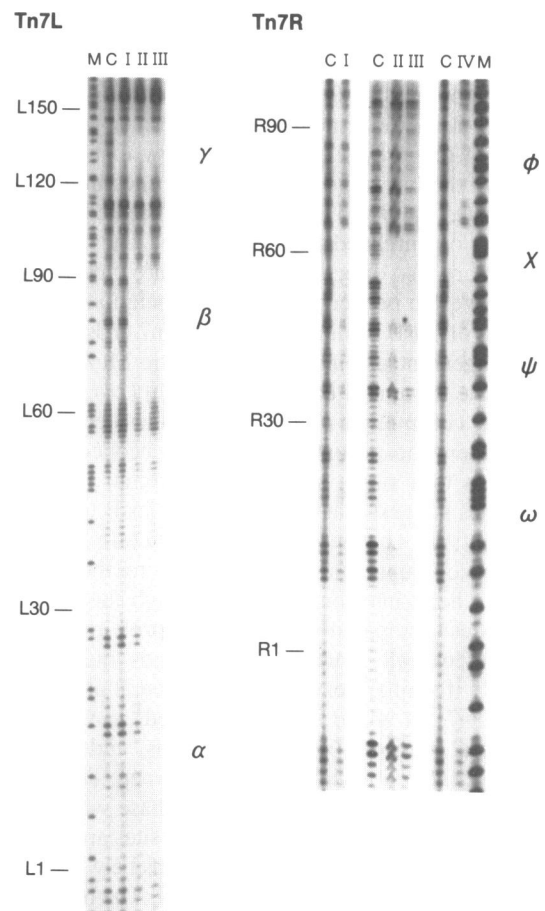
Figure 7. *in situ* phenanthroline-copper footprinting of TnsB complexes. Gels containing the TnsB-DNA complexes decribed in Figure 7 were subjected to phenanthroline-copper attack and extracted DNA analyzed on DNA sequencing gels. The left panel (marked Tn7L) displays DNAs from Tn7L complexes. Lane M, A + G sequencing reaction as markers; lane I, complex LI; lane II, complex LII, lane III, complex LIII and lane C, no TnsB. The right panel (marked Tn7R) displays DNAs from Tn7R complexes: lane C, no TnsB; lane I, complex RI; lane II, complex RII; lane III, complex RIII, lane IV, complex RIV and lane M, A + G sequencing reaction as markers. In both panels, the nucleotide positions in Tn7L and Tn7R are indicated with L1 and R1 being the terminal nucleotides of each end. The positions of the TnsB binding sites (α, β and γ in Tn7L and ϕ, χ, ψ and ω in Tn7R) are shown.

Thus the multiple species observed in band shift assays are protein-DNA complexes that contain TnsB molecules distributed among its various binding sites. TnsB interaction with the ends is thus characterized by the progressive and sequential occupancy of its binding sites. Our observations indicate that the binding sites in each end differ in their apparent affinity for TnsB.

We have previously shown that the ends of Tn7 are functionally distinct (9). For example, miniTn7 elements containing two Tn7R segments transpose whereas elements containing two Tn7L segments do not. One explanation for this difference could be that the ends differ in their affinity for TnsB. However, we have observed no substantial difference in the apparent affinity of TnsB for Tn7L and Tn7R (data not shown). In experiments performed with crude extracts, the apparent affinity of TnsB for Tn7R appeared to be modestly higher than for Tn7L (4). The basis of this apparent difference with purified TnsB remains to be determined.

DISCUSSION

We have examined the interaction of the Tn7-encoded transposition protein TnsB with the ends of the transposon using several high-resolution methods. The results presented here, together with our previous work (5), show that TnsB binds to three sites in the left end of Tn7 and four sites in the right end. The Tn7L sites, α, β and γ, are directly oriented and separated by segments of unrelated DNA sequence. The Tn7R sites, ϕ, χ, ψ and ω, are also directly oriented but are closely juxtaposed with some overlap between the edges of neighboring sites. The arrangement of TnsB binding sites on the Tn7 ends is reminiscient of the arrangement of binding sites for the MuA transposition protein on the termini of bacteriophage Mu (28). In both cases, distinct arrays of multiple protein binding sites occur at each end. In many other transposons, an element-encoded transposition protein binds to a single site at each terminus (29−34).

Our studies suggest that the binding of TnsB to the Tn7 ends may substantially change the overall structure of these DNA segments. We found that TnsB binding to the Tn7L γ site introduces a bend into this DNA. Although the degree of this bend is modest compared to some other protein-induced bends (10,26), the overall structure of the ends may be changed considerably if similar bends are introduced at all TnsB binding sites. The Tn7L α and β sites are 'phased', i.e. are separated by an integral number (6) of helical turns, and thus are well-positioned for possible interactions between TnsB molecules bound at these sites; the γ site is not in phase with the adjacent β site (5.3 turns). In Tn7R, the four TnsB sites are phased, lying two helical turns apart. Possible interactions between the various TnsB binding sites have yet to be evaluated. The binding of TnsB to multiple sites at both ends of Tn7 and the ability of TnsB to bend DNA suggest that TnsB may promote the formation of a highly organized protein-DNA complex involving Tn7L and Tn7R. Such higher order complexes are known to mediate other recombination reactions (25−37).

Sequential Occupancy of the TnsB Binding Sites

We have found that the TnsB binding sites within each end of Tn7 are sequentially occupied by TnsB in an ordered fashion as the concentration of TnsB is increased. A notable feature of this ordered binding is that sites inside the ends of the element are occupied first by TnsB and the sites that are adjacent to the transposon termini are filled only after TnsB occupies the other inner sites. We interpret this ordered occupancy to reflect different apparent affinities of TnsB for its various sites. The relative contributions of intrinsic binding affinity and potential interactions between the TnsB binding sites to the apparent affinity of TnsB for each site remain to be established.

In Tn7L, the innermost γ site has the highest apparent affinity for TnsB and thus is occupied first. With increasing TnsB, the β site and then the terminal α site are also occupied. Occupancy of the γ site is likely critical to recombination because the transposition of miniTn7 elements lacking the γ region is severely impaired (9). The TnsB sites in Tn7R are also occupied in sequential fashion although the order of occupancy is not apparently as fixed as in Tn7L. The inner χ site has the highest apparent affinity for TnsB and thus is occupied first. Our studies suggest that TnsB subsequently occupies either the terminal ω site or the ψ site which appear to have similar apparent affinites for TnsB. After the χ, ψ and ω sites are filled, the innermost

ϕ site is filled. Occupancy of the ϕ site is not essential to transposition (9).

We imagine that the nucleoprotein structure active in recombination contains TnsB bound at multiple sites in each end including the terminal binding sites where strand breakage and joining must occur. Although it is not known whether TnsB actually executes breakage and rejoining or assists protein in this function, it seems likely that the binding of TnsB to the termini must play a critical role in end recognition and utilization, recombination occuring only when the terminal TnsB sites are filled. An attractive hypothesis is that the sequential binding of TnsB to the ends plays an important role in controlling transposition, no strand breakage occuring until a complete 'transposome' is assembled at the ends of the transposon (38−40).

TnsB binding to Tn7R also likely has a role in regulating the expression of the Tn7-encoded *tns* genes. The promoter of an operon that includes *tnsA* and *tnsB* and also perhaps the other *tns* genes (27,41,42) overlaps the ϕ TnsB binding site. Other studies support for the hypothesis that the interaction of TnsB with Tn7R does indeed repress *tns* gene transcription (6,7,41). The observation that the ϕ site has the lowest relative affinity for TnsB suggests that this protein can act as negative regulator of gene expression only when its concentration actually exceeds that required for recombination.

Recognition of DNA by TnsB

We have used several high-resolution techniques to examine the interaction of TnsB with a single binding site to begin to dissect how this protein may contact DNA. Using protection from
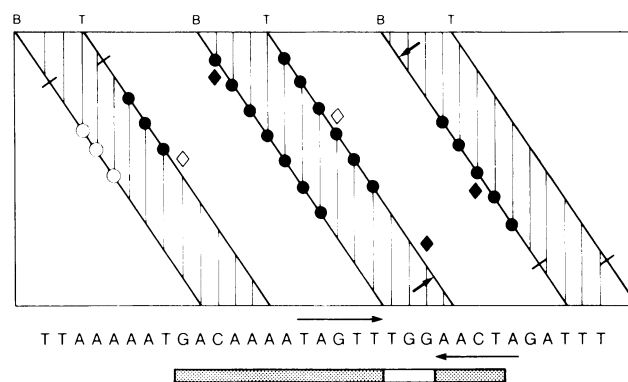
Figure 8. The interaction of TnsB with a single binding site. A DNA helix is diagrammed in planar representation (19). The position of each base pair is indicated by a vertical line across the minor groove. The bottom (B) and top (T) strands are indicated. The nucleotide sequence of the top strand of the γ TnsB binding site is shown, horizontal arrows marking its 5 bp inverted repeats. The gray boxes below the sequence indicate the positions highly conserved in sequence identity among the TnsB binding sites; the open box indicates the 3 bp spacer between the conserved inverted repeats. Marked on the DNA helix are positions where alterations are observed in the presence of TnsB using a variety of probes. Filled circles indicate likely positions of close contact between TnsB and DNA as revealed by hydroxyl radical footprinting and by binding interference by DNA modified by hydroxyl radical attack, i.e. missing nucleosides. Open circles indicate positions protected from hydroxyl radical attack but whose absence does not obviously effect TnsB binding. Bases at which alkylation interferes with TnsB binding are marked with diamonds; filled diamonds indicating a strong block to TnsB binding and open diamonds a modest block to binding. The extent of the γ binding site as defined by protection against DNase I attack is indicated by lines across the phosphodiester backbone; positions of hypersensitivity to DNase I attack are indicated by arrows (5).

hydroxyl radical attack and several binding interference methods, we have detected contacts between TnsB and the Tn7L γ binding site that encompass 26 bp. These contacts extend widely from one side of the helix to another over two major and two minor grooves (Figure 8). The strongest contacts are observed within three clusters of nucleotides located within the portion of the binding site that is most highly conserved in sequence identity among all TnsB binding sites (5). A prominent feature of the TnsB consensus sequence is a 4 bp inverted repeat located at one side of the binding site. Our studies here indicate that symmetrical nucleotides within the inverted repeat are in close contact with TnsB. The inverted repeats flank a spacer segment that remains exposed to solvent in the presence of TnsB. The inverted repeat region is asymmetrically flanked by a highly conserved 7 bp sequence that also appears to make specific contacts with TnsB.

In addition to the strong contacts that TnsB makes in the highly conserved region, markedly weaker interactions are observed outside this segment. These weaker contacts lie at one edge of the binding site in a region of little homology among the TnsB binding sites and may represent non-specific interactions between TnsB and DNA. Alternatively, they may provide the critical features that differentiate the TnsB binding sites. We are intrigued by the possibility that this apparently non-conserved region may provide a special determinant for recognition of the extreme termini of the transposon (5).

What type of protein structural motif(s) in TnsB is used for DNA recognition? Our analysis of TnsB binding to DNA modified by alkylation suggests that specific recognition of DNA by TnsB involves contacts within the major groove. We have observed that a TnsB derivative containing only the N-terminal 200 amino acids can bind specifically to the ends of Tn7 (data not shown). Flores et al. (42) have noted that this region of TnsB contains a segment of weak similarity to the helix-turn-helix DNA binding motif. Understanding how TnsB does recognize DNA and what changes in DNA structure it may promote will help reveal the role(s) of this protein in recombination.

ACKNOWLEDGEMENTS

REFERENCES

1. Berg,D.E. and Howe,M.M. (1989) (eds.) Mobile DNA, American Society for Microbiology, Washington, D.C..
2. Barth,P.T., Datta,N., Hedges,R. and Grinter,N.J. (1976) J. Bacteriol, **125**, 800–810.
3. Craig,N.L. (1989) In Berg,D.E. and Howe, M.M. (eds.), Mobile DNA. American Society for Microbiology, Washington, D.C., pp.211–226.
4. McKown,R.L., Waddell,C.S., Arciszewska,L.K., and Craig,N.L. (1987) *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 7807–7811.
5. Arciszewska,L., McKown,R.L. and Craig,N.L. *J. Biol. Chem.* (in press).
6. Rogers,M., Ekaterinaki,N., Nimmo,E. and Sherratt,D.J. (1986) Mol. Gen. Genet., **205**, 550–556.
7. Waddell,C.S. and Craig,N.L. (1988) *Genes & Develop.*, **2**,137–149.
8. Kubo,K. and Craig,N.L. (1990) *J. Bacteriol.* **5**, 2774–2778.
9. Arciszewska,L., Drake,D, and Craig,N.L. (1989) *J.Mol.Biol* **207**, 35–52.
10. Kim,J., Zwieb,C., Wu,C. and Adhya,S. (1989) *Gene*, **85**, 15–23.
11. Tabor,S. and Struhl,K. (1987) In Ausubel, F.M. et al. (eds.), Current Protocols in Molecular Biology., Greene Publishing Associates and Wiley-Interscience., New York, New York, pp.3.5.7–3.5.10.
12. Tullius,T.D. and Dombroski,B.A. (1986) *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 5469–5473.
13. Hayes,J.J. and Tullius,T.D. (1989) *Biochemistry*, **28**, 9521–9527.
14. Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–559.
15. Kuwabara,M.D. and Sigman,D. (1987) *Biochemsitry* **26**, 7234–7238.
16. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
17. Chalepakis,G. and Beato,M. (1989) Nucleic Acids Res., **17**, 1783.
18. Koudelka,G.B., Harrison,S.C. and Ptashne,M. (1987) *Nature (London)* **326**, 886–888.
19. Siebenlist,U. and Gilbert,W. (1980) *Proc. Natl. Acad. Sci. USA.*, **77**, 122–126.
20. Prentki,P, Chandler,M. and Galas,D. (1987) *EMBO J.* **6**, 2479–2487.
21. Hatfull,G., Noble,S. and Grindley,N.D. (1987) *Cell*, **49**, 103–110.
22. Salvo, J.J. and Grindley, N.D.F. (1988) *EMBO J.* **7**, 3609–3616.
23. Robertson,C.A. and Nash,H.A. (1988) *J.Biol.Chem.* **263**, 3554–3557.
24. Thompson,J.F. and Landy,A. (1988) *Nucleic Acids Res.* **16**, 9687–9705.
25. Mertens,G., Klippel,A., Fuss,H., Blocker,H., Frank,R. and Kahmann, R. (1988) *EMBO J.*, **7**, 1219–1227.
26. Wu,H-M. and Crothers,D.M (1984) *Nature (London)* **308**, 509–513.
27. Gay,N.J., Tybulewicz,V. and Walker,J.E. (1986) *Biochem J.*, **234**, 111–117.
28. Craigie,R., Mizuuchi,M. and Mizuuchi,K. (1984) *Cell*, **39**, 387–394.
29. Ichikawa,H., Ikeda,K., Wishart,W. and Ohtsubo,E. (1987) *Proc. Natl. Acad. Sci. USA.*, **84**, 8220–8224.
30. New,J., Eggleston,A. and Fennewald,M. (1988) *J. Mol. Biol.* **201**, 589–599.
31. Zerbib,D., Jakowec, M., Prentki,P., Galas,D. and Chandler,M. (1981) *EMBO J.* **6**, 3163–3169.
32. Derbyshire,K., Hwang,L. and Grindley,N. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8049–8053.
33. Makris,J., Normann,P. and Reznikoff,W. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2224–2228.
34. Huisman, O., Errada,P., Signon,L. and Kleckner,N. (1989) *EMBO J.* **8**, 2101–2109.
35. Thompson, J.F., and Landy, A. (1989) In Berg,D.E. and Howe,M.M. (eds.), Mobile DNA, American Society for Microbiology, Washington, D.C., pp.1–22.
36. Pato,M. (1989) In Berg,D.E. and Howe,M.M. (eds.), Mobile DNA, American Society for Microbiology, Washington, D.C., pp. 23–52.
37. Stark,W.M., Boocock,M.R. and Sherratt,D.J. (1989) *Trends Genet.*, **5**, 304–309.
38. Craigie,R. and Mizuuchi,K. (1987) *Cell* **45**, 793–800.
39. Surrette,M., Buch,S. and Chaconas,G. (1987) *Cell* **49**, 253–262.
40. Kuo,C-F., Zou,A., Jayaram,M., Getzoff,E. and Harshey,R. (1991) *EMBO J.* **10**, 1585–1591.
41. Orle,K.A. and Craig,N.L. (1990) *Gene*, **96**, 1–7.
42. Flores,C., Qadri,M.I. and Lichtenstein,C. (1990) *Nucleic Acids Res.*, **18**, 901–911.