Purification and characterisation of the TnsB protein of Tn7: a transposition protein that binds to the ends of Tn7

Yue Tang, Conrad Lichtenstein[1] and Sue Cotterill*
Department of Biochemistry and [1]Centre for Biotechnology, Imperial College of Science, Technology & Medicine, Exhibition Road, London SW7 2AZ, UK

ABSTRACT

Tn7, a large bacterial transposon encodes 5 proteins required for its transposition. We report a rapid and easy purification of one of these proteins, TnsB, from an overexpression strain. This protein was shown to bind to the ends of Tn7, in a bandshift assay, in two distinct stages as a function of protein concentration. DNaseI footprinting at each end of Tn7 showed that the TnsB recognition sequence, a set of 22 bp repeats, plus Tn7 termini are protected. Binding of TnsB appeared cooperative but was only observed above a threshold concentration of protein. ATP and Mg^{2+} had no effect on the pattern of protection, nor did addition of other Tn7-encoded proteins. Hydroxyl radical footprinting, performed at the right end, showed that TnsB binds preferentially to one side of the DNA helix.

INTRODUCTION

Tn7 is a large transposable element (14 kb) found in bacteria which codes for resistance to three antibiotics, trimethoprim, streptomycin and spectinomycin (for a review of structure and properties see 1).

Although the transposon has a number of similarities with other bacterial transposons (e.g. inverted repeats at each end of the transposon, duplication of target sequences accompanying transposition, possession of target immunity), Tn7 possesses two features that set it apart from all other bacterial transposons: the first relates to the frequency and specificity of transposition. Most bacterial transposons jump with a rather low frequency to random sites in the genome. Tn7 is capable of transposing in this mode ('cold site' transposition), however it is also able to transpose with very high efficiency (at least 100 times greater) to a specific site in the genome of many bacteria. In *E. coli* this site is in the transcriptional terminator of the glucosamine synthase (*glmS*) gene and is called the Tn7 attachment site, *attTn7*. Transposition of Tn7 usually occurs in a single orientation relative to *attTn7* irrespective of whether the site is contained in the bacterial genome or carried on a plasmid. This latter property gives Tn7 transposition some similarities to bacteriophage lambda integration.

The second difference concerns the transposition proteins themselves. Although almost all bacterial transposons code for their own transposition proteins, in most cases this is one or at most two proteins. Tn7 however codes for 5 different transposition proteins referred to as TnsA, B, C, D, E. TnsA, B, C and D are required for the *attTn7* transposition, and TnsA, B, C and E are required for the transposon to hop into cold sites. Why Tn7 should require over twice as many transposon-encoded proteins as other systems is not known. However there does not seem to be any redundancy of function, since deletion analyses have shown that TnsA, B and C are absolutely essential, and TnsD and E must be present for hot and cold site transposition respectively (1–4).

The roles of each of these proteins in both hot and cold site transposition is not yet well defined. TnsD and TnsB are known to bind to the *attTn7* site and the ends of Tn7 respectively (5, 6). A *tnsD*-dependent factor in crude extracts appears to bind to a single sequence in *E. coli attTn7* 20 bp upstream of the point of insertion. A *tnsB*-dependent factor in crude extracts appears however, to bind to multiple sites at both ends of the molecule. As with most transposons the very ends of Tn7 are perfect inverted repeats (in this case eight base pairs long). At the right end of Tn7, this sequence is followed by four overlapping direct repeats of a 22 bp sequence motif. This motif is also present (in the opposite orientation) at the left end as three direct repeats. Here the repeats are not overlapping but are separated by 42 and 31 bp (7). It is the 22 base pair direct repeats at both ends of the molecule which seem to constitute the binding sites for the TnsB protein. Complete sequence analysis of the region coding for the transposition proteins (4) shows that the sequences of TnsD and TnsB each bear a putative helix turn helix domain which is consistent with their DNA binding functions.

Activities associated with TnsA, C and E have not yet been identified. The sequence analysis shows that TnsA and TnsE also seem to contain DNA binding domains, however DNA binding by these proteins has yet to be demonstrated. TnsC seems to contain an ATP binding site, and our recent studies with purified TnsC protein suggest that it possesses an ATPase activity (Ganju *et al.* man. in prep.).

One way to define the role of each of the Tns proteins in Tn7 transposition is to study the activities of the purified proteins

individually and identify effects of their interactions with each other *in vitro*. To this end we have cloned, overexpressed (Flores *et al.* man. in prep.) and purified TnsA, B, C, and D (Ganju *et al.* man. in prep.). In this paper we report the purification of the overexpressed TnsB protein. We also report a characterisation of the binding of TnsB to DNA, and some data relating to the interaction of the pure TnsB protein to other purified Tn7 proteins.

MATERIALS AND METHODS

Bacterial strains and clones

For protein production the complete TnsB gene was cloned downstream of the T7 promoter in the pET vector system, and the cloned protein was expressed in strain BL21 (8). The full details of the constructs and the cloning procedures used are described elsewhere (Flores *et al.* manuscript in prep.) Protein expression was induced by the addition of isopropyl thio-β-D galactoside (IPTG) to cells grown at 37°C to an OD_{590} of 0.4, and the cells were then allowed to grow for a further 4 hours before harvesting.

DNA manipulations

General methods for DNA endlabelling, restriction endonuclease cleavage etc. were performed as described by Sambrook *et al.* (9).

Preparation of affinity column

Two oligonucleotides of sequences 5' amino alkyl ATATGTGGGCGGCGGACAATAAAGTCTTAAACTGAA-CAAAATAGATCTAAACTATG and 5'CATAGTTTAGAT-CTATTTTGTTCAGTTTAAGACTTTATTGTCCGCCCACA were synthesised using a Pharmacia Gene Assembler. They were mixed together in water and heated to 95°C and then allowed to cool to room temperature slowly. The annealed substrates were attached to cyanogen bromide activated sepharose as described in the Pharmacia booklet 'Affinity chromatography principles and methods'. Analysis of the columns revealed that in most cases 85% of the input oligonucleotides bound to the column. The columns were washed well, equilibrated in the relevant buffers, and used for the purification.

DNA substrates

The right end fragments of Tn7 used for the band shifting and the DNA footprinting were prepared from plasmid pMR11 (2). This contains 'miniTn7' i.e. replacement of the central 11.4 kb of Tn7, with the chloramphenicol resistance gene. Right end fragments were released by cutting with HindIII and BalI to give a 230 bp fragment. Left end fragments were prepared from pYT2 (pMR 11 from which the central region of mini Tn7 has been removed between the HincII and PstI sites). In this case a 456 bp fragment was released using HindIII and EcoRI. In all cases the fragments were labelled by end filling with ^{32}P deoxynucleotides and the Klenow fragment of *E. coli* polymerase I.

Band shifting studies

Basic band shifting studies were performed in buffer A (50 mM Tris HCl pH 8.0, 100 mM KCl, 1 mM DTT, 5% glycerol) plus 1 mM EDTA, 0.5 µg polydIdC and 10 µg BSA. The protein and DNA samples were mixed in the buffer and incubated at 4°C for 30 min. They were then loaded onto either a 5% polyacrylamide gel and run at 4°C in TBE for 3 hrs at 200 v,

or on a 1.5% agarose gel, and run at 4°C in the same buffer at 120 v for 4 hrs. The gels were dried down and the position of the bands visualised by autoradiography. Amounts of protein and DNA used, and additions to the basic mix are described in individual figure legends.

DNA footprinting

Footprinting was carried either by the DNase I method or using hydroxyl radicals. DNase I footprinting studies were carried out in the same buffer as that used for the bandshifting except that it contained 1 mM $MgCl_2$, but no EDTA. The protein and DNA samples were mixed in the buffer (20 µl) and incubated at 4°C for 30 min. The position of the protein was localised by adding 3 units of DNase I to the reaction. It was incubated at 25°C for 7 min, and stopped by the addition of 80 µl of a solution containing 12.5 mM EDTA, 12.5 µg/ml proteinase K. For hydroxyl radical footprinting the DNA and protein were mixed as for the DNase I footprinting, except that glycerol was omitted from the buffer. The footprinting was achieved by the addition of 5 µl 0.03% hydrogen peroxide, 5 µl 0.4 mM iron II sulphate, 0.8 mM EDTA, and 5 µl 10 mM sodium ascorbate and incubating at 25°C for 3 min. In this case reactions were quenched by the addition of 5 µl 50% glycerol. For both DNase and hydroxyl radical footprinting, samples were phenol extracted, then ethanol precipitated twice. Finally they were redissolved in formamide sequencing dye (100% formamide, 10 µg/ml bromophenol blue, 10 µg/ml xylene cyanol) , boiled for 2−3 min and then loaded onto a 8% sequencing gel, and run at 47 W for 2.5 hours. The gels were dried down, and visualised using autoradiography. Again amounts of protein and DNA used, and any variation of conditions are described in the figure legends.

Protein crosslinking studies

These were carried out using DTBPA (N-(4-azidophenyl-thio)pthalimide purchased from Pierce) as the crosslinking reagent. Reactions were carried out in a final volume of 600 µl in 20 mM sodium phosphate pH 7.0, 100 mM KCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$. The crosslinker was added to a final concentration of 6 µg/ml. Samples were exposed to 340 nM UV light at 4°C for 15 min in the dark to allow the crosslinking to occur. Then, methanol precipitated, and analysed using poly-acrylamide gel electrophoresis, and Western blotting with anti-TnsB antibodies (Ganju *et al.* unpublished data).

RESULTS

Purification of TnsB

Since the protein was expressed under a T7 promoter we obtained a very high level of expression (about 50% of total cellular protein). This resulted in much of the protein being insoluble. It was found possible to alter the amount of soluble protein by variation of the time and the temperature of incubation, and the length of induction of the protein in the presence of IPTG. Under optimal conditions approximately 5% of the TnsB was soluble. We carried out the purification only on the soluble fraction of the protein.

After the cells were broken by sonication in buffer A, the protein was purified to homogeneity using only two steps (Figure 1). Firstly an ammonium sulphate precipitation, where the protein precipitated at 40%. The precipitate was collected by centrifugation at 15 K for 20 min, and the protein dialysed against buffer A. The protein was then loaded onto an affinity column,
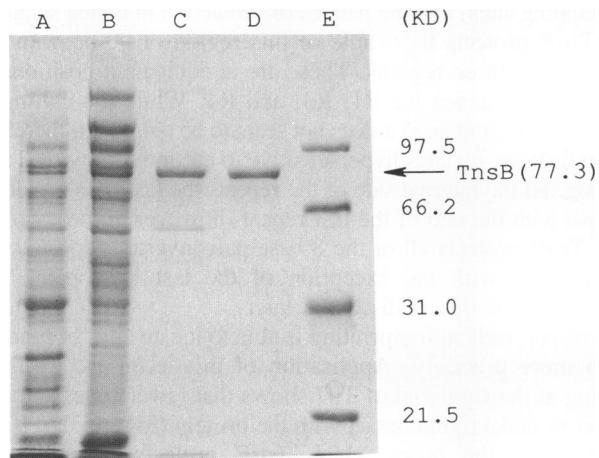
Figure 1. Purification of TnsB protein. Polyacrylamide gel of fractions from the purification of TnsB. The position of TnsB is marked with an arrow. Lane A, soluble fraction of crude extract after sonication; B, precipitate after 40% ammonium sulphate cut; C, 1 M salt elution from affinity column; D, 3 M salt elution from the affinity column; E, markers of molecular weights as indicated.

Figure 2. Bandshifting experiments with the left (LE) and the right (RE) ends of Tn7. LE is 468 bp, RE is 270 bp. In both cases the hatched box represents the 8 bp terminal inverted repeat (TIR) and the solid boxes the 22 bp TnsB binding sites. Polyacrylamide gels showing the bandshifts were carried out as described using left end (A) and right end (B) fragments. All reactions contained 0.2 ng DNA, and protein equivalent to A: 0 ng; B: 30 ng; C: 150 ng; D: 300 ng; E: 900 ng; and F: 1500 ng. Li and Ri denote the unbound DNA. Lii−vi and Rii−vii denote bandshifts.

in the same buffer. The column was washed in buffer A containing 400 mM NaCl. The TnsB eluted in two different batches at 1 M NaCl and 3 M NaCl. The difference between the protein obtained under these two different elution conditions is not clear. However the one that eluted at 3 M was essentially homogeneous (only one band at 77.3 kD is visible on the gel), whereas the 1 M elution contained several impurities. All subsequent experiments described were therefore performed with the protein eluting at 3 M. This procedure was very reproducible and yielded 1.3−1.5 mg of TnsB (3 M fraction) from 1.6 g cells.

The protein was dialysed into buffer A and stored at −20°C. In these conditions it was stable for at least 6 months.

Analysis of binding of TnsB to the ends of Tn7

Bandshifting experiments. Figure 2 shows the patterns obtained from a titration of TnsB with fragments from the left and right ends of Tn7.

Several different bandshifts can be seen for each fragment, and the particular bandshifts observed vary with the amount of TnsB protein added. In addition both fragments appear to show two distinct phases to the bandshifting. The first stage requires relatively low concentrations of protein, and appears to correlate with the number of binding sites in the fragment. The second stage is only observed at very much higher (6−10 fold higher) protein concentrations.

For the left end, at lower protein concentrations, it is possible to see three strong distinct bandshifts. The lower one of these (Lii) is chased sequentially into the two upper bands (Liii then Liv), as the protein concentration is raised from 30 to 150 ng. Addition of a further 150 ng of protein has little effect on the bands, and only band Liv is visible even at long exposures of the gel. At very high concentrations of protein (900 ng to 1500 ng) two additional bands appear (Lv and Lvi).

A similar titration with the right end produces very similar results except that the phase of binding observed at lower protein concentrations seems to contain four rather than three distinct complexes (Rii, Riii, Riv, and Rv).

The protein:DNA ratio required to see band shifts at either end is high. The earliest band shifts are seen at ratios of
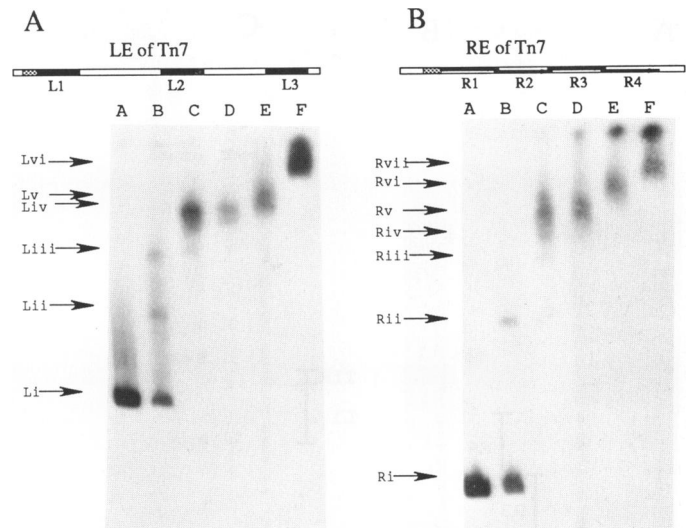
protein:DNA (mol/mol) of about 500:1. All the DNA has been shifted to Rv and Liv by a ratio of 3000:1. To produce the higher bandshifts, protein:DNA ratios in the region of 90,000:1 to 150,000:1 are required.

In the presence of magnesium at 1 mM, the affinity of the protein for the DNA at the lower end of the titration appears to be marginally higher, however the pattern of bands obtained does not alter (data not shown).

If crude TnsB containing extracts, or partially purified TnsB protein (after ammonium sulphate precipitation), is used in the bandshifting experiment, the bandshift obtained is very different. A very high molecular weight band is obtained, even at low concentrations of TnsB protein, and no intermediate bands are seen. Whether this is the equivalent of one of the high bands seen in the presence of TnsB alone is not clear, since the resolution in this part of the gel is not high. This high molecular weight band is most likely to be caused by interaction with a host protein. However it is not observed if the equivalent experiment is carried out with extracts from the host bacterial strain, or a strain which overproduces a different Tn7 protein. It therefore seems that the host factor requires the presence of TnsB to interact with the DNA. The significance of this observation, and whether the host protein has any relationship to the transposition reaction *in vivo* is unknown.

Area of footprinting by TnsB on the right and left ends. Figure 3 shows the results of footprinting experiments carried out on fragments from the two ends of Tn7. For both ends footprinting was carried out on each strand. However the results from only one strand are shown in each case.

For the right end, it was found possible to use both DNase I and also hydroxyl radical footprinting to get a well defined image of the sites protected by the TnsB protein. The data obtained from DNase footprinting is shown in Figure 3A, and diagrammatically in Figure 4A. The right end of Tn7 has four
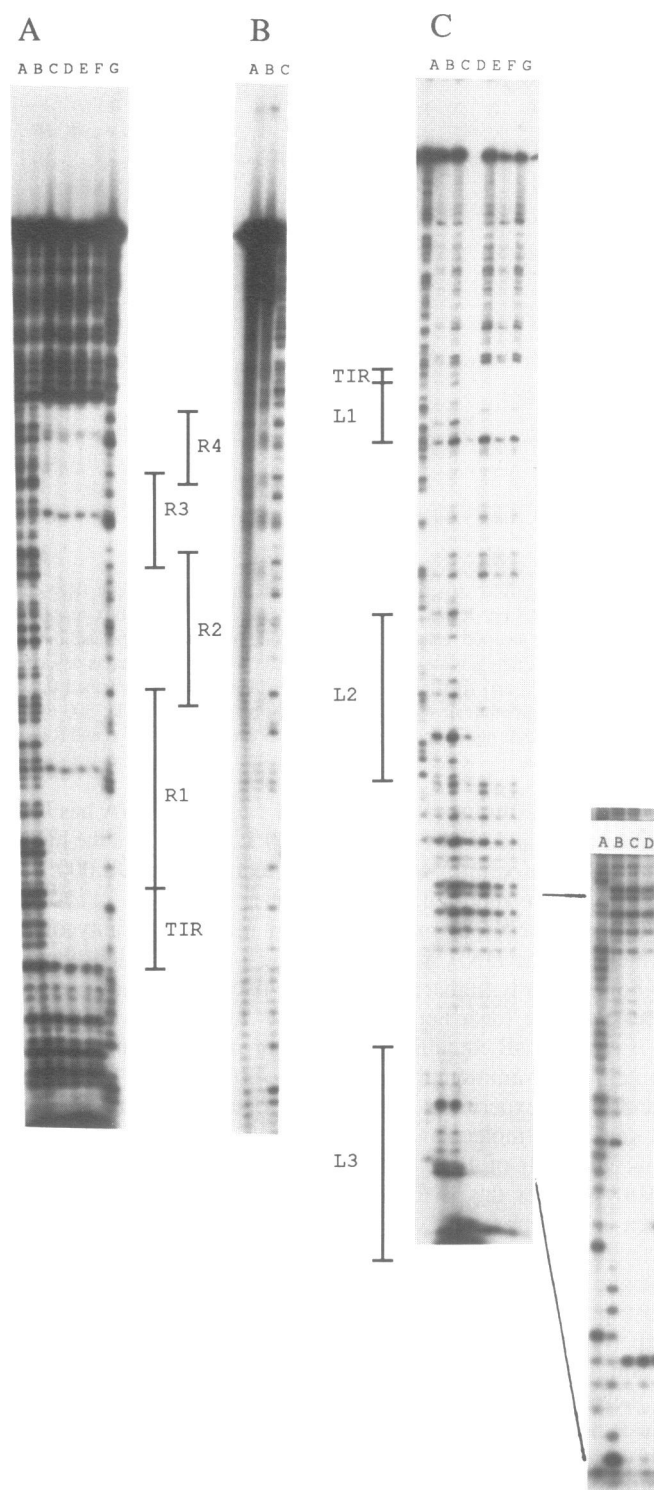
Figure 3. Footprinting of the right and left ends of Tn7. A DNase I footprinting of the right end. Reactions contained 0.5 ng DNA and protein equivalent to A: 0 ng; B: 15 ng; C: 30 ng; D: 150 ng; E: 300 ng; and F: 600 ng. G: a ladder to locate the position of bands inside the fragment. B Hydroxyl radical footprint of the right end. Reactions contained 0.5 ng DNA and A: 0 and B: 60 ng protein.C: the same ladder as above. C DNase I footprint of the left end. A is the ladder for sizing. Reactions contained 1 ng DNA and protein corresponding to B: 0 ng; C: 15 ng; D: 30 ng; E: 150 ng; F: 300 ng and G: 600 ng. Inset to Fig C shows more detail of the footprinting of L3. In this case a 15% polyacrylamide gel was used. A. shows the ladder; The other lanes contain 0.5 ng DNA and protein equivalent to B: 0 ng; C: 30 ng and D: 60 ng.

overlapping sites, and the pattern of protection obtained suggests that TnsB protects the whole of this region of 90 bp with the exception of three regions. These are at equivalent positions in the repeat sequence for R1, R3, and R4. While the cutting at these sites is maintained it does not seem to be enhanced, therefore this end shows no sites hypersensitive to cutting following TnsB binding. At the internal side of the repeats the last base protected is flush with the end of the last repeat, however at the terminal side, TnsB protects all of the 8 base pair inverted repeat at the end of Tn7 with the exception of the last base pair. The significance of this is discussed later.

Hydroxyl radical footprinting is able to locate sites of binding much more precisely. Application of this technique to TnsB binding at the right end of Tn7 shows that fewer bases actually appear to make tight contact with the protein (Figure 3B). If the positions of the bases which TnsB protects strongly, are superimposed over the structure of the DNA (Figure 4A), it can be seen that they are all inside the repeats, and all lie along the same face of the DNA helix.

DNase I footprinting of the left end shows three discrete sites protected (see Figure 3C and diagram in Figure 4B). For the two internal repeats (L2 and L3) the footprint covers the area of the direct repeat, and also some of the surrounding DNA. The footprinting pattern for the L1 repeat again extends into 8 bp terminal repeat at the end of the Tn7. As for the right end, only the last base remains unprotected. Binding of TnsB protein to the left end induces three hypersensitive sites.

For the left end it was not possible to obtain a hydroxyl radical footprint from either strand, despite repeated attempts. The reasons for this are not clear: one possible explanation is that the proteins bound to the separated sites of the left end fragment are themselves more exposed to the effects of hydroxyl radicals, than protein bound at the right end, where the sites of binding overlap.

The footprint was not significantly affected by the addition of larger amounts of Mg^{2+}, or by the addition of ATP.

Effect of TnsB concentration on footprinting. Titration of TnsB protein into the footprinting reaction should give some indication of the dissociation constant of the protein. It should also reveal whether the binding shows any cooperativity. Initially experiments were carried out using low concentrations of DNA (0.5−2.5 ng of fragment), and titrating the protein concentration (Figure 3). A sharp cutoff point was observed at a concentration of protein corresponding to 0.75 μg/ml protein. Concentrations of protein higher than this caused complete footprinting, whereas those below did not footprint. These results suggested cooperative binding. However the protein concentration at which footprinting could no longer be detected appeared to be the same for both the right end and the left end of Tn7. The pattern also did not vary with DNA concentration over a 5 fold range. Since it was not clear whether the observations of these experiments represented an accurate picture of the nature of the interaction between TnsB and the DNA, we carried out a similar experiment but in this case the protein concentration was kept constant, and above the critical concentration, and the DNA concentration was varied. The results of such an experiment with the fragment from the right end of Tn7 are shown in Figure 5. Again the binding appears to be lost from all sites simultaneously, therefore suggesting a cooperative interaction. TnsB continues to footprint the DNA completely right down to a protein to DNA ratio of 8.8:1. This therefore suggests that the maximum number of TnsB

molecules bound per site to give complete footprinting cannot be more than 2.2.

Similar experiments with the left end of Tn7 suggested that binding to this end was also cooperative, and titrated out at a similar protein:DNA ratio (data not shown).

Critical concentration of TnsB required for footprinting. At low concentrations of protein, the critical concentration of protein required for footprinting was independent of the end of Tn7 used. It was also unaffected by the protein:DNA ratio over quite a wide range. This raised the possibility that the observed cutoff point of footprinting was not dependent on the kinetics of the interaction between the protein and the DNA, but only on some property of the protein.

To test this we carried out a footprinting experiment with sets of two tubes. The two components of the set were identical except that the volume in which the reaction was carried out in one tube was three times that in the other tube. Therefore protein:DNA ratios were kept constant, but the absolute concentration of protein was varied. The results of such an experiment are shown in Figure 6.

Lanes C and D have the same amount of protein and DNA but the volume of the reaction in lane C is three times that in lane D. However while the sample loaded in lane D is competent to carry out footprinting that in lane C is not. This therefore suggests that the interaction observed depends on the absolute concentration of the DNA and/or protein present rather than their ratio. The earlier observation of the independence of the footprinting on the end of the Tn7 used, and the absolute DNA concentration, suggests that the important parameter is the protein concentration.

Crosslinking studies. The simplest explanation of the above data would be if there was some requirement for a specific aggregation state of the TnsB to see the footprinting. We therefore carried out crosslinking experiments to test the apparent extent of aggregation of the TnsB protein at concentrations both permissive and non permissive for footprinting. The experiment was carried out in both the presence and the absence of DNA. Higher order aggregates of the protein were observed by this technique, however no difference could be detected between the permissive and the non permissive concentrations of the protein. There was also no effect of carrying out the reaction in the presence of DNA.

TnsB binding to partial sites. It was already clear that the protein would bind tightly to incomplete Tn7 ends, since the purification method involved affinity purification using a column to which was attached an oligonucleotide containing only two sites.
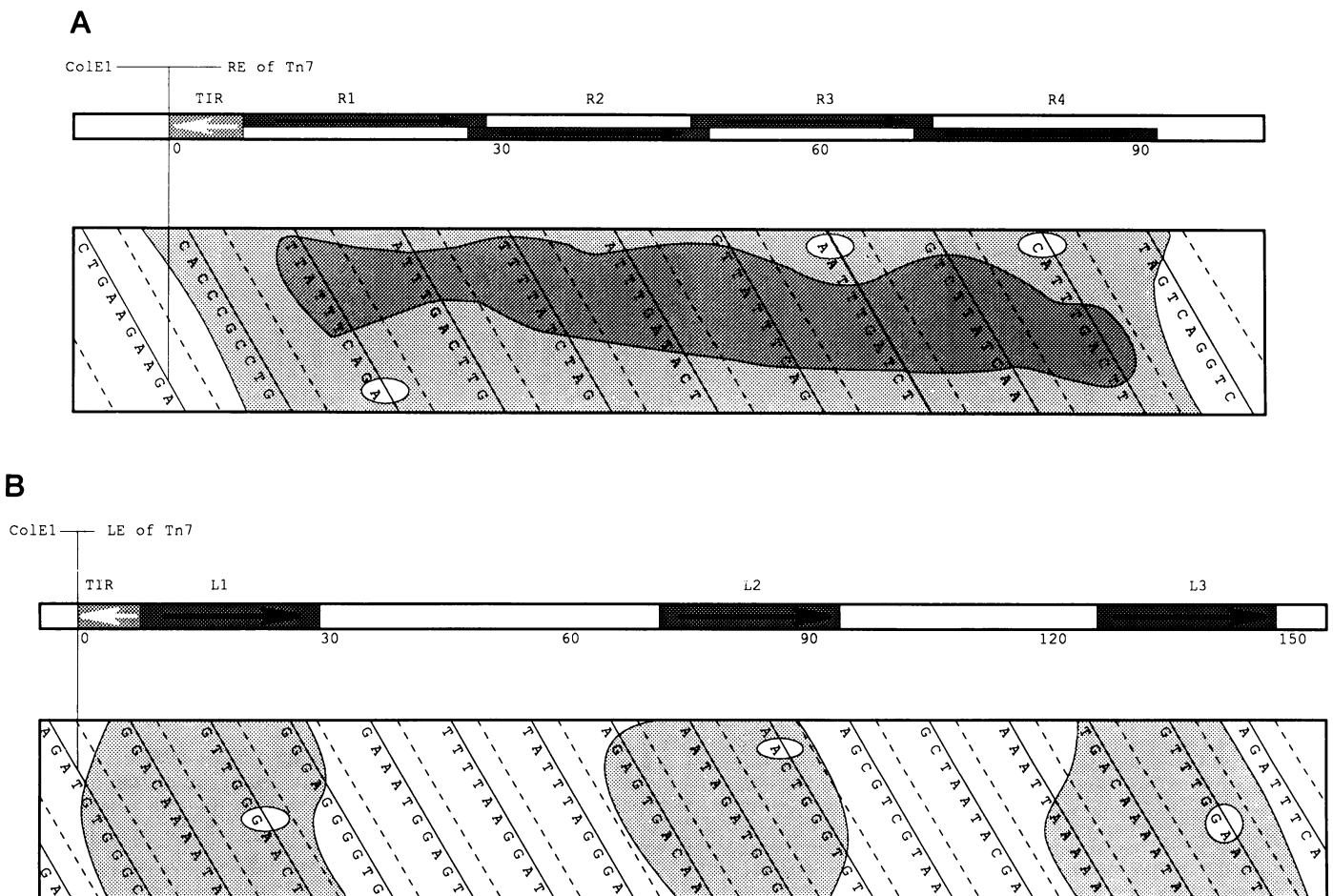

Figure 4. Diagramatic representation of the binding of TnsB to the right and the left ends of Tn7. The box represents the DNA helix opened out flat. Light grey areas are protected from DNase I alone. Dark grey areas are protected from both DNase I and hydroxyl radical. White areas represent unprotected or hypersensitive areas. The position of the protected areas relative to the direct repeats at the right and left end of Tn7 are shown on the marker at the side.
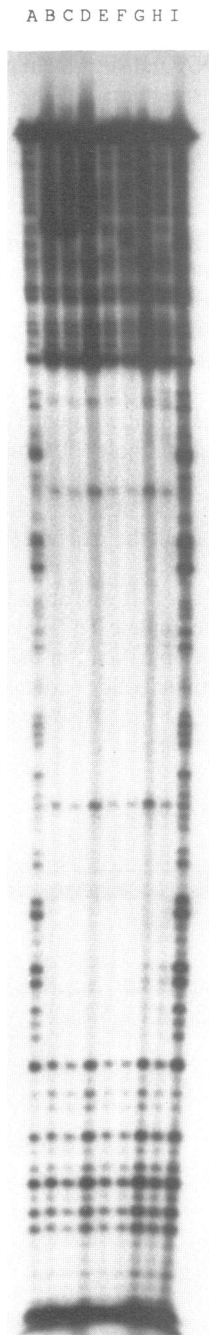
A B C D E F G H I

Figure 5. DNase footprinting at low protein:DNA ratios with the right end of Tn7. Lane A contains 0.5 ng DNA and 0 ng protein. All other lanes contain 1500 ng protein and protein:DNA ratios corresponding to A, 100:1; B, 80:1; C, 60:1; D, 40:1; E, 20:1; F, 12:1; G, 8:1; H, 4:1; I. (mole/mole)

However since the footprinting data suggested that the binding was cooperative, we were interested to see if the presence of partial repeats altered the footprinting pattern in any way. Partial fragments containing 2.5 and 1.5 repeats were released from the right end by cutting with BglII, and of 2 and 1 repeats were released from the left, by cutting with HincII. With half repeats no footprinting was observed. However the footprints of the intact repeats on these partial ends were indistinguishable from that observed for the same region on the intact end, even at fairly low concentrations of TnsB. Therefore even though the reaction
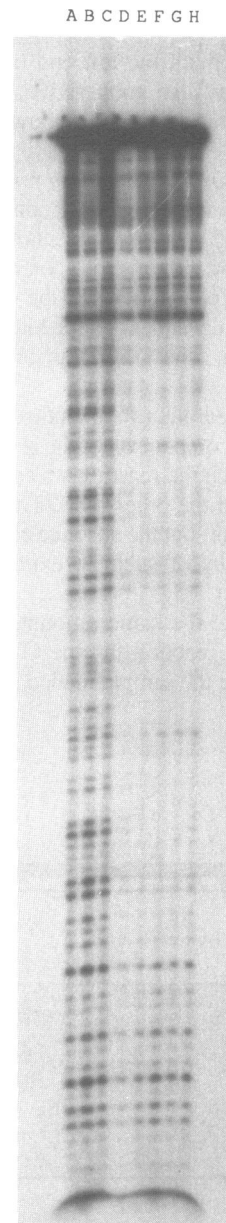
A B C D E F G H

Figure 6. Experiment to show effect of absolute protein concentration on DNase footprinting of the right end of Tn7. All reactions contained 0.5 ng DNA. Tubes contained TnsB protein corresponding to A/B. 6.6 ng; C/D. 19.8 ng; E/F. 30 ng; and G/H. 300 ng. Lanes A, C, E, and G were carried out in a volume of 60 μl Lanes B, D, F and H were carried out in a volume of 20 μl.

appears to be cooperative, the binding of TnsB protein at one site does not qualitatively alter the binding at adjacent sites.

Effect of other Tn7 proteins on the interaction between TnsB and DNA. Preliminary experiments have been carried out to look at the effect of the addition of TnsA, C and D, and combinations thereof, on the binding of TnsB to DNA. TnsA, C and D were added as ammonium sulphate fractions from overproducing strains. The characterisation of these strains is described elsewhere, but all of them had been shown to be active in an *in vitro* transposition system (Ganju *et al.* manuscript in prep.). The results of those analyses are shown in Figure 7. None of the other proteins, either singly or in combination, appear to have
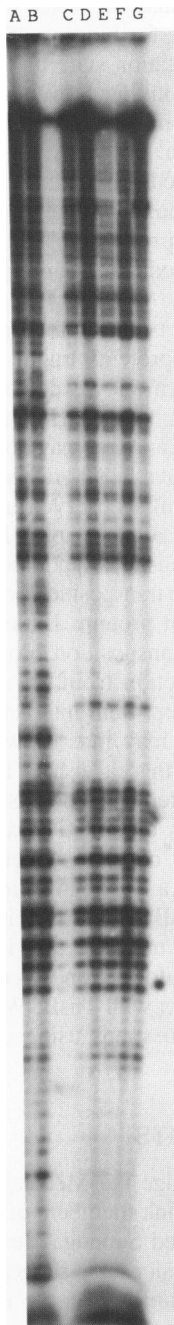
Figure 7. Effect of addition of other Tns proteins on the footprinting of the left end of Tn7. All lanes contained 1 ng DNA. The protein contents of the other lanes are; A: 0 protein; B: 10 μg BL21; C: 30 ng TnsB only; D: 30 ng TnsB plus 10 μg TnsA; E: 30 ng TnsB plus 10 μg TnsC; F: 30 ng TnsB plus 10 μg TnsD; G: 30 ng TnsB plus 10 mM ATP.

any noticeable effect on the TnsB footprinting pattern, or the critical concentration of TnsB required for footprinting. In addition there was no observable effect on the bandshifting (data not shown). Complete titrations to show the effect of pure TnsA, C, and D proteins on the affinity of the TnsB for its substrate are at present underway.

This result was not affected by the addition of extra Mg^{2+} and/or ATP to the reaction. Nor was there any effect of adding unlabelled DNA containing the *attTn7* site (the TnsD binding domain) into the reaction (data not shown).

DISCUSSION

In this paper we have presented a fast and efficient protocol for the purification to homogeneity of the TnsB protein of Tn7. We have also characterised the binding of the TnsB protein to its binding sites in the two ends of Tn7.

All predicted binding sites for TnsB in the right and the left end of Tn7, appear capable of binding TnsB. Therefore the protein can bind to three sites in the left end of Tn7 and four sites in the right end. The area protected from DNase I in each of these regions covers about three turns of the helix. More detailed footprinting of the right end using hydroxyl radicals suggests that the protein binds down one side of the DNA only. This may be necessary to introduce a particular DNA conformation. Alternatively it may leave this face free for approach by other proteins. From our data we cannot conclude whether the same would be true of the left end of the molecule. The footprinting at the outermost repeat at both left and right ends extends into the inverted repeat at the terminus of the Tn7. In both cases however footprinting stops one base short of the ends. This may be related to the cleavage reaction of the ends. TnsB alone cannot cleave the ends of Tn7, at least under all conditions so far tested. However it is possible that TnsB binding in this fashion could direct precise cleavage of the terminus by another protein, or perhaps itself carry out the cleavage in combination with other proteins. It is not yet clear which protein is involved in this reaction, and whether the process involves protein−protein, or protein−DNA interactions. However analysis of Tn7 protein sequence reveals a region of similarity in the TnsA and B proteins, the nature of which could constitute a protein−protein interaction site. In the TnsB protein this occurs immediately adjacent to the putative DNA binding site. It would therefore be well placed to direct cleavage of the ends by a combination of the TnsA and B proteins.

Within the large regions covered by the DNase I footprints there are bases that are not apparently protected by TnsB. At the left end of Tn7, these appear as hypersensitive sites. This is often indicative of protein induced bending of the DNA. These sites all lie down one face of the DNA, therefore the overall effect would be for them to reinforce each other and introduce a continuous bend into the DNA. For unprotected sites at the right end of the molecule there is no enhancement of the cleavage. This may be because there is no bending introduced here, or alternatively the different arrangement of sites at this end of the molecule (i.e. overlapping, rather than separated) might mask effects of the protein on the DNA structure. Again all unprotected bases lie along one face of the helix, thus would interact positively in bending the DNA. From footprints only three of the four sites at the right end show an unprotected base. However this site is not cut in the protein free DNA control, therefore this result is probably just an artefact caused by the preference of the DNase I for particular bases or regions in the DNA.

The patterns of bands seen in bandshifting experiments are quite complicated, and for neither end can the pattern be simply correlated with the binding of TnsB to individual sites inside that end. For the three sites at the left end there are several ways that the sites could be filled: If binding is not cooperative, and all sites are equivalent, as the protein concentration is increased there would be filling of one then two and then three sites. This should give three or seven shifted bands depending on how the mobilities of the different filled sites compare. An obligate order of binding would give only three bands. For cooperative binding,

there should be only one band in addition to free DNA. We would therefore expect to see either one, three, or seven bandshifts. We actually observe five different bandshifts. In addition, the kinetics of appearance of these bands is strange—the first three bands can be seen to be chased into each other at low protein concentrations, but the formation of the higher bandshifts requiring the addition of much more protein.

Following similar arguments for the right end there should be 1, 4, or 15 bandshifts. In this case six are observed. For the right end there appear to be four lower bands that are chased into each other, and then two higher bands.

The simplest explanation of the bandshifting data, is that bands observed at lower concentrations represent sequential filling of the three or four sites, at the left or right end of the molecule respectively. Higher bands therefore represent higher order structures—either DNA:protein, or protein:protein interactions. Higher order interactions would be consistent with a role of the TnsB protein, in transposition, of bringing the two ends of the transposon together prior to insertion. Protein:protein interactions are also suggested by the twin peaks of TnsB from affinity column. The 1 M salt eluant could represent protein:protein interactions, TnsB bound tightly to the recognition sequence may not elute until 3 M salt.

Bandshifting suggests a sequential filling of sites. For each end there are the same number of bandshifts as repeats. This could suggest an order of binding to the repeats. Alternatively the resolution of the technique may not be high enough to distinguish differences in mobility between complexes with the different sites. Data obtained from footprinting however, suggests cooperative binding, since binding to all sites appears and disappears simultaneously. Another discrepancy in the two sets of data is that the binding observed during bandshifting is weak—a protein:DNA ratio (mol/mol) of at least 500:1 is required to observe any bandshifting, whereas for footprinting, complete coverage is still observed even at a protein:DNA ratio of 8:1. This probably reflects the need for protein−DNA complexes to hold together throughout the whole of the electrophoresis for bandshifting, whereas for footprinting they need only be stable over the 2 minutes of DNase treatment. Due to the high salt concentration required to elute TnsB protein from the affinity column the footprinting data probably more accurately reflects the true situation. The observed differences in cooperativity are harder to rationalise; perhaps a more shallow titration of protein:DNA ratios in the footprinting experiment would reveal an obligate order of binding. In any case the differences in affinity of TnsB for its sites cannot be large, since bandshifting data suggests that all sites are filled over less than a 5-fold increase in protein concentration.

None of the data accumulated so far has allowed accurate determination of the dissociation constant for binding of TnsB to DNA. This situation may be further complicated by differing affinities of TnsB for individual repeats. However, full footprinting is observed at protein:site ratios of about 2 (mol/mol). Therefore, at the primary level of binding, there cannot be more than 2 molecules bound per site. Higher order structures such as those observed in the bandshifting, may contain much higher levels of TnsB. However it is not possible to assess this from footprinting data, as there is no apparent change in the footprint at high protein:DNA ratios. In addition it is not clear whether these high molecular weight complexes are functional in the transposition reaction, or if they just reflect some opportunistic property of the TnsB protein e.g. random aggregation.

The footprinting data also suggested that some property of the TnsB alone was limiting for binding. The simplest explanation is that this is the aggregation state of the protein. Preliminary experiments with crosslinking however showed no difference between the aggregation state at limiting, and non limiting concentrations of protein. This may suggest that TnsB needs to be very highly aggregated for binding. Alternatively the result may be an artefact of the technique, since crosslinking cannot distinguish between productive interactions leading to footprinting, and non productive aggregation of proteins due to their 'stickiness'.

It was hoped that effects on the footprinting and bandshifting of TnsB by other Tn7 proteins might give some clues about interactions between them. However no changes are observed. This does not rule out positive interactions. The DNase I footprints are quite large, and this may obscure subtle differences. Also differences may be at a kinetic rather than a qualitative level. For bandshifting, the shifts observed are already quite large. Additional binding of another protein may therefore not significantly alter the bandshift. In addition, interactions may not be observed between pure transposition proteins, they may depend on the presence of a host protein. Little is as yet known about the interactions of the transposition proteins with host factors. However if the 1 M fraction from the affinity column is truly representative of protein:protein interactions, then some of the proteins co-eluting here may interact with TnsB directly.

The techniques used in this paper have been able to give precise information about the interaction of TnsB with DNA. The data obtained has been mainly quantitative, and leaves the details of other more quantitative questions open. For instance: these experiments do not give a clear indication of the stoichiometry of binding between TnsB and DNA; it is not clear how the aggregation state of TnsB may affect the interaction of the protein with the DNA; also no inforination was obtained about how the other Tns proteins interact with TnsB. We are currently looking into these and related questions using other more appropriate techniques.

ACKNOWLEDGEMENTS

REFERENCES

1. Craig,N.L. (1989) In Berg and Howe (eds.), *Mobile DNA.* American Society for Microbiology, Washington, D.C. p. 211−226.
2. Rogers,M., Ekaterinaki,N., Nimmo,E. and Sherratt,D. (1986) *Mol. Gen. Genet.* **205**, 550−556.
3. Waddell,C.S. and Craig,N.L. (1988) *Genes Dev.* **2**, 137−149.
4. Flores,C.C., Qadri,M.I. and Lichtenstein,C.P. (1990) *Nucl. Acids Res.* **18**, 901−911.
5. Waddell,C.S. and Craig,N.L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3958−3962.
6. McKown,R.L., Waddell,C.S., Arciszewska,L.K. and Craig,N.L. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7807−7811.
7. Lichtenstein,C. and Brenner,S. (1982) *Nature (London)* **297**, 601−603.
8. Studier,F.W., Rosenburg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) In Goeddel, D.V. (ed.), *Methods in Enzymology* **185** Academic Press, p. 60−89.
9. Sambrook,S., FritschE.F. and Maniatis,T. (1989) (eds.) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.