TRANSPOSITION OF Tn7

A thesis submitted for
the degree of
Doctor of Philosophy
at the
University of Glasgow

by

Elizabeth Morrell

Institute of Genetics
University of Glasgow
Church St.
Glasgow

December 1990

The research reported in this thesis is my own and original work except where otherwise stated and has not been submitted for any other degree.

TABLE OF CONTENTS

Abbreviations

Chemicals

APS	-	ammonium persulphate
ATP	-	adenosine triphosphate
BSA	-	bovine serum albumin
DTT	-	dithiothreitol
EDTA	-	ethylenediaminetetraaceticacid (disodium salt)
EtBr	-	ethidium bromide
EtOH	-	ethanol
IPTG	-	isopropyl B-D thiogalactopyranoside
SDS	-	sodium dodecyl sulphate
SDS-PAGE	-	SDS-polyacrylamide gel electrophoresis
TEMED	-	NNN'N' tetramethyl ethylenediamine
Tris	-	tris (hydroxymethyl) amino ethane

Phenotype

X^r	-	resistance to X
X^s	-	sensitivity to X
ori	-	origin of replication

Units

m	-	10^{-3}
u	-	10^{-6}
n	-	10^{-9}
p	-	10^{-12}
f	-	10^{-15}

bp	-	base pair
kb	-	kilo base pairs
Da	-	Dalton
A	-	Amps
V	-	Volts
Ci	-	Curies
^oC	-	degrees Centigrade
g	-	grammes
l	-	litres
m	-	metres
M	-	molar
min	-	minutes

Acknowledgements

I would like to thank my supervisor, Dave Sherratt, for his help and advice during this project, and for encouraging me to leave sooner rather than later. Grateful thanks to Mary, who manages to hold the whole place together and yet still remains cheerful despite it all.

Many thanks to the Tn7 group - Angela, Stephen, and especially Karen the postdoc - and all members of the sixth and seventh floors both past and present, for lanes on gels, evenings on port, and all other assistance with my Nightmare on Church Street. Eternal love and gratitude to the boys in MY lab for going to the shops for my lunch (and chocolate), and for utter equanimity in the face of catenated expletives and thesis hysteria: in particular, to Sean for always doing the washing up, and to Dave for catching a sheep at six o'clock in the morning.

Eternal love and gratitude (extended) to Martin, for excitement about proteins and unshakeable faith in my competence, and especially to Marshall for doing everything except write my thesis.

Thanks to Sue for knowing about forms and photocopiers, to Margaret and Linda for being nice to me even when I was determined to be in a bad mood, and to the prep. room and media ladies for keeping my bugs sterilely housed and fed.

Many thanks to my parents and our Jane for financial assistance and being on the end of a phone (and buying the phonecards), to Lorna, Maryanne, Amy, May, Kim and Elspeth, and Dave's cats for letting me live in their houses, and finally to the choir of S. Mary's Cathedral, for friendship, support, and evenings in the pub to take my mind off it all.

This work was supported by the SERC.

SUMMARY

One of the transposition proteins of Tn7, TnsB, was overexpressed, and purified substantially. The protein has a subunit molecular weight of around 85 kDA, which is consistent with prediction from the DNA sequence, and appears to exist as a monomer in solution. The identity of the polypeptide was confirmed by N-terminal sequencing, which also demonstrated that the site of translational initiation is the first AUG in the reading frame, and that there is no N-terminal processing of the protein.

A Tn7-end specific DNA binding activity copurifies with this protein, and this activity is found only in tnsB-containing cells. This activity was further characterised using non-denaturing electrophoresis; the pattern of bands seen is complex. Apparent binding constants and rates of association and dissociation were determined; the binding activity was found to have a 2000-fold specificity for Tn7 ends in competition experiments, and to induce a bend in the DNA on binding. The biological implications of these results are discussed. Two host-encoded proteins, IHF and FIS, were tested for binding to Tn7 ends, but no convincing evidence was found for such an interaction.

The binding sites for TnsB were located by footprinting, to a repeated 22 bp motif found at both ends of Tn7. The bound site shows regions of enhanced sensitivity to DNaseI cleavage, indicating a structural change in the DNA. A potential inverted repeat is proposed to be involved in favouring the structural change.

The role of TnsB in transposition is proposed to be the recognition and bringing together of the transposon ends. A linkage difference, of around -2, was seen when a linear mini-Tn7 was recircularised in the presence of TnsB, compared with in its absence, which was dependent on there being two transposon ends on the same DNA molecule.

The results demonstrate that there may be some kind of interaction between the ends, mediated by TnsB. Similar experiments with singly nicked plasmid substrates did not show any trapped linkage; these results can be explained by proposing that all the induced supercoils are segregated into the domain between the two transposon ends.

Transposition assays were reconsidered with a view to carrying out mutageneses of transposition components. An assay based on papilliation was designed and tested, and this may have potential to be developed as a large-scale screen.

Finally, an experiment was designed in which heteroduplex transposons were used to determine whether Tn7 is replicated during transposition. Strain and transposon constructions are described, and the implications of results from this type of experiment are discussed.

CHAPTER 1

INTRODUCTION

1.1 General Introduction

Transposable elements can be defined as discrete DNA sequences that can move without permutation of their sequence between sites in the genome. They were first reported by Barbara McClintock following observations of instability in pigmentation of maize kernels (McClintock, 1950), and were called 'controlling elements' because of their phenotypic effects due to altered gene expression. The significance of the discovery was not fully appreciated until transposable DNA was also discovered in bacteria in the 1960's. There were two parts to this discovery. Firstly, bacteriophage Mu (Taylor, 1964), and insertion sequences (Jordan et al.,1968) were detected by their ability to cause mutation, and electron microscopy of heteroduplexes showed that these mutations were associated with physical gain of a segment of DNA. Subsequently, in the 1970's, acquisition of antibiotic resistance with a concomitant increase in the size of the recipient DNA was observed (Hedges and Jacob, 1974; Kleckner et al.,1975).

Since then, many more bacterial elements have been identified and characterised, in both Gram-negative and Gram-positive species. They confer a variety of phenotypes, including drug resistance (eg Tn5; Tn3), toluene degradation (eg Tn4653) and heavy metal resistance (eg Tn501). DNA transposition has also been studied in eukaryotic organisms, including yeast (eg Ty1), Drosophila (eg P-elements, retrotransposons such as copia, foldback), nematodes (eg Tc1 of C.elegans) and plants (eg Tam elements of Antirrhinum majus). The work has been facilitated by the advent of molecular biology techniques, with DNA sequencing being particularly important in identifying and delineating new elements, and in establishing common features.

As the subject of this thesis is the bacterial transposon Tn7, the rest of this Introduction will mainly

be concerned with prokaryotic transposable elements. However, some of the general features described, and the ideas and concepts involved, are equally applicable to transposons in higher organisms. Transposition in both prokaryotic and eukaryotic organisms has recently been extensively reviewed in 'Mobile DNA' (1989: Howe and Berg, Eds.).

NOTE on terminology: the term 'transposon' was originally used to describe mobile elements which contained drug resistance genes, such as Tn3 and Tn10, and to exclude elements such as the insertion sequences and Mu. For the purposes of this thesis, however, I shall use the terms 'transposon' and 'transposable element' interchangeably.

1.2 General properties of transposons

The termini of a transposon define the transposing unit, and are highly conserved between isolates of the same element. The two ends of a transposon contain related sequence elements. Commonly this is in the form of an inverted repeat, which confers symmetry on the transposon. In some transposons the two ends contain common sequence motifs, which may be arranged differently in the two termini: eg Mu (Craigie et al.,1984) and Tn7 (Lichtenstein and Brenner, 1982) in E.coli, Tn554 in S.aureus (Murphy et al.,1985), and the foldback element in D.melanogaster (Potter, 1982). The conserved sequences at the ends of a transposon are absolutely required for transposition (eg Tn10: Way and Kleckner, 1984; IS1: Gamas et al., 1985).

For the classical bacterial transposons, and some of the eukaryotic elements, a common consequence of transposition is the generation of a 3-13 bp direct duplication of the target sequence flanking the inserted transposon (Grindley and Sherratt, 1978). The transposon-specific size of the duplication indicates that it is dependent on transposon-encoded functions. The direct repeats are not required for further transposition, but

are believed to be generated as a consequence of staggered cleavage of the target; this feature is incorporated into models for transposition mechanisms (eg Berg, 1977; Shapiro, 1979).

A second class of element in bacteria includes Tn554 of S.aureus (Murphy, 1989), and the conjugative transposons such as Tn916 (Scott et al.,1988; Caparon and Scott, 1989). These do not generate a target duplication, and transpose by a mechanism related to phage lambda integration, with staggered breaks made at the element's ends and in the target. This alternative form of mobile element, more closely related to bacteriophage than to other transposable elements, will not be considered further.

The size of most transposons falls in the range 0.7-40 kb, although elements as large as 70 kb have been reported (Tsuda et al.,1989). At least some (and in the case of small elements, probably all) of the coding capacity is taken up with genes which allow the element its mobility. The enzymatic activities encoded are termed the 'transposase'. Although the precise nature of the activities encoded will depend on the mode of transposition (eg whether the intermediate is DNA or RNA), common features will be requirements for recognition of, and cleavage at, the ends of the transposon, for cleavage at the target site, and for ligation of the transposon ends into the cleaved target.

1.3 Prokaryotic transposons

Transposable elements in bacteria have traditionally been classified according to their genetic organisation. Not all elements can be fitted into this classification (eg Tn7) (and the classification should perhaps be extended to four classes to include the conjugative transposons).

Perhaps a more useful classification is in terms of whether the predominant mode of transposition is

replicative or conservative; as expected, the classes defined in this way are very similar to the original form; this second classification is a mechanistic one, and the transposition mechanism will determine the genetic structure as well as the observed transpositional properties. Previously unclassified transposons can be readily accommodated. However, the bacterial transposons are discussed here in their classical groups as a convenient way of describing the different types. The structures of the different classes are shown in figure 1.1, and some of the consequences of transposition in figure 1.2.

1.3.1 Class I: insertion sequences and composite transposons

Insertion sequences (IS's: reviewed by Galas and Chandler, 1989) are small (0.7-2 kb). Much of the coding capacity of the element is generally taken up with encoding the transposition functions. The transposase coding regions are characterised by their compact genetic organisation, with multiple overlapping reading frames, the expression of which is regulated by frameshifting. The termini of insertion sequences are inverted repeats, of 9 (IS50) to 30 (IS2) bp.

Composite transposons are larger than IS elements, and the additional coding capacity is generally taken up in encoding accessory determinants which are not required for transposition. They are composed of two IS elements, in either relative orientation, flanking the accessory gene, such as a drug resistance gene; eg Tn10: tetracycline resistance gene flanked by IS10 elements (Kleckner, 1989); Tn5: IS50 elements flanking a gene conferring resistance to kanamycin (Berg, 1989). Transposition functions are provided by the IS elements. The coherence of the composite transposon (tendency to remain as a unit) could be due to mutational inactivation of parts of the IS elements (eg IS50L in Tn5: Rothstein et

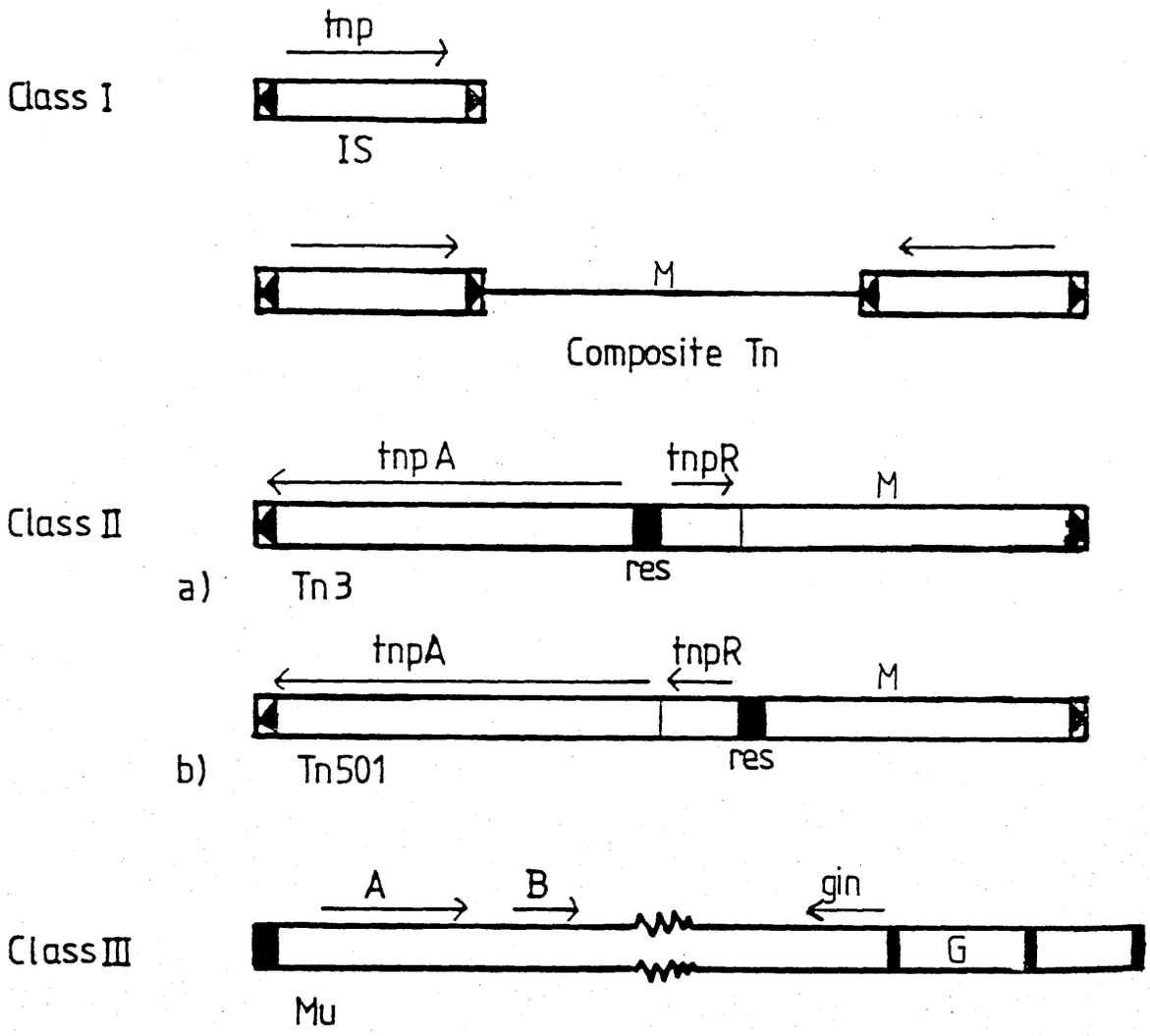


Figure 1.1 Illustration of the organisation of the transposable elements classes I-III, discussed in the text.
Arrows indicate the direction of transcription; the inverted repeats at the transposon ends are shown as black triangles. Blocks indicate the site-specific recombination sites res (Tn3 resolution) and gix (Mu inversion).
(From Grindley and Reed, 1985)

M marker gene, eg antibiotic resistance
G invertible G region

al., 1980), inhibition of transposition at the internal ends by marker gene transcription, interactions with cellular factors (eg dam methylase at the outside end of IS10: Roberts et al., 1985) or combinations of these.

The products of insertion sequence transposition are generally simple insertions into a new target site (figure 1.2A). In some cases, a 'cointegrate' structure is observed, in which the donor and target replicon are fused, joined by directly repeated copies of the element (figure 1.2B). For example, IS903 forms cointegrates at a frequency of around 1% (Weinert et al.,1984); for IS1, the cointegrate frequency is higher (Ohtsubo et al.,1980; Galas and Chandler, 1982). The structure of the cointegrate indicates that the element must have replicated during transposition, whereas the simple insert could have been formed with or without replication.

Composite transposons are able to transpose in the same way as their parent IS elements; the ones about which most is known generally produce simple insertions. In the case of Tn10, it has been demonstrated that the transposon is not replicated during transposition (Bender and Kleckner, 1986; see chapter 7). The composite transposon has a greater repertoire of transposition events than an IS element; in addition to movement of the whole transposon, the IS elements might transpose independently, and the presence of two IS elements on a plasmid creates two possible transposable segments - one containing the drug resistance gene, and the 'inverse' transposon containing the plasmid origin (figure 1.2C).

Intermolecular simple insertion and cointegration both have intramolecular analogues (reviewed by Kleckner, 1981). Simple insertion within the same molecule gives rise to deletion inversions or adjacent deletions (figure 1.2D; Tomcsanje et al.,1990), depending on the orientation of the target. The target is invariably within the transposon (i.e. between the transposing ends); this selectivity is consistent with models for conservative

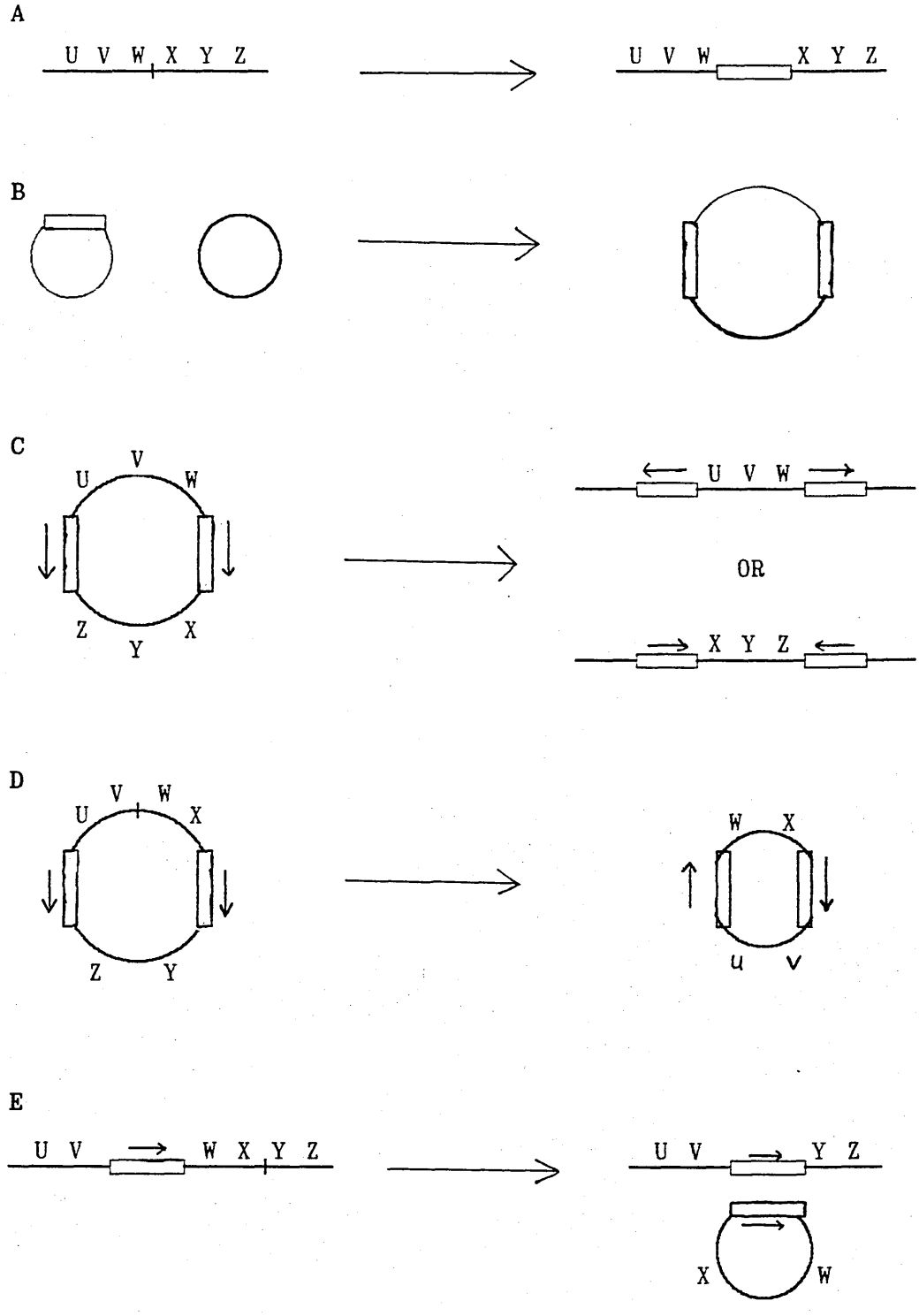


Figure 1.2 Transposon-mediated rearrangements
Letters U-Z are used to indicate the configuration of genes flanking the transposable element. The transposon is shown as a block. In A and B the target DNA is shown with a heavy line. Where appropriate, the target site is indicated (⊣). Arrows indicate orientation of the transposon where this is relevant.

A. simple insertion
B. cointegrate
C. inverse transposition
D. deletion inversion
E. adjacent inversion

transposition.

1.3.2 Class II: Tn3-like transposons

Class II elements, of average size 5 kb, are larger than IS elements, and, like composite transposons, encode accessory genes such as antibiotic (eg ampicillin: Tn3) or heavy metal (eg mercury: Tn21) resistance determinants. However, the accessory gene in these elements is an integral part of the transposon. The class II transposons have terminal inverted repeats of 38-40 bp, and mostly generate a 5 bp target duplication (reviewed by Sherratt, 1989). Two subclasses of type II elements can be distinguished on the basis of their genetic organisation and their ability to cross-complement.

Intermolecular transposition occurs by a two-step, replicative pathway (Arthur and Sherratt, 1979), mediated by two transposon-encoded enzymes. Transposase (the product of the tnpA gene) mediates transposition to generate a cointegrate. The second enzyme, resolvase (encoded by the tnpR gene) then regenerates the donor and target replicons, each carrying a transposon, by site-specific recombination at a site (res) within the transposon.

The intramolecular analogues of cointegration are adjacent inversions or deletions, depending on the target orientation (figure 1.2E); the target is always outwith the transposon ends (Bishop and Sherratt, 1984).

A phenomenon which was first observed for Tn3-like elements is transposition immunity, whereby replicons containing a copy of the transposon are poor targets for subsequent transposition events (Robinson et al.,1977). A single 38 bp transposon end is sufficient to cause the effect, i.e. the same sequences that are required for transposition (Kans and Casadaban, 1987), although it has been shown for Tn1000 that transposition and immunity can be uncoupled (Wiater and Grindley, 1990b). The mechanism of transposition immunity is not clear, but the phenomenon

is clearly important in limiting the number of transposon copies within a cell.

1.3.3 Class III: transposing bacteriophage

The best characterised members of this class are bacteriophage Mu (Mizuuchi and Craigie, 1986) and its heteroimmune relative D108 (reviewed by Pato, 1989). Elements in this class are large (around 40 kb), as they code for the structural proteins required for the extracellular stage of the phage lifecycle. Transposition is an integral part of the lifecycle; formation of a lysogen on infection is by a simple insertion, that has been shown to be conservative (Ackroyd and Symonds, 1983; Harshey, 1984). The phage replicates by multiple cycles of transposition *via* unresolved cointegrates, which results in the packaged phage containing random host sequences attached to the ends.

Two transposition functions are encoded; Mu A is required for all transposition events, and contains the transposase activities. Mu B, whilst not being absolutely required, stimulates transposition 10-fold, and is responsible for target site selection and the phenomenon of transposition immunity. Experiments *in vitro* have suggested that the presence of Mu A protein bound to phage ends causes Mu B to dissociate from that DNA molecule (at the expense of ATP), so the DNA is less efficiently utilised as a target (Adzuma and Mizuuchi, 1988, 1989).

Transposition generates a 5 bp target duplication. The ends of Mu do not consist of inverted repeats, but contain a common repeated motif which is the binding site for transposase (Craigie et al.,1984).

1.4 Models of transposition

Models of transposition mechanisms are of two types. Non-replicative, simple insertions can be explained by 'cut and paste' models (Berg, 1977). In these models, the transposon is excised from the donor by double-stranded

cleavages (which may or may not occur simultaneously), and ligated into the cleaved target. The target cleavage is staggered between the two strands; DNA synthesis to repair the gaps generates the target site duplication.

Transposition of Tn10 is believed to use this type of pathway, for a variety of reasons. The elegant heteroduplex experiment of Bender and Kleckner (1986) demonstrated genetically that both strands of the transposon are maintained in the primary transposition product, indicating that only limited replication of the transposon has occurred, if at all (discussed more fully in chapter 7). Tn10 promotes adjacent deletion and deletion inversion rearrangements (figure 1.2) (the structure of which can most easily be explained by a cut-and-paste model) and does so in vitro using substantially purified transposase (Benjamin and Kleckner, 1989). In addition, Tn10 has been shown to induce the SOS response (Roberts and Kleckner, 1988), presumably because of the formation of double stranded breaks in the DNA (Morisato and Kleckner, 1984), and transposase mutants have been isolated which promote the excision of a linear transposon fragment (Haniford et al.,1989).

Replicative transposition models for cointegrate formation were proposed in 1979 by Shapiro, and by Arthur and Sherratt; the strength of these models is that they allow for the production of simple insertions as an alternative product of the same pathway (figure 1.3). They also predicted the structure of intramolecular replicative events before they were demonstrated experimentally.

The choice between replicative and conservative transposition is made at (3) in figure 1.3, i.e. whether there is a second round of strand cleavage and religation (6,7), or continuation of replication through the transposon (4,5). The ratio of simple inserts to cointegrates generated by a given transposon could be determined by the efficiency of replisome formation at the termini (Grindley and Reed, 1985). Evidence from in vitro

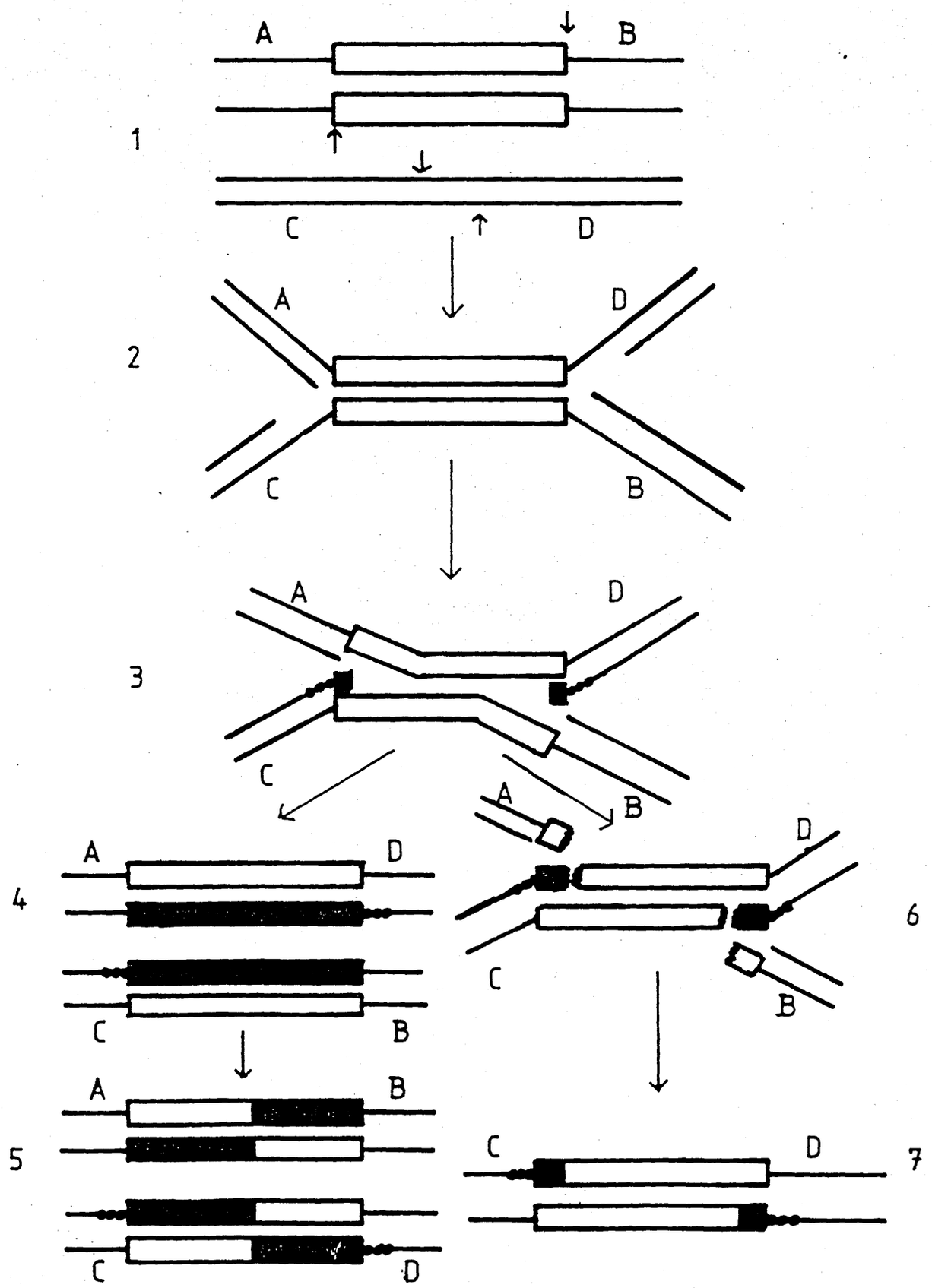

1
A
B
C
D
2
A
D
C
B
3
A
D
C
4
A
D
C
B
5
A
B
C
D
A
B
6
D
C
B
7
C
D

Figure 1.3 The Shapiro/Arthur and Sherratt model of transposition

1) The transposon is cleaved at the 3' ends, and the target is cut with a transposon-specific stagger with recessed 3' hydroxyl groups.
2) The transposon ends are transferred to the protruding ends of the target, forming two replication forks (3).
4) Replication may continue through the element, followed by sealing of the replicated ends; this generates a cointegrate, which can be resolved by site-specific recombination (5).
6) Alternatively, replication may terminate, after repair synthesis across the target site staggered break. A second round of strand breakages and religations generates a simple insert after repair of gaps.
(after Grindley and Reed, 1985)

experiments on phage Mu supports this model; uptake of labelled precursors showed that the polarity of strand exchange is as shown in figure 1.3, and a transposition intermediate with a structure consistent with the one proposed has been isolated (Mizuuchi, 1984; Craigie and Mizuuchi, 1985; reviewed by Mizuuchi and Craigie, 1986).

1.5 The transposase enzyme

Transposases are large, complex enzymes responsible for mediating the transposition reaction, i.e. strand cleavage at the transposon ends and at the target site, and religation. In order to do this, the transposase might be expected to recognise and bind to sequences in the ends of the transposon, and this has now been demonstrated for several elements, eg phage Mu (Craigie et al.,1986), Tn1000 (Wiater and Grindley, 1988), Tn3 (Ichikawa et al.,1987; New et al.,1988), and IS1 (Zerbib et al.,1987, 1990b). Where footprinting data are available, the binding site is found to lie within the inverted repeat (Tn1000, Tn3), or, in the case of MuA, to include sequences which are found repeated in the two ends.

Mutational analyses of the ends of several transposons have been carried out (IS903: Derbyshire et al.,1987; IS10: Huisman et al.,1989; Tn21: Martin et al., 1989; IS1: Zerbib et al.,1990). The data from these studies indicate that the inverted repeat of the transposons can be considered to consist of two domains. A domain involved in recognition and binding by transposase begins around bp number 6 of the inverted repeat, and includes most of the rest of the inverted repeat. The terminal 3 bp of the element can be mutated without affecting binding, and these sequences are thought to be involved in 'guiding' the transposase to the terminus, and in mediating cleavage.

The transposases are also found to have non-specific DNA-binding activity, with a lower affinity than that of specific binding; it is suggested that the non-specific

activity is required for searching the DNA to find the transposon ends, or for target-site binding (Ichikawa et al.,1987).

In the case of Mu, an *in vitro* system exists, which uses purified components; the transposase (MuA) provides all the enzymatic activities required for the cleavage and religation reactions listed above. The IN proteins of two retroviruses have also been shown to be both necesary and sufficient for *in vitro* integration (MoMLV: Craigie et al.,1990; ASLV: Katz et al.,1990). The same is likely to be true for other elements' transposases.

The level of transposase is generally rate-limiting for transposition *in vivo*, and mechanisms for transposition control are generally directed towards control of transposase expression (Kleckner, 1989; Berg, 1989; see section 1.7). Some of the bacterial transposons exhibit the phenomenon of the '*cis*-acting' transposase, whereby the transposase acts preferentially on transposon ends close to the transposase coding region (eg Tn10: Morisato et al.,1983. IS903: Grindley and Joyce, 1980). It is likely that this is an additional control mechanism, so that the presence of multiple copies of the transposon, all expressing transposase, will not result in an elevated level of transposition. In the case of IS903, it has been shown that *cis* action is the result of lability of the protein (Derbyshire et al.,1990). The phenomenon has only been observed in prokaryotes, where it is presumably dependent on coupled transcription/ translation.

There is some sequence similarity between the Tn3 family transposases (Mahillon and Lereclus, 1988), but they do not cross-complement to a great extent, and there is little obvious similarity to other types of elements. Hybrid transposases have been generated by homologous recombination between Tn21 and Tn501 transposases genes, allowing the specificity determinants to be located (Evans and Brown, 1987).

Some transposases from both Gram-positive and Gram-

negative elements have been found to show strong homology to retroviral integrase protein. Examples include the IS3 class of insertion sequences (Fayet et al.,1990) and Tn552 of S.aureus (Rowland and Dyke, 1990).

Purification of transposases has proved to be problematic, as most of them are insoluble in moderately high salt concentrations, and tend to form aggregates (eg Tn3: Ichikawa et al.,1987; Slatter, 1987. Tn10: H. Benjamin, personal communication).

1.6 The relationship between transposons and the cellular genome

The role played by transposons in the cell is still a matter of debate. Given the metabolic load imposed on the cell by the presence of a transposon, it can be argued that the transposon must confer a compensatory advantage. For transposons which carry drug resistance genes, the advantage to the cell is apparent, and the spread of these genes correlates well with the intensive medical use of antibiotics this century (Datta et al.,1981). However, the selective advantage is conferred by the gene: maintenance of the rest of the transposon is not necessary. A more sophisticated version of this model is that the transposon-borne accessory genes are not always required; by maintaining them on transposons (or other extrachromosomal entities like plasmids or phage) the size of the gene pool is increased, with the resistances being available when necessary, without significantly increasing the genome size of the individual (Campbell, 1981).

In chemostat experiments, strains containing Tn5 or Tn10 were found to out-compete isogenic strains lacking the transposon. In the case of Tn10 the advantage was due to insertion, probably at a particular site (Chao et al.,1983); with Tn5, the effect did not require transposition (Hartl et al., 1983).

In a changing environment in the wild, it might be that the rearrangements mediated by transposons confer a

selective advantage, in generation of genetic diversity on which selection can act (Campbell, 1981). The selection for transposon maintenance is then at the population level. The diversity created could be simply due to transposition, and its consequences for target gene expression. In addition, the rearrangements mediated by transposon activity might create new combinations of genes, or bring genes under new control regimes. It has also been suggested that transposons act as portable regions of homology (Kleckner and Ross, 1980) on which cellular recombination can act to generate new configurations with altered patterns of gene expression. In all these cases, the linkage of the transposon to a favourable mutation acts to ensure the element's survival.

An alternative, although not exclusive, hypothesis is the concept of 'selfish DNA' (Doolittle and Sapienza, 1980; Orgel and Crick, 1980), which proposes that a DNA sequence that increases its probability of survival in the cell needs no further phenotypic justification for its continued existence. The way in which a transposon ensures its survival is by its ability to overreplicate; simultaneous elimination of all copies is then unlikely. Increases in copy number can readily be understood for elements which transpose replicatively. Conservative transposition can also lead to overreplication; for example, Tn5 and Tn10 use dam methylation to couple their transposition to cellular replication, so that transposition occurs predominantly immediately after the passage of a replication fork, when the element has been passively copied by the host (Roberts et al.,1985; Reznikoff, 1982). Transposition to a site ahead of the replication fork ensures further replication; the Ac/Ds elements of maize are observed to transpose in this way (reviewed by Fedoroff, 1989) As conservative transposition generally destroys the donor molecule, transposition after replication also ensures survival of the element, although the host has lost a cell generation.

1.7 Control of transposition

Whatever the role of transposition is in the cell,

uncontrolled transpositional overreplication will be deleterious to the cell, and so ultimately also to the transposon, due to the high level of mutagenesis and the metabolic load imposed. Hence to minimise the selective disadvantage of transposons, systems for the control of transposition have evolved, and these are apparent for both prokaryotic and eukaryotic transposons.

In the case of Tn5, regulation is by a repressor protein which is an N-terminally truncated version of transposase, and inhibits transposase activity (Isberg et al.,1982). For Tn10, expression of transposase is regulated at the levels of transcription (Simons et al.,1983; Davis et al.,1985; Roberts et al.,1985), and translation (Simons and Kleckner, 1983), and by its preferential cis-action (1.5: Morisato et al.,1983). Control of Tn3 and phage Mu transposition is by transcriptional repression of transposase expression (Gill et al.,1979; Craigie et al.,1984), and the phenomenon of immunity.

The insertion sequence IS1 regulates expression of its transposase by frameshifting (Sekine and Ohtsubo, 1989). Two ORF's of IS1, insA and insB, are encoded on the same strand, and both are required for transposition; a -1 frameshift in the 3' end of insA allows expression of an insA-insB fusion protein which has transposase activity. The frameshift region contains a sequence identified as a 'frameshift signal' for retroviral Gag-Pol frameshifting. InsA has been shown to bind to the IS1 ends (Zerbib et al.,1987, 1990b); the fusion transposase protein maintains the helix-turn-helix domain of InsA, which may be responsible for binding activity in the transposase. InsA also acts as an inhibitor of transposition (Zerbib et al.,1990a); repression activity of a truncated 'transposase' is reminiscent of the IS50 'inhibitor' mechanism. In the case of IS1, control by InsA might be a result of repression of IS1 transcription, or of competition with transposase for the IS1 ends.

The transposition of P-elements in Drosophila provides another example of regulation. Transposition occurs only when P-elements are introduced into a strain which did not previously contain any; two models proposed are that the presence of multiple copies of the P-elements titrates out transposase (Simmons and Bucholz, 1985), or that the elements encode a repressor (i.e. similar to Tn5) (reviewed by Engels, 1989). In addition, P-element transposition is restricted to the germline, and this is mediated by germline-specific splicing of the mRNA to generate active transposase only in this tissue (Laski et al.,1986).

1.8 *in vitro* transposition systems

In order to study the biochemical details of the transposition reaction, several groups have attempted to set up cell-free transposition systems. The best characterised are those of phage Mu (Mizuuchi et al.,1983; and Tn10 (Morisato and Kleckner, 1987; Benjamin and Kleckner, 1989). Retroviral integration *in vitro* has recently been described (Craigie et al.,1990; Katz et al., 1990).

The high transposition frequency exhibited by phage Mu in both lysogenic and lytic phases, made it a preferred candidate for *in vitro* characterisation. In addition, the reaction requires only two phage-encoded proteins, both of which have been overexpressed and purified (Craigie and Mizuuchi, 1985a; Chaconas et al., 1985). The original system used extracts from MuA and B overproducing strains, and a crude cell extract, under conditions which allowed replication of the *E.coli* origin *in vitro*. The reaction required ATP and magnesium, and demonstrated that cointegrate formation involved replication of the transposon DNA. Initial experiments assayed full transposition into a phage target; more recent work, using purified A and B proteins and host factor (HU: Craigie et al.,1985) assayed the formation of an intermediate, seen

by gel electrophoresis, which is a fusion product of donor and target DNAs, and was shown to function as an intermediate on incubation with a replication-competent extract (Craigie and Mizuuchi, 1985b).

The in vitro system has proved to be a powerful tool, in allowing isolation and characterisation of the transposition intermediate, determination of the polarity of strand transfer (Mizuuchi, 1984), defining the substrate and enzymatic requirements for transposition (Craigie et al.,1985; Mizuuchi and Mizuuchi, 1989), and providing insight into the details of the transposition mechanism, such as the mechanism of bringing together the transposon ends and sensing their relative orientation (Craigie and Mizuuchi, 1986), and the basis of immunity (Adzuma and Mizuuchi, 1988, 1989).

The Tn10 system is more recent (Morisato and Kleckner, 1987), and assays an intramolecular transposition event which gives rise to transposon circles. No replication of the transposon is involved, and ATP is not required. The reaction requires magnesium, and two host factors (IHF and HU), and responds to dam methylation in the same way as the in vivo reaction. Experiments using strand-specific probes have allowed the structure of the products to be determined; strand interruptions are observed at the 5' ends of the transposon, which is consistent with the transposition models shown above (Benjamin and Kleckner, 1989).

1.9 Transposon Tn7

Tn7 is an unusually large transposon (14 kb), which confers resistance to trimethoprim, streptomycin, and spectinomycin. First isolated from plasmid R483 by transposition of its drug resistances (Barth et al.,1976: Barth and Datta, 1977), it was originally known as TnC, and its first reported use was in insertional mutageneses of plasmid RP4 (Barth and Grinter, 1977; Barth et al.,1978). It was apparent from this early work that Tn7

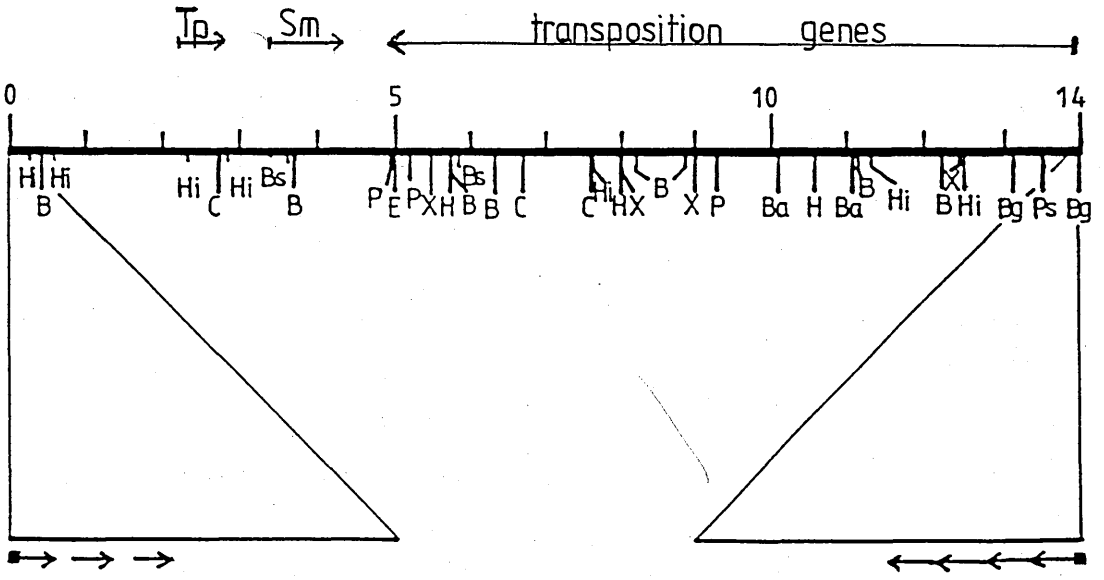


Figure 1.4 Restriction map of Tn7
The map is taken from Gosti-Testu et al., 1983. The organisation of the repeated sequences at the ends is shown below the map; arrows represent the 22 bp motif, and solid squares the 8 bp terminal inverted repeat. An outline of the genetic structure is indicated above the map, with arrows showing the direction of transcription.

Tp dhfrI (trimethoprim resistance)
Sm aadA (streptomycin/spectinomycin resistance)

B=BclI Ba=BamHI Bg=BglII Bs=BstEII C=ClaI E=EcoRI
H=HindIII Hi=HincII P=PvuII Ps=PstI X=XbaI

had interesting transpositional properties. Firstly, all the chromosomal insertions isolated in E.coli mapped to a single site (Barth et al.,1976), in marked contrast to the low degree of site selectivity exhibited by other transposons. The chromosomal insertion site is known as the 'hotsite' or attTn7; use of the latter terminology is not intended to imply any analogy with the phage λ att site.

Secondly, although insertions into plasmids did not display such marked site specificity, they all occurred in a particular orientation with respect to the plasmid restriction map (Barth and Grinter, 1987; Barth et al., 1978; Moore and Krishnapillai, 1982). The reason for this is not known. It may be that the transposon recognises some orientation-specific feature of the plasmid; it is unlikely to be replication, as some of the plasmids replicate bidirectionally, nor transcription, for a similar reason. Flores (1991) suggests that the phenomenon might be a result of the conjugation process by which the products were isolated; this is a testable hypothesis.

1.10 Tn7 transposition functions

Complementation studies by Mark Rogers (1986; Rogers et al.,1986), and insertional mutagenesis by Waddell and Craig (1988) identified five transposition functions, accounting for 9 kb of the coding capacity of Tn7. The experiments also demonstrated that transposition can be complemented efficiently in trans. The genes were named tnsA-E, for transposon seven, to avoid implications concerning the activities of the gene products, and they correlate with the open reading frames determined by Flores et al. (1990). The genes can in some cases be correlated with proteins produced from cloned fragments of Tn7 (Brevet et al.,1985), and with earlier data regarding Tn7 transposition requirements (Hauer and Shapiro, 1984; Smith and Jones, 1984, 1986; Ouartsi et al.,1985)

Not all the functions are required for all

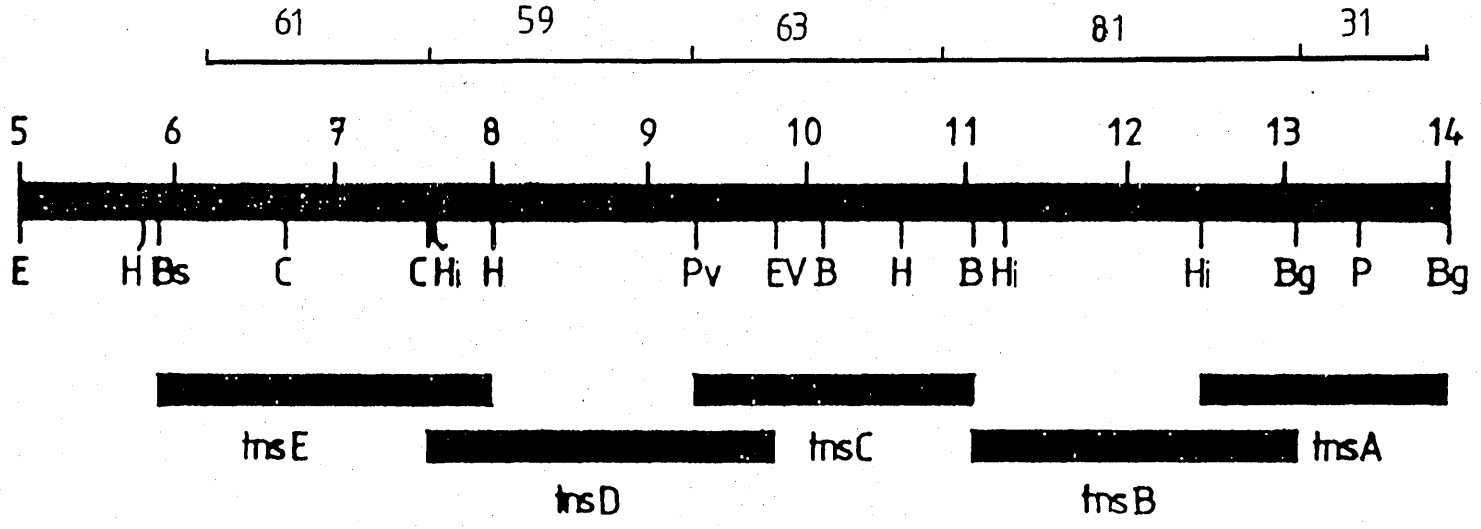


Figure 1.5 Organisation of Tn7 transposition functions
The figure represents the right-hand 9 kb of Tn7. The lower boxes show the limits of the genes as defined by complementation (Rogers, 1986; Rogers et al.,1986). The upper lines indicate the open reading frames in the DNA sequence (Flores et al.,1990; Flores, 1991); the predicted molecular weights of the encoded proteins are shown in kDa.

B=BamHI Bc=BclI Bg=BglII C=ClaI E=EcoRI Ev=EcoRV
H=HindIII Hi=HincII Pv=PvuII P=PstI

transposition events, as shown earlier by Smith and Jones (1984) and Hauer and Shapiro (1984). Two pathways of Tn7 transposition are observed, which require overlapping subsets of tns genes. All transposition events require tns ABC. High frequency transposition to the hotsite additionally requires tnsD, whereas tnsABC+E promotes transposition at 100-fold lower frequency to random sites unrelated to attTn7.

The different target choices of the two pathways immediately suggest roles for TnsD and TnsE, in recognition and selection of specific (TnsD) and non-specific (TnsE) insertion sites. Indeed, a tnsD-dependent DNA-binding activity specific for attTn7 has been detected (Ekaterinaki, 1987; K.McCurrach, personal communication; Waddell and Craig, 1989).

The roles of TnsABC are less clear. TnsC has been shown to interact with TnsD in binding to the hotsite, in an ATP-dependent manner (N.Craig, personal communication). The combination of ATP hydrolysis and target site choice is reminiscent of the activities of MuB. TnsB has been shown to bind to the ends of Tn7 (Ekaterinaki, 1987; McKown et al.,1987; this thesis) suggesting a role in transposon end recognition (chapters 4 and 5).

The tns genes are all transcribed in the same direction (Flores et al.,1990), from a promoter P1 in the right-hand end of the transposon (Ekaterinaki, 1987). A second promoter is found upstream of tnsB (Ekaterinaki, 1987). Mutations in tnsA show polarity on tnsB (Waddell and Craig, 1988), indicating that they are co-transcribed. There is no evidence for polar effects between the other tns genes, suggesting that they are independently transcribed; however, no promoters have been detected in functional assays, and the reading frames are very closely spaced or overlapping, which is characteristic of genes that are translationally coupled (Nomark, 1986).

1.11 The attTn7 site

Transposition to the chromosomal hotsite occurs at a

high frequency (around 1%), and is both site- and orientation-specific. The site lies at minute 84 on the E.coli chromosome, between the genes glmS and phoS (Gay et al.,1986) and has been cloned and sequenced (Lichtenstein and Brenner, 1981, 1982). The cloned site retains site and orientation specificity. The point of insertion lies in the glmS rho-dependent terminator (Gay et al., 1986); when Tn7 has inserted, the glmS transcript terminates at a rho-independent terminator in the right end of the transposon (Gay et al.,1986; Gringauz et al., 1988). The external transcript terminates before the tns coding region, so the tns genes are not expressed from this transcript; a similar control mechanism is seen with Tn10 (Davis et al.,1985). Tn7 insertion does not interrupt the glmS coding sequence, and so has no phenotype; this may explain why such a high transposition frequency can be tolerated.

Deletion studies to define the sequences required for hotsite activity have shown that all the required sequences lie to the glmS side of the point of insertion, and do not include the 5 bp target sequence that is duplicated as a result of Tn7 insertion (McKown et al, 1988; Qadri et al.,1989; see figure 1.6). The tnsD-dependent binding activity recognises these upstream sequences (Waddell and Craig, 1989). This 'action at a distance' is reminiscent of type I restriction enzymes, which cleave at a site other than their binding site. Tn7 chromosomal hotsites have been found in many other species of bacteria. These include Pseudomonas aeruginosa (Caruso and Shapiro, 1982), Caulobacter crescentus (Ely, 1982), Xanthomonas campestris pv. campestris (Turner et al., 1984), Vibrio parahaemolyticus (Thomasian and Voll, 1989) Pseudomonas fluorescens (Barry et al.,1986), and several others. Tn7 inserts into several sites in one Vibrio species, and shows a preference for one site which results in ilv auxotrophy (Thompson et al.,1981). The hotsites from Klebsiella pneumoniae and Serratia marcescens have been sequenced, and are found to have homology to glmS of E.coli over the region required for

att site activity (I.Qadri, personal communication).

If no att site is available, tnsABC+D directs Tn7 to 'pseudo-att' sites, with some sequence homology to the required region of attTn7, at a lower frequency than to attTn7 (Kubo et al., 1990). tnsABC+E - mediated transposition is to sites unrelated to the hotsite.

Because the required sequences lie to one side of the point of insertion, transposition of Tn7 to the site does not disrupt those sequences. However, secondary insertions into an occupied hotsite have not been detected. This could be an immunity effect analogous to that of Tn3; Tn7 does show immunity, but the strength of the effect is variable (Hauer and Shapiro, 1984; Arciszewska et al., 1989). Alternatively, the effect may be due to structural changes in the DNA following insertion.

1.12 The termini of Tn7

Each terminus of Tn7 has a 28 bp inverted repeat, the outside 8 bp of which are perfectly conserved between the left and right ends (Lichtenstein and Brenner, 1982). Within the repeat lies a 22 bp motif, which is further repeated, giving a total of four copies in the right end and three in the left. Hence the ends of the transposon are asymmetric with respect to this motif, which is reflected in functional asymmetry; two minimal right ends (70 bp) in inverted orientation are capable of transposition, whereas two left ends (166 bp) are not (Arciszewska et al.,1989). There is further asymmetry in the organisation of the 22 bp repeats, which are contiguous in the right end, but separated by spacers of different lengths in the left end. Another difference was noted by Flores et al (1990); a strongly conserved AT dinucleotide in the right end copies is GG in the left end ones.

TnsB has been shown to repress P1 in vivo (Ekaterinaki, 1987; Rogers et al.,1986). The -35 region of P1 lies within the final 22 bp copy, and it has been proposed that these repeats are the binding site for

'transposase', (specifically, for TnsB). This is supported by the work presented in this thesis.

The right end of Tn7 also contains a dam methylation site, and potential IHF binding sites. No differences in transposition frequency have been detected in strains mutant in these proteins (Ekaterinaki, 1987).

1.13 Tn7 drug resistances

The trimethoprim resistance gene carried on Tn7 encodes a type I dihydrofolate reductase enzyme, which confers resistance to trimethoprim because of its lower affinity for the toxic analogue than the chromosomal enzyme (Fling and Elwell, 1980). The gene is 0.5 kb long, and the product has a subunit molecular weight of 18 kDa and exists as a dimer (Novak et al.,1983). Acquisition of Tn7 is the most common response in the wild to selection with trimethoprim (Pulkkinen et al.,1984; Steen and Skold, 1985; Kraft et al.,1986); the drug is used extensively in both human and veterinary medicine to treate kidney and other urinary tract infections, so Tn7 has a medical importance.

The spectinomycin/streptomycin resistance is conferred by aadA, encoding an adenylyltransferase enzyme which detoxifies the antibiotics by modification (Fling et al.,1985; Hollingshead and Vapneck, 1985).

The resistance genes are flanked by conserved sequences, which are also associated with resistance genes on other transposons and on plasmids (Cameron et al., 1986). There is a highly conserved 59 bp inverted repeat at both ends of the aadA gene, and the same sequence is present when aadA is located in combination with other resistance genes in other elements. An open reading frame to the left of the trimethoprim resistance gene has been shown to have homology with the lambda Integrase family of site-specific recombinases (Sundstrom et al., 1988).

This arrangement of conserved sequences and integrase gene is found surrounding various combinations of drug resistances in many plasmids and transposons, including

Tn21, Tn1696, pSa, R46, R388, and Tn7 (Cameron et al.,1986; Hall and Vockler, 1987; Sundstrom et al.,1988; Stokes and Hall, 1989; Sundstrom and Skold, 1990).

It has been proposed that this is a system for the insertion of drug resistance 'cassettes' into transposons and plasmids, to generate novel combinations. The 59 bp repeat might be the recombination site. RecA-independent cointegrate formation by this system has been observed for Tn21 (Martinez and de la Cruz, 1988, 1989).

1.14 Aims of this project

A long term objective of the work on Tn7 is to set up a cell-free system, to examine the biochemical details of the transposition process; at the time when this project was started, no such system existed for Tn7, although now Nancy Craig's group have created one (personal communication). To facilitate *in vitro* work, it was decided to overproduce and purify each of the *tns* gene products. Chapter 4 describes the purification of one of the proteins, TnsB; in chapters 4 and 5 the DNA binding activity of the protein is further characterised.

However, it was also felt that there was much information concerning transposition of Tn7 which we did not have, and which would facilitate the design of an *in vitro* system by eliminating some of the variables to be considered. It had not been directly demonstrated whether Tn7 replicated during transposition or not; the preliminary stages of an experiment to answer this question are presented in chapter 7.

Having decided to characterise further some of the features of Tn7 transposition *in vivo*, the major problem was to find a means of assaying transposition in a meaningful way, appropriate to the questions being asked. Chapter 6 discusses the question of *in vivo* transposition assays, and describes three alternative ways of assaying transposition, and their potential applications both *in vivo* and *in vitro*.

CHAPTER TWO

MATERIALS AND METHODS

2.1 **Bacterial strains.** The bacterial strains used were all derivatives of Esherichia coli K-12 and are listed in Table 2.1.

2.2 **Plasmids.** The plasmids used and constructed in this study are listed in Table 2.2.

2.3 **Bacteriophage.** The bacteriophage lambda derivatives used in this study are listed in Table 2.3.
Phage P1 was a generous gift from Dr. Millie Masters, and phage P2 from Dr. Noreen Murray.

2.4 **Tn7 derivatives** are listed in Table 2.4

2.5 **Synthetic oligonucleotides** were synthesised on an Applied Biosystems 391 PCR-Mate oligonucleotide synthesiser.

A) **R1 22 bp motif** (chapter 4): two complementary oligonucleotides, generating XbaI and SalI cohesive termini.
Top strand:
5'-CTAGAGACAATAAAGTCTTAAACTGAAG-3'
Bottom strand:
5'-TCGACTTCAGTTTAAGACTTTATTGTCT-3'
(Bold type represents Tn7 sequences.)

B) **R1 plus Tn7 8 bp terminal inverted repeat** (chapter 4): strategy as above.
Top strand:
5'-CTAGATGTGGGCGGACAATAAAGTCTTAAACTGAAG-3'
Bottom strand:
5'-TCGACTTCAGTTTAAGACTTTATTGTCCGCCCACAT-3'

C) **HindIII site mutagenesis oligonucleotide** (chapter 7): one self-complementary oligonucleotide.
5'-AGCTGAATTC-3'

Table 2.1 Bacterial Strains

Strain	Genotype	Source
DS801 (AB1157)	thr1, leu6, hisG4, thi1, ara14 proA2, argE3, galK2, sup37, xyl15, mtl1, tsx33, str31	D.Sherratt
DS902 (AB2463)	DS801, but recA13, arg^+	D.Sherratt
DS941	DS801, but recF143, supE44, lacZΔM15, $lacI^q$	D.Sherratt
DS953	DS941, but sup^o	D.Sherratt
DS959	DS941, but lambda lysogen	D.Sherratt
DS942	DS941, but Δlac-pro	D.Sherratt
DS947	DS942, but sup^o	D.Sherratt
MR4	DS903::Tn7-1 (DS903 is DS801, but recF143)	M.Rogers
C600	hsdR, supE, lacY1, tonA21	Promega
C600 hfl	C600 but hfl::Tn10	Promega
C1400	supE, supF, hsdS, met, recA	K.Murray
N3331	DS801 but sfiA::Mu*d*(Ap^rlac)	R.Lloyd
NK7515	Packaging strain BHB2688, but mutL::Tn5	N.Kleckner
NK7516	Packaging strain BHB2690, but mutL::Tn5	N.Kleckner
RH302	W1331, but mutL::Tn5	N.Kleckner
EM3	DS953 attTn7::galK	Chapter 6
EM5	DS959 attTn7::galK	Chapter 6
EM1	DS947, but mutL::Tn5	Chapter 7
EM2	EM1, but P2 lysogen	Chapter 7
EM11	EM1 but lambda lysogen	Chapter 7
EM12	EM2 but lambda lysogen	Chapter 7

Table 2.2 Plasmids

Plasmid	Size(kb)	Marker	Origin	Comments	Source
pUC18	2.7	Ap	ColE1	Cloning vector	1
pACYC184	4.2	TcCm	P15A	Cloning vector	2
pBEND2	2.5	Ap	ColE1	Cloning vector	3
R388	32.2	Tp	IncW	Conjugative plasmid	4
pKO1	3.9	Ap	ColE1	Promoter probe vector	5
pKO500	3.9	Ap	ColE1	Promoter probe vector	5
pKL500	3.9	Ap	ColE1	pKO1 with pUC18 polylinker and Plac	5
pCB104	4.3	Cm	λdv	Cloning vector	C.Boyd
pMR100	5.6	Km	λdv	Ptac expression vector	M.Rogers
pMR78	3.4	Tp	ColE1	Ptac expression vector	"
pEAL1	5.2	Tc	P15A	attTn7 in pACYC184	6
pEAL::Tn7	19.2	TcTpSpSm	P15A	Tn7 in pEAL1	
pEA305	7.5	Ap	ColE1	cI overproducer	7
pEN300	33.3	Tp	IncW	attTn7 from pEAL1 in R388	E.Nimmo
pNE200	2.9	Ap	ColE1	203bp Tn7 RE, in pUC8	N.Ekaterinaki
pMR11	5.5	ApCm	ColE1	Tn7-1 constructed in pUC18	M.Rogers
pMR80	3.0	Ap	ColE1	280bp attTn7 in pUC18	"
pMR203	5.6	Tp	ColE1	Tn7 HindIII(8-5.8) in pMR78 (tnsE)	"
pMR204	4.8	Tp	ColE1	Tn7 0.12-HincII(1.5) in pMR78 (tnsA)	"
pMR205	6.7	Tp	ColE1	Tn7 0.12-HindIII(10.6) in pMR 78 (tnsAB)	"
pMR207	5.9	Tp	ColE1	Tn7 BglII(13.1)-HindIII(10.6) in pMR78 (tnsB)	"

pMA1441	3.0	Ap	ColE1	Tn3 *res* in SmaI of pUC18	M.Boocock
pNK1079	14.5	ApTc	ColE1	Tn10-*lacZ*$^{+}$ in pBR322 derivative	N.Kleckner
pNK1080	14.5	ApTc	ColE1	pNK1079 but Tn10-*lacZ*$_{am}$	"
pLM1	3.1	Ap	ColE1	Tandem dimer of Tn7 RE, in pUC18	Chapter 4
pLM2	2.9	Ap	ColE1	Tn7 RE from pNE200, in HincII of pUC18	"
pLM3	2.9	Ap	ColE1	pLM2 but opposite orientation	"
pLM5	2.7	Ap	ColE1	Oligo.1 in SalI-XbaI of pUC18	"
pLM8	2.6	Ap	ColE1	Oligo.2 in SalI-XbaI of pBEND2	"
pLM11	3.6	Ap	ColE1	pMR11 with Cm^{r} deleted (Tn7-4 in pUC18)	Chapter 5
pLM110	5.9	Tc	P15A	Tn7-4 transposed into pEAL1	"
pLM111	8.0	TcCm	P15A	Tn7-1 transposed into pEAL1	"
pLM112	13.2	Tc	P15A	EcoRI deletion of 5kb Tn7L from pEAL::Tn7	M.Rogers
pLM113	6.0	Tc	P15A	PstI(13.5)-BstEII(10) deletion from pLM112	Chapter 5
pLM414	6.6	Cm	λdv	EcoRI LE/Cm^{r} fragment from pLM111,in pCB104	"
pLM8071	5.8	ApCm	ColE1	Tn7-1 transposed into pMR80	"
pLM8075	4.0	Ap	ColE1	EcoRI *att*/RE/Cm^{r} deletion from pLM8071	"
pLM50	3.9	Ap	ColE1	EcoRI-HindII *att* from pMR80 in EcoRI-HindIII of pKO1	Chapter 6
pLM50::Tn7	17.9	ApTpSpSm	ColE1	Tn7 transposed into pLM50	"
pLM50::Tn7-1	6.7	ApCm	ColE1	Tn7-1 transposed into pLM50	"
pLM50::Tn7-5	3.6	Ap	ColE1	Tn7-5 transposed into pLM50	"
pKM2	15.4	ApSpSm	ColE1	Tn7-3 constructed in pLM50::Tn7	K.McCurrach

Table 2.2 continued

Plasmid	Size(kb)	Marker	Origin	Comments	Source
pLM55	7.0	Cm	λdv	att/galK (EcoRI-NdeI) from pLM50, plus downstream att (DraI-EcoRI) from pEAL1, in EcoRI of pCB104	Chapter 6
pLM60	6.6	Km	\dv	EcoRI att from pEAL1, in EcoRI of pMR100	"
pLM601	12.3	KmCmAp	λdv,ColE1	pLM60 linearised at SalI, into pMR11 linearised at SalI of Tn7-1	"
pLM602	12.3	KmCmAp	λdv,ColE1	pLM601 but opposite orientation	"
pLM79	7.4	Ap	ColE1	BamHI-NcoI $lacZ^+$ from pNK1079, in BamHI of pMA1441	"
pLM80	7.4	Ap	ColE1	As pLM79, but $lacZ_{am}$	"
pLM800	9.9	ApCm	ColE1	BamHI-PstI $lacZ_{am}$ from pLM80, in SalI of pMR11 (Tn7-lac80)	"
pLM801	9.9	ApCm	ColE1	pLM800 with polylinker HindIII replaced with EcoRI	"

References:

1: Yanisch-Perron et al.,1985
2: Chang and Cohen, 1978
3: Zweib et al., 1989
4: Datta and Hedges, 1972
5: McKenney et al., 1981
6: Lichtenstein and Brenner, 1981
7: Amann et al., 1983

Tn7 coordinates are given in kb from the right-hand end.

Table 2.3 Bacteriophage

Phage	Description	Source
EMBL4	Replacement vector	Murray, 1983
EMBL4 lib	*T.brucei* library in EMBL4	J. Mottram
gt10	Insertion vector	Murray, 1983

Table 2.4 Tn7 derivatives

Transposon	Markers	Size	Description	Source
Tn7	Tp Sm Sp	14	Canonical Tn7	
Tn7-1	Cm	2.8	RE - PstI (536 bp); LE - HincII (168 bp); *cat* from pACYC184	M.Rogers
Tn7-3	Sm Sp	11.5	Tn7 with 2.5kb AvaI (Tp^r and adjacent gene) deletion	K.McCurrach
Tn7-4	-	2.1	Tn7-1 with *cat* deleted	Chapter 5
Tn7-5	-	2.15	Tn7-4 with Ptac at SalI site	Chapter 6
Tn7-*lac*79	Cm $lacZ^+$	7.3	Tn7-1 with *lacZ* at SalI site	Chapter 7
Tn7-*lac*80	Cm $lacZ_{am}$	7.3	As Tn7-*lac*79, but $lacZ_{am}$	Chapter 7

2.6 **Proteins** used were generous gifts from the following:

FIS	Regina Kahmann (Berlin)
IHF	George Szatmari
Tn3 resolvase	Martin Boocock
TnsD extract	Karen McCurrach

2.7 **Chemicals.**

CHEMICALS	SOURCE
General chemicals, biochemicals and organic solvents	BDH, May and Baker, Sigma
Media	Difco, Oxoid
Agarose	BRL
X-gal, IPTG	BRL
Radiochemicals	NEN
10x restriction enzyme buffers	BRL, Boehringer Mannheim
5x ligation buffer	BRL
Nucleotides	Boehringer Mannheim

2.8 **Buffer solutions.**

A) DNA Electrophoresis

10x E buffer: 242g Tris, 82g sodium acetate, 18.6g $Na_2EDTA.2H_2O$, made up to 5 litres in tap water, adjusted to pH 8.2 with glacial acetic acid.

10x TBE buffer: 109g Tris, 55g boric acid, 9.3g $Na_2EDTA.2H_2O$, made up to 1 litre in distilled water; pH is 8.3.

10x TBE buffer (sequencing gels): 121.1g tris, 55g boric acid, 9.3g $Na_2EDTA.2H_2O$, made up to 1 litre in distilled water; pH is 8.3.

Single colony gel loading buffer: 2% ficoll, 1% SDS, 0.01% bromophenol blue, 0.01% orange G in buffer E. RNaseA

(boiled) was added, to 10ug/ml.

Polyacrylamide gel loading buffer: 1% ficoll, 0.1% SDS, 0.02% orange G, 0.01% bromophenol blue in distilled water.

4x Horizontal agarose gel loading buffer (K mix): 25% sucrose, 0.2mg/ml protease K, 0.01% bromophenol blue in distilled water.

B) Protein electrophoresis

Laemmli gels

4x running gel buffer: 1.5M Tris/HCl pH 8.8, 0.4% SDS.

4x stacking gel buffer: 0.5M Tris/HCl pH 6.8, 0.4% SDS.

10x Laemmli gel buffer: 30g Tris, 144g Glycine, 10g SDS, made up to 1 litre in distilled water.

Sample buffer: 625mM Tris/HCl pH 8.0, 3% SDS, 5% 2-mercaptoethanol, 10% glycerol, 0.01% bromophenol blue. Also used with 1% SDS, for protein samples with high concentrations of KCl.

Tris/Trycine gels

Anode buffer: 0.2M Tris pH 8.9.

Cathode buffer: 0.1M Tris, 0.1M tricine, 0.1% (w/v) SDS; pH should be 8.25, adjusted with tricine.

Gel buffer: 3M Tris, 0.3% SDS; pH 8.45.
(The running gel contains 1/3x gel buffer, and the stacking gel, 1/4x.)

C) *in vitro* DNA manipulation

10x restriction buffers: the recommended buffers provided with the restriction enzymes were used; stored at 4°C or over the long term at -20°C. Klenow polymerase and CIP

were used in standard restriction buffers.

5x ligase buffer: BRL; stored at -20°C.
T4 polymerase and kinase buffers: as described in Sambrook et al. (1989).

1x TE buffer: 10mM Tris/HCl, 1mM EDTA; pH 8.0

D) Other solutions

Phenol: all phenol contained 0.1% 8-hydroxyquinoline, and was buffered against 0.5M Tris/HCl (pH 8.0).

Alkaline lysis plasmid preparation solutions
Birnboim-Doly I: 50mM glucose, 25mM Tris/HCl pH 8.0, 10mM EDTA.
Birnboim-Doly II: 0.2M NaOH, 1% SDS: made fresh.
Birnboim-Doly III: 5M potassium acetate: equal volumes of 3M CH_3COOK and 2M CH_3COOH, pH will be 4.8.

B-lactamase assay solutions
Iodine: 0.32N iodine, 1.2N potassium iodide, diluted 1:19 in 2M sodium acetate, pH4.2.
Penicillin: 2.4mg/ml benzylpenicillin in 0.1M phosphate buffer, pH7 (Na_2HPO_4:NaH_2PO_4 = 30.5:19.5).

2.9 Culture media.

L-Broth: 10g tryptone, 5g yeast extract, 5g NaCl, made up to 1 litre in distilled water and adjusted to pH 7.5 with NaOH. Supplemented with 2% glucose (from 20% stock solution).

L-Agar: as L-Broth with the addition of 15g/l agar.

Iso-sensitest Broth: 23.4g Iso-sensitest Broth made up to 1 litre in distilled water.

Iso-sensitest Agar: 31.4g Isosensitest Agar made up to 1 litre in distilled water.

MacConkey agar: 40g MacConkey agar base (Difco) made up to 1 litre in distilled water, supplemented with 1% relevant carbon source (lactose, galactose).

Minimal agar: 25ml D&M salts, 75ml 2% agar in distilled water; supplemented with 2% glucose, 20ug/ml thiamine (vitamin B1), and amino acids as required for each strain (at 40ug/ml, or 1% casamino acids).

4x Davis and Mingioli minimal salts (D & M salts): 28g K_2HPO_4, 8g KH_2PO_4, 1g sodium citrate, 0.4g $MgSO_4H_2O$, made up to 1 litre in distilled water.

Phage buffer: 7g Na_2HPO_4, 3g KH_2PO_4, 5g NaCl, 0.25g $MgSO_4$, 15mg $CaCl_2.2H_2O$, 1ml 1% gelatin, made up to 1 litre in distilled water.

2.10 **Sterilisation.** All growth media were sterilised at 120°C for 15 minutes; supplements and buffer solutions at 108°C for 10 minutes and $CaCl_2$ at 114°C for 10 minutes.

2.11 **Indicators.**

A) X-gal (5-bromo-4-chloro-3-indolyl-B-galactoside) was used in conjunction with the host strain DS941 and the pUC vectors, providing a screen for plasmids with inserts in the polylinker. Clones containing inserts were generally white; clones lacking inserts were blue. X-gal (40mg/ml in DMF) was stored at -20°C and added to L-agar to a final concentration of 20ug/ml.

B) **Crystal violet** was used to screen for chloramphenicol sensitivity; colonies which were sensitive grew white, and those which were resistant grew purple.

2.12 **Antibiotics.** The antibiotic concentrations used throughout for both liquid and plate selection were as follows:

Antibiotic	Source	Stock solution	Selective
Ampicillin (Ap)	p	5mg/ml (water)	50ug/ml
Tetracycline (Tc)	p	1mg/ml (10mM HCl)	10ug/ml
Chloramphenicol (Cm)	p	2.5mg/ml (ethanol)	25ug/ml
Kanamycin (Km)	p	5mg/ml (water)	50ug/ml
Streptomycin (Sm)	c	10mg/ml (water)	100ug/ml
Streptomycin	Tn7	0.5mg/ml (water)	5ug/ml
Trimethoprim (Tp)	p/Tn7	5mg/ml(50% ethanol)	50 ug/ml
Spectinomycin (Sp)	Tn7	2.5mg/ml (water)	25ug/ml
Rifampicin (Rif)	c	5mg/ml (methanol)	50ug/ml
Naladixic acid (Nal)	c	2mg/ml (1M NaOH)	20ug/ml*
Valine (Val)	c	4mg/ml (water)	40ug/ml

c: chromosome p: plasmid

*: for mutation rate assays, naladixic acid was used at 40ug/ml.

All stock solutions were stored at 4°C.

Antibiotics were added to molten agar which was precooled to 55°C.

2.13 **Growth conditions.** Liquid culture for transformation, DNA preparations or <u>in vivo</u> transposition assays were routinely grown in L-broth at 37°C with vigorous shaking. Growth on both L-agar and minimal plates were used. Antibiotics were used as required. Plates were generally incubated overnight at 37°C, or 30°C for MacConkey plates. Bacterial strains were stored in 50% L-broth, 20% glycerol and 1% peptone at -20°C. -70°

2.14 **Plasmid DNA isolation.**

Large scale DNA preparation (Birnboim and Doly, 1976; as modified in this laboratory).

200ml cultures of stationary phase plasmid-containing cells were harvested by centrifugation (12,400g, 10 min, 4°C). The pellet was resuspended in 4ml of solution 1. 8ml of solution 2 was added and the solution left on ice for a further 5 min. 6ml of solution 3 was then added, gently mixed and the cell debris and chromosomal DNA removed by centrifugation (39,200g, 30 min, 4°C). The plasmid DNA was precipitated from the supernatant with 12ml isopropanol for 15 minutes at room temperature. The DNA was pelleted at 27,200g for 15 min at 20°C and was further purified by banding on a CsCl/EtBr gradient. The DNA was resuspended in 2ml of TE buffer and added to 100-270ul of a 15mg/ml ethidium bromide solution. 5g of CsCl were dissolved in 3ml of water and added to the DNA/EtBr solution. The gradients were centrifuged in a Beckman Ti70 fixed angle rotor at 200,000g for 16 hours at 25°C. Where two bands were visible, a lower supercoiled plasmid band and an upper nicked DNA band; the lower band was removed using a 1ml syringe. The ethidium bromide was removed by repeated butanol extractions and the DNA was precipitated with ethanol and resuspended in 1x TE. The DNA was then ready for use.

Small scale preparations: method of Holmes and Quigley (1981). RNaseA was added to restriction digests of this DNA, to 50ug/ml, before loading on a gel. Alternatively RNaseA was added to the gel staining solution.

2.15 **Bacteriophage lambda DNA isolation.**

Large scale preparations followed the method given in Sambrook et al. (1989), except that phage particles were banded by equilibrium centrifugation in 0.71g/ml CsCl.

Small scale preparations were made from plate lysates (Sambrook et al.,1989). DNaseI and RNaseA were added to 4ml phage, to a final concentration of 1ug/ml, and incubated for 30 minutes at room temperature. 4ml of a 20% PEG/2.5M NaCl mix was added, and phage particles

precipitated, on ice, for 30-60 minutes. Phage were pelletted by centrifugation (27 200 g, 15 minutes, 4°C), resuspended in 0.5ml phage buffer, and incubated with SDS (0.1%) and EDTA (5mM) at 70°C for 15 minutes. The solution was extracted with phenol, phenol/chloroform, and chloroform, and the DNA was precipitated with isopropanol and resuspended in 50ul TE.

2.16 Ethanol precipitation of DNA.

2.17 Restriction of DNA.

2.18 Filling-in of restriction enzyme-cleaved termini.

2.19 Calf intestinal phosphatase (CIP) treatment.

2.20 Ligation of DNA fragments. These routine DNA manipulations were carried out by standard methods, as described in Sambrook et al., (1989)

2.21 Gel electrophoresis: agarose gels.

A) Single colony gel analysis.

Using this technique, the plasmid content of an isolate can be observed without the need to purify the DNA. A single transformant was patched out (1cm square) on a selective plate and grown overnight. Using a toothpick, a large scrape of cells was collected and resuspended in 150ul of single colony gel buffer. The cells were left to lyse at room temperature for 15 minutes. Cell debris and chromosomal DNA was spun down in an Eppendorf microfuge for 15 min at 4°C, and 50ul of the supernatant was loaded onto an agarose gel.

B) Horizontal agarose gels.

0.7-1.2% agarose gels were used.

Agarose powder was dissolved at 100°C in buffer E and precooled to 55°C prior to use.

Horizontal gels were used to analyse restriction digests, products of ligation experiments, and for single colony gel analysis.

Gels were usually run for 15-18 hours at 1.5V/cm in gel

tanks containing 3 litres buffer E, and then stained in 0.6ug/ml ethidium bromide. The DNA was visualised on a 254nm wavelength UV transilluminator.

2.22 Gel electrophoresis: polyacrylamide gels.

A) Polyacrylamide restriction gels.

A variety of acrylamide concentrations were used, depending on the sizes of the fragments of interest, as described in Sambrook et al. (1989). All gels were made using a stock solution of 30% acrylamide : 0.8% bis-acrylamide in distilled water. Vertical gel kits were used, with 1.5mm spacers. The gel apparatus was sealed with 0.6% agarose in 1xTBE, which was also the running buffer. The gels were run at room temperature in 1xTBE at a constant current (25-30 mA), for 2-3 hours. DNA bands were visualised under 254nm UV illumination after staining in 0.6 ug/ml ethidium bromide for 20 min.

B) Non-denaturing polyacrylamide gels.

These gels were used to separate protein:DNA complexes. 5% polyacrylamide gels were used; vertical gel kits were sealed with 0.6% agarose in the appropriate TE running buffer. The gels were prerun at 15V/cm for 30-90 min at 4^{o}C. After loading (with the power connected), the gels were run for 2-3 hours at 15V/cm, at 4^{o}, with the buffer being recirculated between compartments. When the gel run was complete, if labelled fragments were used, the gel was transferred to filter paper and dried under vacuum. Bands were visualised by autoradiography of a sheet of Kodak S1 film overnight. Non-radioactive gels were stained with ethidium bromide, as for polyacrylamide restriction gels.

Gel conditions

'standard' conditions: 50mM Tris/Glycine (pH 9.4), 0.1mM EDTA.

'pH 8.2' conditions: 10mM Tris/HCl (pH 8.2), 0.1mM EDTA.

C) Polyacrylamide sequencing gels.

Sequencing and footprinting reactions were analysed on 8% high resolution polyacrylamide/urea gels, as described in the 'M13 cloning/dideoxy sequencing instruction manual' published by BRL. Gels were prerun for 30 min and run for 2-3 hours at 40W. Samples were denatured prior to loading (100^{o}C, 3 min). After the gel run was complete, gels were fixed in 10% acetic acid for 30 min, dried under vacuum and autoradiographed.

2.23 **Gel electrophoresis: protein** SDS-PAGE followed the method of Laemmli (1970). Gels were stained with Coomassie Blue (0.1%, in 50% MeOH, 10% AcOH: destain 10% MeOH, 10% AcOH) or with silver (Morrissey, 1981).

2.24 **Photography of gels.** Gels were photographed using Polaroid type 67 land film or using a Pentax 35mm SRL loaded with Ilford HP5 film. Both cameras were fitted with a Kodak Wratten filter No.23A.

2.25 **Autoradiography** was carried out as described in Sambrook et al. (1989), using Kodak XS1 or Fuji RX100 Xray film.

2.26 **Extraction of DNA from agarose gels.** Gels were stained with ethidium bromide, the bands visualised on a UV transilluminator, and the band of interest excised using a scalpel. Various techniques were used to extract the DNA. Electroelution was carried out using an IBI UEA unidirectional electroelutor, following the protocol provided. 'Geneclean' (Bio101) was used according to the manufacturer's directions. The centrifugation method of Heery et al.(1990) was also used, for fragments of <1 kb.

2.27 **Extraction of DNA from polyacrylamide gels.** The band of interest was excised from a gel as above. The DNA was extracted by electroelution (see above), or by the 'crush

and soak' method of Maxam and Gilbert (1977).

2.28 Purification of synthetic oligonucleotides.

The oligonucleotide was removed from the support and deprotected according to the protocol provided by Applied Biosystems; the solution was evaporated to dryness under vacuum, and the DNA resuspended in TE. Where necessary, the full length oligonucleotide was purified (to remove shorter, prematurely terminated products) by denaturing gel electrophoresis as described in Sambrook et al. (1989).

2.29 Labelling of DNA fragments.

A) 5' end-labelling by T4 kinase

Isolated DNA fragments (0.5-10pmol ends) that had been treated with CIP were end-labelled as described in Sambrook et al. (1989), and the DNA purified by phenol extraction and ethanol precipiation.

B) 3' end-labelling by the Klenow fragment of DNA polymerase I

Appropriately restricted DNA or purified DNA fragment were end-labelled by filling in recessed 3' ends with the Klenow fragment of DNA polymerase I. The reaction contained 1-200ug/ml DNA (0.5-2 pmol of ends), 50nM unlabelled nucleotides, 10uCi alpha (^{32}P)dNTP, restriction buffer (REact2, REact4, and 0.5xREact3 were used), and 0.5-2 units Klenow enzyme. After incubation at 16^{o}C for 90 min, the reaction was stopped by phenol extraction, and the DNA was ethanol precipitated.

3) Random priming followed the protocol of Feinberg and Vogelstein (1984); the unincorporated nucleotides were removed on 1ml Sephadex G-75 or G-50 columns (Sambrook et al., 1989).

2.30 **DNase I nicking of DNA.** DNase I was diluted in 10mM Tris/HCl pH 8.0, 10mM $MgCl_2$, 50% glycerol (v/v) and stored at -20°C. DNA samples in 50 mM Tris/Cl (pH 8.2), 1 mM EDTA, 10mM $MgCl_2$, and 300 ug/ml ethidium bromide, were treated with 2 ug/ml DNase I for 30 minutes at 30°C. Reactions were quenched with EDTA (20 mM), and ethidium bromide and DNaseI were removed by extraction with phenol/chloroform or butanol; the DNA was then precipitated with ethanol.

2.31 **Transformation with plasmid DNA.** Genetic transformation introduced plasmid DNA into different host strains. A standard $CaCl_2$ protocol was used (Sambrook et al., 1989). The presence of the correct plasmid in the transformant colonies was confirmed by the single colony gel electrophoresis procedure (2.21A).

2.32 **Transduction with phage P1.** Generalised transduction was used to move alleles between strains, and followed the method of Miller (1972).

2.33 **Conjugation.** R388 and its derivatives are conjugative plasmids, and were moved between strains by plate mating according to Miller (1972). Conjugative, or 'mate-out' transposition assays were performed by the same method; exconjugants were scored for the presence of plasmid markers (trimethoprim) and transposon markers, and the proportion of products bearing the transposon marker was termed the transposition frequency.

2.34 **Phage lambda packaging, and infection of** E.coli. Phage DNA was routinely packaged using Gigapack Gold packaging extracts (Stratagene), using the manufacturers protocol for the packaging reaction and subsequent infection into E.coli.
Packaging extracts defective in mismatch repair were prepared from the strains NK7515 and NK7516 (both mutL::

Tn5), by the method described in Sambrook et al.(1989).

2.35 **Testing of UV sensitivity.** Tenfold dilution series (10^{0}-10^{-6}) were prepared from late log phase cultures, and 10ul spotted onto agar plates. The plates were exposed to UV irradiation (17.5 ergs/sec/m^2), for a known period of time, and incubated overnight in the dark.

2.36 **Mutation rate assays.** The rates of mutation of strains to resistance to rifampicin, naladixic acid, and valine, were determined as described by Glickman and Radman (1980).

2.37 **Protein quantitation.**

A) Bradford's assay: as described (Bradford, 1976).

B) Bicinchoninic acid (BCA) assay: as described by Smith et al.(1985). Initial protein assays on crude extracts were done with Bradford's reagent. However, TnsB was insoluble in the acidic conditions of the Bradford assay, and tended to form aggregates; protein estimates during the purification were carried out by the BCA assay.

2.38 **Galactokinase assays** were carried out as described by McKenney et al.(1981). The results were corrected for plasmid copy number by assaying B-lactamase activity. 0.5ml of the cell lysate used for GalK determination was added to 2.5ml penicillin solution, and incubated at 30°C for 10-50 minutes. 5ml iodine/acetate was added, and incubation continued for a further 10 minutes. The absorbance at 499 nm was read, compared to a control (culture added after the iodine). Activity in units/ml/hour were calculated as:

$$\frac{\Delta A_{499} \times \text{assay volume(ml)} \times \text{time(hours)}}{\text{sample volume(ml)}}$$

and could be converted to specific activity by dividing by the A_{650} of the culture.

2.39 **Purification of TnsB.** TnsB was made from the overproducing strain DS941/pMR207. A small (2.5 ml) overnight culture was inoculated into 400ml isosensitest broth plus trimethoprim, in a 2 litre flask, and the culture grown at 37°C, with shaking, until an A_{600} of 0.5 was reached (3-4 hours). tnsB expression was induced by the addition of an equal volume of the same medium, prewarmed to 37°C, made 1mM in IPTG, and growth was continued for 2-3 hours. The cells were collected by centrifugation (12,400 g, 10 minutes, 4°C), and washed in extraction buffer. After the second spin, the cells were resuspended in 1.5-2 ml extraction buffer per gram wet weight of cells.

Cells were broken open in a French press (2 x 950 psi), and the lysate cleared by centrifugation (39,200 g, 30 minutes, 4°C). The pellet was washed several times in the same buffer, by thorough resuspension and respinning, until no further protein eluted into the supernatant. The pellet was then resuspended in TnsB buffer (around 700ul per 400 ml culture), the solution left on ice for 30 minutes to ensure complete resuspension, and repelletted. TnsB should be in the supernatant, and can be concentrated by reducing the salt concentration; this can be achieved by overnight dialysis against extraction buffer, or by dilution of the solution with extraction buffer with no KCl, to return the salt concentration to 200mM. TnsB was recovered as a precipitate, and was resuspended in 0.25-1 ml TnsB buffer; this solution consists of around 75% TnsB, and can be stored in 50% glycerol at -20°C.

The TnsB buffer was replaced by column buffer, by dilution or dialysis, and the sample loaded onto a Superose 6 gel filtration column (Pharmacia). The column was run on a Waters 650E FPLC system, at a flow rate of 0.5ml/minute. 0.5 ml samples were collected. TnsB eluted at an elution volume of 15.6 ml, and the relevant fractions were stored in 50% glycerol at -20°C.

Buffers

Extraction buffer: 150mM Tris/HCl (pH8.2), 200mM KCl, 1mM EDTA, 1mM PMSF, 1mM DTT, 1mM benzamidine. The last three were added fresh to a stock solution, which was stored at $4^{o}C$.

TnsB buffer: as extraction buffer, but made 1.3M in KCl.

TnsB storage buffer: 50% TnsB buffer, 50% glycerol.

Column buffer: 50mM Tris/HCl (pH8.2), 675mM KCl, 1mM EDTA, 0.5mM PMSF, 0.5mM DTT, 0.5mM benzamidine.

2.40 **Gel binding assays.** 1 ul protein samples were added to 0.2-1 ng end-labelled DNA fragment containing the binding site of interest, and 30-50 ng pUC18 carrier, in a volume of 10 ul. The reactions were incubated for 10-15 minutes, at room temperature, then quenched by placing the samples on ice. The samples were loaded immediately onto a 5% polyacrylamide gel, running at 100V. Control incubations were set up as above, but with 1ul TnsB dilution buffer, ie all samples contained the same salt and glycerol concentrations.

Binding conditions (final)

'standard' conditions: 50mM Tris/Glycine (pH9.4), 1mM EDTA, 68mM KCl, 10% (v/v) glycerol.

'pH 8.2' conditions: as above, but 50mM Tris/HCl (pH8.2).

TnsB was diluted into storage buffer.

FIS conditions: 50mM Tris/Glycine (pH9.4), 0.1mM EDTA, 60mM NaCl, 10% (v/v) glycerol. FIS was diluted in 50mM Tris/HCl pH8.2, 1M NaCl, 1mM EDTA.

IHF conditions: as standard or pH 8.2 conditions, but 50mM NaCl instead of KCl. IHF was diluted in 50mM Tris/HCl (pH8.2), 500mM 1mM NaCl, 1mM EDTA, 50% glycerol.

2.41 **Footprinting.** DNaseI footprinting was carried out as described by Galas and Schmitz (1979). Reactions were carried out in a volume of 40ul, and contained 25ug/ml carrier DNA (sheared salmon sperm) and 10-20ng labelled fragment. 4 ul of protein dilution was added. After incubation, the reactions were treated with 4 units FPLC-purified DNase I(Pharmacia) for 1 minute, the cleavage was stopped with excess EDTA, and the reactions were extracted with phenol and precipitated with ethanol before resuspension in formamide loading buffer for electrophoresis.
Hydroxyl radical footprinting was as described by Tullius et al.(1987).
Binding reaction conditions: 50mM Tris/HCl pH8.2, 1mM EDTA, 67mM KCl, with 5mM $MgCl_2$ and 10% glycerol for DNaseI reactions. Reactions were incubated for 15 minutes at 37^{o}C (DNase I) or room temperature (hydroxyl radical) prior to cleavage.
For both types of footprint, in order to reduce degradation of the DNA, it was stored in ethanol or as a dried pellet after the experiment, and resuspended in loading buffer immediately before electrophoresis.
Sequence standards were generated by base-specific cleavage reactions (Maxam and Gilbert, 1977), using the protocol published by I.B.I. (1987 catalogue).

2.42 **Experiments with DNA ligase.** Binding reactions were carried out in 20ul (nicked substrates) or 100ul (linear substrates), as for gel binding assays (2.40). Reaction conditions were as pH 8.2 conditions for gel binding assays, with the addition of 10mM $MgCl_2$, 1mM DTT, and 0.5-1mM ATP, and 50mM NaCl for reactions with linear substrates. 1ul of protein dilution was added. After 15 minutes incubation at room temperature, DNA ligase was added (0.5 units per nicked substrate reaction, 2.5 units per linear substrate reaction) and incubation continued for 30-50 minutes. Reactions were stopped by adding SDS or

protease K, or were precipitated with ethanol or extracted with butanol where the volume needed to be reduced.

2.43 **Protein N-terminal sequencing.** A TnsB sample was run on Tris/Tricine SDS-PAGE (5% stacking gel, 10% running gel), and the protein was electroblotted onto Problott membrane (Applied Biosystems), according to the manufacturer's instructions. The 85 kDa band was located by Coomassie staining, and the band excised. The sequence was determined on an Applied Biosystems 477A automated protein sequencer (Department of Molecular Palaeontology, University of Glasgow).

2.44 **South-Western blotting** was carried out as described by Miskimins et al.(1985) with slight modifications; the filters were blocked using 0.02% boiled BSA instead of Blotto (Roth et al., 1988), and with 100ug/ml sheared salmon sperm DNA. Binding reactions were carried out overnight at 4^{o}C, or for one hour at room temperature, and the probed blots were washed for 15 minutes in binding buffer containing 50mM NaCl (0.3M NaCl for resolvase filters).

CHAPTER 3

PURIFICATION OF TnsB

3.1 INTRODUCTION

In order to study the biochemistry of Tn7 transposition in detail, we would like to be able to carry out transposition reactions in a cell-free system, using purified components. There are two approaches to this. The components can be purified independently and reconstituted into an active system. Alternatively an *in vitro* reaction can be achieved in a crude cell extract, and proteins etc. not required can be purified away from this. The alternative approaches are compared in chapter 8. We have taken the former approach, which has the advantage that each of the transposition functions can be characterised independently during the purification stage of the work.

The product of the *tnsB* gene was a prime candidate for purification, for two reasons. Firstly, the polypeptide had previously been observed on gels, which was not the case for other Tn7 proteins (eg TnsA) when this project was started. Brevet et al. (1985) used strains containing fragments of Tn7 cloned into plasmids, labelled the cells with ^{35}S-methionine, and looked at whole cell lysates on SDS-PAGE. A protein of 85 kDa was observed in cells containing the 2.5 kb BglII (13.1 kb) to HindIII (10.6 kb) fragment. This result is consistent with the complementation data (Rogers, 1986; Rogers et al., 1986) and with DNA sequencing (Flores et al., 1990), which define the BglII-HindIII region as containing *tnsB*. Mark Rogers (1986) used minicells (Adler et al., 1967) to identify proteins coded by subclones of Tn7 carried on plasmids; again an 85 kDa protein was found to be encoded by the same region of the transposon. A second band, corresponding to a 57 kDa protein, was observed only in the presence of both the *tnsB* and *tnsE* genes, and this may be a *tnsE* dependent shortened form of TnsB (or *vice versa*).

More recently, Orle and Craig (in press) used antibodies to Tn7 proteins to probe Western blots of cell

extracts. For cells containing tnsB, a protein with an apparent molecular weight of around 85 kDa is detected, as well as a series of smaller polypeptides. They suggested that these are degradation products of TnsB.

Secondly, an assayable activity was known for TnsB. Work by Ekaterinaki (1987) demonstrated tnsB-dependent binding to the right end of Tn7, both in vivo (tnsB-dependent repression of P1 in promoter probe vectors) and in vitro, using gel retardation assays (Fried and Crothers, 1981; Garner and Revzin, 1981). Crude extracts of cells containing plasmids with cloned tnsB, on incubation with a labelled Tn7 end fragment, caused retardation of that fragment in non-denaturing gel electrophoresis, and gave rise to a series of retarded complexes. McKown et al. (1987) have also reported such a tnsB-dependent DNA binding activity in crude cell extracts. For use in the purification of TnsB, the gel shift assay has been developed and improved (see section 4.5), so that estimates to allow some quantitation of the binding activity can be made. The assay also allows us to observe non-specific binding to control DNA fragments, and so assess the extent of contamination by other DNA-binding proteins during the purification.

This chapter describes the overexpression of tnsB in E.coli, and subsequent purification of the protein. The DNA binding activity described above was used as an assay to follow the purification, and experiments to determine which polypeptides are responsible for the binding activity were carried out. The results suggest that the complexity of the sequences at the Tn7 termini is reflected in multiple interactions between the DNA and TnsB, and probably some host-encoded proteins.

RESULTS AND DISCUSSION

3.2 Overexpression of TnsB

The overexpressing plasmid used to produce TnsB is

pMR207; this is based on a pBAD-derived P_{tac} expression vector pMR78, which has a cer site inserted into it to increase plasmid stability (Summers and Sherratt 1984); the beta-lactamase gene has been disrupted by insertion of the dihydrofolate reductase gene (trimethoprim resistance) from R388. The plasmid carries the pUC8 polylinker. To construct pMR207, the 2.5 kb BglII (13.1 kb) to HindIII (10.6 kb) Tn7 fragment containing the complete tnsB gene was cloned into pMR78 cut with EcoRI and HindIII (table 2.2; M. Rogers, personal communication).

The insert in pMR207 includes the C-terminal 18 amino acids of TnsA; it also includes the -10 region of the proposed promoter P2 (Rogers et al., 1986) but in transcriptional fusions these sequences did not promote reporter gene expression, probably due to the deletion of the -35 region (Ekaterinaki, 1987). The construct also includes the N-terminal 132 amino acids of TnsC, if tnsC translation begins at the first ATG codon in the open reading frame. However, this clone does not complement tnsA or tnsC defects.

pMR207 differs from the expression clone used by Ekaterinaki (1987) in that the protein produced lacked 7 amino acids from the C terminus of TnsB; this deletion does not affect the protein's ability to mediate transposition (Rogers, 1986).

Preparations of the protein were made in strain DS941 containing pMR207; this strain contains a chromosomal $lacI^q$, and so is able to repress the tac promoter to some extent. Control preparations were made from DS941 containing pMR78.

3.3 Binding activity in crude extracts of DS941/pMR207

Crude cell extracts were prepared from DS941/pMR207, and were assayed for binding to Tn7 RE sequences by native gel electrophoresis (figure 3.1). The DNA used is pNE200 (table 2.2, figure 4.1), which carries 205 bp of the right

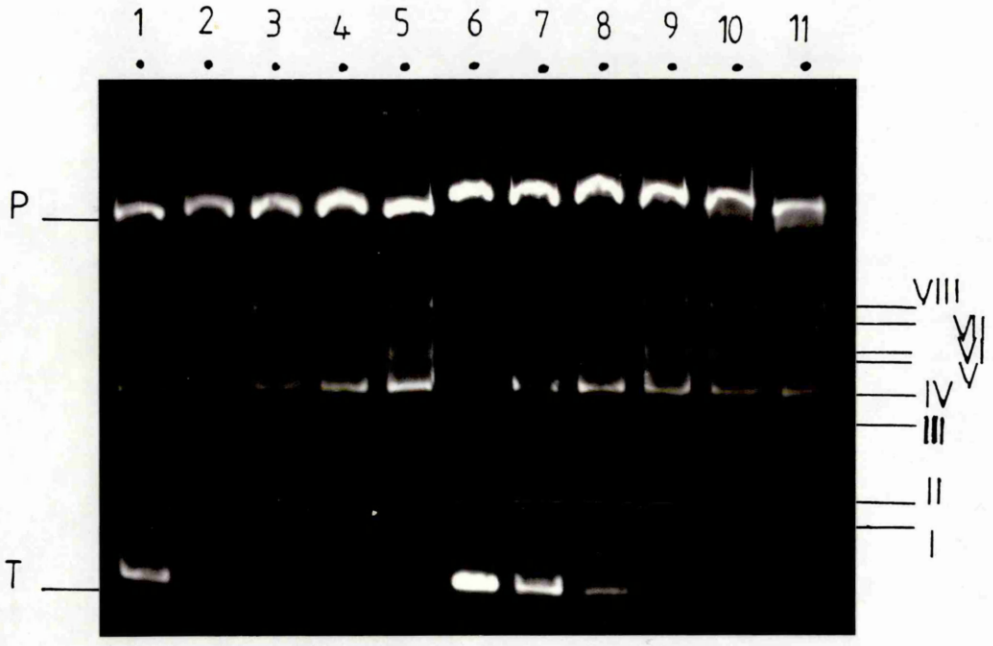


Figure 3.1 DNA-binding activity in crude extracts of DS941/pMR207

10 ng pNE200 (cut with EcoRI and HindIII) was incubated with 2 ug crude extract for the time indicated (lanes 1-5) or for 15 minutes with the indicated amount of extract (lanes 7-11), at room temperature, and the reactions run on a 5% non-denaturing polyacrylamide gel.

Binding conditions: 50mM Tris/HCl (pH8.2), 1mM EDTA, 20mM KCl, 10% glycerol, 150ug/ml sheared salmon sperm 'carrier' DNA.

Gel conditions: 10mM Tris/HCl (pH8.2), 1mM EDTA; 4^{o}C.

lane 1	15 sec	lane 7	1 ug
2	3 min	8	1.5 ug
3	6 min	9	2 ug
4	9 min	10	3 ug
5	12 min	11	4 ug
6	no extract		

Complexes are indicated I-VIII.
T fragment containing Tn7 right end
P pUC8 vector fragment

The DNA substrates used in this and all subsequent binding experiments are end-labelled, and contain the right-hand end (205 bp) of Tn7, on a 235 EcoRI-HindIII fragment.

end of Tn7 on an EcoRI-HindIII fragment; the pUC8 vector fragment remains in the assay as an internal control for non-specific binding.

The pattern of complexes seen in figure 3.1 was slightly different from those reported earlier (Ekaterinaki, 1987). This could be due to the slight C-terminal deletion in previous preparations, or changes in reaction conditions. Eight complexes were distinguishable, with IV being predominant. As the protein concentration was increased, more strongly retarded complexes were seen (VII, VIII), and the amount of free DNA gradually decreased until >90% was in the bound form; at this level of protein there was also some binding to the pUC fragment, and a smear of non-specific retardation of the Tn7 fragment. Control extracts from DS941/pMR78 showed no specific DNA binding activity (data not shown).

3.4 Optimisation of TnsB expression

By SDS-PAGE, a band corresponding to the predicted molecular weight of TnsB (81 kD; Flores et al., 1990) was seen in cell lysates of DS941/pMR207 on induction with IPTG (figure 3.2 lanes 1-4, 8-10); this band was absent from the control strain (lane 7). However, initial experiments gave variable levels of induction. To reduce this variability, some of the factors potentially responsible were investigated.

One factor is the growth stage at which a growing culture is induced, and the length of time of induction. DS941/pMR207 cells inoculated straight into 1mM IPTG failed to grow; IPTG was therefore added once the culture was growing. To ascertain the optimal time of induction, samples were taken from growing cultures of DS941/pMR207 at various growth stages, induced with IPTG, and protein expression monitored by SDS gel electrophoresis (figure 3.2). Growth rates were also determined (figure 3.3).

The growth curve of the main cultures (figure 3.3A) shows that the presence of either plasmid caused a slight

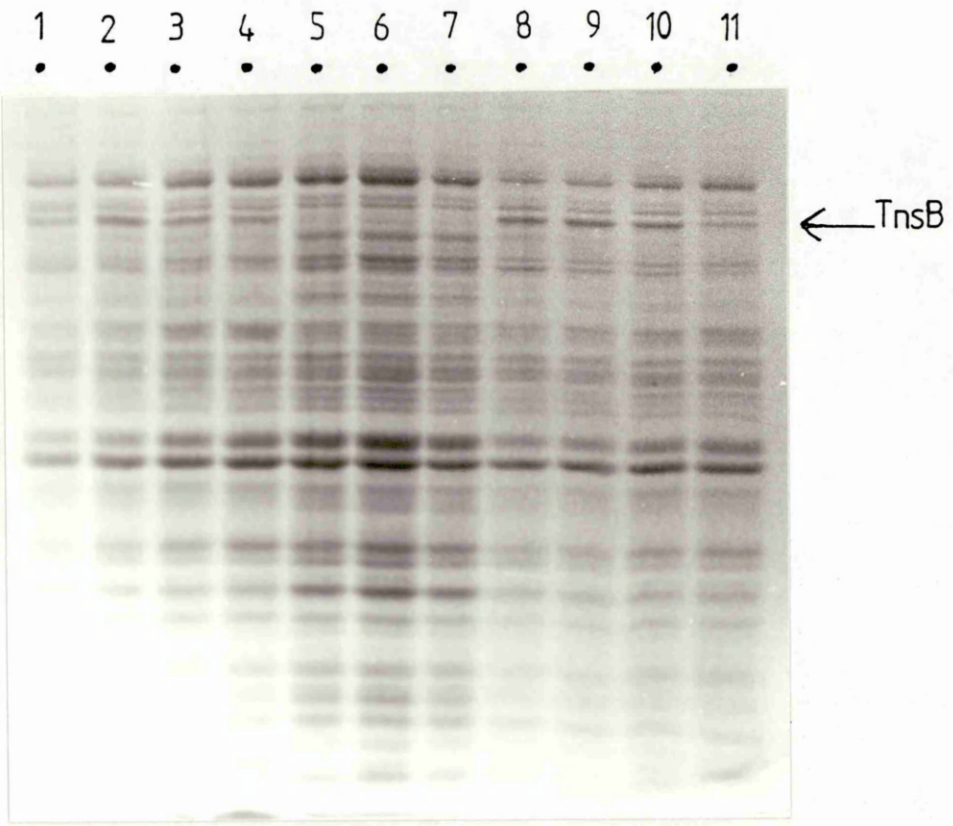


Figure 3.2 Induction of tnsB expression
DS941 and DS941/pMR207 were grown in isosensitest broth (with 50ug/ml trimethoprim for plasmid selection). 5ml samples were taken when the A_{600} reached 0.4 and 0.7, and induced by adding 5ml prewarmed isosensitest broth made 1mM in IPTG; samples were withdrawn at the times indicated and the cells loaded onto an SDS-polyacrylamide gel after boiling with SDS and 2-mercaptoethanol.

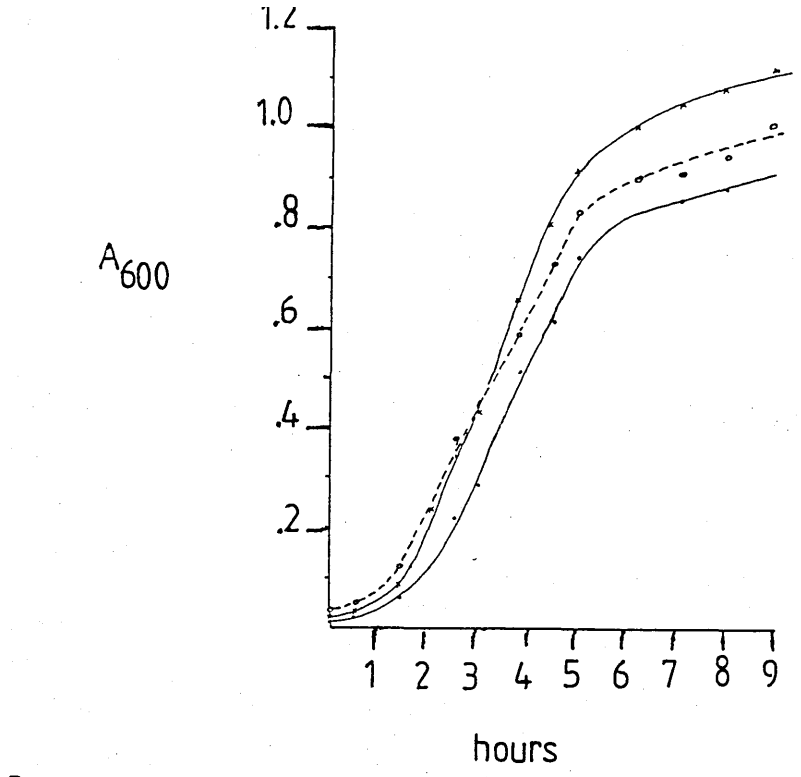
lane 1-4 DS941/pMR207: induced at A_{600}= 0.69, for 1, 2, 3, 5 hours
5 DS941/pMR207: uninduced
6,7 DS941 : uninduced, induced
8-11 DS941/pMR207: induced at A_{600}= 0.4, for 1, 2, 3, 5 hours

reduction in the growth rate, which might be due to the plasmid itself, or to the presence of trimethoprim. The difference in growth rates was slight, and it became no more marked on the addition of IPTG (figure 3.3B). The lack of effect is surprising, as presumably much of the transcriptional and translational machinery of the cell is sequestered in expressing the induced plasmid gene. There was no difference between cells containing pMR78 and pMR207; the growth curves superimpose, indicating that overproduction of TnsB is not toxic to the cells.

Protein profiles for various expression regimes show that production of TnsB was greater from induction at lower cell densities (OD_{600} 0.4, 0.6) than at later stages in growth (OD_{600}= 0.8, 1.0). As the growth rate decreases and cells enter stationary phase, there are many changes in cell physiology; for example, the level of negative supercoiling of the DNA decreases (Balke and Gralla, 1987; Dorman et al.,1988), and synthesis of stable RNA is reduced; the amounts of many proteins change with growth conditions (Pedersen et al.,1978), as do levels of toxins and other secondary metabolites (Connell et al.,1987). It is likely that such changes in physiology and gene expression are responsible for the observed reduction in the response to IPTG induction.

Prolonged induction failed to increase the amount of TnsB produced, as the cells appear to reach a maximum internal level of TnsB that can be tolerated. 2-3 hours was chosen as the optimal length of time for induction. As a compromise between efficiency of TnsB production and total biomass, OD_{600}=0.5 was selected as the point at which the culture was induced.

Another significant factor was aeration. Production of full length TnsB (as observed on protein gels) was poor under conditions of good aeration. Optimal expression was seen with growth of 400 ml cultures in 2 litre flasks with moderate shaking, i.e. conditions of largely anaerobic growth. Indeed, under conditions of high aeration, smaller



B.

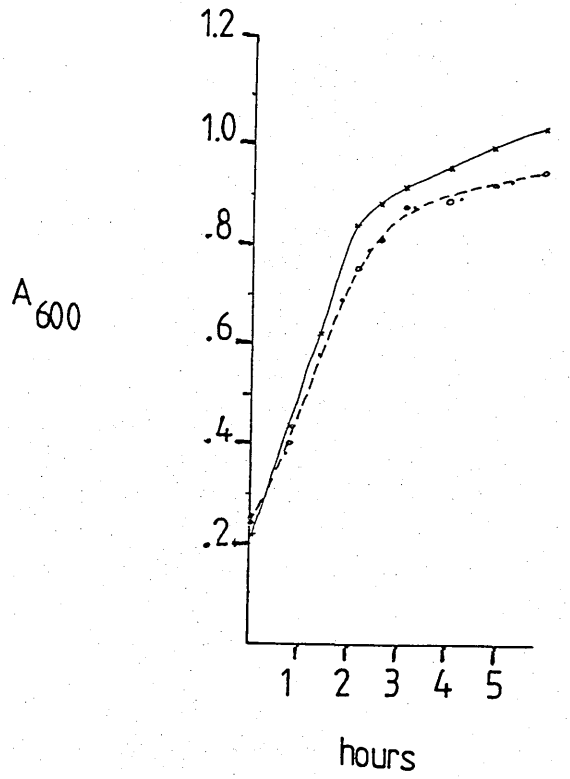


Figure 3.3 Growth curves of DS941 expression strains
A. Growth of uninduced cultures.
100ml cultures were grown in isosensitest broth (with 50 ug/ml trimethoprim for plasmid selection) at 37°C, with shaking. 1ml samples were taken at the times indicated, and the absorbance at 600 nm read.
B. Growth of induced cultures.
Cultures were induced when the A_{600} reached 0.4, by addition of an equal volume of prewarmed isosensitest broth made 1mM in IPTG, and growth was monitored as above.

For clarity, the graphs are plotted with a linear vertical axis.

X DS941 ○ DS941/pMR78 ● DS941/pMR207

peptides were seen on induction with IPTG (data not shown; A. Gawthrop, personal communication). It is not known whether these are synthesised de novo, or are degradation products of a longer polypeptide.

Plasmid supercoiling is known to increase under aerobic growth conditions (Dorman et al.,1988), and it is likely that this is responsible for the effect on TnsB production. To establish a causal relationship between supercoiling and tnsB expression, it would be interesting to see whether other conditions which alter supercoiling have the same effect on TnsB production; for example, growth in conditions of high osmolarity, which also leads to an increase in supercoiling (Higgins et al.,1988).

3.5 Purification of TnsB

3.5.1 Purification by differential solubility

Having improved the expression of TnsB from DS941/pMR207, the protein was produced at a sufficiently high level (5% of total cell protein; figure 3.5) to attempt purification. The process was facilitated by having two complementary assays for the protein; visualisation on SDS-PAGE, and the gel retardation assay.

The initial purification step is based on the insolubility of the tnsB gene product in DS941/pMR207 extracts (see figure 3.4). Under conditions of moderately high salt (200 mM KCl), cells were broken open using a French press and the resultant lysate cleared by centrifugation (I). A tnsB dependent polypeptide, with a subunit molecular weight of around 85 kDa, was found predominantly in the pellet, along with cell membranes, membrane proteins, and chromosomal DNA. This protein was solubilised from the pellet with 1.35 M KCl (II; figure 3.5 lane 5), and could be concentrated by reprecipitation on reducing the salt to 0.2 M, by dilution or dialysis (IV; lane 6). The reason for the apparently low yield of fraction IV in the gel shown was that the protein was not fully resuspended when the sample was taken.

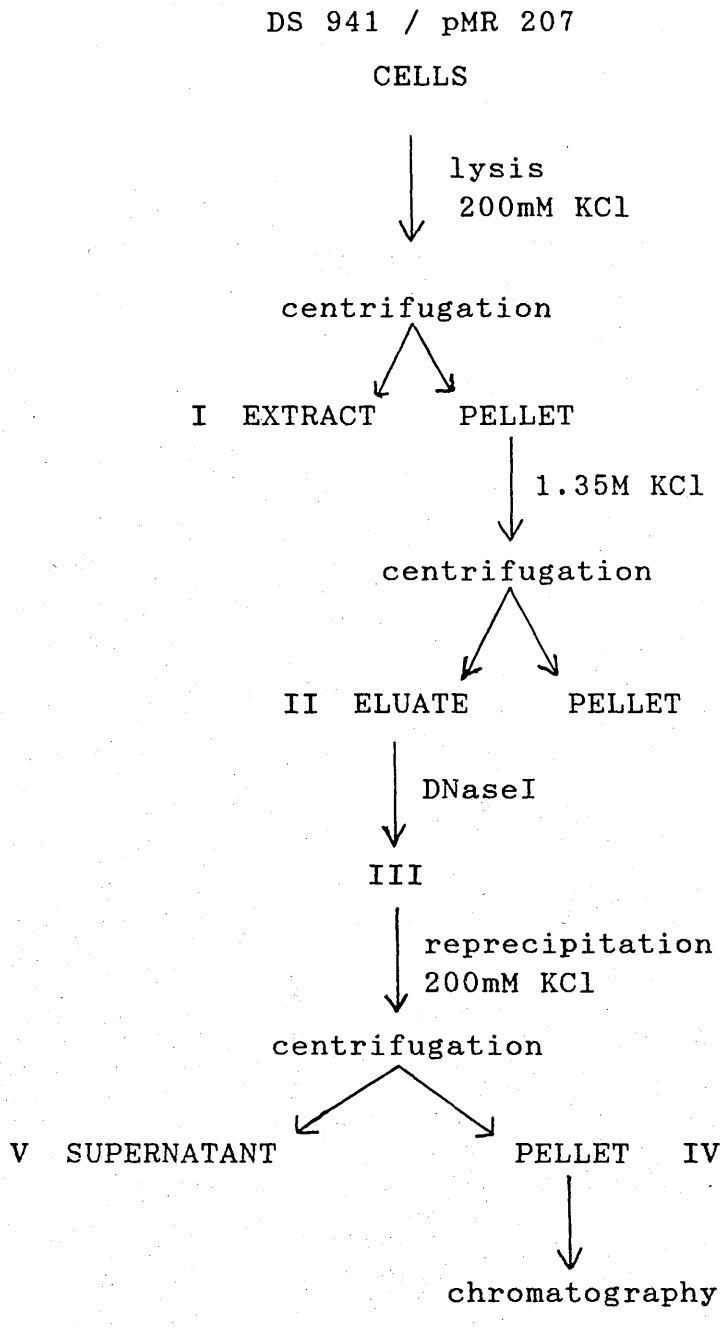


Figure 3.4 Purification of TnsB: flow chart
The purification is described in the text. Numbers I-V are used to designate fractions in figures 3.4 and 3.6.

Much of the DNA binding activity was found in the fractions which contain the 85 kDa protein (II, III, IV; figure 3.6). In these assays, a 216 bp NdeI-HindIII fragment of pUC18 was present to act as an internal control for non-specific binding. The amount of non-specific binding decreased from a high level in the cleared extract, to a much lower level in fractions II, III and IV.

3.5.2 Evidence that the protein is TnsB

The evidence at this stage suggesting that the 85 kDa protein was TnsB is as follows:

1. its expression was induced by IPTG, suggesting that it is under the control of the tac promoter (figure 3.2);
2. it was specific to pMR207-containing cells, and was absent from the control strain carrying pMR78 (figure 3.5);
3. Tn7 end-sequence binding activity copurified with the protein (fractions II, III, IV; figure 3.6).

Extracts from the control strain containing pMR78 showed much less Tn7 end binding activity (figure 3.6); some extract-dependent bands were seen in some gels, but these are likely to be due to the presence of contaminating host-derived DNA binding proteins in these preparations. These complexes appeared only at high protein concentrations equivalent to those giving non-specific binding for pMR207 extracts.

Whilst it has not been demonstrated that the polypeptide described is TnsB, it has several of the properties expected of TnsB; in the interest of brevity, the 85 kDa protein will be referred to as 'TnsB' in this section.

3.5.3 Contaminating species in this preparation

The membrane eluate (fraction IV) comprised predominantly TnsB, as judged by Coomassie staining of gels. Silver staining has greater sensitivity and a different pattern of

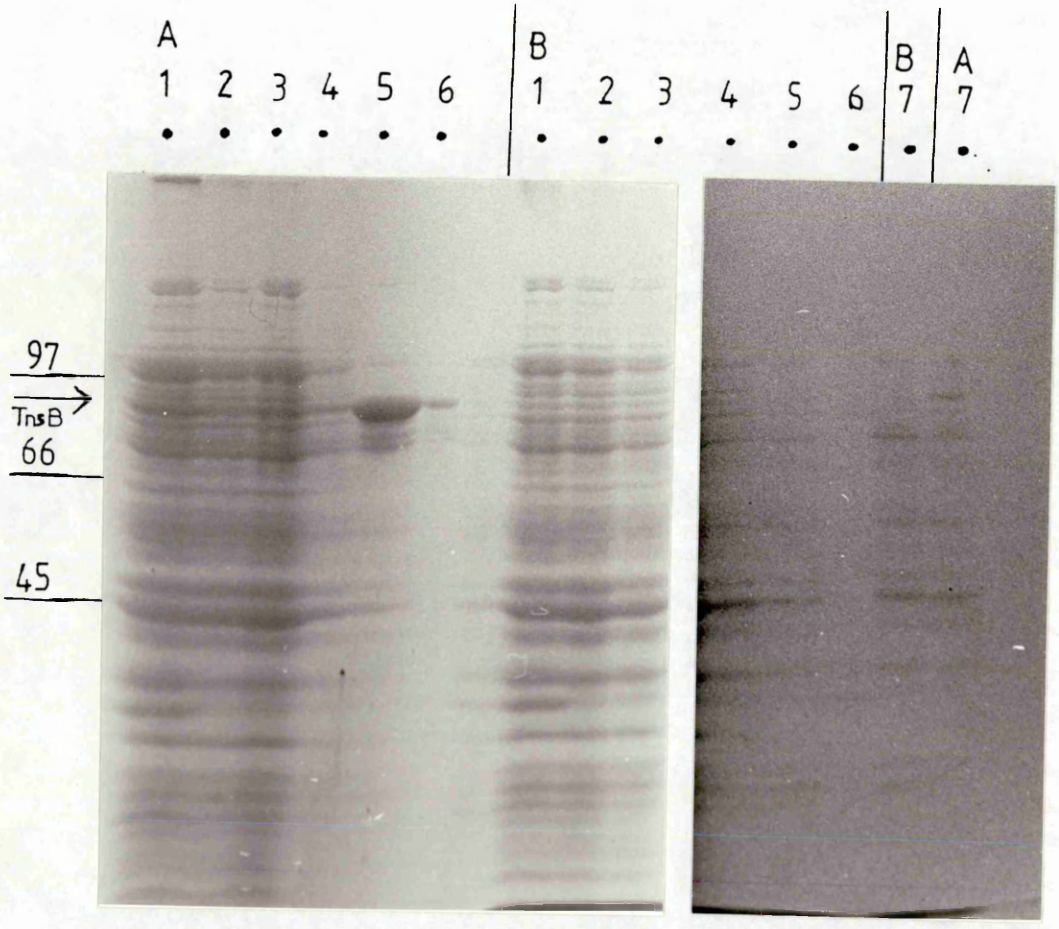


Figure 3.5 Purification of TnsB
DS941/pMR207 was grown and induced, and the cells harvested as described in Materials and Methods, and fractionated as described in the text (see figure 3.4 for fraction numbers).

		% of total
1	whole cells: induced	
2	cleared extract (I)	0.1
3,4	washes of membrane pellet	1
5	pellet eluate (II)	1
6	low salt precipitate from II (IV)	1
7	low salt supernatant (V)	1

A. DS941/pMR207 B. DS941/pMR78

Molecular weight markers are indicated in kDa

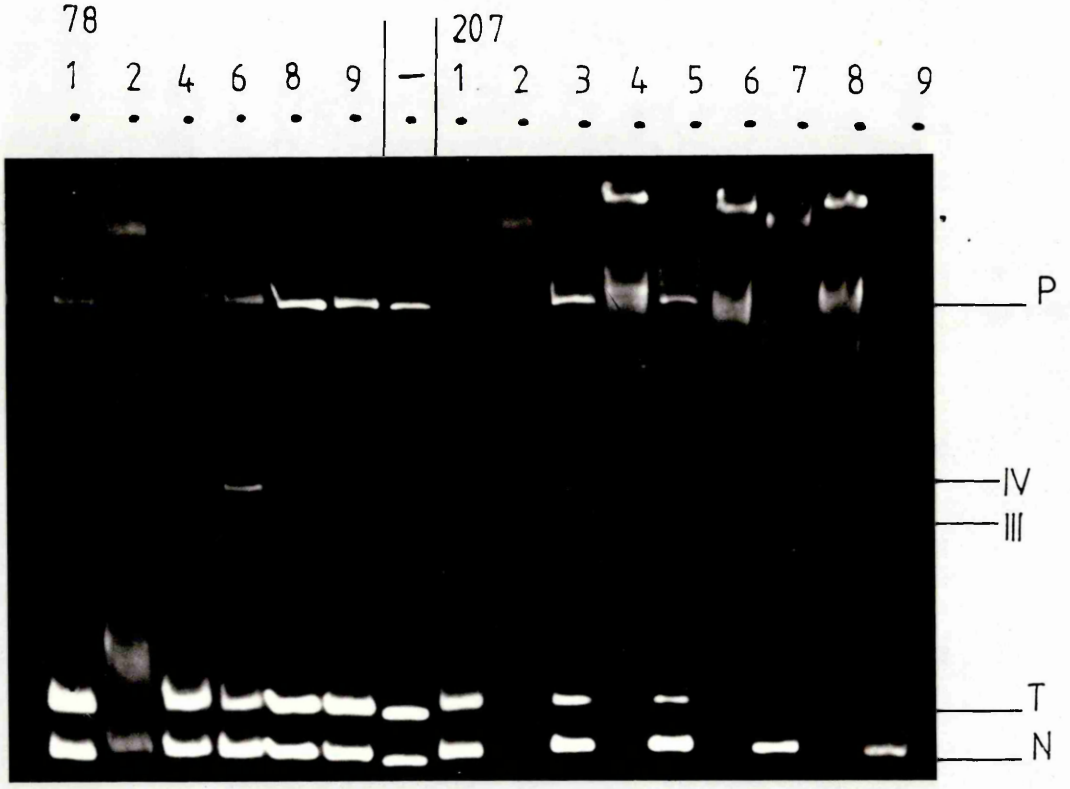


Figure 3.6 DNA-binding activity in TnsB fractions
5ng labelled pNE200 was incubated with the indicated fraction, for 15 minutes at room temperature.
Binding conditions: 50mM Tris/Glycine (pH9.4), 1mM EDTA, 68mM KCl, 10% glycerol, 5ug/ml pUC18.
Gel conditions: 50mM Tris/Glycine (pH9.4), 0.1 mM EDTA.

lane 1,2 extract (I): 2^{-4}, 2^{-1}
3,4 pellet eluate (II): 2^{-3}, 2^{0}
5,6 DNaseI-treated eluate (III): 2^{-3}, 2^{0}
7,8 low salt reprecipitate (IV): 2^{-4}, 2^{-1}
9 low salt supernatant (V): 2^{0}

78 DS941/pMR78 extracts 207 DS941/pMR207 extracts
T Tn7 right end fragment N NdeI-HindIII fragment
P pUC8 vector fragment

The pairs of II, III, and IV reactions contain equivalent amounts of the total protein, given the concentration step between III and IV. For comparison, the proportion of total protein used in lanes 2 (2^{-4}x frn.I) is equivalent to that in lanes 3, 5, and 7.

2^{-1}, 2^{-3}, 2^{-4} refer to two-fold dilutions.

staining, and brought up more, smaller species; TnsB remained about 50% of the total protein. Hence the insolubility of TnsB gave an extremely high level of purification. This was judged to be sufficient for characterisation of the binding activity, and much of the work presented in chapters 4 and 5 used preparations of this or similar purity.

The major contaminating species in these preparations was DNA, visualised by ethidium bromide staining of protein gels. The DNA appeared to follow TnsB through the fractionation; in control DS941/pMR78 preparations, the concentration of DNA eluted from the membrane pellet was much lower. Indeed, the protocol for purifying TnsB used by the Craig group includes use of an affinity column loaded with non-specific DNA (L.Arciszewska, personal communication). Precipitation of DNA and associated proteins is a common feature in purification protocols for many DNA binding proteins, eg Tn3 resolvase (Stark et al., 1989).

In this context it is interesting to note that the total amount of protein precipitated in fraction IV is much higher for $TnsB^+$ preparations; the contaminating bands are not visible in the equivalent pMR78 tracks. The phenomenon may be DNA dependent; cellular proteins which are bound to the DNA will copurify with it, and hence with TnsB, specifically enriching these fractions in DNA binding proteins. The equivalent pMR78 fractions are almost devoid of these contaminants, because DNA is not being precipitated in the absence of TnsB. Alternatively, these smaller proteins might be shorter forms of TnsB, either synthesised de novo or produced as a result of breakdown of TnsB.

It is possible that the DNA is responsible for maintaining the solubility of TnsB, even at high salt concentrations. The same is seen in the case of Mu B protein, which is extremely insoluble and requires the equivalent of its own weight in DNA to keep it in solution

(Chaconas et el., 1985). In attempts at further purification of TnsB by removing the DNA, the solubility properties of the protein were altered, and in some cases the protein was degraded (data not shown).

3.6 Chromatographic purification

3.6.1 Gel filtration by Superose 6

Given the high salt concentration required to maintain TnsB in solution, it was thought that ion exchange chromatography would not be a profitable route to further purification. It is likely that few proteins would bind to the column, and the only effect would be one of dilution. Alternatively, any proteins that do bind at this high salt concentration may be very difficult to elute. In addition, if it is believed that fraction IV is specifically enriched for DNA binding proteins, any step relying on ionic interactions would not be expected to have a great resolution.

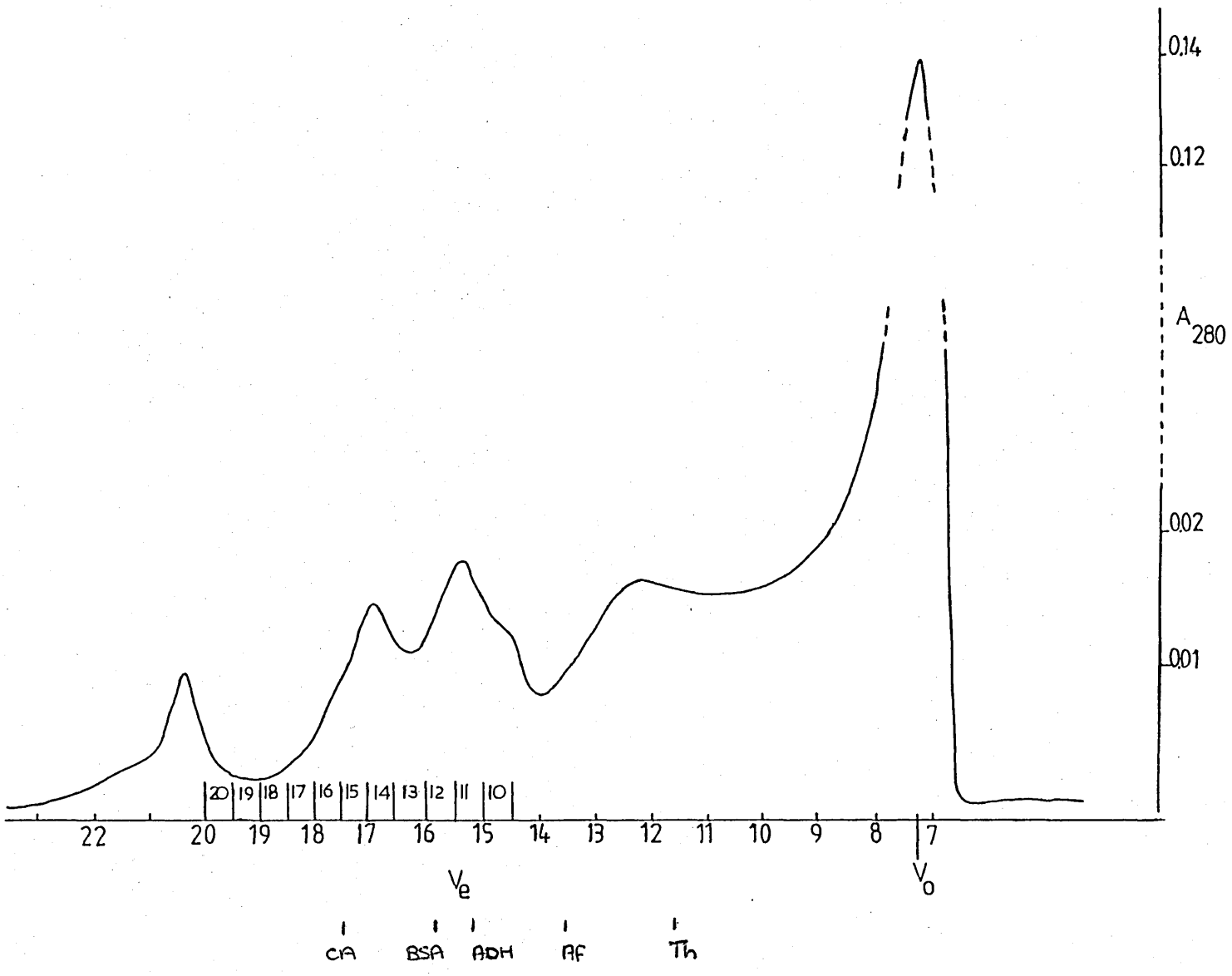
Gel filtration was chosen as the next purification step as it does not have the difficulties associated with ion exchange, and provides additional information regarding the native molecular weight of the protein in solution. Initial experiments, however, indicated that at the high salt concentration of the protein buffer, hydrophobic interactions caused proteins to be retained on the column (data not shown). The salt was therefore reduced to 675 mM, in which the protein was just soluble.

The matrix used for gel filtration was Superose 6 (Pharmacia), using a Waters 650E FPLC system. To calibrate the column, a standard curve was constructed using a set of protein size standards, run under the buffer conditions to be used for TnsB separation. The partially purified TnsB preparation (fraction IV) was dialysed into a buffer containing 675 mM KCl; insoluble material was removed by centrifugation and the supernatant loaded onto the Superose column; the A_{280} trace is shown in figure 3.7. Fractions were collected as indicated, and some of these

Figure 3.7 Purification of TnsB: gel filtration
Superose 6 column: buffer conditions were 50 mM Tris/HCl pH 8.2, 675 mM KCl, 1 mM EDTA: flow rate 0.5 ml/min: chart rate 0.5 cm/min.

A_{280} profile for elution of TnsB (fraction IV).
Elution volume is indicated below the horizontal axis, in ml. The void volume of the column, 7.3 ml, is marked.
Fraction numbers (0.5 ml fractions) are indicated above the axis.

The proteins used to calibrate the column were carbonic anhydrase (29 kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), apoferritin (443 kDa) and thyroglobulin (669 kDa), all at 4mg/ml.


0.14
0.12
A280
0.02
0.01
20 19 18 17 16 15 14 13 12 11 10
22 20 19 18 17 16 15 14 13 12 11 10 9 8 7
Ve
Vo
CA
BSA
ADH
AF
Th

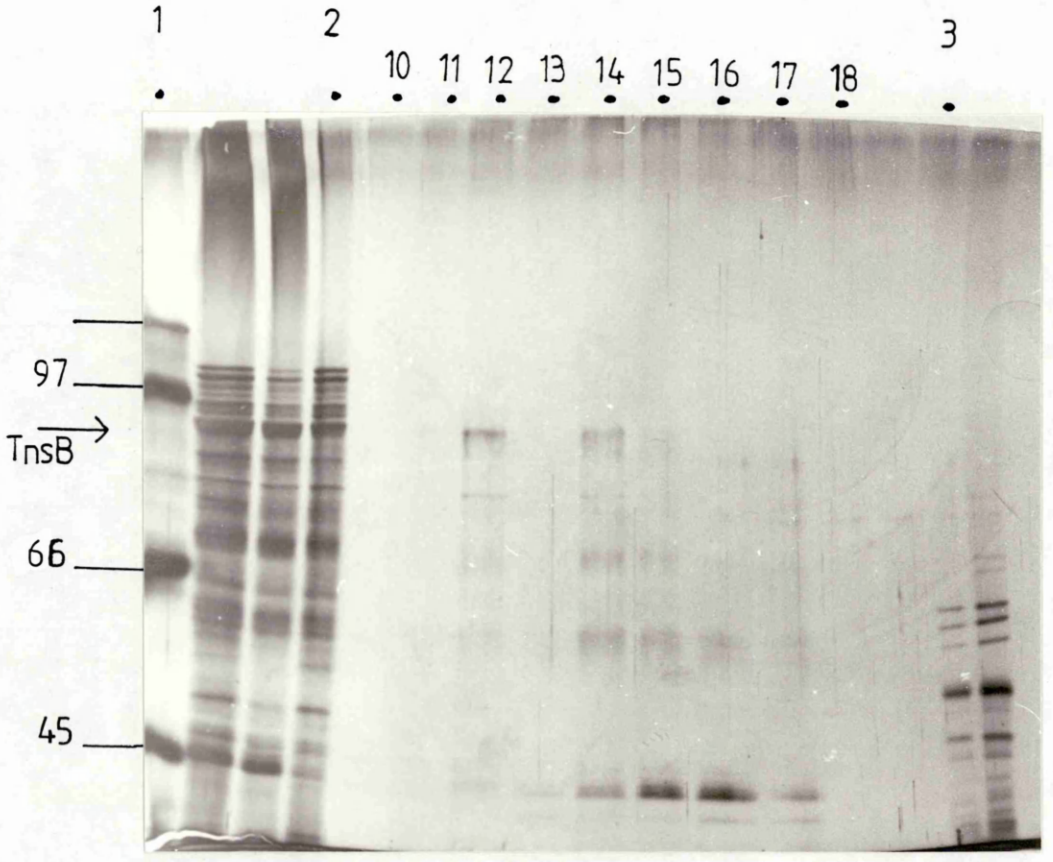


Figure 3.8 SDS-PAGE of column fractions 10-20
10ul (2%) of the indicated fraction was loaded.

lane 1 molecular weight markers
2 starting material (fraction IV)
3 IHF (partially purified)
10-18 Superose fractions, as indicated

Marker sizes are indicated in kDa

are shown on SDS-PAGE in figure 3.8; those fractions not shown contained no bands visualised by silver staining.

A major peak eluted at around the void volume of the column (7.3 ml). This material had a high absorbance at 260 nm (A_{260}/A_{280} = 2), and so is likely to be largely DNA as well as some high molecular weight protein. This was borne out by the silver stained gels, where DNA was seen as a smear in the stacking gel and at the top of the running gel (data not shown).

The next major peak had an elution volume of 15.6 ml, and it was these fractions which contained an 85 kDa protein (figure 3.8 lanes 11 and 12). From the standard curve, 15.6 ml elution volume corresponds to a molecular weight of 80-95 kDa. This range includes the predicted molecular weight of a TnsB monomer (81 kDa; Flores et al., 1990), suggesting that the monomer is the predominant form in solution under these conditions.

Fractions 14-18 contained progressively smaller species, with fractions 15, 16 containing proteins of 40-50 kDa, and fractions 17, 18 proteins of 12-20 kDa. A small late peak again had a high absorbance at 260 nm, and is likely to be small DNA fragments and RNA.

TnsB was still visible in fractions 15 and 16, ie it trails on the low molecular weight side of the elution peak. The apparent retardation in the column may be due to some hydrophobic interactions with the column matrix. About 30% of the visualised TnsB was included in this trailing edge, which fell into fractions containing other predominant species; as these impure fractions were not included in quantification, this partly accounts for the low yield of pure protein from the column step.

Interestingly, if the salt concentration was reduced by dilution rather than dialysis, the behaviour of the protein was slightly altered. In this case, all of the TnsB was in the supernatant, on centrifugation after reducing the salt, whereas following dialysis 30-50% of the TnsB was insoluble. The Superose trace for a diluted

preparation was slightly different, with the peaks being less well resolved and more difficult to interpret. TnsB elution occured at a smaller elution volume more consistent with the predicted size of a dimer (162 kDa). Thus it would appear that the multimerisation state of the protein is highly dependent on the preparation method used, and may be related to the solubility of the protein under those conditions.

3.6.2 DNA binding activity in the column fractions

The DNA binding activity of selected column fractions was assayed using band shift gels (figure 3.9). Fractions 11 and 12, which contained a TnsB-sized band by SDS-PAGE, contained a DNA binding activity (lanes 4-7). Although the level of activity is low, the binding appears to show some specificity for the Tn7RE fragment over the 216 bp control fragment.

Fractions 14-18 also had DNA binding activity. Although in some cases it can be ascribed to the presence of some TnsB, the amount of binding was too great to be simply explained in this way. In addition to the high level of activity, the binding was highly specific; 80% of the Tn7RE fragment was bound by 30 ng of fraction 16, for example, with very little loss of the pUC18 small fragment (figure 3.9 lanes 14 and 15). The pattern of bands produced was different between the two sets of fractions; binding reactions with fractions 14-18 display an evenly spaced ladder of bands, whereas the pattern for fractions 11 and 12 is more like that seen previously.

There are several hypotheses to explain these observations. Perhaps the simplest is that these smaller polypeptides in the later fractions are fragments of TnsB, due to degradation, for example by contaminating proteases. This provides an explanation for their binding activity, as well as for their absence from pMR78-containing cell extracts. Orle and Craig have also detected proteolytic fragments of TnsB, using anti-TnsB

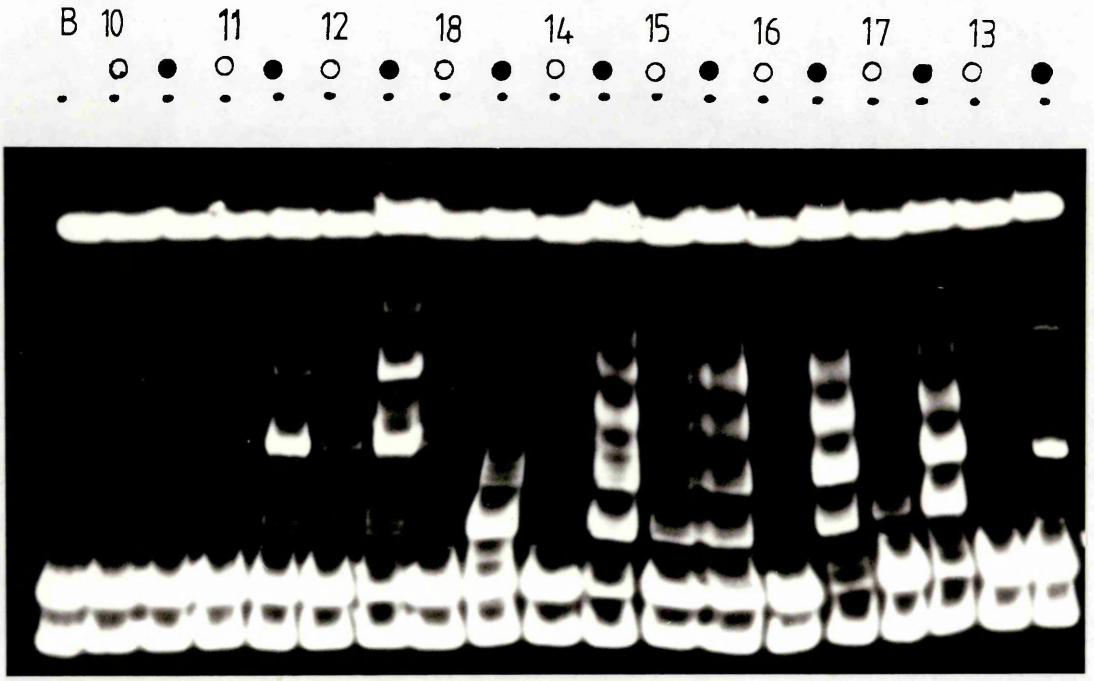


Figure 3.9 DNA-binding activity in the column fractions
1ul samples of the fractions was incubated with end-labelled pNE200, as previously described (Methods, figure 3.6).

Fraction numbers, and protein dilutions, as indicated:

B no protein ● neat sample O 2-fold dilution

Protein concentration of fractions (ug/ml: by BCA assay):

10	ND	13	ND	16	50
11	14	14	30	17	100
12	20	15	30	18	ND

ND - not detectable

antibodies (in press). It is not unprecedented for deletion products of TnsB to be able to bind DNA; the C-terminally deleted protein used by Ekaterinaki (1987), and a 20 kDa N-terminal fragment of TnsB (A. Gawthrop, personal communication) have both been shown to have binding activity. It is possible that the species responsible for the majority of the binding activity is a cleavage product, in binding assays and in transposition in vivo. However, it is perhaps surprising to find so many fragments of a protein which retain activity.

An observation relevant to this hypothesis is that during a recent preparation of TnsB, the fraction expected (from its elution volume) to contain TnsB, after concentration and gel electrophoresis showed a predominant band at around 50 kDa, as found in fractions 14-16. Its appearance at an elution volume of 15.6 ml suggests that the polypeptide originated from full length TnsB. Although the N-terminal sequence of the 50 kDa species would show whether it originated from TnsB, this was not thought to be a priority, especially as there was no more of that perticular sample to use in DNA-binding assays, and no evidence that it was the same as the protein observed previously.

An alternative explanation is that these are different proteins, presumably host-encoded, which have binding activity with an apparent specificity for Tn7 right ends. This again would not be unprecedented, as many cellular proteins are known to have specific binding sites on phage, transposons and plasmids, eg IHF on transposon gamma-delta (Wiater and Grindley, 1990), phage lambda att sites (Craig and Nash, 1984), and pBR322 (Prentki et al., 1987); FIS in the DNA inversion systems (Koch and Kahmann, 1986; Hubner et al.,1989) and ArgR on the cer site of ColE1 (Stirling et al.,1988).

IHF is a strong candidate for involvement in Tn7 transposition, as there are potential IHF binding sites in the right end (see chapter 4). To test whether IHF was

involved in any of these binding activities, a sample of partially purified IHF was run on a gel beside the fractions (figure 3.8 lane 13, and data not shown). None of the bands visible were common to IHF and any of the fractions. In addition, the binding activity of the fractions, with and without added IHF, was compared. No significant differences in the pattern of complexes were seen, suggesting that IHF is not one of the proteins in fractions 14-18 (data not shown).

It would also be possible to test FIS, and other host proteins, in the same way, ie to attempt to reconstitute the binding pattern of the partially purified TnsB, by adding back purified proteins to the column fractions. To complement this work, the identities of the major proteins in these fractions could be determined by N-terminal sequencing, and comparison to sequences in the databases. As discussed above, sequencing would also clarify whether the source of the proteins was TnsB.

3.6.3 Stability of the purified protein

Various attempts were made to concentrate the TnsB-containing fractions. This was desirable for binding experiments and for N-terminal sequencing. However, this could not be achieved; the dilute fraction was stable when stored at 4^{o} or $-20^{o}C$, but no full length protein was recovered from any of the concentration techniques attempted (precipitation in low salt, centrifugation in Centricon columns). The instability is likely to be due to the removal of the DNA by the column; the protein may be aggregating and coming out of solution. There might also be a proteolytic activity which has become active at this stage.

This will need to be overcome, in order to make concentrated, pure TnsB for further biochemical characterisation, and, ultimately, for use in an <u>in vitro</u> transposition system.

Table 3.1 Purification of TnsB

fraction	ml	mg	mg TnsB[a]	U[b]	U/mg
high salt (II)	2.8	43	21	22400	521
DNaseI (III)	2.8	39	19	22400	574
precipitation (IV)	1.0	15	11	32000	2133
storage (4°C)	1.0	5.2	2.6	8000	1538
dialysis[c]	0.1	0.19	0.095	600	3158
Superose 6:					
fraction 11	0.5	0.007	0.007	250	35714
12	0.5	0.01	0.01	707	70711

a. Estimated as the percentage of total protein, as seen on SDS-PAGE; the accuracy is limited by the linearity of staining with Coomassie / silver, of different proteins.

b. One unit is defined as the amount of protein required to shift 50% of the DNA into complexes, as assayed by gel retardation. An alternative unit, defined as the amount of protein required to give a binding pattern in which the intensities of bands III and IV were equal (chapter 4) gave very similar results.

c. From this stage, only 10% of the total was used.

3.7 Quantitation of purification

A purification table is shown in table 3.1. The overall yield, allowing for the fact that only 10% of the total was loaded onto the column, was around 0.8% Much of the loss occurred during storage at 4°C, when the total protein fell to 35%, with further loss of TnsB into the insoluble fraction after dialysis. The yield of TnsB from the column step was poor, although the other proteins were efficiently recovered; it is not known whether the loss occured on the column (possibly by hydrophobic interactions with the matrix), or by degradation after the sample had eluted.

The units defined are somewhat arbitrary, but serve to provide an estimate of the extent of purification, of around 135-fold. The increase in the total number of units, and in the number of units per milligram of TnsB, can be explained if the contaminating species are inhibiting binding, eg by other proteins competing for the site, or by large amounts of chromosomal DNA binding TnsB non-specifically.

3.8 Identity of the purified polypeptide: N-terminal sequencing

Strictly, the protein referred to as 'TnsB' is in fact a tnsB dependent protein of 80-85 kD, which copurifies with a binding activity specific for Tn7 end sequences. Whilst these are the properties expected for the tnsB gene product, this does not prove the identity of the polypeptide.

To identify the purified protein, its N-terminal sequence was determined. A sample was loaded onto an SDS-polyacrylamide gel (using a Tris/Trycine buffering system to avoid glycine contamination), and transferred to Problott membrane (Applied Biosystems); the 85 kDa protein band was excised and the N-terminal sequence determined (see Materials and Methods).

The first 15 amino acids were determined, and the

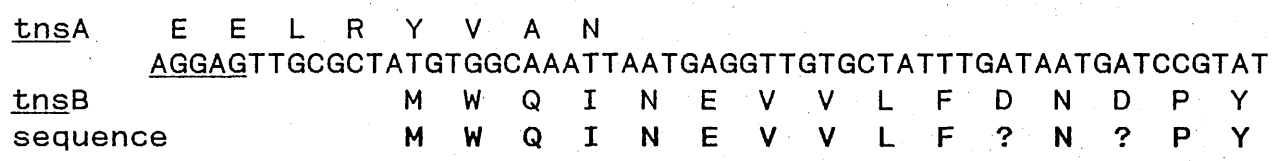


Figure 3.10 N-terminal sequence
The N-terminal sequence obtained is shown, compared to the DNA sequence (taken from Flores et al.,1990), and the predicted amino acid sequence. The proposed ribosome binding site for tnsB is also indicated.

sequence is shown in figure 3.10, compared to that predicted from the DNA sequence in the region around the start of tnsB. The match of the sequences over the region determined, confirmed the identity of the purified protein as the product of the tnsB gene. There are two positions which were difficult to assign (residues 11 and 13). The predicted amino acid, aspartate in both cases, is not inconsistent with the data from the sequencing reaction.

The amino acid sequence demonstrates that translation begins at the first AUG codon in the reading frame, rather than at any downstream start codons. This overlaps with the end of the tnsA gene; such reading frame overlaps are often associated with translational coupling of gene expression (Nomark et al.,1983). There is no N-terminal processing of the protein, at least in the overproducing strain; the protein sequence shows the presence of all the residues predicted by the gene, including the N-terminal methionine.

3.9 Identity of the binding activity

To identify the protein species present in the complexes, two approaches were taken. The first was so-called 'South-Western blotting' (Bowen et al., 1980; Miskimins et al., 1985; Roth et al., 1988). Protein separated by SDS-PAGE is transferred to a nitrocellulose filter, which is then probed with a labelled DNA fragment carrying the putative binding site; bound DNA is detected and located by autoradiography.

There are several difficulties with this technique, which may contribute to any failure to produce positive results. The experiment relies on refolding of the protein upon transfer to the membrane; TnsB (or other protein(s) responsible for the DNA binding) may not renature efficiently, or may not do so in an active conformation. Related to this is the question of multimerisation; if a particular quaternary structure is required, again this may not reform on the filter. Several modifications have

been introduced into the standard Western blotting procedure to facilitate renaturation (Miskimins et al., 1985), which include carrying out the blotting procedure in the absence of SDS.

In addition, there may be a solubility problem, as previously discussed for TnsB preparations. Once the SDS is removed the protein may aggregate in the gel, or on the filter. Once bound to the filter, the protein may have reduced flexibility or mobility; any conformational change required during the reaction, or a precise spacing or spatial arrangement of protein needed for effective binding, may not be able to take place under these conditions.

The proteins in the samples have been subjected to denaturing electrophoresis. If TnsB, for example, required a host protein or other cofactor for efficient binding, the two species will have been separated in the gel; the transferred protein, even if it were efficiently renatured, would be unable to bind DNA.

Despite several attempts, no bound DNA was detected, for any stage of the purification, on blots probed with the Tn7 right end fragment. As a positive control, Tn3 resolvase was run on the same gel, blotted, and probed with a labelled res site; a faint band at 20 kDa was visible, despite a high background (data not shown). This indicates that the gel ran normally (also shown by the biotinylated markers), and that the blotting procedure was effective, at least for small proteins.

The absence of any observable DNA binding activity may be due to any combination of the points noted above. Another factor is the binding conditions used. These have not been exhaustively tested. 15 minutes binding at room temperature, and an overnight incubation at 4^{o}C, both failed to demonstrate binding. The salt concentration used was 50mM M NaCl, which is less than that recommended by Miskimins et al. (0.1 M); this reduction was because TnsB binding may be sensitive to high salt, but it may be that

the salt is required to maintain the protein in a state competent to bind DNA. We should also consider the non-specific 'carrier' DNA, used to block non-specific binding proteins and potential binding to the filter. Too high a concentration may also block specific binding. As TnsB appears to have a fairly high affinity for general DNA, the levels of carrier recommended for other proteins may be too high in this case.

A complementary approach was taken, whereby complexes visualised on a native (Tris/Glycine) gel were cut out, and the slices loaded into the wells of an SDS-polyacrylamide gel. This method was used by Straney and Crothers (1987) to identify the proteins present in ternary DNA/LacI/RNA polymerase complexes.

The difficulty with this experiment is in loading sufficient protein onto DNA and into complexes, to have enough to see on a protein gel. The amounts of complex in a standard gel assay are of the order of fmol; this is less than 0.1 ng protein for a species binding a single monomer, which is too low to be seen by standard staining methods. The advantage in the lac system is the limited number of complexes; the protein is concentrated in a few bands. For TnsB the large number of complexes means the protein is spread through a large volume of gel, so the increase required in DNA and protein concentrations is much greater.

Initial experiments using relatively low concentrations of DNA and protein gave no visible bands on silver staining of the protein gel, as predicted by estimates similar to those above (data not shown). The result of using high concentrations of components, however, was that individual complexes could no longer be identified in the band shift gel. Rather, the Tn7 fragment was strongly retarded, presumably in higher order complexes, or possibly trapped in protein aggregates (data might be shown). It was then not clear that any protein detected was responsible for the formation of specific

complexes. Under none of the conditions tested was a high level of specific complex formation achieved, although not all possible variations have been tested. One suggestion is to try adding polyanions, such as polyglutamate, which may reduce aggregation (Darby and Creighton, 1990). Formation of the large complexes may be reduced by adding more carrier, but this will also lower the amount of protein available to form complexes, which returns to the question of detection limits.

The experiment would be an informative one, both for identification of the proteins in the complexes, and to allow estimates of the stoichiometries of the complexes. One solution to the detection/aggregation problem would be to use radioactive protein, labelled *in vivo* with ^{35}S-methionine. This would allow all proteins which bind to the Tn7 end to be observed, and the components of each complex identified. Further quantitative work would probably be best accomplished by using detection with antibodies against TnsB (or other DNA-binding protein of interest).

3.10 CONCLUSION

TnsB, the largest of the Tn7 transposition proteins (81 kDa) has previously been shown to have a DNA binding activity specific for the ends of Tn7, both *in vivo* and *in vitro* (Ekaterinaki, 1987; McKown et al., 1987). In order to characterise the binding activity further, and determine the role of the TnsB protein in transposition, the *tnsB* gene was overexpressed in *E.coli* and the gene product purified. Purification was followed by observing the protein on SDS-PAGE, and by gel retardation assays for binding activity.

The first purification step took advantage of the insolubility of the *tnsB* gene product, and allowed purification of TnsB substantially. . Use of this preparation in further binding experiments is described in

chapters 4 and 5.

Further purification, by preparative gel filtration, gave fractions containing TnsB of greater purity, as judged by silver staining of protein gels. The elution volume of the protein showed that the monomer is the predominant form in solution, under these conditions.

DNA binding activity copurified with the 85 kDa TnsB. Following gel filtration, fractions containing smaller species (45, 20 kDa) were also found to contain a high level of binding activity which was specific for the Tn7 right end. It is not known whether these are degradation products of TnsB, or host-encoded proteins which bind to the ends of Tn7. N-terminal sequencing of the polypeptides would be informative, in identifying the proteins and their source.

The 'TnsB'-dependent pattern of binding, as assayed by the gel retardation assay, is likely to be the result of interactions between TnsB and these other proteins, and this may reflect the situation during <u>in vivo</u> transposition of Tn7. It would be interesting to investigate the binding properties of these proteins further. Experiments using gel retardation assays would be informative, for determining the affinity of the binding for Tn7 ends and other DNA sequences, for detecting cofactor requirements, and for examining any structural changes induced by binding; the binding sites of the proteins can be determined by footprinting techniques. Experiments to address the role of these proteins in transposition can also be carried out; for example, are they required for enzymic activities on the transposon DNA ? Or for bringing together of transposon ends ? This last suggestion can be examined using ligation experiments such as those described in chapter 5, to detect interactions between the transposon termini, as well as by modified gel shift experiments.

The identity of the purified polypeptide was confirmed by N-terminal sequencing. The sequence also

determined the translational start of TnsB, as the first AUG in the tnsB open reading frame. There is no N-terminal processing of the protein, as all residues, including the N-terminal methionine, are present in the sequence.

The activity of TnsB in binding specifically to the ends of Tn7 suggests that the role of TnsB in transposition is to recognise and bind to the transposon termini. A secondary effect of this binding is repression of P1, the major Tn7 promoter, which lies in the right end. It is not known whether TnsB is directly responsible for cleavage at the transposon termini. The most likely role for TnsB, given its binding activity at both ends of the element, is in bringing together the two ends to form a 'transpososome'. This hypothesis is developed more fully in chapters 4 and 5, which describe experiments to characterise both the binding activity and the recognition site, and to detect interactions between the ends mediated by TnsB.

CHAPTER FOUR

BINDING OF TnsB TO THE RIGHT-HAND END OF Tn7

4.1 INTRODUCTION

During non-denaturing electrophoresis, DNA bound into stable complexes with protein has a reduced mobility, and separates from the free DNA. This method for studying protein/DNA interactions was first described for binding to the lac promoter by purified CAP protein (Garner and Revzin, 1981) and binding of purified lac repressor to its operator (Fried and Crothers, 1981). The technique has since been applied to many prokaryotic and eukaryotic DNA binding proteins, both purified (eg Tn3 resolvase: Bednarz, 1989; MetJ: Phillips et al., 1989; TrpR: Carey, 1988; OBP100: Baumroker et al., 1988) and in crude extracts (Strauss and Varshavsky, 1984; Ekaterinaki, 1987). Binding assays of this type have been a particularly profitable route to purification of eukaryotic transcription factors (Carthew et al., 1985).

The band retardation system complements the filter binding assay traditionally used to study specific DNA/protein interactions, and to analyse kinetic and thermodynamic parameters. The gel assay also allows these parameters to be estimated (Fried and Crothers, 1984a, b). It has the additional powerful feature that different bound species are resolved, rather than included together in the 'filter bound' class. Hence much qualitative information is obtained, which would be lost in the filter binding assay; a change in the distribution of the complexes, for example, with little change in the percentage bound, would not be detected by filter binding but would be apparent in the gel assay. This is particularly useful when the interactions of several proteins with a DNA fragment are being observed, as all species can be detected (eg lacI/RNA polymerase/operator complexes: Straney and Crothers, 1987). Structural changes in the complexes can also be visualised, eg the shift from closed to open complex during transcription initiation (Straney and Crothers, 1985).

By observing the number, and relative concentrations, of complexes formed, inferences can be made concerning the nature of each complex, and how the protein binds; models of binding can be tested by their ability to predict the number and characteristics of complexes formed. This approach has been used effectively for binding of Tn3 resolvase to its res site (Bednarz, 1989). The gel assay also allows these questions to be approached directly, as the complexes can be analysed independently. Footprints of individual complexes, for example, can be obtained by carrying out the cleavage reactions in the gel, then isolating the DNA for denaturing gel analysis (eg Kuwabara and Sigman, 1987). In addition, the stoichiometry of each complex can be determined (eg Kolb et al., 1983; Carey, 1988).

The gel assay has another advantage in that more than one DNA fragment can be observed in the same reaction, more readily than by filter binding assays. During a protein purification, for example, the extent of non-specific DNA binding activity can be followed by including a control DNA fragment (see chapter 3). Band disappearance assays, using a mixture of labelled fragments, can determine which fragments contain a binding site of interest (Craigie et al., 1984), and to compare directly the relative affinities of two sites (Kolb et al., 1983). In addition, binding assays containing two specific fragments of different sizes can be used to detect intermolecular complexes, as the mixed complex will have an intermediate mobility (Kramer et al., 1987).

An area of research in which band shift gels have proved valuable is that of DNA bending. Bending induced by binding of a protein to the DNA can be distingiushed from sequence-determined intrinsic bending. DNA does not behave as an isotropic rod, but shows sequence dependent variations in flexibility in different planes (Travers and Klug, 1987). For example, work on sequence determinants of nucleosome positioning (Drew and Travers, 1985; Satchwell

et al., 1986) showed that (A+T) runs tend to occupy the compressed minor groove on the inside of the bend, with (G+C) runs on the outside where the minor groove widens. Data of this type allow predictions to be made concerning the behaviour of a given DNA sequence when constrained. The ease with which different sequences are formed into nucleoprotein complexes has biological importance; for example, in the effects nucleosome positioning can have on gene expression (Almer et al., 1985).

In some cases, the DNA sequence imposes an intrinsic bend on the DNA, in the absence of external constraints. An example of intrinsic bending is the kinetoplast minicircle DNA of L.tarentolae (Marini et al., 1982). Bending is detected by the anomalous mobility of the DNA in polyacrylamide gels; bent fragments are retarded, as they experience greater resistance to their passage through the gel matrix. Runs of A or T greater than 3 bp, with a periodicity of 10-11 bp, are sufficient to confer detectable curvature on a DNA fragment (Koo et al., 1986).

The centre of a DNA bend can be determined by gel electrophoresis. A set of circularly permuted fragments containing the proposed bend is generated and their gel mobilities are compared. Fragments containing a bend at the centre are more strongly retarded in the gel than fragments with a bend near the end. A plot of the relative mobilities allows the position of the bend centre - the position of maximum retardation - to be estimated (Wu and Crothers, 1984).

Bending can also be induced by binding of a protein to the DNA. This is detected by observation of mobility variations in the complexes, in a circular permutation experiment as described above. Examples of proteins which induce bends include CAP (Wu and Crothers, 1984), IHF (Prentki et al., 1987; Robertson and Nash, 1988), Tn3 resolvase (Brown et al.,unpublished) and gal repressor (Kuhnke et al., 1989). Some proteins, such as lambda repressor (Griffith et al., 1986) and lac repressor (Wu

and Crothers, 1984) have been shown not to bend DNA detectably, although Zwieb et al. (1989) have observed bending by lacI using different constructs. The fact that not all proteins exhibit bending in this experiment shows that the mobility variations that are observed are not simply due to the changed position of the bound protein on the DNA fragment.

4.1.1 Equilibria and kinetics

For an equilibrium of the type

$$P + D \rightleftharpoons PD$$

then the dissociation constant, K_D, is

$$\frac{[P].[D]}{[PD]} \qquad (1)$$

where P, D, and PD represent protein, DNA site, and complex, respectively. If the protein is in excess, then $[P] \simeq [P]_o$.

When 50% of the DNA is bound,

$$[D] = [PD], \quad \text{so} \quad K_D = [P] \simeq [P]_o \qquad (2)$$

ie the dissociation constant is equal to the protein concentration at which half of the DNA is bound. The data can be plotted in various ways:

a) from (1), the ratio of bound DNA to free DNA

$$\frac{[PD]}{[D]} = \frac{[P]}{K_D}, \Rightarrow \log\left(\frac{\text{bound}}{\text{free}}\right) = \log[P] - \log K_D \qquad (3)$$

hence a plot of log(b/f) against log[P] is linear, with an x-axis intercept of $-\log K_D$, and a gradient of 1 for simple binding of one protein molecule per DNA fragment; for binding of more than one protein molecule, the gradient is ≥ 1.

b) from (1), the proportion of total DNA bound

$$\frac{[PD]}{[D_o]} = \frac{[P]}{[P]+K_D} \qquad (4)$$

which is analogous to the Michaelis-Menten equation. A plot of proportion of DNA bound against [P] is hyperbolic for simple binding; if there is cooperativity of binding of more than one protein molecule, the graph becomes sigmoid.

There are several caveats to making and using an estimate of the dissociation constant based on data from the gel binding assay. The binding reactions have been subjected to electrophoresis through polyacrylamide; the complexes seen on the gel may not be truly representative of the situation in solution. The complexes which survive electrophoresis may be only the most stable ones produced, or may have been modified by interactions within the gel matrix to produce a stable form. A similar concern applies to filter binding experiments, in that the protein/DNA interaction may be perturbed when the complex binds to the filter. However, the gel assay allows the amount of free DNA to be measured, providing that the complexes are stable for as long as it takes the DNA to enter the gel (around 80 seconds: Fried and Crothers, 1981).

In order to determine a dissociation constant, the system must be at equilibrium. In practice, this is likely to be the case after an incubation time equivalent to several dissociation halflives; during this time the complexes will have had the opportunity to dissociate and reform several times. For TnsB, the binding pattern appears to be stable within five minutes, and does not change during an extended incubation of over an hour, so we can be confident that equilibrium is indeed reached within the standard 15 minutes incubation.

To calculate a dissociation constant, the protein concentration must be known. Bradford and BCA assays measure the total polypeptide concentration in the solution, but this value is not necessarily the concentration of available, active protein. Hence the stated protein concentration may be an overestimate, giving an artificially high dissociation constant. This is

particularly relevant in the case of TnsB, where some of the protein might be aggregated, or bound non-specifically to DNA. Further considerations relevant to systems where many undefined complexes are produced, as with TnsB, are discussed in the Results section.

For a reaction of the type shown above, with the protein in excess, if the rate of the forward reaction is directly proportional to the concentration of the components, then

$$\text{rate} = -\frac{d[D]}{dt} = k_f.[P].[D] \qquad (5)$$

where k_f is the forward rate constant. Separating the variables and integrating gives

$$\ln \frac{[D]}{[D]_o} = -k_f.[P].t \qquad (6)$$

and a plot of $\ln ([D]/[D]_o)$ against t is linear, with gradient $-k_f.[P]$ (the pseudo first order rate constant).

From (6), an equation for the halftime of the reaction can be derived:

$$t_{1\backslash 2} = \frac{\ln 0.5}{-k_f.t} \qquad (7)$$

Similarly, in the general case for dissociation of the nth protein protomer from a multiply bound complex, ie

$$P_nD \rightleftharpoons P_{n-1}D + P \ ,$$

if the probability of dissociation is constant, then

$$\text{rate} = -k_b.[P_nD]$$

and equations for the rate constant and the halflife can be derived as above.

At equilibrium, the forward and back rates are equal, i.e. $k_b.[P_nD] = k_f.[P_{n-1}D].[P]$

from which

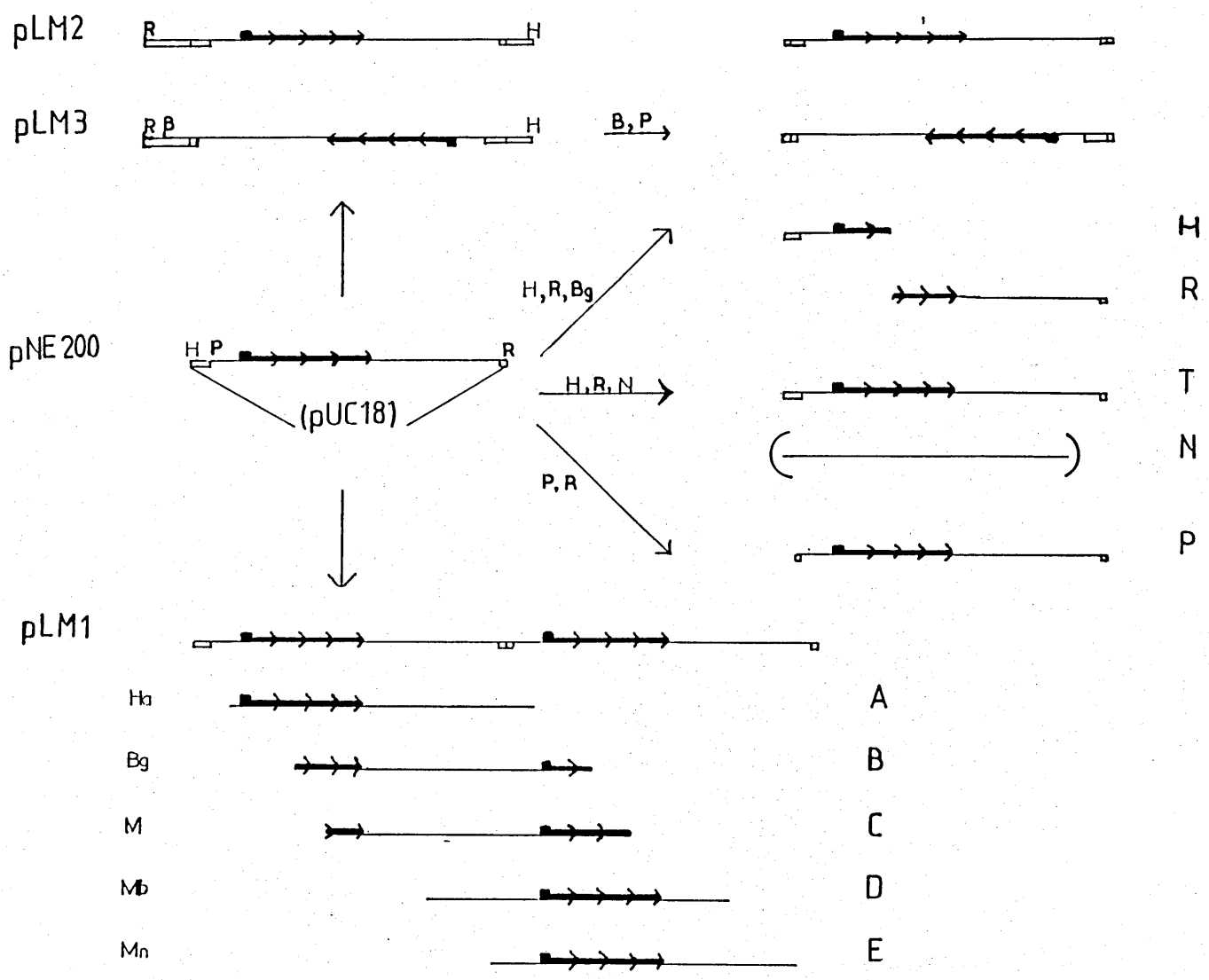
$$\frac{k_b}{k_f} = \frac{[P_{n-1}D].[P]}{[P_{nD}]} = K_{D(n)} \qquad (8)$$

Hence for a reaction at equilibrium, once two of the values K_D, k_f and k_b are known, the other can be calculated if this equation holds, ie the reaction rates have first order dependence on reactant concentrations, and the values used refer to the same step in the reaction; this is an important consideration if multiple species are produced, and the relationships between them are unknown.

The retardation of complexes during electrophoresis could be due to various factors, such as the larger size or increased mass of a complex compared to naked DNA, or structural changes in the DNA. Charge differences are minimal, as the slight charge on the protein is swamped by the DNA's negative charge. The molecular weight of a 300 bp fragment is around 200 kDa, so the binding of a small DNA binding protein such as a FIS dimer (24 kDa) causes a weight increase of only 10%, although for larger proteins such as TnsB the weight increase may contribute more significantly. The major contribution to retention, then, is likely to be apparent size, via both the increased volume due to the protein, and structural changes such as bends.

In chapter 3, the existence of a tnsB-dependent binding activity, specific for Tn7 ends, was demonstrated. This chapter describes the application of the gel retardation assay to examine the interactions of TnsB with the Tn7 ends in more detail, using a substantially purified TnsB preparation (chapter 3), and improved gel conditions to optimise complex detection. The ends of Tn7 contain a repeated 22 bp motif, and the observation that

A


pLM2
R
H
pLM3
R B
H
B, P
pNE 200
H P
R
(pUC18)
H, R, Bg
H, R, N
P, R
H
R
T
N
P
pLM1
Hn
Bg
M
Mb
Mn
A
B
C
D
E

B

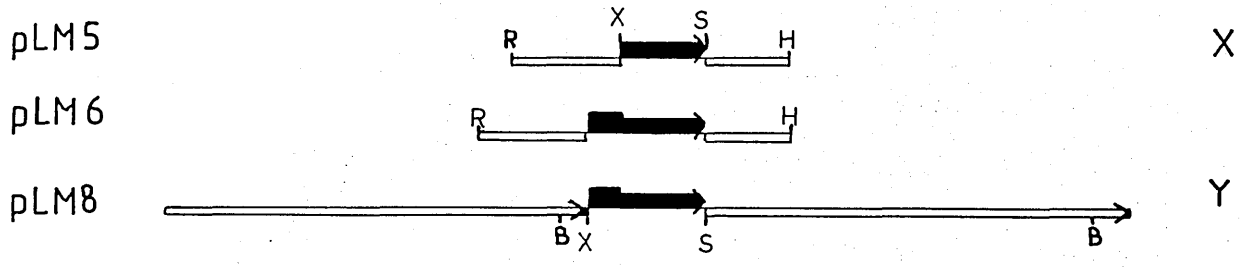


Figure 4.1 DNA fragments used in binding experiments

A. Plasmids and DNA fragments containing 205 bp of the right end of Tn7, derived from pNE200 by the restriction digestions and subclonings indicated. pLM2 and pLM3 contain the HindIII-EcoRI fragment from pNE200 cloned into the HincII site of pUC18. pLM1 contains a tandem dimer of the Pst-EcoRI fragment. [pNE200: ref Ekaterinaki, 1987; see table 2.2]

B. Plasmids containing cloned synthetic oligonucleotides.

B=BamHI; Bg=BglII; H=HindIII; Ha=HpaII; M=MaeI; Mb=MboII;
Mn=MnlI; N=NdeI; P=PstI; R=EcoRI; S=SalI; X=XbaI.

—— Tn7 sequences

═══ pUC18 polylinker and other sequences

■ Tn7 terminal 8 bp

⟶ 22 bp repeat

⟹ pBEND repeated polylinker

an isolated copy of this repeat binds TnsB suggest that the 22 bp repeat contains the TnsB binding site. Advantage is taken of the qualitative data available, to look for structural variations, complexes involving more than one DNA molecule, and effects of cofactors and other proteins; kinetic and thermodynamic parameters are determined, and the biological implications of the values are discussed.

RESULTS AND DISCUSSION

4.2.1 Protein preparations

The TnsB preparations used in this chapter are high salt eluates from the membranous pellets after cell lysis (fraction IV; see chapter 3). Judged by Coomassie staining of protein gels, about 75% of the protein in this fraction is TnsB. As discussed in chapter 3, the eluate also contains other proteins which have been shown to bind to the right end of Tn7; the binding properties observed, then, could be due to interactions involving these as well as TnsB. However, the data from chapter 3 indicate that the predominant component is TnsB; it is the major species in the preparation, and the Superose-purified fraction contains binding activity which gives a band pattern related to that of fraction IV in the gel assay. In addition, the binding activity is dependent on tnsB, and is highly specific for the Tn7 ends. It is possible that some of the other proteins are degradation products of TnsB.

4.2.2 DNA fragments used in these experiments

These are shown in figure 4.1. The letters designating the various fragments are used in the text and the figure legends. Unless otherwise indicated, all native gel assays used ^{32}P end-labelled DNA fragments; other gels were stained with ethidium bromide as normal.

4.3 Titration of TnsB binding

A titration of the TnsB preparation against a DNA fragment containing the right-hand end of Tn7 (Tn7 RE) is shown in figure 4.2. As before, digestion of the plasmid with NdeI releases a 216 bp pUC18-derived fragment in addition to the Tn7 fragment, which acts as a control for non-specific binding. The small fragment and the pUC18 vector fragment remained unretarded until the added protein reached 0.5-1 ug (50-100 ug/ml), whilst the Tn7RE fragment was 90% bound at this concentration.

The retarded bands showed a characteristic pattern, which changed consistently as the protein concentration increased. Complexes I, III and IV were the first to appear, with IV becoming predominant (lane 4). At higher protein concentrations, more severely retarded complexes appeared, at the expense of lower complexes and free DNA (complex bands V-VIII and above; lanes 5-8). This is consistent with the idea of 'chasing-up' into higher order complexes with more protein units bound per DNA molecule.

As the protein concentration increased, the relative intensities of the complex bands changed. At lower protein concentrations, band IV was strong, but became less so, and the relative intensity of band III increased, at higher protein levels (compare lanes 7 and 8). The increased mobility at higher protein concentration cannot be explained by simple binding of additional protein; one possible explanation is that there is a structural change in the complex which allows it a higher mobility in the gel. One can envisage, for example, a bend induced by binding of one protein unit being straightened out by interactions with a second.

Alternatively, it may be that complex III is a non-productive species that cannot go on to form higher order complexes. IV could become less intense by further retardation in higher complexes, whilst the level of the dead-end complex simply increases as a function of protein concentration. Differentiation between these two

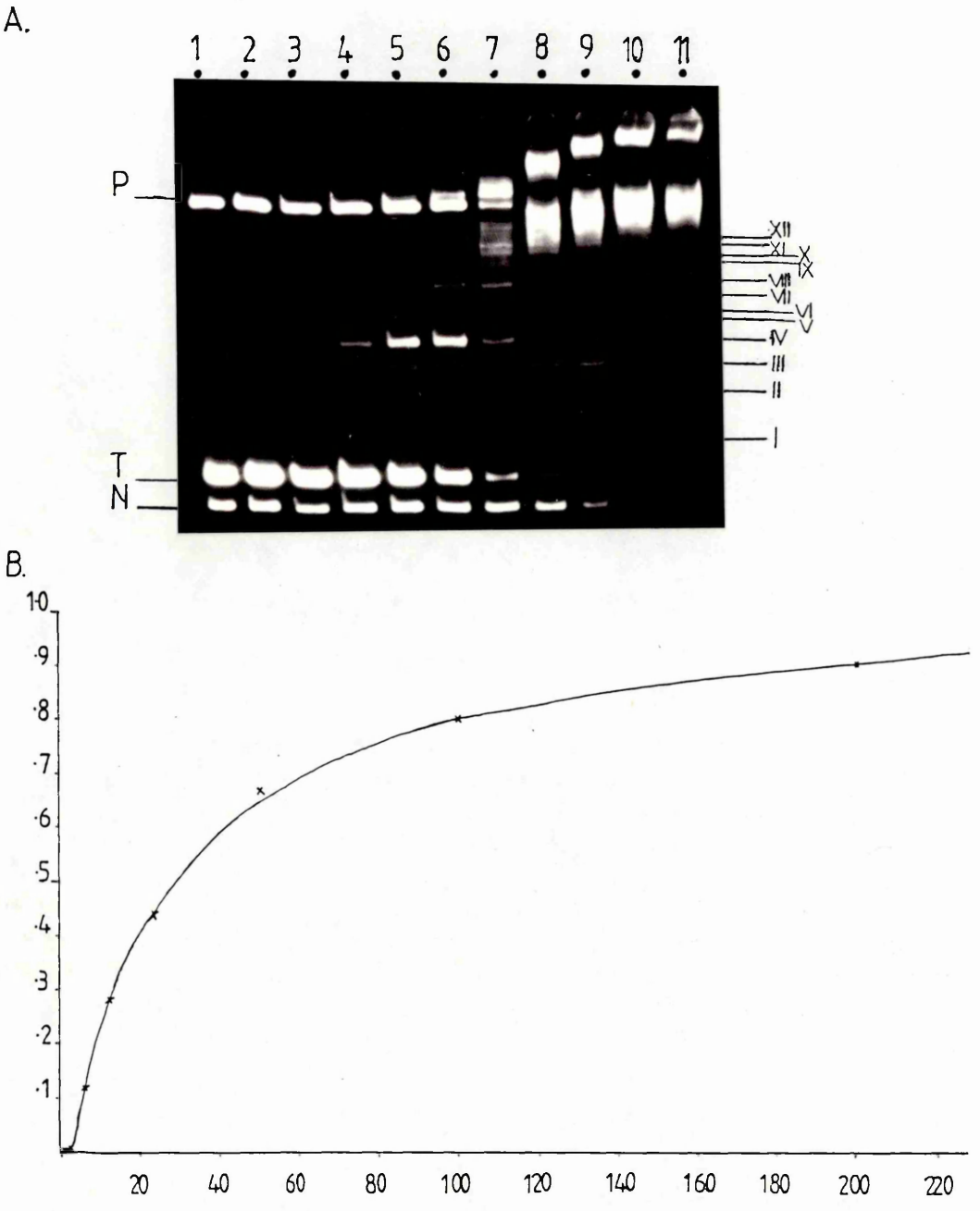


Figure 4.2 Titration of TnsB binding to Tn7 right end
A. 5 ng end-labelled pNE200 (0.5 ng Tn7 fragment) was incubated, for 10 minutes at room temperature, with TnsB extract (see Materials and Methods). 50 ng supercoiled pUC18 (5 ug/ml) was included as carrier.
Binding conditions: 50mM Tris/Glycine (pH9.4), 1mM EDTA, 68mM KCl, 10% glycerol ('standard' conditions).
Gel conditions: 50mM Tris/Glycine (pH9.4), 0.1mM EDTA.

lane	protein	lane	protein	lane	protein
1	0	5	125 ng	9	2 ug
2	16 ng	6	250 ng	10	4 ug
3	31 ng	7	0.5 ug	11	8 ug
4	62 ng	8	1 ug		

Identifiable complexes are indicated I-XII.

P = pUC18 vector	2.4 kb
T = Tn7 right end fragment	245 bp
N = NdeI-HindIII pUC18 fragment	216 bp

B. Plot of the ratio of [bound DNA]/[total DNA], against the concentration of protein added. (Molar)

hypotheses will require experiments to trace the source of the DNA in each complex during a titration. A related suggestion is that complex III contains the 216 bp control fragment; the shift from IV to III is seen at the protein concentration at which the non-specific fragment begins to be retarded (lanes 8-11), and it is likely that such complexes would be unable to generate higher forms. Binding to the control fragment is discussed in the next section.

The decrease of free DNA with increasing protein is plotted in figure 4.2B. The Tn7 fragment is almost completely bound with 2 ug (200ug/ml) protein (lane 9); if TnsB is 75% of the protein present, this is 18 pmoles of monomer, which is a vast excess over the 5 fmoles of DNA fragment present (20 fmoles of 22 bp sites). It is not clear why such a vast excess is required, but this is not unprecedented in these native gel assays, eg trpR (Carey, 1987), Tn3 resolvase (Bednarz, 1989). It is possible that TnsB is not fully soluble under the binding conditions, so that only a proportion is available in solution to bind to DNA; there may also be some inactivation of the protein during the preparation, again reducing the total available binding activity. Thirdly, a significant proportion of the protein might be bound to chromosomal DNA (see section 4.7.3) which could make it unavailable for binding.

4.4 Determination of the DNA species present in the 'complexes'

The novel bands in these gels have been referred to as complexes; however it could be argued that these are products of a DNA modification reaction, such as ligation or nucleolytic digestion. It was therefore necessary to demonstrate that the DNA in these bands remained the same size as the unbound DNA. To do this, the complexes were dissociated after native electrophoresis, by soaking the gel in detergent (0.25% SDS). The gel was then run in the second dimension (figure 4.3). The ease with which the

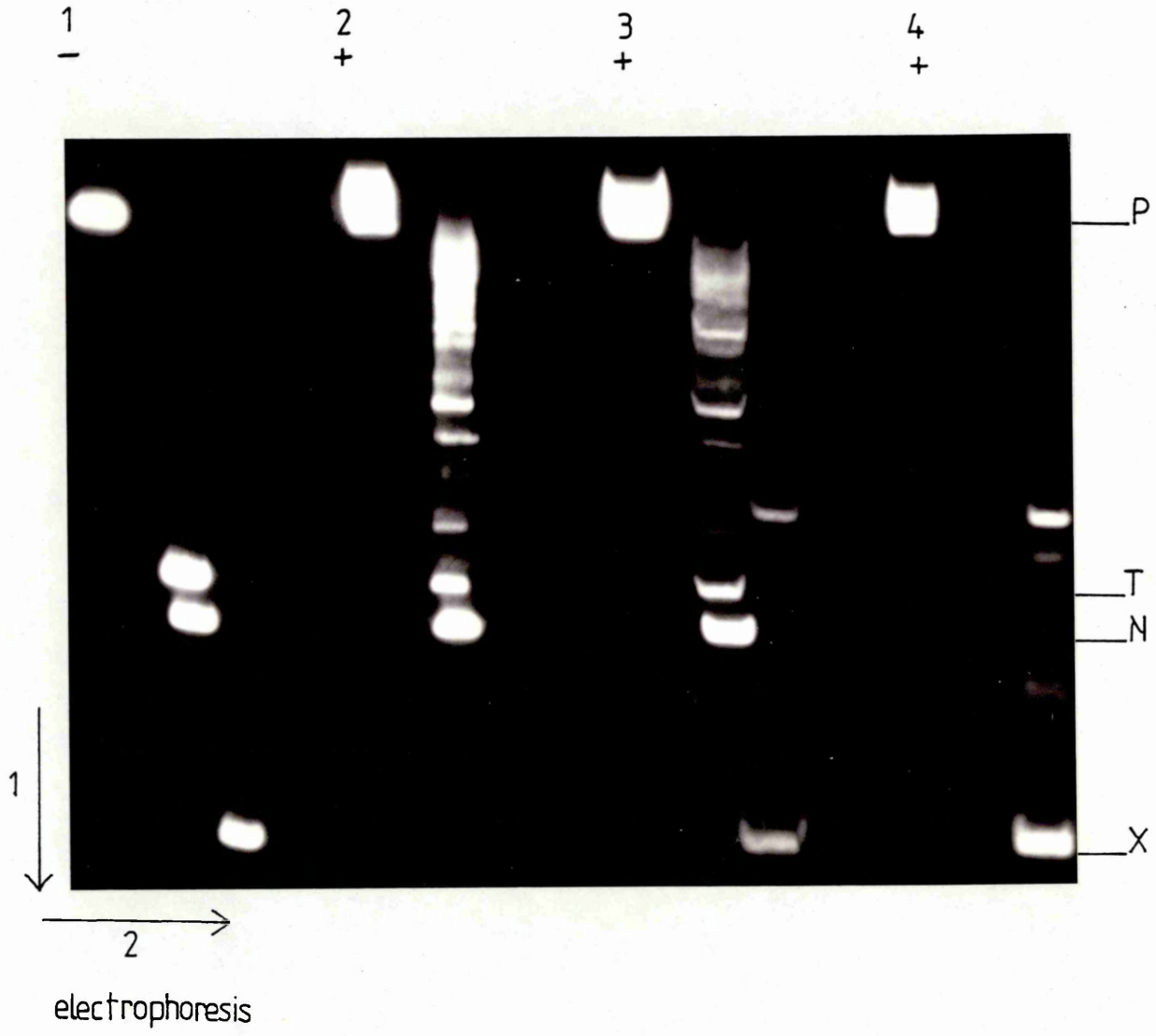


Figure 4.3 Two dimensional electrophoresis

Binding reactions were electrophoresed for three hours under native conditions at 4°C. The gel was then soaked in running buffer containing 0.25% SDS, and electrophoresis in the second dimension carried out in the same buffer at room temperature.

Binding conditions: 50mM Tris/Glycine (pH9.4), 1mM EDTA, 68mM KCl, 10% glycerol, 5ug/ml pUC18 carrier.
Gel conditions: 50 mM Tris/Glycine (pH9.4), 0.1 mM EDTA.

Lane 1 pNE200 and pLM5
2 pNE200
3 pNE200 and pLM5
4 pLM5

−: no TnsB extract +: 0.5 ug TnsB extract

P pUC18 vector (2.5 kb) T Tn7RE fragment (238bp)
N pUC18 fragment (216 bp)
X BamHI fragment from pLM5, containing first 22 bp unit
P,T,N are from pNE200; all DNA fragments are end-labelled.

complexes were dissociated, by a moderate detergent concentration, suggests that there is no covalent attachment of the protein to the DNA.

Whilst the control fragments fell on the expected diagonal, the majority of the labelled species from the TnsB-dependent bands moved as a column with the same mobility as the unbound Tn7RE DNA. This confirms that these bands do indeed represent complexes rather than new pieces of DNA, and suggests that at least under these conditions TnsB has no enzymic activities which chemically modify the DNA to affect its mobility.

The exception was the band corresponding to complex III (lanes 2 and 3). DNA from this band appeared to be migrating faster than that in the other complexes. It is unlikely that this was an artefact due to uneven current distribution in the second dimension, as it has happened identically in both lanes. The mobility of this band was the same as for the free 216 bp control fragment. There was also some 245 bp Tn7RE fragment present, seen most clearly in lane 3.

One hypothesis is that this DNA originates from the 245 bp Tn7RE fragment, with a TnsB-dependent cleavage removing about 30 bp. Such a cleavage would not be in a biologically relevant position at the 3' end of the fragment; at the 5' end it could conceivably be at the terminus of Tn7, generating a 207 bp fragment. However, in footprinting experiments no such cleavage by TnsB alone was detected (section 5.2.3).

A blotting experiment could be carried out to determine the source of this fast-moving fragment. If 'band III' does indeed contain the 216 bp DNA, there are two possibilities as to the nature of the complex. One is that this is an intermolecular complex involving both fragments. The experiment in lane 3 was designed to look specifically for intermolecular complexes, between two DNAs carrying 22 bp repeats. None were visible, and such complexes have not been detected by other means (section

4.11). Given this, it seems unlikely that we should detect intermolecular complexes with a DNA for which the protein has a lower affinity. However, such an interaction might have a biological relevance, if the transposase has to search for transposon ends among the cellular DNA. When the 216 bp fragment was tested alone in binding assays, no binding was detected (data not shown). It may be that the Tn7RE fragment is required for this binding, which would have implications for the searching mechanism (Richet et al., 1988).

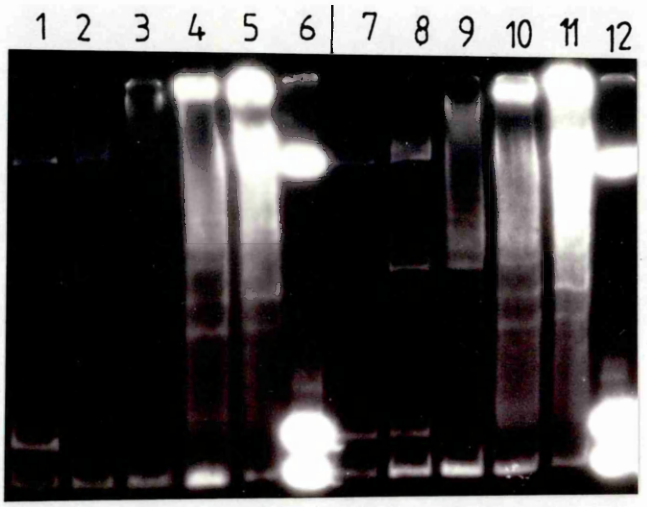
Alternatively, independent complexes of the two DNAs may be comigrating. In this case, the protein bound to the 216 bp fragment could be TnsB or a host factor in the protein preparation.

The behaviour of 'complex III' relative to the other complexes, as a function of protein concentration, was consistent. I do not feel that its unknown nature means that it should be disregarded as an indicator of available binding activity. For this reason I have made use of the pattern of complexes, including the 'IV-III shift' described above, in binding activity estimates in both chapters 3 and 4.

4.5 Effects of reaction and gel conditions

The initial experiments indicating that TnsB possessed sequence specific DNA-binding activity were carried out in Tris/HCl (pH 8.2) -buffered polyacrylamide gels (Ekaterinaki, 1987). However, it was found that the stability of the complexes and the sharpness of the bands were improved by increasing the pH to 9.4. The gels in figure 4.4A compare pH 8.2 and 9.4 for both binding conditions and gel running conditions. The pH of the reaction buffer had little effect on the pattern seen; however the high pH gel gave improved resolution. Similarly, if the binding reactions were carried out in 1X KGB buffer (McClelland et al., 1988), the same pattern of bands was observed (figure 4.4C lanes 1,2); the rationale

A. I



II

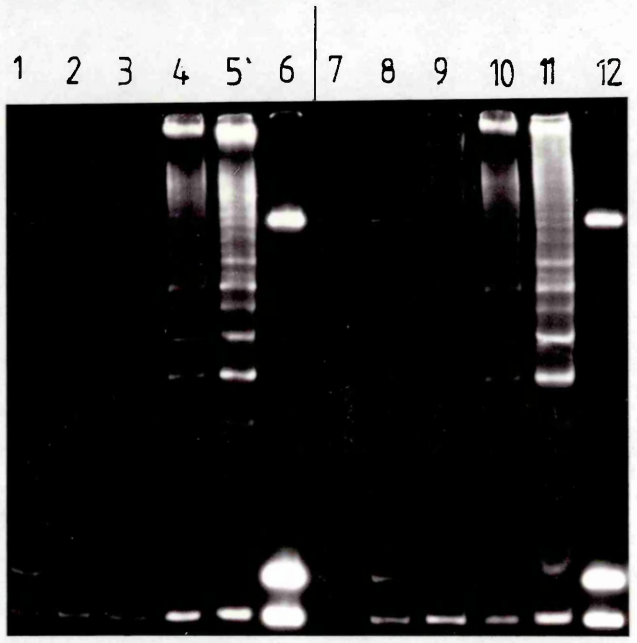


Figure 4.4 Effects of reaction conditions

A. Variation of pH.
Lanes 1-6 pH 8.2 reactions
7-12 pH 9.4 reactions

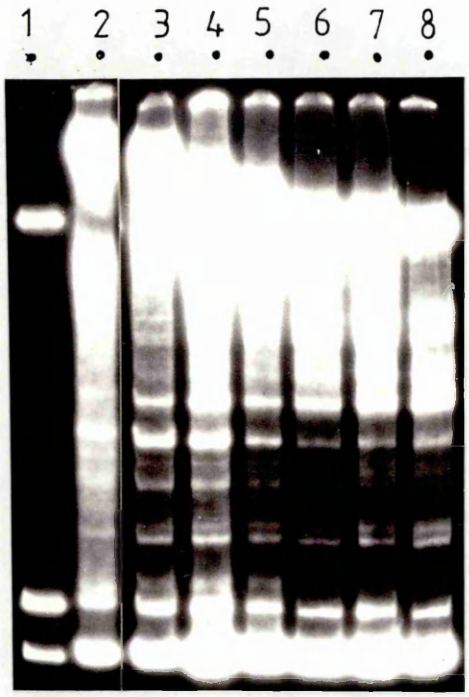
Lane	TnsB	Lane	TnsB	Lane	TnsB
1,7	20 ng	3,9	312 ng	5,11	5 ug
2,8	78 ng	4,10	1.3 ug	6,12	0 ug

Gel I 10 mM Tris/HCl (pH8.2), 0.1 mM EDTA
Gel II 50 mM Tris/Glycine (pH9.4), 0.1 mM EDTA

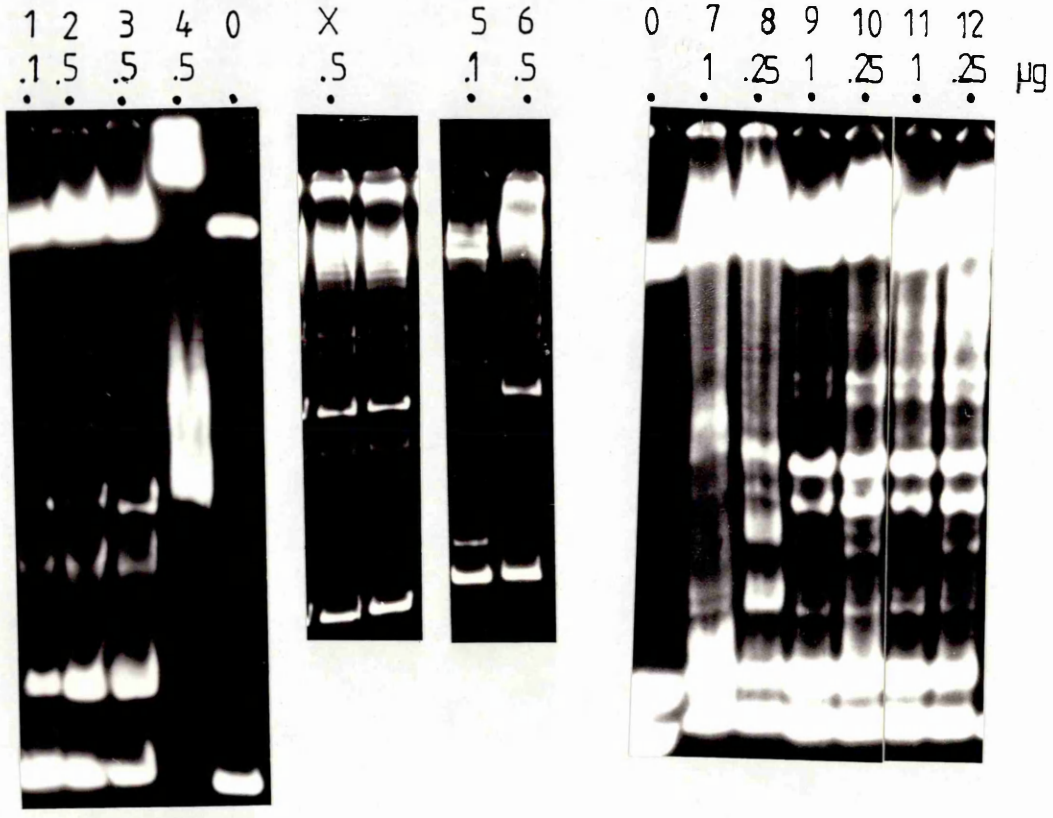
B. Effect of carrier concentration
Binding conditions: pH 9.4, 10 minutes, room temperature
Gel conditions: pH 9.4, 4°C
DNA substrate is pNE200 cut EcoRI, HindIII, NdeI, and endlabelled
Lane 1 no extract
2 0.5 ug extract
3-8 0.5 ug extract plus: 3 3 ug/ml pIC
4 10
5 30
6 100
7 300
8 1 mg/ml

pIC = poly(dIdC).(dIdC)
poly

B.



C.



C. Effects of temperature and divalent cations
All reactions were incubated under standard conditions except:

lanes 0	no protein		
1,2	KGB buffer	7,8	10mM $ZnCl_2$
3	5mM $CaCl_2$	9,10	4°C
4	10mM $MnCl_2$	11,12	37°C
5,6	10mM $MgCl_2$		

In lanes 1-4 the DNA substrate is pLM8 cut with BamH1 (see figure 4.1); otherwise pNE200 cut EcoRI, HindIII, NdeI. All end labelled.
The amount of protein added is indicated in ug.

X standard, pH9.4

behind testing this buffer was that anionic species such as polyglutamate have been observed to assist the solublising of insoluble, overproduced proteins (Darby and Creighton, 1990), and that glutamate is a physiological anion. The lack of variation with reaction buffer allows us to be confident in using the high pH buffer despite its non-physiological nature; most of the binding experiments presented were done under pH9.4 conditions, and run on Tris/Glycine (pH9.4) gels, and these are referred to as 'standard' conditions in the text and figure legends.

Some examples of the variations in experimental conditions tested are shown in figure 4.4B and C. The resolution of complexes in the gel was improved by reducing the salt concentration. In practice, this was reduced to 50-65 mM KCl; this residual level was due to the high salt concentration in the protein storage buffer, required to maintain TnsB in solution. The total amount of salt added to the wells was reduced by using a small reaction volume (10 ul, compared to the original 20-25 ul).

Carrier DNA was included in the reaction to absorb any non-specific DNA-binding activities. The effect of omitting carrier is shown in figure 4.4B; extensive retardation of the pUC control band was observed, with much of the radioactivity being caught in the well. As the carrier concentration was increased, this effect was reduced, with little shift of the control band at a carrier concentration of 10ug/ml. The gel shown used the double stranded synthetic oligomer poly(dIdC).poly(dIdC); the same effect was observed using pUC18 (figure 4.5).

Early experiments used sheared salmon sperm DNA as carrier (eg figure 3.1). Later work, however, used supercoiled pUC18 as the non-specific DNA (unless otherwise indicated). There are two reasons for this:

i) the sequence of the plasmid is known, so all the DNA sequences in the reaction are defined, in contrast to the unknown nature of the salmon sperm DNA.

ii) supercoiled plasmid runs as a single band, only a short distance into the gel. Hence the carrier is present only in the binding reaction and in the wells of the gel; once the sample has electrophoresed into the gel the carrier can no longer participate in any protein-DNA interactions. Sheared DNA, on the other hand, runs as a variable smear down the gel, and so has an unpredictable effect on protein and its interactions with DNA during the running time of the gel.

Carrier was included at 30-50 ng/reaction, ie 3-5 ug/ml. This is substantially less than the amounts used standardly in our laboratory in similar experiments, eg with Tn3 resolvase (Bednarz, 1989). In the case of resolvase, binding to the res site is highly specific, and the absolute value of the association constant is high; the vast excess of carrier DNA is required to take up some of this excess binding activity to allow complexes to form, rather than large non-specific complexes.

The effect of temperature on the binding reaction is shown in figure 4.4C. Over the range 6-37^{o}, very little change in the resulting pattern of complexes was observed. The order in which the components (protein, carrier and labelled fragment) were added to the reaction also had no effect (data not shown).

The effect of various divalent cations on the binding activity was examined using the gel assay. No significant effect, either quantitative or qualitative, was seen with any of the cations tested (figure 4.4C). Perhaps the most significant of these results are for magnesium and zinc. Many DNA metabolic enzymes require Mg^{2+}, for example DNase I, restriction enzymes and ligase. The absence of such a requirement for TnsB binding is consistent with the lack of evidence to date for any DNA metabolic activity for isolated TnsB.

One class of DNA-binding proteins is the so-called 'zinc finger' proteins (Miller et al.,1985: Klug and Rhodes, 1987). These are named for the structural motif responsible for binding; a zinc atom is tetrahedrally coordinated by cysteine and histidine residues, and the amino acids in between form a 'finger' which interacts with the DNA. The protein sequence motif found in these proteins consists of pairs of cysteine and histidine residues (or two pairs of cysteines in some examples), separated by 12 amino acids, which are largely polar or basic. TnsB does not contain this motif, which is consistent with the absence of any effect on adding zinc to the binding reaction.

The effect of manganese was not dependent on TnsB, as the same result was observed in a no-protein control (data not shown).

4.6 Affinity and specificity of binding

An apparent K_D for the binding reaction can be estimated from the plot in figure 4.2B, as described in the Introduction to this chapter. Assuming TnsB to be 75% of the protein present, and that it binds as a monomer, K_D is estimated as around 280 nM TnsB. As a comparison, the K_D for Tn3 resolvase measured by this assay is around 10 nM, and for TrpR under ideal conditions is around 0.5 nM.

Some of the difficulties associated with such estimates have been covered in the Introduction to this chapter. Further considerations relevant to TnsB are discussed here. The apparent K_D calculated here is approximately equivalent to the K_D for a single protein binding to a free DNA molecule; however, its value is also affected by the binding of further proteins to such singly bound species. It is therefore likely to be an underestimate, as any cooperativity in such multiple binding will make the binding appear stronger.

Each DNA fragment contains four 22 bp repeats; the calculated K_D is an average across all singly bound

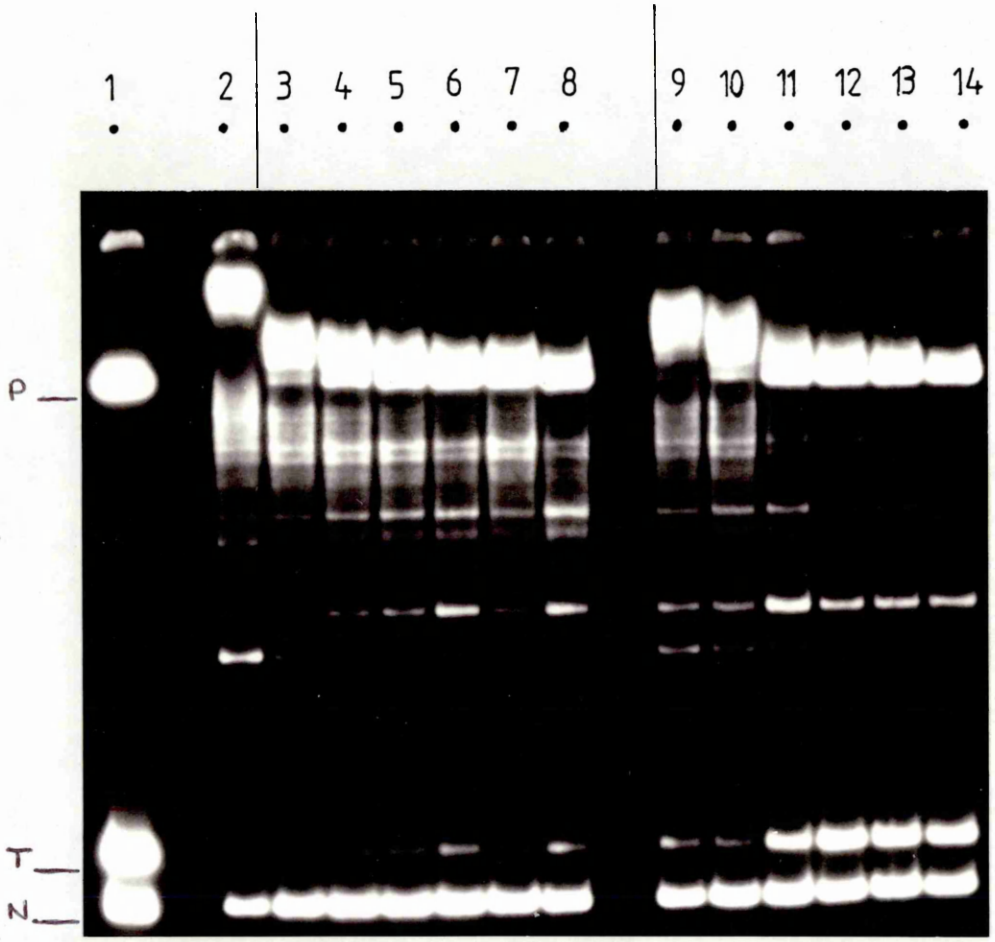


Figure 4.5 Competition experiments

Binding reactions including the indicated type and concentration of supercoiled competitor were carried out under standard conditions, with 1 ug extract (0.9 uM TnsB) per reaction. Labelled DNA is 0.5 nM, and supercoiled pUC18 is included at 2 ug/ml as carrier in all tracks.

Lane 1 no extract

2 1 ug extract

3	+ 100 ng pUC18	9	+100 ng pLM2
4	200 ng	10	200 ng
5	400 ng	11	400 ng
6	1 ug	12	600 ng
7	700 ng	13	800 ng
8	1.2 ug	14	1.2 ug

species. With multiple sites, the proportion of DNA fragments bound is not necessarily the same as the proportion of site occupancy; when 50% of the DNA is bound, it is likely that more than 50% of the sites are occupied. The stoichiometries of the complexes are not known, so we cannot make any statements about site occupancy.

To investigate the specificity of the binding activity for Tn7 end sequences, unlabelled DNAs were added to the binding reaction, either as linear fragments or as supercoiled plasmids (figure 4.5). The presence of pUC18 up to 70 fold molar excess over the labelled fragment had very little effect on the binding observed, although non-specific binding to the vector fragment was effectively competed by a tenfold excess. A shift from band III to band IV predominance was observed, which was earlier shown to correlate with an effective reduction in available binding activity; however, there was little change in the amount of free DNA.

In the presence of pNE200 (or pLM2, pLM3: see figure 4.1) the same effect was seen on the vector retardation (compare lanes 2, 9-14). However, as the concentration of pLM2 increased, the amount of unbound DNA increased, until 85% of the radioactivity migrated in the unbound band. The strongly retarded complexes IX-XII were the first to disappear, at 50-fold excess (lane 11), with V-VIII fading at higher competitor concentrations. The order of disappearance seen is consistent with our view of these as complexes in which the DNA is loaded with multiple protein units. The ratio of bands IV:III increased gradually from approximately 1:10 (lane 2) to around 10:1 (lane 14); this ratio change, and the percentage of unbound DNA, correlate with a reduction in available binding activity of around tenfold in titration experiments (cf. figure 4.2).

These data show that pLM2 is a more effective competitor for the binding activity than pUC18, indicating specificity in the binding of TnsB. This is in agreement

with the low level of binding to the 216 bp control fragment.

By defining a 'binding unit' in terms of the amount of specific fragment bound, ie

(fmoles Tn7RE) x % bound

the total DNA binding activity can be estimated. The number of units per reaction increased from 49.5 (90% binding of a tenfold excess of fragment; lane 9) to 215 (lane 13). When these are plotted against the total number of moles of Tn7RE fragment, the curve is hyperbolic, tending towards an asymptote at 230-250 fmol binding units, or 1 pmol of 22 bp repeats.

The TnsB monomer concentration was 900 nM, or 9 pmol/reaction. This 9-fold molar excess is less of a discrepancy than the 1000-fold suggested by figure 4.2, and indicates that under conditions of limiting protein, between 3% and 11% of the TnsB present is able to bind to the Tn7RE fragment. If we take a value of 10%, then the value of K_D can be reduced 10-fold, to 3×10^{-8}, which is more consistent with the examples given above.

In the gel shown, pUC18 at a 120-fold excess competed as well as pNE200 at 20-30 fold excess. Assuming that the total number of non-specific binding sites is equal to the number of basepairs of non-specific DNA (this assumption is valid if the binding sites are in excess), then we can calculate the number of specific (4/pLM2) and non-specific (2700/pUC18) binding sites; equating the extents of binding in the situations above:

$$(125 \times 2700)K_N + (1x4)K_S = (35 \times 2700)K_N + (31x4)K_S$$

where K_N and K_S are the affinities of non-specific and specific sites respectively. The ratio of the affinities, K_S / K_N, is 2025; ie there is a 2000-fold specificity for the Tn7RE sequences. Given the K_D calculated above, 3×10^{-8} M, the dissociation constant for non-specific DNA is

then around 6 x 10^{-5} M. The biological implications of the observed affinity and specificity are discussed in section 4.7.3.

4.7 Kinetics of TnsB binding

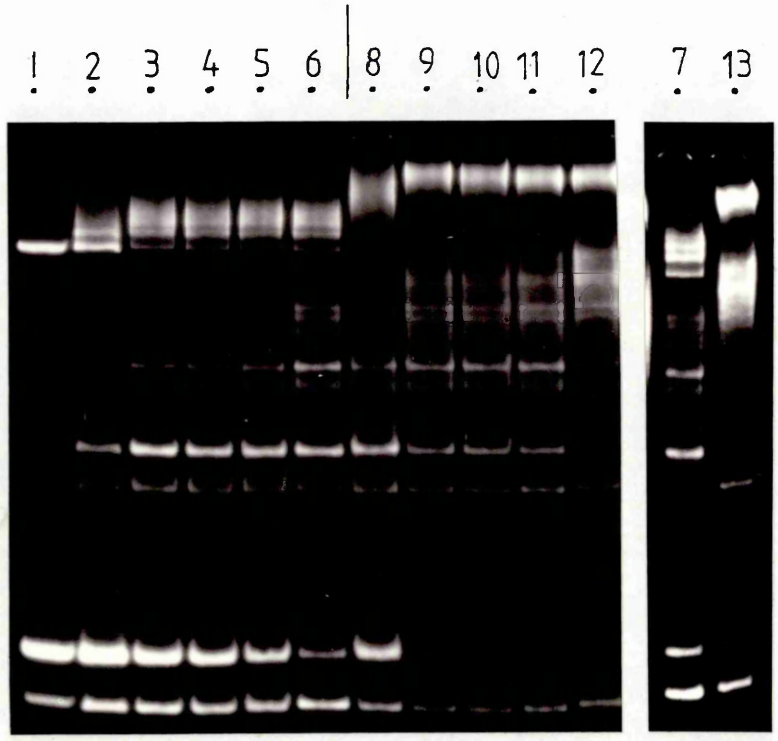
4.7.1 Association rate

Experiments to examine the kinetics of binding of TnsB to the Tn7 right end are shown in figure 4.6. The time course for binding at the lower protein concentration (0.25 ug; 25ug/ml) showed a gradual loss of free DNA; the radioactivity reappeared in higher complexes, eg VII, X and beyond, with very little change in the lower complexes over time. Binding appeared to be complete within 5 minutes (compare lanes 6 and 7). The pattern of relative intensities of the complexes remained fixed throughout the experiment.

At the higher protein concentration (1ug; 100ug/ml), the initial binding was perhaps slightly faster than for 25ug/ml, reaching a state with 80% bound within 1 minute. The pattern of complexes was as for the lower protein concentration. This remained static with little visible change until t=5 minutes, when there was a change in the distribution of complexes. The intensity of band IV decreased, with more counts appearing in band III and in indistinguishable complexes near the top of the gel. The pattern change was associated with further loss of free DNA. There was little further change over the next 5 minutes (lane 13), or indeed within a further half-hour (data not shown).

From this gel, an initial association rate of 2 pM/sec can be estimated, as the rate of loss of unbound DNA. Unbound DNA is used as a measure rather than complexes, because the free DNA band represents precisely the proportion of DNA unbound at the start of electrophoresis; complexes may have dissociated during gel running, allowing the DNA that was in them to run faster, producing a smear which may not be detected. The rate is

A.



B.

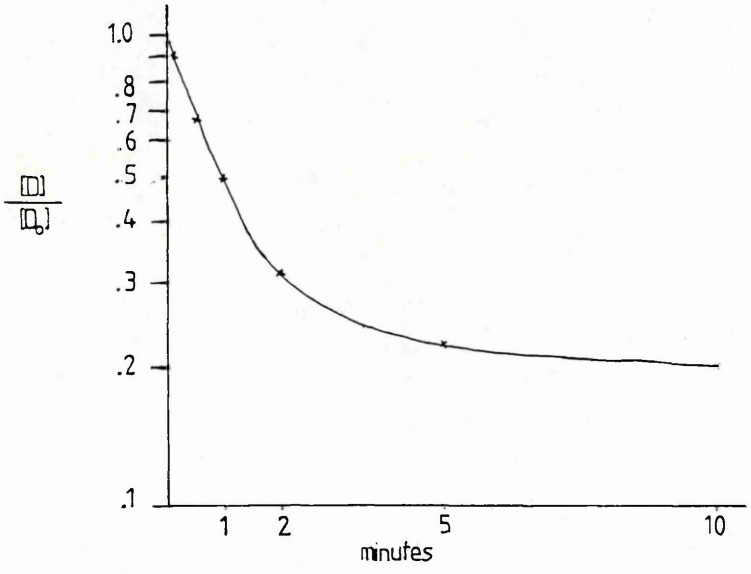


Figure 4.6 Kinetics of binding: association rate

A. Binding reactions with the indicated protein concentrations were incubated at 6°C, and samples withdrawn at the times shown and loaded onto a running gel (100 V).

Binding and gel conditions: standard.

DNA: pNE200 cut EcoRI, HindIII, NdeI, and end labelled (figure 4.1)

Lane 1 no extract

2,8	0 min	5,11	2 mins
3,9	0.5 min	6,12	5 mins
4,10	1 min	7,13	10 mins

lanes 1-7: 0.25 ug protein; lanes 8-13: 1 ug protein

B. Plot of the proportion of free DNA (logarithmic scale) against time, for 0.25ug protein.

taken from the beginning of the experiment rather than later, as there will be little dissociation at this stage. The measure of rate of loss of free DNA is an estimate of the rate of initial production of singly bound species; as the complex concentration is low early in the time course, the majority of interactions will be of this type. Binding of second, third (etc.) proteins to bound DNA do not contribute to rates measured in this way, but these will only be a minority of the interactions occurring.

A plot of log [free DNA] against time is shown in figure 4.6B. The linearity of the graph for early time points shows that the association rate is proportional to [DNA], as the protein is in excess, so its concentration can be regarded as constant. A first-order association rate constant of 2.5×10^5 M^{-1} s^{-1} is estimated from the graph. The plot becomes non-linear at later time points because complexes begin to dissociate as a free:bound equilibrium is established.

The proportion of pUC18 to become bound in the first minute is similar to that for the Tn7RE fragment, i.e. the association rate constants are similar. This could suggest that both reactions occur by the same mechanism; ie there is no apparent enhancement of the rate of binding to specific sites due to, for example, capture by a non-specific site followed by direct transfer.

4.7.2 Rate of dissociation

The rate of dissociation of the Tn7 RE / TnsB complexes was examined in experiments such as that shown in figure 4.7. Tenfold dilution of a binding reaction with a buffer containing an excess of supercoiled pLM2, reduced the likelihood of any dissociated protein re-associating with the labelled DNA, so allowing the off-rate to be observed independently of binding.

During the first minute after dilution, most of the non-specific binding to the vector dissociated (lane 2), with reassociation being quenched by the effectively

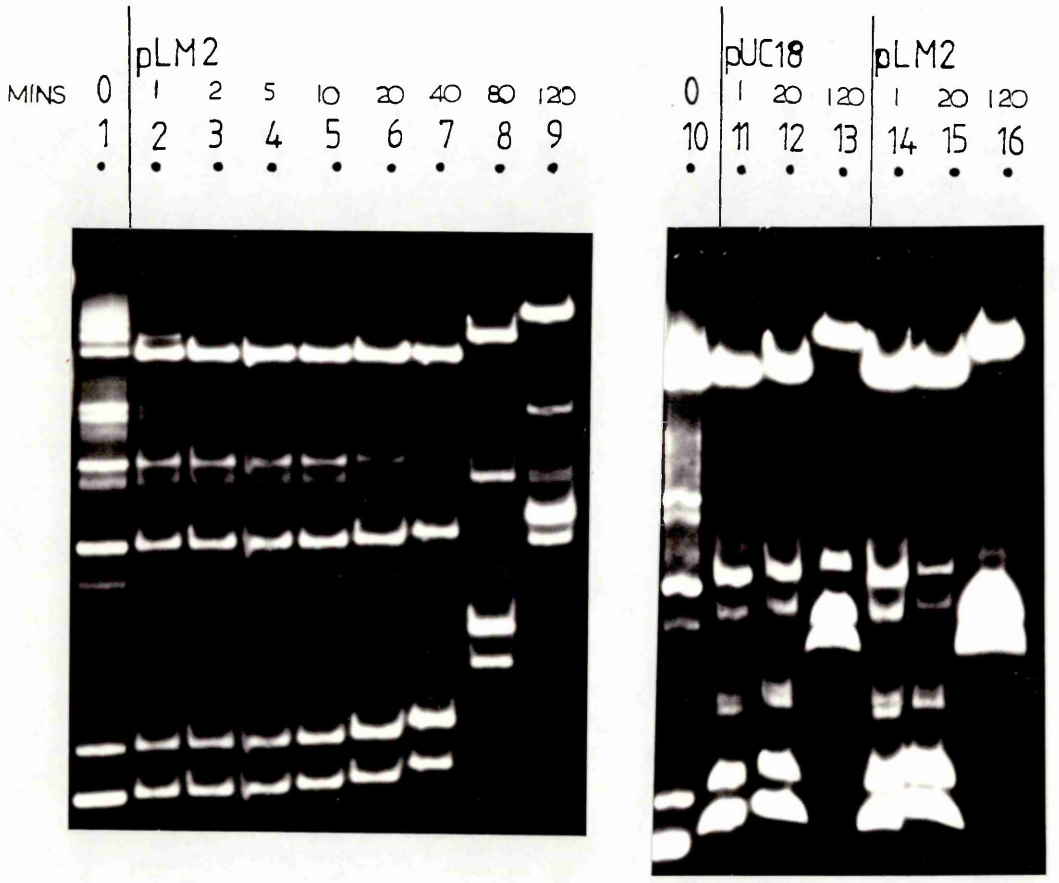


Figure 4.7 Kinetics of binding: dissociation rate.
After 15 minutes' incubation at 37°C, the binding mix was diluted tenfold in reaction buffer containing 2 ug competitor (as indicated). Samples were removed at the times shown (in minutes) and loaded onto a running gel.

lanes 1-9 0.5ug protein 10-16 1.0ug protein

0: reaction mix before dilution

Reaction conditions: standard, carrier 5 ug/ml pUC18
Gel conditions: standard, 200V running voltage

DNA: pNE200 cut EcoRI, HindIII, NdeI, and endlabelled.

excess 'carrier' in the reaction. This short lifetime was in marked contrast to the apparent stability of the more specific Tn7 end complexes (lanes 3-8). The higher order complexes IX-XII disappeared within 5 minutes. Complexes VII and VIII also began to disappear, after 10-20 minutes, and were barely detectable after 80 minutes.

The concentration of unbound DNA also increased gradually, at the expense of the complexes. The rate of change of concentration of free DNA can be estimated as around 4×10^{-14} M s^{-1}. It is likely that the source of the free DNA is dissociating singly bound species, but that has not been demonstrated. The advantage of looking at the free DNA is that its concentration is determined only by complex dissociation, whereas the concentration of any complex is dependent on its own dissociation as well as that of higher complexes.

A dissociation rate constant for the dissociation of singly bound species cannot be determined, as the concentration of the dissociating species is not known. A rate constant calculated from the rate of increase of free DNA concentration, would be effectively the rate constant for a limiting step in the dissociation pathway. The rate limiting step is not necessarily the dissociation of singly bound complexes; it is more likely to be dissociation of multiply bound species, which will be stabilised by protein/protein interactions.

However, a dissociation rate constant for singly bound species can be estimated from equation (8). Taking the K_D and k_f values above, k_b is 7×10^{-3} s^{-1}, with a halflife for these species of around 10 seconds. Similarly, the rate constant for dissociation of non-specifically bound TnsB is 15 s^{-1}, giving a halflife of 0.05 s.

The solution halflives calculated are in marked contrast to the stability of the complexes during electrophoresis. There is little smearing forward of the bands, which would indicate dissociation in the gel. It is

not known whether the effect is a genuine increase in halflife, or due to enhanced reassociation. A 'cage' effect has been proposed to explain the apparently enhanced stability of protein-DNA complexes in gels, whereby any dissociated protein is held in close proximity to the DNA, and the locally high protein concentration may then favour reassociaton. The nature of the stabilisation could be addressed by experiments like those of Fried and Crothers (1981). After allowing the complexes to enter the gel, cold DNA containing the binding site was loaded in the same lane. As the naked DNA overtakes the complexes, it can compete for any free protein, and this will be observed as smearing forward of the labelled complex band.

The tenfold dilution in this experiment would be expected to give an equilibrium binding pattern similar to that in lane 4 of figure 4.2, as this is the final protein concentration. However, some account must be taken of the large amount of Tn7 end fragment present. By comparison with figure 4.5, this amount of additional binding sites would be predicted to give a further reduction in effective protein concentration of around eightfold. As the pattern of complexes has not reached a state consistent with this concentration, it appears that even after 120 mins the dissociation has not reached an equilibrium, i.e the dissocation of the higher complexes is slow.

A similar experiment was carried out adding an equivalent excess of pUC18 DNA to the dilution mix. This was to determine whether the addition of extra DNA <u>per se</u> had any effect on the dissociation; a concern was that the protein was not fully soluble in the binding reaction, and that this solubility could be altered by DNA (see chapter 3). Under these conditions, the apparent amounts and distributions of the complexes remained constant for over an hour, whereupon a slow increase of unbound DNA and complexes I was observed over the next hour (lanes 11-13). This dissociation is occurring at about one quarter of the

rate of that seen in the presence of excess binding sites.

The apparent differences in stability of the complexes under the two conditions can be explained in two ways. The apparent reduced rate of dissociation in the absence of competitor might indicate that although dissociation is occurring at the rate seen with competitor, there is also some reassociation. The fourfold difference then suggests that 75% of all dissociated proteins are replaced.

Alternatively, the rate seen in the latter experiment may be closer to the 'true' rate, with the competitor causing an enhanced off-rate. One mechanism for this would be the direct removal of proteins bound to the labelled DNA, by the added competitor. Given the role we propose TnsB to have in transposition, in bringing together transposon ends, such intermolecular interactions would not be surprising, and cannot be excluded.

4.7.3 Biological implications of equilibrium and kinetic data

In order to mediate transposition, the 'transposase' must locate, and bind to, the transposon ends, amongst a vast excess of chromosomal DNA sequences present. Its ability to do this will be a function of its affinity for its site relative to that for non-specific DNA, and will be dependent on the total concentration of DNA. We can use the estimates of equilibrium constants to model the behaviour of the protein <u>in vivo</u>, bearing in mind that the true situation in the cell is more complex, with other general DNA-binding proteins competing with specific proteins for the DNA.

For non-specific binding,

$$\frac{[B]}{[B]^{o}} = \frac{K_D}{[D]+K_D}$$

where $[B]$ and $[B]_o$ are the free and total TnsB concentrations, K_D is the non-specific dissociation

constant (6×10^{-5} M), and [D] is the concentration of potential binding sites; assuming all random sequences to have a similar affinity for TnsB, then the number of sites is equal to the number of basepairs in the cell. This is perhaps an overestimate, due to occlusion of potential sites by binding of proteins like HU, or unavailability of DNA due to structure; however, given this assumption, the number of non-specific sites is 10^7, or 10^{-2} M . Hence

$$\frac{[B]}{[B]_o} = 6 \times 10^{-3}$$

ie only 0.6% of the protein is free.

For binding to the 22 bp repeats,

$$\frac{[S]}{[S]_o} = \frac{K_s}{K_s+[B]}$$

where [S] and $[S]_o$ are the free and total concentrations of specific site, and K_s is the specific binding constant (3×10^{-8} M). So, for 10% of the binding sites to be occupied ($[S]/[S]_o = 0.9$), a TnsB concentration of 3.3×10^{-9} M, or 4 molecules per cell, is required. Similarly, to occupy 99% of the sites, TnsB must be at 3×10^{-6} M, or 3600 molecules per cell. However, as shown above, most of the protein is bound to non-specific DNA, so the total protein concentration,

$$[B]_o = \frac{[B]}{6 \times 10^{-3}}$$

which is 670 molecules of TnsB per cell for 10% occupancy (ie one every 15 kb of the chromosome), and 5×10^5 per cell for 99% occupancy (one every 20 bp, which would completely coat the chromosome). The levels do not appear to be as high as this *in vivo* in a cell carrying a single copy of Tn7; using data from transcriptional and translational fusions (Ekaterinaki, 1987) it can be estimated that from a single copy of *tnsB*, there would be 390 molecules of TnsB in the cell.

If TnsB at that concentration is to find, and bind to, the transposon ends, additional mechanism(s) must operate. One such mechanism is cooperativity, whereby the binding of subsequent proteins is facilitated compared to binding of the first. An example where cooperativity is important is in binding of lambda repressor to its operator sites (Griffith et al.,1986). An alternative mechanism is that of '<u>cis</u>-acting' transposase, as observed for IS10 (Morisato et al.,1983) and IS903 (Grindley and Joyce, 1980). Because transcription and translation occur simultaneously in bacteria, the local concentration of 'transposase' is elevated close to its site of action, and loss by diffusion to chromosomal DNA is reduced. Tn7 is not believed to show this phenomenon, as <u>trans</u> complemented transposition is as efficient as <u>cis</u> action. However, the data can also be explained if transposase acts preferentially in <u>cis</u> and transposition is limited by other factors.

Another way to increase specificity is to increase equally the number of both specific and non-specific contacts, by doubling the size of the protein and its binding site; ie by binding as a dimer. K_D and K_S would then be squared, as would their ratio. However, a K_D as low as 3.6×10^{-9} means that the protein would spend all of its time bound to non-specific DNA, and would rarely find a binding site; the problem has become one of kinetics rather than thermodynamics.

4.8 Binding to partial right-end fragments

The binding experiments described so far all use a 'complete' Tn7 right end, which contains the four contiguous 22 bp repeats and is longer than the minimal transposition-competent end (70 bp; Arcizewska et al. 1988). The pattern of bands seen is very complex; it was of interest to create 'deleted' fragments, to look at the behaviour of a reduced number of 22 bp repeats. It was hoped that the binding pattern would be altered in a way

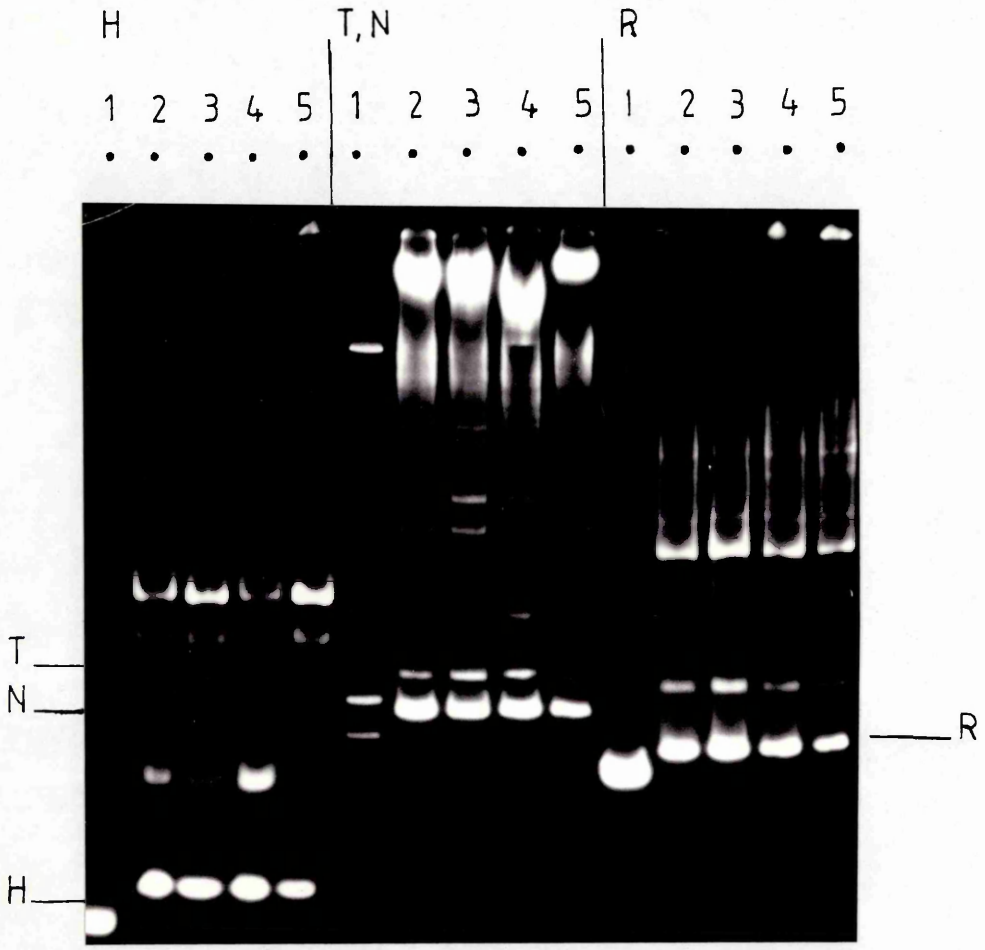


Figure 4.8 Fragments containing 1.5 and 2.5 22 bp repeats
pNE200 was digested with EcoRI, HindIII and BglII, and the two small fragments were purified and end-labelled, and incubated with TnsB extract at room temperature for 10 minutes, in pH 9.4 buffer. 3 ug/ml pUC18 was included as carrier.

Lane 1 no extract
2 0.5 ug TnsB extract
3 1 ug
4 2 ug
5 4 ug

The labelled DNA fragments were derived from pNE200:
T EcoRI-HindIII fragment containing 'complete' Tn7RE
N HindIII-NdeI control fragment
H HindIII-BglII fragment containing 1.5 22 bp repeats
R BglII-EcoRI fragment containing 2.5 22 bp repeats

A.
Oligonucleotide 1: R1 (pLM5)

5'- CTAGAGACAATAAAGTCTTAAACTGAAG -3'
3'- **TCTGTTATTTCAGAATTTGACTTCAGCT** -5'

Oligonucleotide 2: R1 plus 8 bp IR (pLM8)

5'- CTAGATGTGGGCGGACAATAAAGTCTTAAACTGAAG -3'
3'- **TACACCCGCCTGTTATTTCAGAATTTGACTTCAGCT** -5'

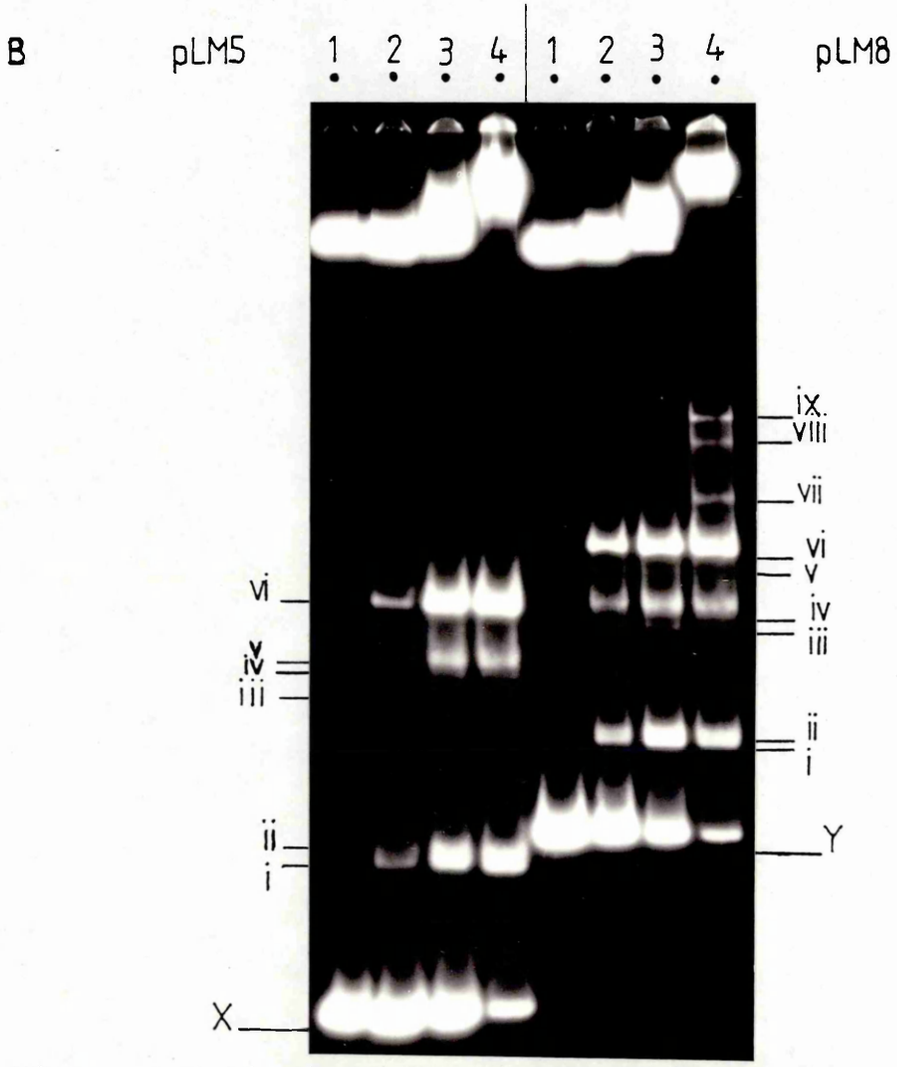


Figure 4.9 Binding to a single 22 bp repeat
A. Sequences of the oligonucleotides cloned in pLM5 and pLM8. Oligonucleotide 1 consists of the first right-end 22 bp repeat (R1).
Oligonucleotide 2 is R1 plus the 8 bp terminal inverted repeat of Tn7.
Tn7 sequences are shown in bold type.
B. pLM5 and pLM8 were digested to release fragments X and Y (figure 4.1), which were end-labelled and used in binding experiments as above (carrier 7.5 ug/ml).

Lane 1	no extract	3	500 ng
2	125 ng TnsB extract	4	1 ug

Binding and gel conditions: standard

which would allow us to define the contribution due to the different regions of the transposon end. These experiments would also confirm our belief that it is the 22 bp repeats to which TnsB binds.

Cleavage of pNE200 with BglII in addition to EcoRI and HindIII cuts the second repeat in the middle, generating two fragments carrying 1.5 and 2.5 repeat units (fragments H and R; figure 4.1). The former fragment (78 bp) gave rise to three strong and two faint, strongly retarded protein-dependent bands in the gel assay (figure 4.8, lanes 1-5). The pattern resembled that seen for the right end fragment in bands I, III, and IV.

For fragment R (2.5 sites) a plethora of bands was observed (lanes 11-15). The predominant bands again looked similar in distribution to I, III and IV. These correlations are insufficient to conclude that the bands are analogous, without looking at the stoichiometries of the components; however it does suggest that there is a loading pattern common to the three substrates.

Binding to a single 22 bp repeat was also investigated. For this experiment, complementary oligonucleotides were synthesised, consisting of the first RE 22 bp copy (R1) flanked by SalI and XbaI restriction sites (figure 4.9A). These were annealed and cloned into the polylinkers of pUC18 and pBEND2 (pLM5, pLM7). A second oligonucleotide was synthesised which contained R1 plus the extreme terminal 8 bp which are in perfect inverted repeat at the two ends of Tn7; this was cloned in the same way (pLM6, pLM8).

Binding to these substrates is shown in figure 4.9B. Both substrates showed a titration similar to the standard right end (figure 4.2), with respect to both the specific fragment and the vector. There was perhaps a greater proportion of free DNA at a given protein concentration, for these isolated repeat fragments, than for the complete end (eg lanes 4, 8; cf lane 9 of figure 4.2). This suggests that there is cooperativity with the four

repeats, which is absent on the isolated sites.

The single 22 bp repeat (fragment X of pLM5; figure 4.1) gave rise to six distinguishable protein-dependent bands (lanes 2-4). This was the same pattern as seen for the purified annealed oligonucleotide (data not shown), so we can be confident that these bands are due to complex formation with the cloned sequence rather than with the polylinker. The pattern was also very similar to that seen with 1.5 repeats (figure 4.8); by analogy with the isolated repeat, it is possible that the lowest complex for fragment H is in fact a doublet, and that there is a faint band below the upper doublet which is not visualised in figure 4.8, making a total of seven bands for 1.5 repeats.

The fragment carrying the additional terminal 8 bp gave a very similar pattern of bands, with three more slowly migrating bands at high protein concentration (lanes 5-8).

The fact that specific complexes were seen with an isolated 22 bp repeat confirms our belief that this at least contains the motif recognised by TnsB. The very similar patterns of complexes seen for the two clones, which differ only in the presence of the terminal 8 bp, suggest that binding to this 8 bp, if it indeed occurs, does not grossly alter the structures of the complexes formed, and that these 8 bp are less important for binding than the 22 bp motifs. This is borne out by footprinting (section 5.2; L. Arciszwska, personal communication), and is consistent with experiments on binding of transposases of other elements (discussed in chapter 1).

The large number of bands seen in these gels raises the question of the nature of the complex in each band. As discussed in the Introduction, inferences can be made from the number and distribution of the complexes, concerning the nature of each complex. Given the four 22 bp repeats, if each binds a single protein 'unit' and all combinations are separable, then we would predict 2^4 bands, i.e. 15

complexes plus free DNA. This would be consistent with the 12 identifiable complexes seen in figure 4.2; the variation in intensity may then be a consequence of certain conformations being preferred over others.

Other scenarios can be envisaged, and predictions made concerning the number of bands expected. Allowing occupation to be random or ordered, in one or two steps per repeat unit, and with or without 'shuffling' which generalises binding between sites or halfsites (Bednarz, 1989; Prentki et.al., 1987), the predicted number of bands ranges from 4 to 256, for four 22 bp sites. In the same way it is possible to predict the patterns for the 'deleted' fragments. However, to date none of these predictions match the observed data precisely.

Although it is useful to consider the different ways in which TnsB could be binding to its site, we should be wary of concluding too much from this speculation. The exact number of species is not truly defined, for any of the fragments. It is likely that the gels give an underestimate of the number of complexes formed; this could be due to comigration, or to dissociation of some of the complexes in solution or in the gel. However, it can also be argued that some of the complexes might be non-specific, or are due to the other protein species present in the TnsB preparation, and so should not be counted. This uncertainty of numbers of bands makes it impossible to draw any firm conclusions about the nature of the complexes; TnsB binding to Tn7 ends is found to be of a complexity beyond our understanding of the gel system. Further data are required in order to allow such conclusions to be reached. This would include experiments to determine directly the stoichiometries of the complexes (Fried and Crothers, 1981; Carey, 1988), and footprinting of isolated complexes to determine precisely which sequences are protected (Straney and Crothers, 1985; Kuwabara and Sigman,1987).

4.9 Bending of the Tn7 right end

As described in the Introduction to this chapter, band shift gels can be used to detect and locate bends and curvature of the DNA (figure 4.10A).

TnsB-induced bending of the Tn7 right end was investigated using a set of circularly permuted fragments of the right end. These were prepared from a tandem dimer of the right end (pLM1), cut with restriction enzymes with a unique site within the insert (figure 4.1).

There was some intrinsic bending in the Tn7RE fragment, shown by the variation in mobility of the naked DNA (figure 4.10B). The mobility plot (figure 4.10C) implicates the third repeat, R3, as the centre of an intrinsic bend. The DNA sequence in this region contains several runs of three or four AT basepairs, which have been shown to be preferentially located on the insides of bends (Travers and Klug, 1987).

These fragments were incubated with TnsB under normal binding conditions, and the products were separated on a native gel (figure 4.11). The variation in mobility of the free DNA was less apparent in this lower percentage gel. The pattern of complexes seen had several interesting features.

Firstly, there was a clear fluctuation in the mobility of analogous complexes for the different fragments. A plot of the mobility of the strongest complexes (III, IV, VII, VIII) as a proportion of the free DNA mobility, is shown in figure 4.11B. The maximum on the graph represents a maximum of retardation; by extrapolation the centres of these fragments (ie the bend centre) can be located between bp 57 and 69, ie in repeat 1. The minimum on the graph is the fragment with the greatest mobility; assuming such fragments have a bend at the terminus, the graph shows the bend to be around bp 62, which lies at the centre of a short region of imperfect dyad symmetry within R1 (see chapter 5).

The fluctuation in mobility strongly suggests that

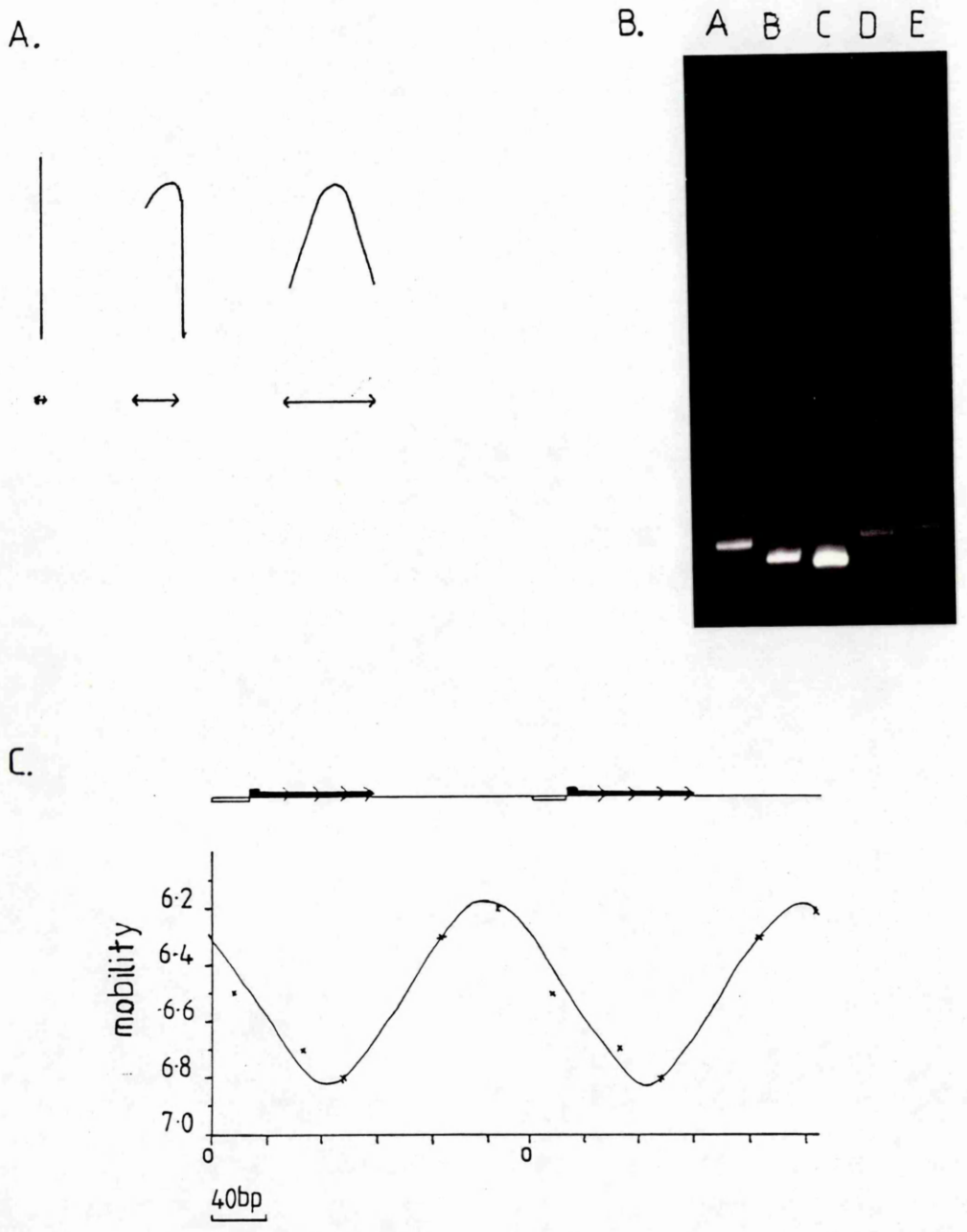


Figure 4.10 Bending of the Tn7 right end: intrinsic bend
A. The basis of the circular permutation experiment. DNA fragments with a bend located in the centre have a greater end-to-end distance, and experience greater retardation in the gel, than do fragments with a bend near the end.
B. Circular permutation of the right end: mobilites of the free DNA. Fragments A-E from pLM1 (figure 4.1) were purified, end-labelled, and run on a 10% polyacrylamide gel (Tris/HCl, pH 8.2) at $4^{o}C$.
C. Plot of the mobilities of the free DNA fragments. The distance migrated is plotted against the distance of the end of the fragment from the PstI site in pLM1 (ie the horizontal position of each point corresponds to the end of the fragment on the representation of pLM1 above the graph).

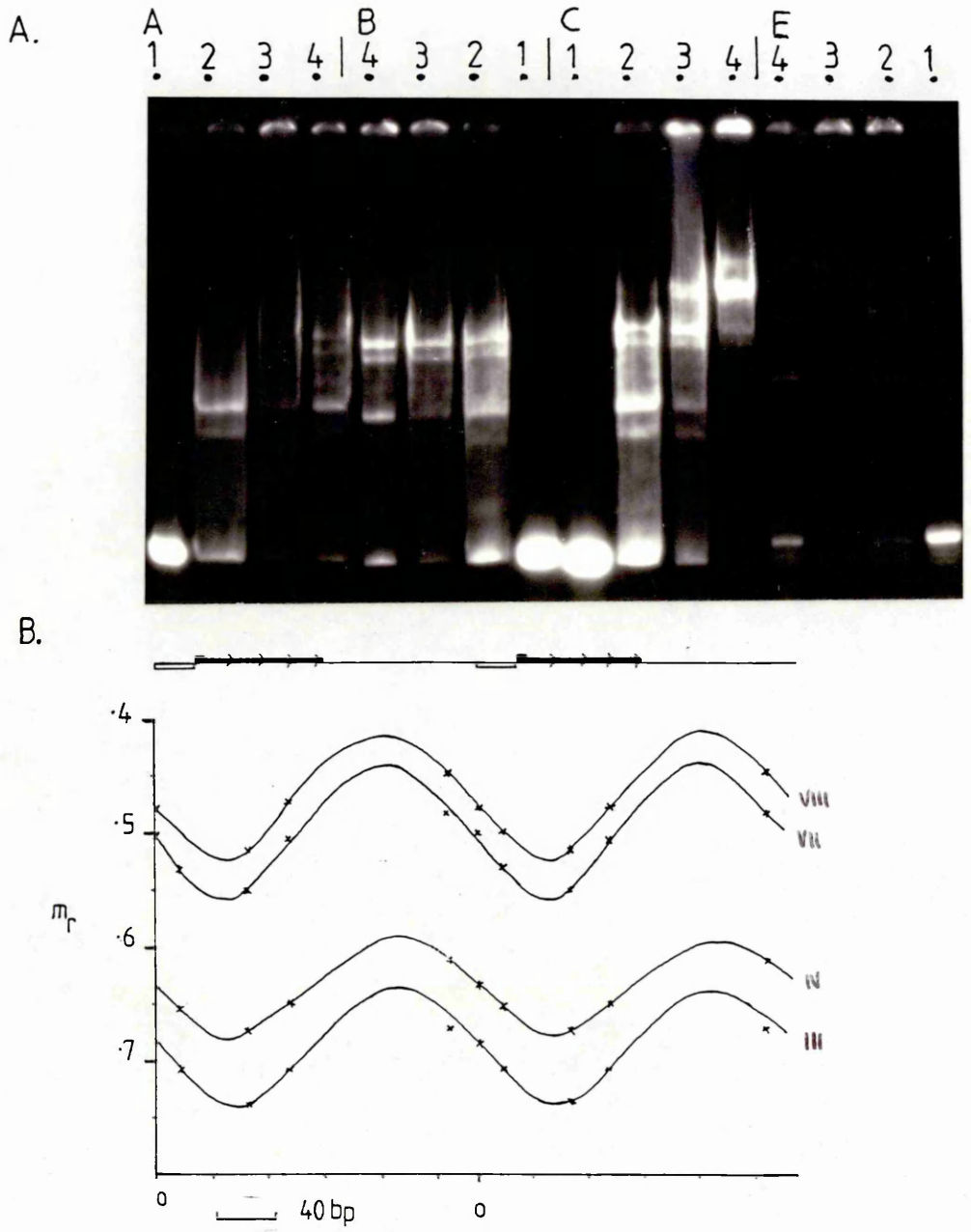


Figure 4.11 Bending of the right end by TnsB
A. Retardation of fragments A-E by TnsB. Binding conditions are pH 8.2, 25 ug/ml sheared salmon sperm DNA carrier, 10 mins incubation at room temperature.

Lane 1 no extract
2 5 ug extract
3 1.25 ug
4 0.3 ug

Gel conditions: pH 8.2

B. Plot of the relative mobilities of complexes III, IV, VII and VIII from A. Mobilities are plotted as a proportion of the mobility of the naked DNA, against the distance of the fragment end from the PstI site. (Fragment P is derived from pNE200 as shown in figure 4.1, and is equivalent to the pLM1 fragments.)
The horizontal axis is plotted as described for figure 4.10.

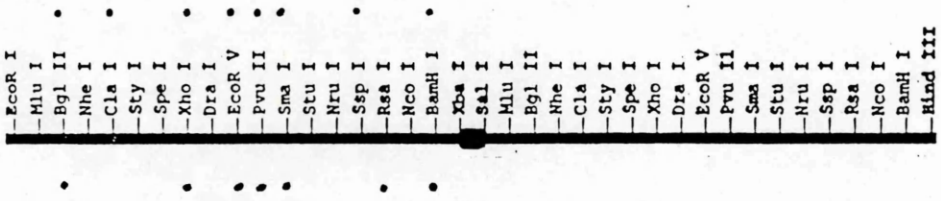
there is bending of the DNA on binding TnsB, with the bend located within the first repeat. The slight variation in the precise bend centre position of the four complexes shown is probably insignificant; however it is not inconceivable that there are real variations in the location of the bend centre in the different complexes.

A second point to note is that the pattern of bands, and the percentage of fragment bound, remained constant for all fragments, and is the same as for the standard right end fragment. The phenomenon is particularly noteworthy for the BglII and MaeI fragments, where the array of 22 bp repeats is disrupted. The constant percentage bound suggests that any cooperative interactions are not dependent on the repeats being contiguous; this is consistent with their spacing in the left end. The fact that the pattern is unchanged is more difficult to explain, especially for the BglII fragment in which the R2 copy is split. The lack of an effect could be fortuitous, if R2 is relatively unimportant for binding; indeed, R2 has the poorest match to the 22 bp consensus (figure 5.3). Alternatively, the constancy of the pattern, taken with the conservation of position of the bend centre in the different complexes, may indicate that the binding observed is a time-average either in solution or during gel running. This has been proposed for other DNA binding proteins in the gel assay, eg Tn3 resolvase (Bednarz, 1989) and IHF (Prentki et. al., 1987).

The synthesised Tn7 terminus cloned in pLM8, consisting of the R1 repeat with the terminal 8 bp, was used to investigate the bending of a single 22 bp repeat. The vector, pBEND2 (figure 4.12A; Zwieb et. al., 1989), contains a tandem repeat of a 120 bp polylinker containing 16 restriction sites; this allows the site of interest to be moved through the fragment in a similar way to the tandem dimer, but avoids cutting into the binding site(s).

The mobilities of DNA fragments of pLM8 showed very little variation in the absence of protein (figure 4.12B),

A.



B.

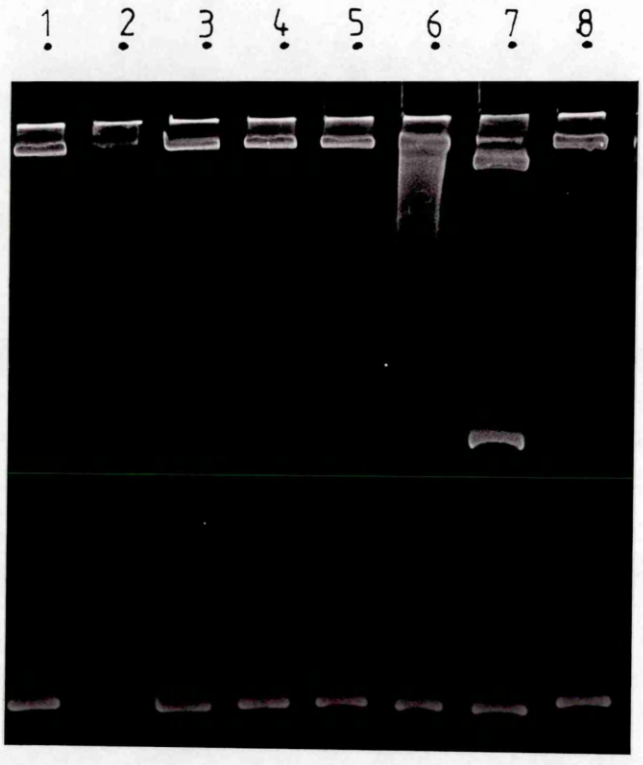
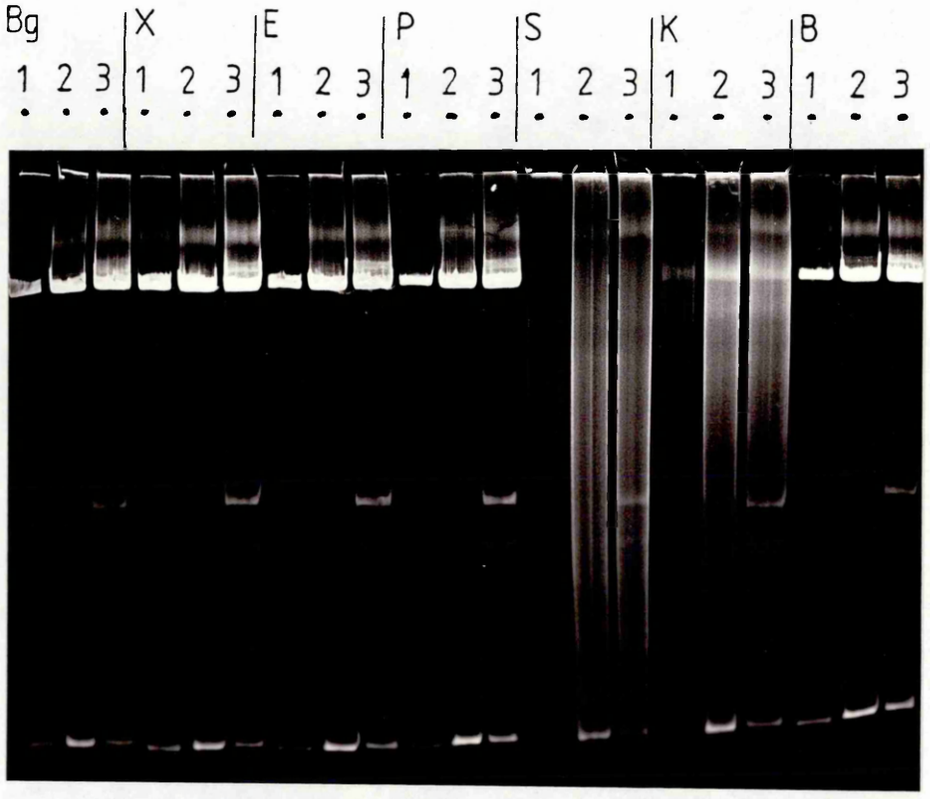


Figure 4.12 Bending of a single 22 bp repeat
A. The tandemly repeated polylinker of pBEND2 (Zwieb et al.,1988). In pLM8, oligonucleotide 2 is cloned between the SalI and XbaI sites. The restriction sites cut to generate the fragments for the gels shown are indicated above (B) and below (C) the map.
B. Fragments from pLM8, run on a 10% polyacrylamide gel at 4°C.

lane 1	BglII	lane 5	PvuII
2	(ClaI)	6	SmaI
3	XhoI	7	SspI
4	EcoRV	8	BamHI



C. pLM8 fragments were incubated with the indicated amounts of protein, and the binding reactions loaded onto a 5% non-denaturing gel.
Binding and gel conditions: standard.

lanes 1 no protein
2 0.125 ug protein
3 0.5 ug protein

Bg	BglII	X	XhoI	E	EcoRV	P	PvuII
S	SmaI	K	KpnI	B	BamHI		

nor in its presence (figure 4.12C). The variations are small enough to be barely distinguishable from fluctuations in the gel running; the available data are not sufficient to conclude that the single site is bent.

4.10 Interactions with cellular DNA binding proteins

The role of host factors in transposition is discussed in the Introduction to this chapter. Two proteins were investigated for binding to Tn7RE : IHF and FIS. IHF was originally isolated as a host factor required for lambda integration, and has been found to be involved in transposition of phage Mu (Surette and Chaconas, 1989) and IS10 (Morisato and Kleckner, 1987). The right end of Tn7 contains two putative IHF binding sites (as defined by Goodrich et al.,1990). Although no differences in transposition frequency were detected between IHF^+ and IHF^- strains in the mate-out assay (N. Ekaterinaki, personal communication) it was of interest to look for any interactions between this sequence and IHF.

FIS is involved in many site-specific recombination reactions, eg the Gin/gix inversion system of phage Mu, where it has an accessory role as a determinant of site alignment. FIS^- strains show a reduced level of Tn7 transposition (N. Craig, personal communication); hence binding experiments were carried out to test whether the role of FIS is in binding to the ends of the transposon, and mediating the formation of a correctly aligned 'transpososome'.

When partially purified IHF was used in the binding reaction, no bands were seen which could be IHF/DNA complexes (figure 4.13, lanes 1-5), nor did IHF have any effect on the binding of TnsB (data not shown). At low carrier concentrations, or with high levels of IHF, all the labelled DNA remained trapped in the wells of the gel (figure 4.13, lane 5). The carrier used in these experiments was sheared salmon sperm DNA, or poly(dIdC). poly(dIdC), as there is a known IHF binding site in the

IHF was a generous gift from G. Szatmari
FIS was a generous gift from R. Kahmann

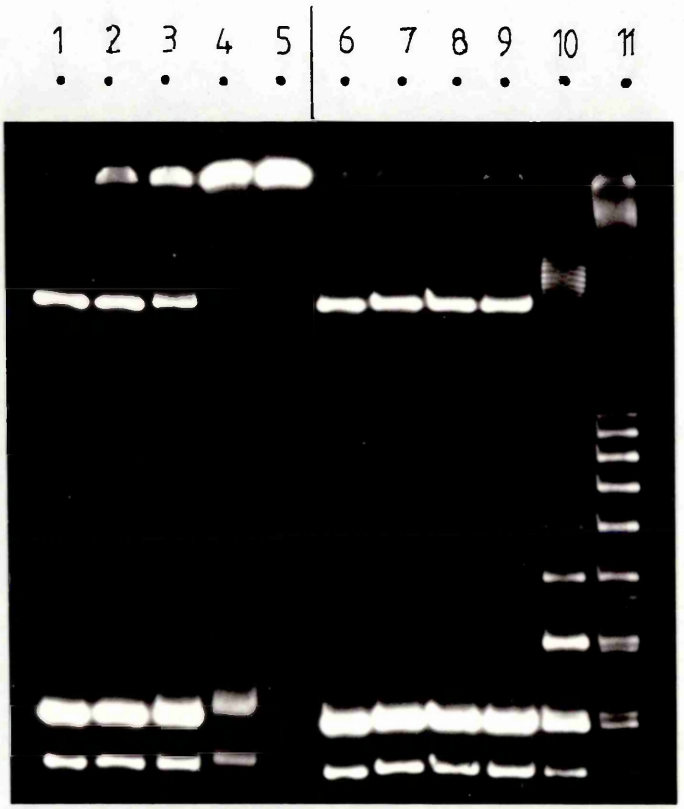


Figure 4.13 Binding of IHF and FIS
Partially purified IHF, and purified FIS, were incubated with labelled pNE200 under the conditions described in section 2.40.

Lane	1	IHF: no protein	6 FIS:	no protein
	2	0.6 ug	7	6.3 ng
	3	1.25 ug	8	12.5 ng
	4	2.5 ug	9	25 ng
	5	5 ug	10	50 ng
			11	100 ng

Gel conditions: standard

The DNA substrate used is pNE200 cut with EcoRI, HindIII, and NdeI, to generate fragments P (EcoRI-NdeI linear pUC8), T (EcoRI-HindIII cloned Tn7 right end), and N (HindIII-NdeI 216 bp pUC8 fragment); the fragments were end-labelled.

origin region of pBR322 and hence in pUC18. It is difficult to say whether there was preferential retention of the pUC18 fragment in these gels, which could be ascribed to this element, or whether the tenfold greater length of the vector over the fragment is sufficient to explain its retention.

In the presence of purified FIS, a regular ladder of bands was visible, with increasing extent of retardation at higher FIS concentrations (lanes 6-11). These bands appeared at the expense of the free Tn7 RE fragment. A second ladder of bands became superimposed on this at the higher protein concentrations, correlating with loss of the 216 bp fragment. The regularity of spacing of these bands is consistent with each acquiring an additional FIS dimer, rather than formation of specific complexes. Non-specific binding by FIS has previously been observed in this laboratory, with a similar result for almost all DNAs tested (G. Szatmari, A. Bednarz, K. McCurrach, S. Colloms; personal communications). However, the affinity of FIS for the Tn7 RE fragment appears to be greater than for the small pUC fragment, as it is bound at a lower FIS concentration; this may indicate some preferential binding, which might nucleate cooperative coating of the DNA by FIS under these conditions. Thus a role for FIS in Tn7 transposition *in vivo* should not be excluded.

4.11 Intermolecular binding

Various experiments were carried out to visualise complexes in which TnsB is bound to more than one DNA molecule. One approach was to carry out binding reactions containing two different sized specific fragments, and look for novel bands. This type of experiment can be extended to include second dimension electrophoresis, as shown in figure 4.3. In no case were any such novel bands detected, using two fragments carrying 22 bp repeats (eg figure 4.3, lane 3; other data not shown). Experiments including left ends also showed no novel complexes.

Another approach was to vary the total amounts of DNA and protein, and their ratio. Experiments with lac repressor showed appearance of looped structures, which were replaced by 'sandwich' structures at higher protein concentrations, under conditions of high concentrations of both components. These types of structures were not detected with Tn7 RE fragments, even with 100-fold higher DNA concentrations.

If, as we propose, the role of TnsB in transposition is to bind to the ends and bring them together, this result is not as expected. However, it may be that such complexes are formed but are too large to enter the gel. Alternatively, they may be unstable either under the binding conditions used, or in the gel. The gel conditions are optimised for the visualising of complexes; these conditions may not favour the stabilising of intermolecular 'synapses'. This is not unprecedented; although Tn3 resolvase is known to form active synaptic complexes with linear DNA, under the same conditions these are not observed on native gels (Bednarz, 1989). In the case of Tn7, it is also possible that such interactions are mediated by another protein, either host- or transposon-encoded, which is absent from these reactions.

4.12 CONCLUSION

This chapter presents gel retardation experiments that investigate the binding of TnsB to a linear DNA fragment containing 205 bp of the right-hand end of Tn7. The fragment includes four copies of a 22 bp motif, present in both left and right ends, which is thought to be the binding site of a 'transposase'. A consistent pattern of complexes was seen, which varied with protein concentration, and the binding showed some cooperativity. The sharpness of the bands was improved by raising the gel pH to 9.4, and reducing the salt concentration in the binding reactions. Changes in reaction buffer pH, in

temperature, or in order of addition of the components did not affect the binding observed, nor did addition of any of the four divalent cations tested.

The gel assay was used to estimate a dissociation constant for the binding of a single protein monomer to free DNA, and to estimate kinetic constants. There are caveats to the making of these estimates. Concerns relating specifically to the retardation assay system are discussed in this chapter. In addition, the calculations make use of parameters whose accuracy is uncertain. One of these is the concentration of active TnsB. The total protein content of the extracts can be measured, but the purity of the preparation, and hence the actual concentration of TnsB, have been estimated by eye from Coomassie-stained SDS gels. Without standardisation of the degree of staining (i.e. per microgram of protein), in conjunction with densitometry, these estimates are somewhat subjective. The purity estimate used here is probably an overestimate; the effect of this would be to lower the K_D. In addition, a value for the amount of TnsB that is active has been assumed; again, this is an estimate based on retardation assay autoradiographs, and should be taken as a guide to the activity rather than a precise measurement.

It has been assumed that TnsB is binding as a monomer; if this was not the case, the concentration estimate would be altered to account for the multimeric state.

A second concern relates to the estimation of the proportion of the DNA that is free compared to that in complexes. By considering the free DNA only, we can be fairly confident that this is representative of the proportion of free DNA when the sample was in the well of the gel. However, the estimates of the proportion of free DNA were made by eye from the autoradiographs, and this leads to concern about the linearity of the response of the film over the range of band intensities seen, and the reproducibility of this type of subjective estimate.

Given these caveats on the estimates for the kinetic and thermodynamic constants, these values should be treated with caution, and taken largely as a guide for the design of future experiments.

Second dimension electrophoresis of a retardation gel was used to demonstrate that the extract-dependent

bands were non-covalent complexes rather than modified DNA fragments. The same experiment showed the presence of a 216 bp pUC18-derived control fragment in one of the complexes; it is not known whether this is an intermolecular complex, or stable binding by TnsB (or another protein) to the control fragment. Further experiments to detect intermolecular complexes gave no firm evidence for their existence in the gel assay.

A set of circularly permuted fragments of the Tn7 right end was used to detect bending of the DNA. The unbound fragments exhibited a sinusoidal variation in mobility, indicating an intrinsic bend centered at the junction of the third and fourth 22 bp repeats. In the presence of protein, the complexes also showed variation in mobility, which allowed location of the bend centre to around the dyad centre of the first 22 bp repeat. It is likely that this represents a time-average of the bending induced by TnsB.

Similar experiments using an isolated 22 bp repeat gave no evidence for bending; however, the fragment was able to bind TnsB, indicating that the motif contains a sequence recognised by TnsB. Further experiments to determine the binding site are described in chapter 5.

The host proteins IHF and FIS, that are candidates for involvement in Tn7 transposition, were tested for binding to the Tn7 right end. No evidence was obtained for specific binding by IHF. A regularly spaced ladder of complexes was observed in the presence of FIS, as has been observed previously for FIS binding non-specifically to DNA. The Tn7 fragment was preferentially bound compared to the control DNA, so a role for FIS in transposition is not excluded.

TnsB, then, binds specifically to the ends of Tn7, and to a lesser extent to non-specific DNA. The binding site for TnsB is likely to be the 22 bp motif repeated in the transposon ends. It is proposed that a function of TnsB in transposition is to recognise and bind to the ends of the transposon, perhaps bringing them together in a synaptic complex. This proposed function, and the precise location of the bound protein, are investigated in experiments described in chapter 5.

CHAPTER 5

THE FUNCTION OF TnsB IN TRANSPOSITION

5.1 INTRODUCTION

The experiments presented in chapter 4 examined the binding of TnsB to the right end of Tn7; some of the results have been shown to hold also for the left end (L.Arciszewska, personal communication). A 22 bp motif, which is found in both ends of Tn7, is able to bind TnsB. The presence of the binding site in both ends of the element leads us to propose that a function of TnsB in transposition is to recognise the transposon ends, bind to them *via* the 22 bp motifs, and so bring the ends of the transposon together, in a protein/DNA complex which is able to interact with the target site and other Tns proteins to mediate transposition.

The technique of DNaseI footprinting, used to detect site-specific binding of proteins to DNA, was first described by Galas and Schmitz (1978), using *lac* repressor binding to its operator as a model system. Tne method has since been applied to many systems, and modifications to the basic technique have been introduced. These include use of alternative cleavage reagents such as the hydroxyl radical (Tullius and Dombroski, 1986; Hatfull et al., 1987) and copper-phenanthroline (Kuwabara and Sigman, 1987). Alternative forms of footprinting are alkylation protection and interference experiments. The former is similar to standard footprinting, with bound protein protecting the DNA from modification; the technique can also be used to monitor binding *in vivo* (Martin et al., 1986). Alkylation interference experiments define the nucleotides required to interact with the protein, and can also provide information about phosphate contacts.

This chapter presents experiments to examine the proposed role of TnsB in transposition. Footprinting experiments are described, which determine the precise sequences bound by TnsB. The results show a TnsB-dependent protected region which includes the 22 bp repeat, and that a short region of dyad symmetry centred at bp 14 of the

repeat unit might be important in binding.

Experiments are also described which test the proposed role of TnsB in bringing together the transposon ends. The proposal is supported by the footprinting data, which demonstrate binding to precisely those sequences that are present in both ends of Tn7. One experimental method for detecting a TnsB-mediated interaction between the ends is based on enhancement of circularisation rate; the bringing together of two transposon termini on a linear fragment containing a mini-Tn7, will pull the fragment's ends into closer proximity, which will enhance the rate of circularisation by DNA ligase. This approach was used by Mukherjee et al., (1988) to demonstrate protein-mediated DNA looping in the origin region of plasmid R6K. Related experiments by Buc's group using the lac repressor detected bending of the lac operator on binding repressor, revealed by an enhanced rate of ligation into circles (Kotlarz et al.,1986). In addition to the ligation enhancement, the technique can also detect any wrapping of the DNA, whether around the protein, or by interwrapping of two sites; the result will be a difference in the number of supercoils trapped by the ligation in the presence or absence of bound protein, and this linkage difference can be visualised by gel electrophoresis (figure 5.1A).

Wrapping of the DNA can also be detected by experiments of the type shown in figure 5.1B. Binding reactions are carried out using nicked plasmid substrates. Wrapping of the DNA around protein, or other structural changes such as underwinding of the helix caused by protein binding, cause changes in the DNA topology that are trapped by subsequent closure of the nicks using ligase, and are detected as a linkage difference on electrophoresis of the closed products. An alternative approach to this experiment is to use supercoiled DNA as the substrate for binding, and then relax it with topoisomerase I; any supercoils trapped by protein binding

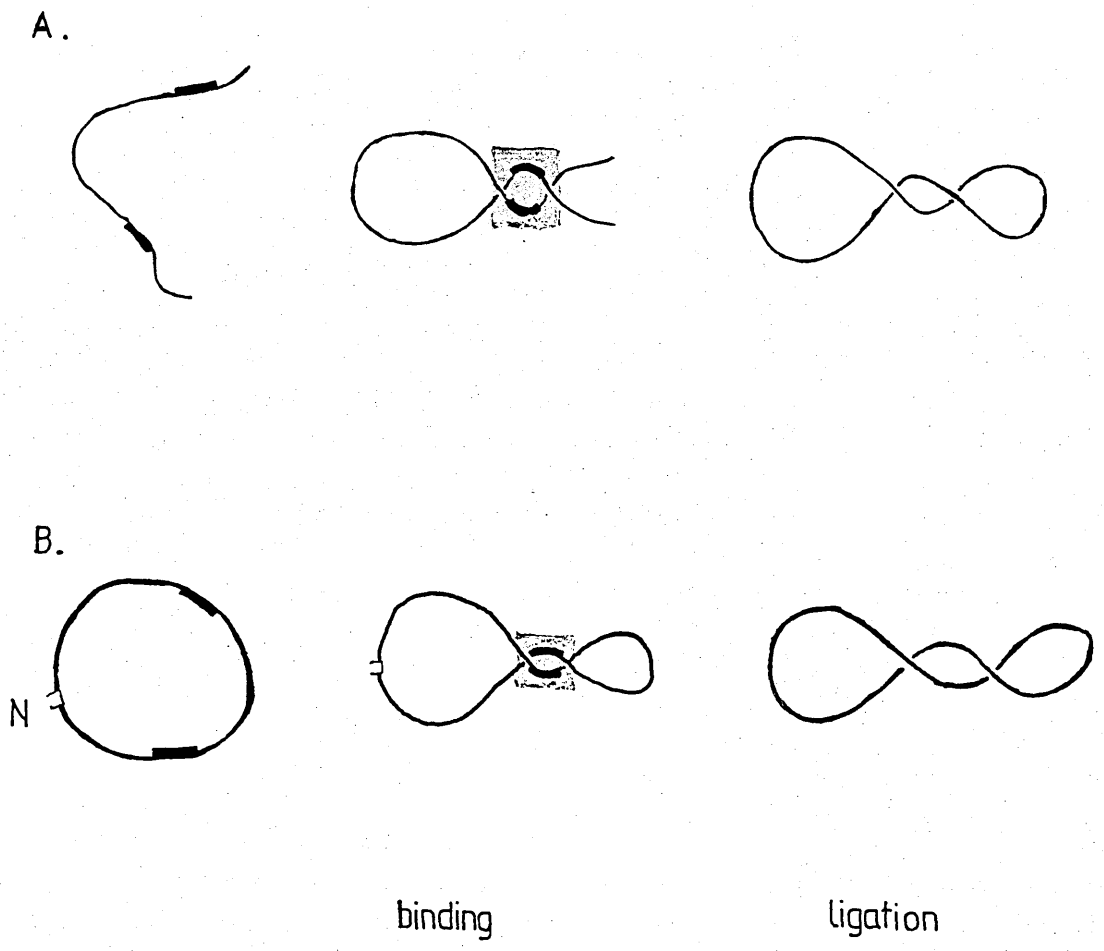


Figure 5.1 Basis of the linkage trapping experiments
In the example shown, the binding of the protein causes a linkage difference of -2 when it binds to the two recognition sites on linear (A) or relaxed (B) DNA substrates; these supercoils are trapped when the DNA ends are sealed by the action of DNA ligase.

N nick in DNA strand

▬ binding site

▩ protein

will be protected from relaxation. This approach is subject to potential artefacts due to incomplete relaxation by topoisomerase, but would be the method chosen for proteins which require supercoiling in order to form synapses.

Both techniques have been applied to Tn3 resolvase-mediated synapsis of res sites, and have demonstrated a linkage change of around -3 (Benjamin and Cozzarelli, 1989; M.Stark, personal communication), which is consistent with the two-step synapsis model for resolvase action (Boocock et al., 1986).

Experiments of both types, using ligase as a probe for structural changes, are described in this chapter. The results support a role for TnsB in mediating interaction between the ends.

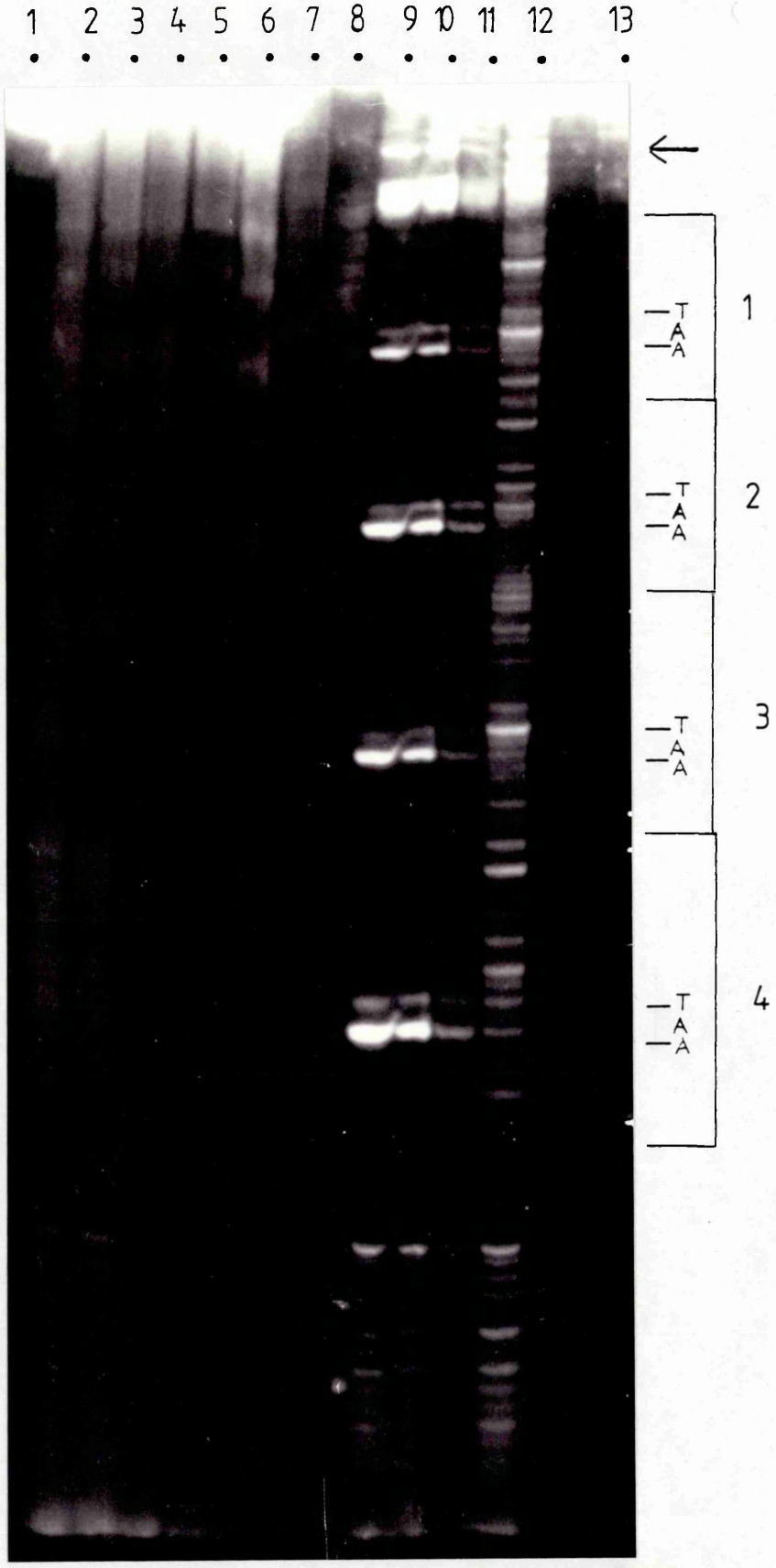
RESULTS AND DISCUSSION

5.2 Determination of the binding site of TnsB: DNA footprinting

In order to determine which sequences within the Tn7 right end fragment were involved in protein binding, protection experiments were carried out in vitro, using partially purified TnsB.

5.2.1 DNase I protection

Protection of DNA from cleavage by DNase I was investigated using standard methods (Galas and Schmitz, 1978). The protection observed was over a large region, including the 22 bp repeats (figures 5.2, 5.3). Within the protected region were several positions of enhanced cleavage. These occurred every 20 bp, ie one in the analogous position in each repeat. The 20 bp spacing also means that the cleavages lie down one face of the DNA helix, assuming standard B-form DNA. The increased reactivity of the DNA suggests a change in structure, possibly a bend or kink (as found for gamma-delta



The four 22 bp motifs are indicated R1-R4; the arrow indicates the end of Tn7.

Figure 5.2 Protection of the right end of Tn7 from cleavage by DNase I and hydroxyl radicals
Footprinting experiments were carried out as described in Materials and Methods. The labelled fragment was the BamHI-PstI fragment of pLM3, (figure 4.1), 3'-end labelled at the BamHI site. The protein sample is the substantially purified fraction IV, as used in chapter 4.

lanes 1-6	hydroxyl radical footprint:
1	no hydroxyl radical
2	no TnsB
3	2.2 ug/ml TnsB
4	8.7 ug/ml TnsB
5	35 ug/ml TnsB (*)
6	133 ug/ml TnsB
lanes 7,8	CT, AG cleavage ladders
lanes 9-13	DNaseI footprint:
9	500 ug/ml TnsB
10	125 ug/ml TnsB (*)
11	31 ug/ml TnsB
12	no TnsB
13	no DNaseI

In the reactions with no cleavage reagent, the TnsB concentration was the one marked (*).
The amount of protein added in the hydroxyl radical experiments was 0.3-21 ug, and in the DNase I experiments, 1.25-20 ug; the two experiments used the same amount of DNA. Hence, the relative protein and DNA amounts are very similar in the two experiments; the concentration units differ due to the different reaction volumes used (hydroxyl radical, 150 ul; DNase I, 40 ul).

resolvase binding to res: Hatfull et al.,1987; Salvo and Grindley, 1988) which alters the reactivity to DNase I by changing the width of the minor groove (Travers and Klug, 1987).

5.2.2 Hydroxyl radical footprinting

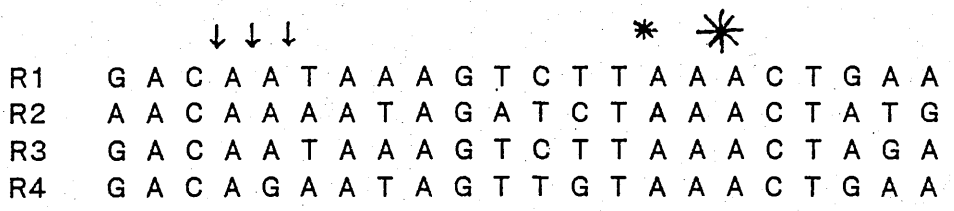
DNase I is a relatively large molecule, with a diameter of 4nm, compared with DNA, with a diameter of 2nm. This means that it can be sterically hindered from attacking phophodiester bonds close to, but not covered by, bound protein. It also has some sequence specificity; the footprint, then, provides a fairly coarse map of the region protected.

The hydroxyl radical, however, is a much smaller species, and can penetrate closer to the bound protein. As it mediates cleavage by attacking the sugar of the nucleotide, it has almost no sequence dependence, giving a very even pattern of cleavage in the absence of protein or structural abnormalities of the DNA (Tullius and Dombroski, 1986; Tullius et al.,1987).

Cleavage of TnsB binding reactions using hydroxyl radicals gave the pattern of protection seen in figure 5.2. In addition to the enhanced cleavages seen with DNaseI, there were additional regions of access, located halfway between the DNaseI-cleaved sequences. The 10 bp periodicity shows that these sequences are on the same face of the helix as the DNaseI-cleaved ones.

Another feature apparent in the hydroxyl radical footprint was that of sequential occupation of the 22 bp repeats. Protection of the first two copies was observed at low protein concentration (2ug/ml; lane 3), but repeat 4 showed only partial protection even at 133ug/ml protein (lane 6). Ordered occupation is likely to be due to a combination of the relative affinities of the sites for TnsB, and potential protein/protein interactions giving cooperative binding, and leading to the unidirectional 'building' of an array. The data presented here are

A. Right-end 22 bp repeats (5'-3')



B. R1: both strands

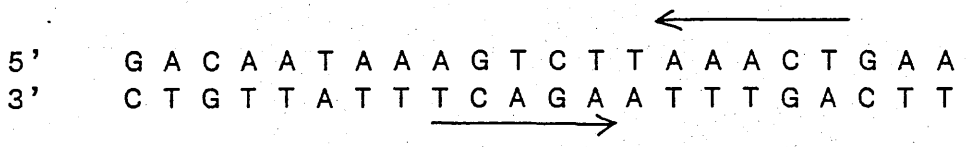


Figure 5.3 22 bp repeats: protection and enhancement
A. The sequences in the 22 bp repeats which are cleaved at an enhanced level by DNaseI (✱), and are cleaved by hydroxyl radicals but not by DNaseI (↓) are indicated above the aligned sequences.

B. Both strands of the R1 repeat are shown, and a potential dyad indicated.

The dyad takes the form: A G (T_3) T (A_3) C T

consistent with the results of Lydia Arciszewska in Nancy Craig's laboratory (personal communication).

5.2.3 Discussion of footprinting data

For both DNase I and hydroxyl radical experimental conditions, no cleavages were observed which can be attributed to a nicking activity of TnsB. In particular, there was no strong cleavage at the end of the transposon, on the strand examined, suggesting that TnsB does not have a cleavage function. However, if transposition proceeded via a Shapiro/Arthur and Sherratt intermediate, current models predict that it would be the other (bottom) strand on which the promary cleavage is made.

The positions of the non-protected regions are indicated over the sequences of the right end 22 bp repeats in figure 5.3. The cleavage pattern does not correspond with the beginning or end of the motif; the strongly cleaved bonds are 3' of nucleotides 14 and 16. Position 14 is also the centre of a region of imperfect dyad symmetry, which can be seen more clearly in figue 5.3B, where the complementary strand is also shown. The correlation of the positions of the enhanced cleavage and a possible dyad sequence suggest that this region might be significant in binding, perhaps in providing symmetry to which a protein dimer can bind.

The stronger cleavage, on the strand tested, does not lie at the centre of the dyad, but 2 bp to the 3' side. However, the helical nature of DNA means that the bond at an equivalent position on the other strand will generally be 2 bp 5' of this, i.e. at the central T residue.

Three models for the binding of TnsB to the 22 bp repeat are shown in figure 5.4. In all cases, the regions cleaved only by hydroxyl radicals are sequences that are accessible, eg between two bound protein molecules, but which DNase I cannot reach due to its large size (i.e. steric hindrance); the enhanced cleavage by DNase I could represent a structural change in the DNA, which makes it

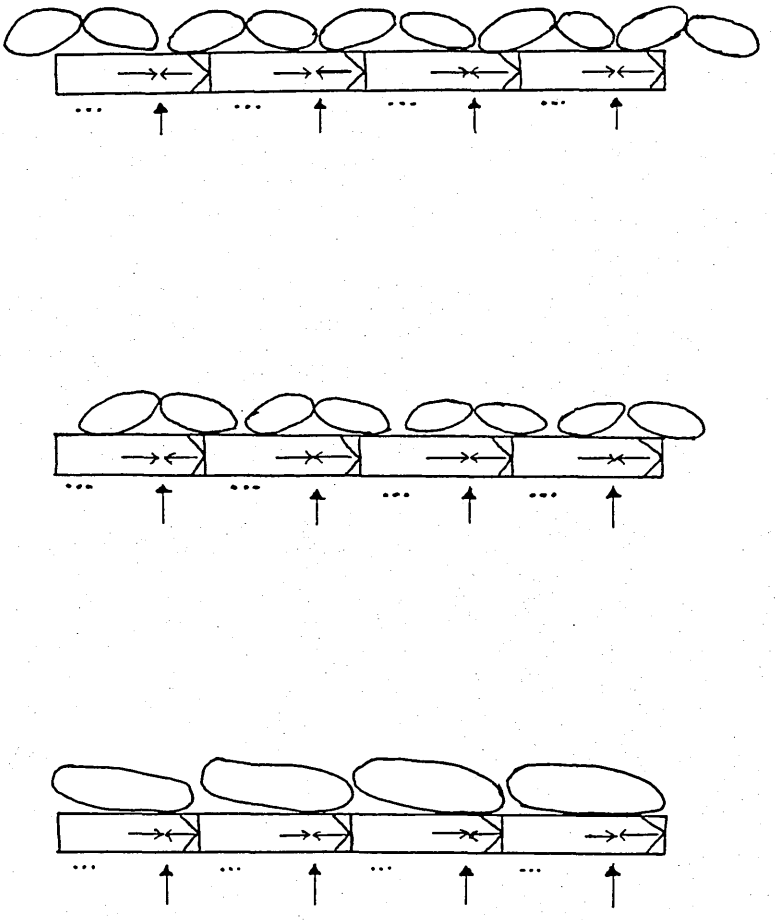
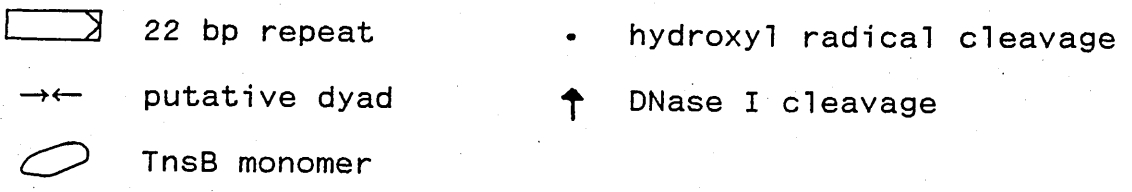


Figure 5.4 Models for TnsB binding
The implications of the different modes of binding illustrated here are discussed in the text. In I and II, TnsB is shown binding as a dimer; in III, a monomer is shown.



susceptible to DNase I attack.

The models differ in what is defined as the 'binding site'. Models I and II assume that the protein binds, perhaps as a dimer, to a symmetrical site, as found for many DNA binding proteins. In I, the symmetry lies at the ends of the binding site, as it does in the case of Tn3 resolvase binding to the subsites of res (Falvey and Grindley, 1987) and comes from the arms of the dyad in adjacent repeat copies. The conserved sequences at the ends of the proposed dyad (AG.....CT) would then occupy symmetrical positions in adjacent major grooves, suggesting that dimer binding could be by monomers recognising and binding in these two adjacent grooves. In II, the dyad is in the centre, within one 22 bp copy.

Given this type of binding, the next question is, how long is the binding site ? If it is the full length of a '22 bp repeat', then protein is binding to non-conserved sequences at the end of the array in the right end, and at the end of each repeat in the left end. Extension of binding to non-conserved sequences has been observed for other proteins, for example ArgR (G.Szatmari, personal communication). However, if binding only involves conserved sequences, then the extreme ends of the repeat are not included. In II, the sequences excluded would be the first five nucleotides, which are the most highly conserved in the motif; the strong conservation makes it difficult to justify excluding this sequence from the proposed binding site.

The third model takes the conserved 22 bp motif as the binding site, with protein wrapped around the helix, leaving one face unprotected. The DNaseI enhanced cleavage can be explained as a structural change as before. The function of the proposed dyad, then, could be in facilitating the protein-induced structural change. In the same way as the bending preferences of the DNA helix contribute to the positioning of nucleosomes (Drew and Travers, 1985), the dyad symmetry of this sequence may

favour a particular conformation of the DNA, which facilitates the binding of TnsB and/or its distortion of the DNA helix which is responsible for the observed enhanced cleavages by DNaseI.

Because this model makes use of the whole 22 bp conserved region, it is currently the model of choice. The proposed binding site is asymmetric, which suggests that the site is bound by a monomer. Protection of a large site is seen for IHF (Craig and Nash, 1985), so as TnsB is much larger than IHF, it is not unreasonable for it to bind over two turns of the DNA helix.

To extend the footprinting work, it is desirable to obtain data on the other strand of the right end, and both strands of the left end. Protection data from the left hand end would be useful because of the non-contiguous organisation of the 22 bp repeats; the extent of protection of the isolated 22 bp repeats, and of flanking sequences, will help to distinguish between the models suggested. Footprinting of the isolated 22 bp repeats in pLM5 and pLM6 would complement the data from complete left end footprints.

The centre-to-centre spacings of adjacent 22 bp repeats in the left end are 56 and 65 bp respectively, which places analogous positions in adjacent repeats on opposite sides of the helix, in contrast to the situation in the right end. Interactions between proteins bound at adjacent sites might introduce distortions into the DNA helix, which could be detected by footprinting reagents.

The sequence in the left end repeats at the position of enhanced cleavage is GG, compared to AT in the right end. It would be interesting to see what effect, if any, this has on the footprint, which might suggest a function for the sequence difference between the ends.

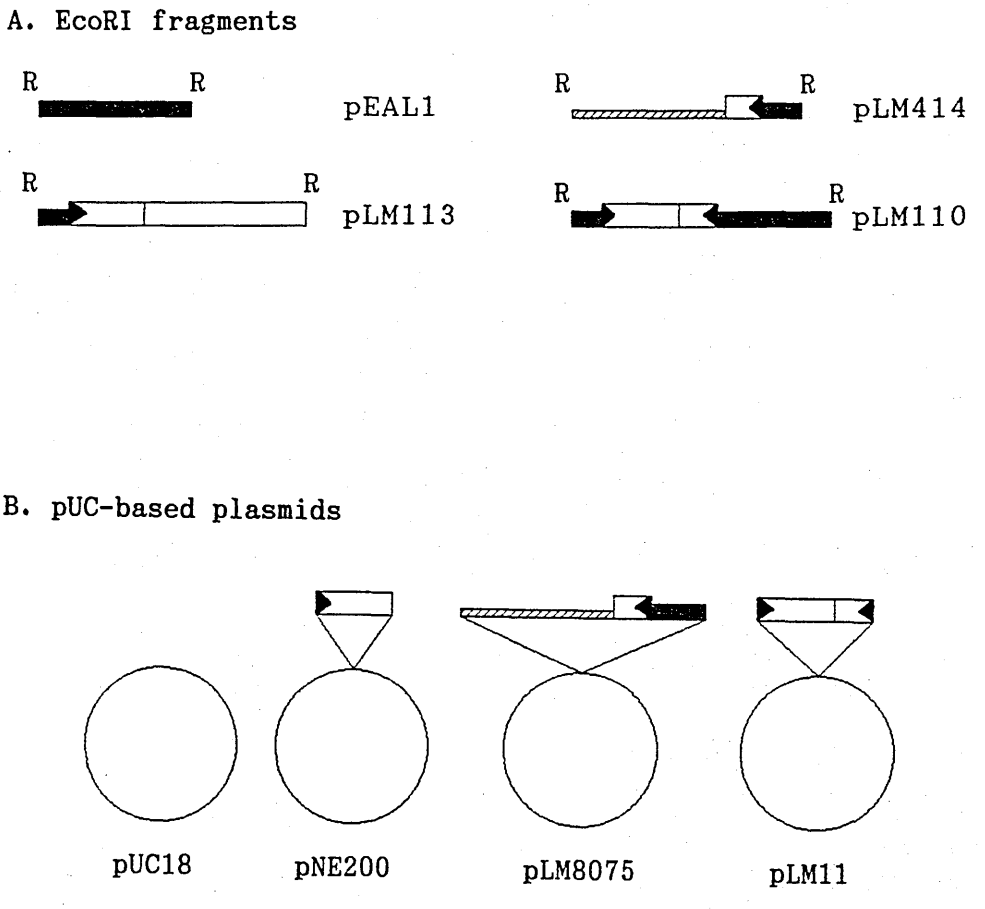


Figure 5.5 Substrates used in ligation experiments
A. The fragments shown were generated by digestion of the indicated plasmids with EcoRI, and contain no Tn7 ends, the right end, the left end, and both ends respectively. In all cases, a 4.2 kb vector fragment is also generated, and this is included in the binding/ligation reaction.
B. The plasmids are all based on pUC18, and contain no Tn7 ends, the right end, the left end, and both ends respectively. The plasmids were nicked with DNase I for use in binding/ligation reactions.

attTn7 sequences
CAT gene sequences
Tn7 ends
R EcoRI site
Tn7 sequences

5.3 Use of DNA Ligase to detect interactions of Tn7 ends

5.3.1 Ligation of linear substrates

The effect of TnsB on circularisation of Tn7-end bearing fragments was investigated, using ligase to reveal any change in ligation rate, or in the trapped supercoiling in the products. The linear substrates used are shown in figure 5.5. The 1 kb EcoRI fragment from pEAL1 is a chromosomal fragment containing the hotsite; the point of Tn7 insertion is 90 bp from the EcoRI site. pLM113, pLM414, and pLM110 contain right, left, and both Tn7 ends respectively, in their natural insertion position in the hotsite; in the latter case the ends are in the correct relative orientation to form a transposon, and this element does in fact transpose.

Binding/ligation reactions with the EcoRI linears were performed as described in Methods. The 4.2 kb vector fragment remained in the reactions as an internal control for ligation. In order to promote intramolecular ligation over intermolecular, the DNA concentration was kept low; this was achieved by using a large reaction volume. A high ligase concentration (2.5 units per reaction) was also found to facilitate circle formation rather than linear multimerisation; the reason for this effect is not clear. KCl from the protein storage buffer was found to inhibit ligation when above 50 mM; the large reaction volume also served to dilute this to 7mM, and sodium chloride was added to 50mM, which maintained protein solubility without affecting ligase activity.

A ligation time course with the EcoRI linear fragments was carried out, and the samples were run on two gels. Figure 5.6 shows a 0.75 ug/ml EtBr gel; under these conditions all the closed circular topoisomers run as a single band, allowing relative rates of closure to be compared. No change in the ligation rate was seen on addition of TnsB, even for shorter time points or with lower temperature ligation (eg lanes 3, 4 or 5, 6). Much of the total circularisation of the Tn7 fragment occurred

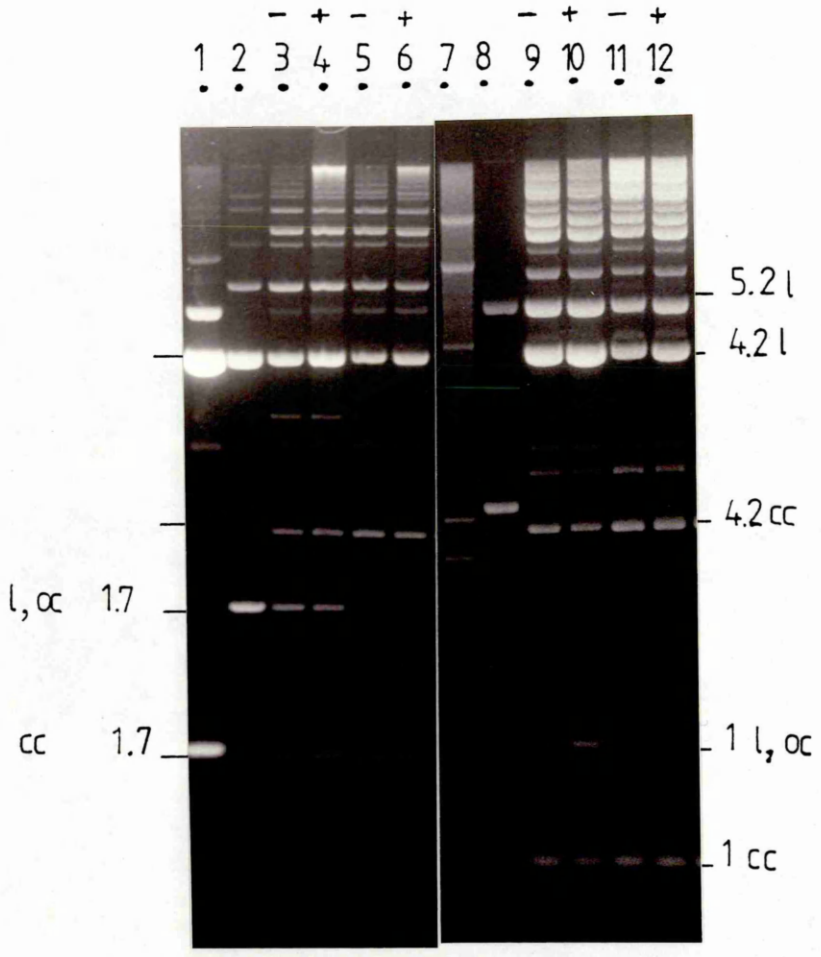


Figure 5.6 Rate of ligation of a linear mini-transposon
Binding/ligation reactions were carried out as described in Methods, using the indicated plasmid substrates cut with EcoRI. Samples were withdrawn at the indicated times, stopped with SDS, precipitated with ethanol, and loaded on a 1.2% agarose gel containing 0.75ug/ml ethidium bromide. TnsB concentration: 20ug/ml. Presence or absence of protein (+, -) is indicated above the figure.

Lane	1	pEAL1 untreated
	2	pLM110 untreated
	3,4	pLM110: 2 min ligation
	5,6	pLM110: 5 min ligation
	7	\ HindIII markers
	8	pACYC184 uncut
	9-12	pEAL1: as lanes 3-6

l linear oc relaxed circle cc closed circles

EcoRI fragments:

pEAL	1.0, 4.2 kb
pLM110	1.7, 4.2 kb

within the first five minutes, with very little further change in the amount of closed circular material after this.

The lack of enhancement of ligation rate may be because TnsB is not acting to bring together the transposon termini, but is binding to the two ends independently. This may be the true role of TnsB in transposition, with another factor being required to bring the ends together. Another explanation is that TnsB requires supercoiling in order to form a synapse. Neither of these interpretations are consistent with the data below.

Alternatively, it may be that the linear substrates used in these experiments are too long, and so the EcoRI ends are too far away from the protein binding sites to be brought significantly closer by interactions between the transposon ends.

The supercoiled products migrated as a single band; this indicates that there is no significant knotting of the products, as knots would separate on this gel, running ahead of the unknotted circles.

The same samples were run on a gel containing 3ug/ml chloroquine; under these conditions, relaxed and slightly negatively supercoiled plasmids run as positively supercoiled (figure 5.7). In the presence of TnsB, the 1.7 kb pLM110 fragment showed a population of more slowly migrating topoisomers; that is, topoisomers with less positive supercoiling in the presence of chloroquine. The linkage shift was of 2-3 topoisomers, ie the binding of TnsB to the transposon termini introduced 2-3 negative supercoils. The population migrating with the same mobility as the blanks was possibly derived from unbound fragments.

With substrates containing only a single Tn7 end, the linkage change in the presence of TnsB was much reduced. There was a slight shift with pLM113 (right end; lanes 10, 11), of less than one negative supercoil; no change

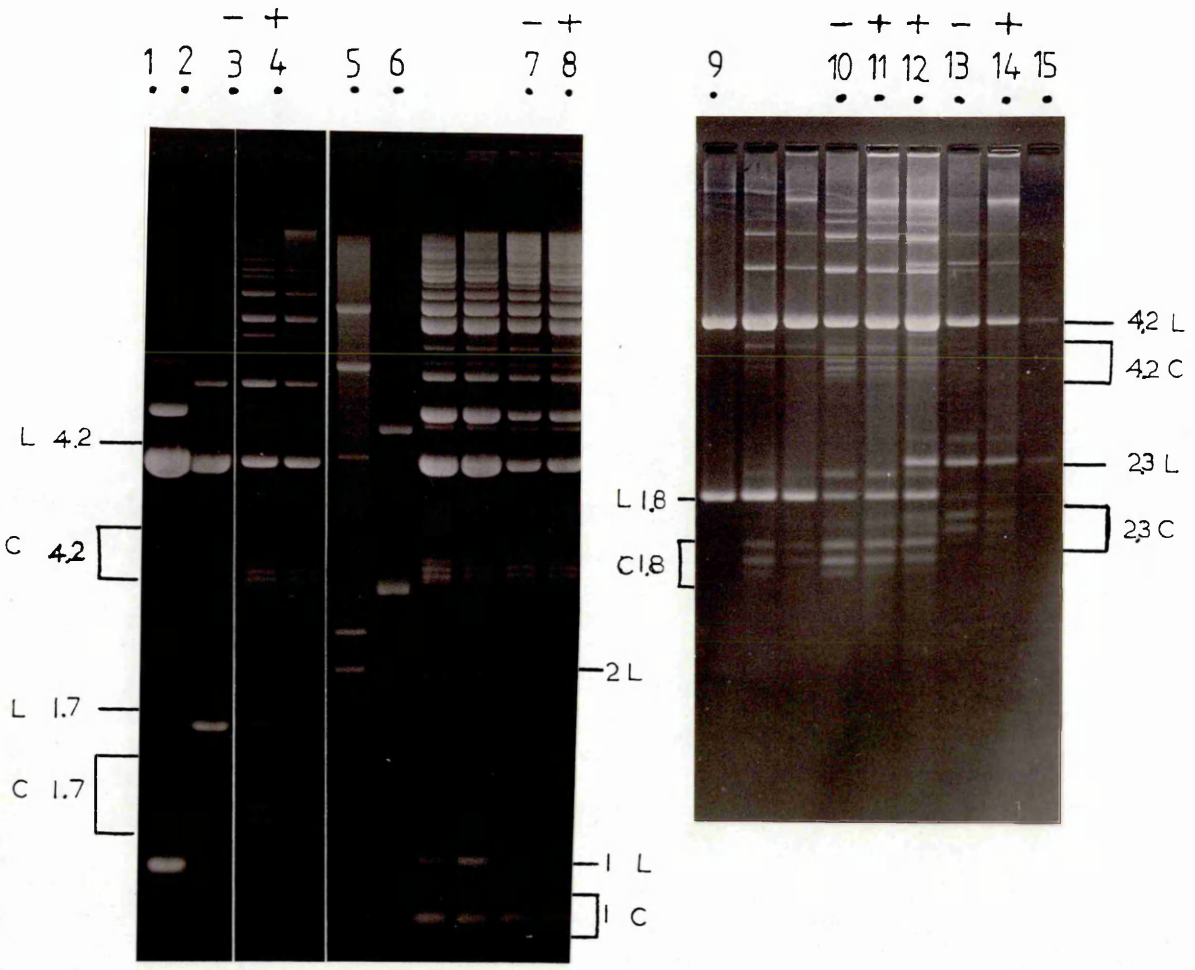


Figure 5.7 Trapped linkage in linear substrates
Binding/ligation reactions were carried out as described (Methods, figure 5.6), using the indicated plasmid substrates cut with EcoRI. The samples were run on a 1.2% agarose gel containing 3 ug/ml chloroquine.
TnsB concentration: 20ug/ml. Presence or absence of protein (+, -) is indicated above the figure.

lane	1	pEAL1 untreated
	2	pLM110 untreated
	3,4	pLM110: 5 min ligation
	5	\ HindIII
	6	pACYC184 uncut
	7,8	pEAL1: 5 min ligation
	9	pLM113 untreated
	10,11	pLM113: 15 minute ligation
	12	pLM113 + pLM414: 15 min ligation
	13,14	pLM414: 15 minute ligation
	15	pLM414 untreated

l linear o relaxed circle c closed topoisomers
Fragment sizes are indicated in kb

EcoRI fragments:

pEAL	1.0, 4.2 kb	pLM113	1.8, 4.2 kb
pLM110	1.7, 4.2 kb	pLM414	2.3, 4.2 kb

was detected for pLM414. This experiment suggests that the majority of the linkage change on circularising the mini-Tn7 comes from interactions between the two ends, rather than from binding and wrapping at the left and right ends independently.

When left and right end-carrying linears were mixed, then treated with TnsB and ligase, the topoisomers produced were all accounted for as products of the two independent binding reactions; there was no linkage change induced in either fragment by the presence of the other (lane 12 and data not shown). There was also no evidence for an enhancement ofintermolecular ligation between the two fragments, nor for catenane formation (data not shown). Thus the interaction between the two ends was only observed when the ends were present on the same molecule. This is the normal situation in a transposition reaction.

A change in trapped linkage was seen only for those fragments containing Tn7 ends; it is therefore unlikely to be due to non-specific binding of protein to DNA.

The ligation experiments described are a powerful method for studying the topology of interactions between DNA sites, and could be extended to examine the interaction between the transposon ends and the target site.

5.3.2 Linkage change in relaxed substrates

Plasmids containing no Tn7 sequences, a Tn7 right end, a Tn7 left end, and Tn7 left and right ends in the correct orientation (pUC18, pNE200, pLM8075 and pLM11 respectively: figure 5.5) were nicked using DNase I in the presence of ethidium bromide (see Methods). These substrates were incubated with TnsB under standard binding conditions; ligase was added (0.25 units/reaction) and the incubation was continued for a further 30 minutes. The samples were electrophoresed on an agarose gel containing 4ug/ml chloroquine to separate the topoisomers; the result is shown in figure 5.8.

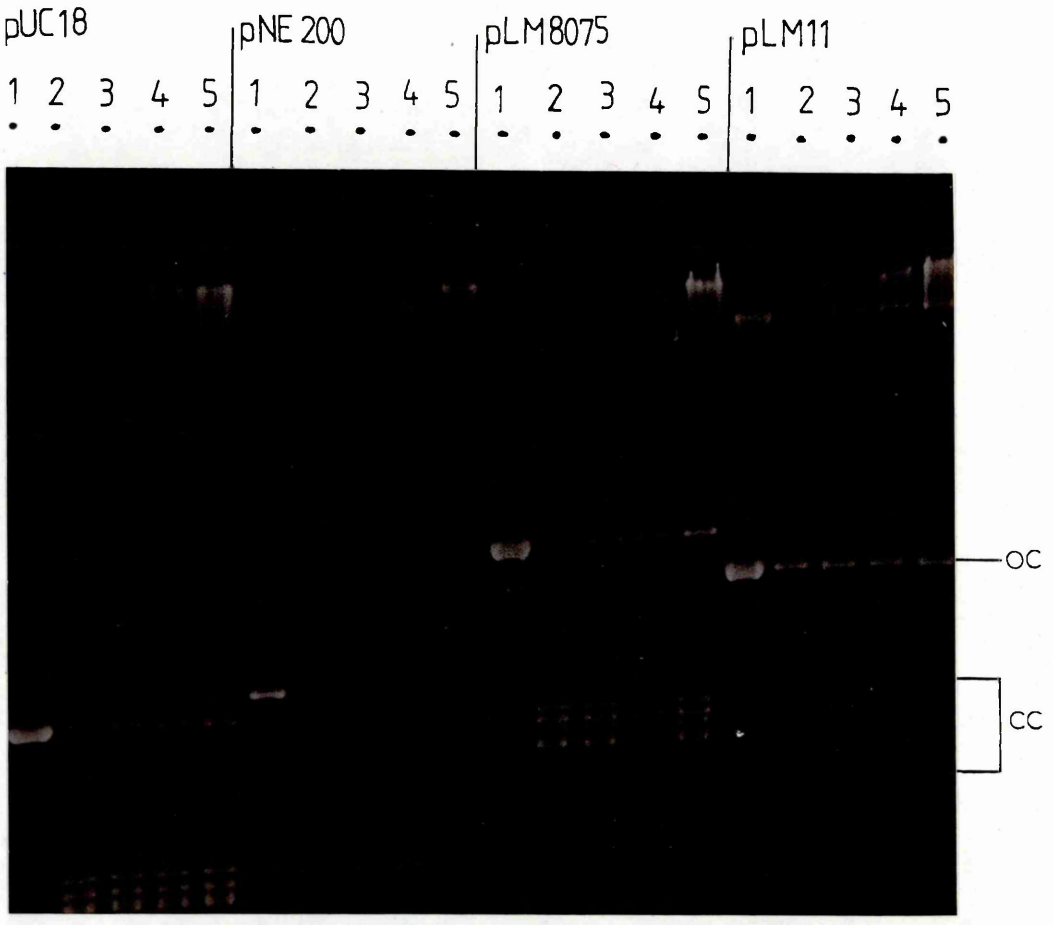


Figure 5.8 Trapped linkage in relaxed substrates
Binding/ligation reactions were carried out as described in Methods, using the indicated plasmid substrates which had been nicked with DNase I. After treatment with ligase for 40 minutes, the reactions were stopped with protease K and the samples were run on a 1% agarose gel containing 4ug/ml chloroquine.

lanes 1 unligated: 2 ug TnsB
2 no TnsB
3 1 ug TnsB
4 2 ug TnsB
5 4 ug TnsB

The bands are identified for pLM11:
oc relaxed circle cc closed topoisomers

There was no change in trapped linkage on binding of TnsB. The amount of DNA in these experiments was equivalent to that with 40-fold excess competitor (figure 4.5, lane 12), so the prediction is that with the same amount of protein (0.5 ug; the lowest used here) we would expect to see around 30% of the DNA bound.

The simplest explanation is that there is no wrapping of the DNA in the complexes, or at either end of the transposon, nor any other topological change (eg underwinding of the DNA helix, or changes in helical pitch). This is not consistent with the results from experiments with linear substrates (section 5.3.1). Alternatively, the protein may be failing to bind to the relaxed substrates. However, similar samples containing some labelled Tn7 RE fragment showed that binding to the small fragment does occur under these conditions, and that this binding is effectively competed by the presence of nicked substrates (data not shown), showing that the relaxed plasmids are being bound by TnsB in the reaction.

The explanation may be in terms of the segregation of the supercoiling between the two domains formed by protein binding. The substrates generally contain a single nick. Given the relative lengths of the mini-Tn7 (0.7 kb) and the vector (2.7 kb), the nick is four times as likely to be in the vector DNA as in the transposon. Generation of supercoiling in the synapse introduces an equal amount of supercoiling of the opposite sense into the plasmid; if these supercoils are segregated into the transposon domain, in 80% of events this domain will be closed, so the introduced supercoils cannot relax, and the net linkage change is zero. The 20% of bound plasmids which do show a linkage change may not be detected, against the background of unbound and unchanged plasmids. With the larger, pACYC184-based substrates, only one binding event in seven would trap the introduced supercoils, so detection of a shift becomes even less likely.

The explanation given above suggests that the

supercoiling introduced into the plasmid by protein binding is all segregated into the transposon domain. This hypothesis makes the prediction that a trapped linkage should be observed for multiply nicked plasmids.

The linkage change observed in the linear fragment, by this reasoning must have been due to nicks in the linear DNA; variation in the quality of the DNA would also explain the variability in the results of this experiment. However, a linear transposon cut within the transposon sequences would be predicted to show a linkage change regardless of nicking, and in a preliminary experiment this appeared to be the case.

5.4 CONCLUSIONS

Footprinting experiments on the right hand end of Tn7 showed that TnsB protects a large region from cleavage by both DNase I and hydroxyl radicals. The protected region included the 22 bp repeats, confirming that this motif is important in binding. A region of access to cleavage reagents was observed, down one face of the DNA helix. Ordered occupation of the 22 bp repeats was observed, with R1 being occupied at low protein concentration, and R4 occupation requiring the highest concentration. No evidence was seen for TnsB-dependent cleavage at the transposon terminus.

Positions of strongly enhanced cleavage by DNase I were observed; the cleavage occurred once in each 22 bp repeat, at the analogous nucleotides (position 14 and 16). The enhancement could be caused by a structural change in the DNA which is induced by the binding of TnsB; for example, the DNA may become kinked or bent. The same nucleotides are part of a region of imperfect dyad symmetry within the 22 bp motif, of the form AGT_3TA_3CT. If we maintain that the whole of the 22 bp repeat forms the binding site for TnsB, then the function of this dyad might be to allow or favour a particular DNA structure, in

a way analogous to the positioning of nucleosomes, which facilitates the observed distortion of the DNA on the binding of TnsB.

This chapter also presents experiments using DNA ligase, which were designed to detect interactions between the ends of Tn7, mediated by TnsB. In experiments where linear mini-transposons were circularised in the presence or absence of TnsB, no enhancement of ligation rate by the protein was observed. However, the circles formed in the presence of TnsB contained 2-3 more negative supercoils than the controls in the absence of TnsB. As the linkage change for linear fragments containing only a single left or right end was one supercoil or less, most of the trapped supercoiling is probably due to interactions between the transposon ends. This supports the proposal that a function of TnsB in transposition is to recognise and bring together the transposon ends.

No evidence was seen for interactions between left and right ends in <u>trans</u>, suggesting that the two ends must be on the same molecule in order to be brought together by TnsB.

In experiments with singly nicked substrates, the presence of TnsB did not result in the trapping of any supercoils when the DNA was resealed by ligase, although it was demonstrated that there was binding to the nicked plasmids in these reactions. The result can be explained by proposing that all the induced supercoils are being segregated into the transposon domain of the plasmid.

CHAPTER 6

ASSAYS OF TRANSPOSITION *IN VIVO*

6.1 INTRODUCTION

Whilst the work presented in chapters 3-5 will contribute to the setting up of an *in vitro* system using purified components, it was felt that much more information was required regarding the natural behaviour of Tn7 *in vivo*. It is notably complex; at least four gene products are required for each event, demanding almost 8 kb of coding capacity in the transposon, and two pathways of transposition are observed, differing in their choice of target site (see Chapter 1; Rogers et al., 1986; Waddell and Craig, 1988). Although we can propose roles in the reaction for *tnsB*, D, and E, we have little further information; when this work was started, even less was known for *tnsA* and C, and their interactions with each other and the host cell (although recent data from Nancy Craig's group answer some of these questions: personal communication to D.Sherratt). For example, little is known about the stoichiometries of the various components in the reaction and the control of activity or gene expression to achieve these, or about host proteins involved, nor have the active domains of the proteins been located. It has not even been demonstrated whether the element transposes replicatively or conservatively. It was therefore decided to attempt further *in vivo* studies of transposition, to clarify some of these points, and so provide information to facilitate the design of *in vitro* experiments.

One important set of data which would be desirable concerns mutageneses. The question which prompted this work was that of host factors involved in Tn7 transposition. This would be interesting in its own right, as well as for *in vitro* work. Many transposition reactions, and other site-specific recombination systems, require host encoded proteins. One of the first to be defined was IHF, required for phage lambda integrative recombination. It is also involved in transposition of phage Mu (Surette and Chaconas, 1989), Tn10 (Morisato and

Kleckner, 1987) and several other transposons. Plasmid ColE1 carries a site-specific recombination system in which the recombinase (XerC; Colloms et al.,1990) and the proposed accessory proteins (Stirling et al.,1987; Stirling et al.,1988) are encoded by the host chromosome. Other examples include dam methylase in temporal control of IS10 and Tn5 transposition (Roberts et al.,1985; Yin et al.,1988), DnaA (Tn5; Yin and Reznikoff, 1987), DNA gyrase (Tn5; Isberg and Syvanen, 1982) and rho termination factor (Datta and Rosner, 1987). The transposition machinery of Tn7, involving five element-encoded genes, is complex; it is interesting to know whether further factors are still required.

Other aspects which would benefit from a mutational analysis include characterisation of the sequences at the ends of the transposon; similar studies with other elements have revealed 'domains' within the termini with different contributions to the reaction (eg IS903; Derbyshire et al., 1987: IS1 ; Zerbib et al., 1990b).

Another target for this approach would be the transposition genes themselves. Other than the mini-Mu insertions used to define the five complementation groups (Waddell and Craig 1988), there have been no reported mutageneses of tnsA-E.

6.2 Assaying transposition

The major obstacle to carrying out such experiments is the problem of assaying transposition in vivo. The current assay in use in the laboratory is the conjugative, or mate-out, assay, which is described in figure 6.1. Although it does not give an absolute measure of transposition frequency, comparisons between assays can be made. The mate-out asssay is the classical measure of transposition, and is used by many groups for assaying transposon activity (eg Tn10: Roberts and Kleckner, 1988; Tn5: Berg, 1983; IS903: Wiater and Grindley, 1990). A modified version scores donor resistance transfer due to

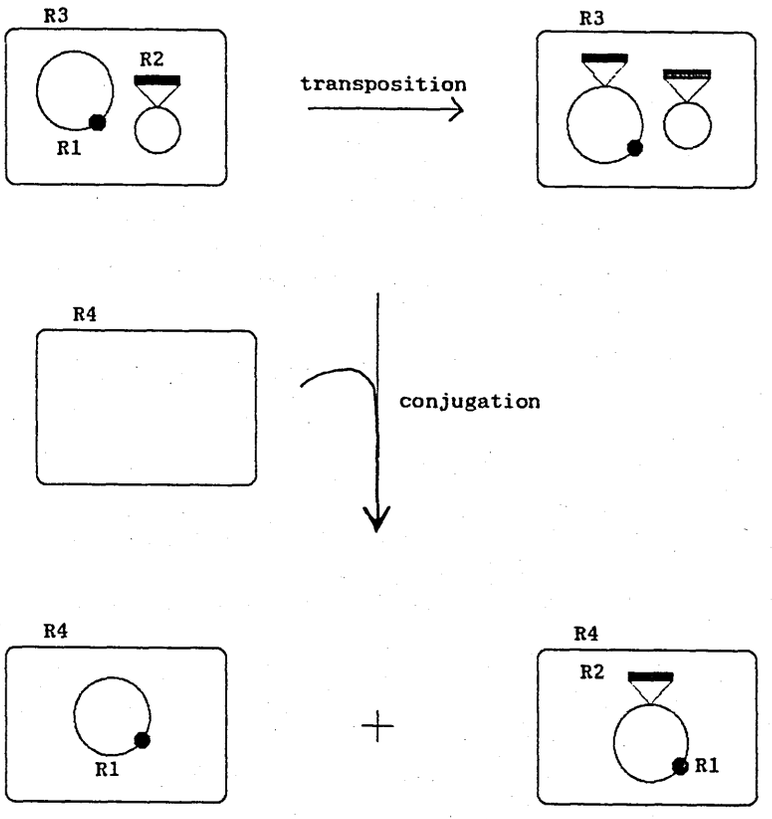
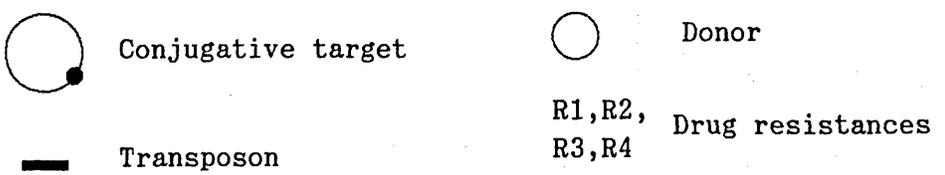


Figure 6.1 The mate-out assay
The assay strain contains a transposon with drug resistance R2 (carried on a non-conjugative, non-mobilisable plasmid or in the chromosome), and a conjugative target plasmid with resistance R1. After growth of the strain to allow transposition, the target plasmid population is sampled by mating with a naive strain which also has a selectable marker (R4), and scoring the proportion of transposon-bearing targets by growth on appropriate selective plates.

$$\text{Transposition frequency} = \frac{R1\ R2\ R4}{R1\ R4}$$



cointegrate formation (eg Tn21: Martin et al.,1989; Tn3: Arthur and Sherratt, 1979).

6.3 Problems with the mate-out assay

There are, however, difficulties in the use of, and interpretation of data from, this assay. These can be considered in several categories; some of these are specific to Tn7, but others have more general validity.

1. "inherent" problems: for example relative stabilities and conjugation efficiencies of target and product, or loss of plasmids by transposition into origin sequences, marker or transfer genes. Mark Rogers (1986) examined two possible sources of error, namely the relative growth rates of cells carrying R388 or pEN300 (table 2.2), with or without Tn7, and the efficiencies of mating of these plasmids. In neither case did the presence of Tn7 appear to have an effect under the conditions used.

The assay also requires transposons and targets on plasmids, often in high copy number, so there is potential for titration, 'multicopy inhibition' (Arciszewska et al., 1989) and immunity effects. It must be assumed in comparing assays that these are remaining constant in all circumstances, which need not be the case.

However, one feature of all these points is that they can be tested for. They can therefore if necessary be controlled for in any given experiment, if the variation is such that comparisons between assays become invalid.

2. Time: by the time exconjugants are scored for transposition, the cells are many generations from the original event; hence we lose information on the timing of transposition during cell growth.

3. Population dynamics: as well as being a time average, the data from this assay are also a population average. The target plasmid is segregating during cell division, with occupied and non-occupied targets presumably partitioning at random, leading to a distribution of cell types within the culture. In

addition, we suffer a "fluctuation" type artefact - it is not possible to distinguish a single, early event from multiple (ie higher frequency) later events. This makes a screen (assaying transposition within a single colony) impossible, given the low frequency of transposition (10^{-5} - 10^{-2}).

4. Plasmid interactions: in the case of Tn7, it is frequently observed that markers carried on the donor plasmid are transferred to the recipient cell (data not shown; K.McCurrach, personal communication). The level of fusions appears to be affected by the Tn7 sequences present; specifically, by a region leftward of the dihydrofolate reductase gene, which contains an open reading frame with homology to the λ Integrase family of site-specific recombinases (Sundstrom et al.,1988; Stokes and Hall, 1990). The same sequences are also present in R388 (Hall and Vockler, 1987), which is the parent plasmid of the conjugative recipients commonly used. Hence the assay system as currently used is far from ideal, with multiple interactions making the data difficult to interpret.

5. Fate of the donor: as transposition is scored only after transfer to naive cells, and transposition products are specifically selected for, the donor molecule is lost during the assay. We thereby lose information as to its fate during transposition. This is a feature common to many current assay methods, hence the dearth of knowledge about this component of the reaction. This becomes particularly important when considering mechanisms of transposition.

6. Practicality: carrying out large numbers of such assays (ie more than 10) consumes large amounts of both time and materials, and can lead to cross-contamination.

7. As a result of the points noted above, the mate-out assay sometimes fails to reveal changes in transposition frequency; for example, FIS^- strains do not show a reduced frequency of Tn7 transposition by this

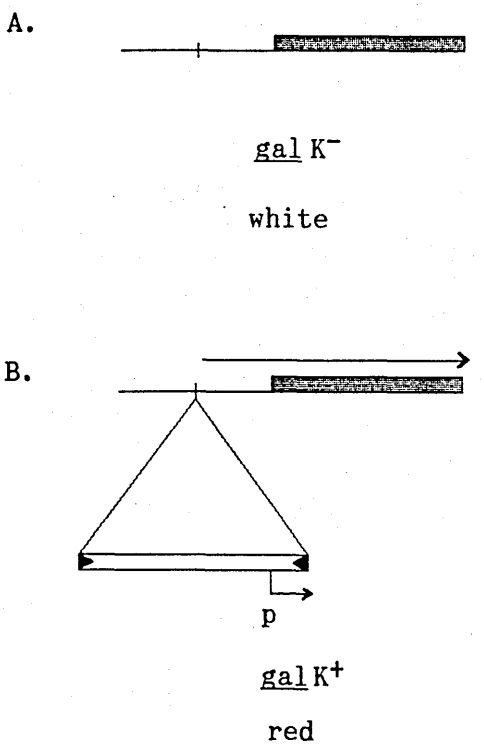


Figure 6.2 The basis of the papilliation assay
A. A promoterless reporter gene (eg galK, coding for galactokinase) located downstream of the attTn7 is not expressed, and the cells are gal⁻ and white on MacConkey/ galactose agar.
B. Insertion of Tn7 into att might provide transcription signals allowing the gene to be expressed; the cells are gal⁺ and red on MacConkey/galactose agar.

—|—	attTn7, and point of insertion
▭	Tn7
▭	galK
P	promoter activity from within Tn7

assay (K. McCurrach, personal communication), but show a much reduced frequency when assayed by a papilliation test (N.Craig, personal communication). Other variations in frequency may have been masked in the same way, eg between IHF^+ and IHF^- strains (N.Ekaterinaki, personal communication).

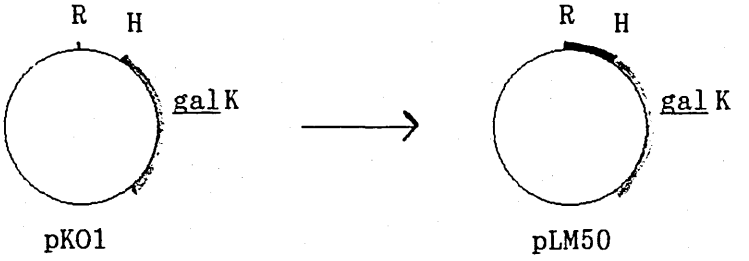
In order to carry out large scale testing of mutant banks, an assay was required which gives a rapid quantifiable screen for changes in transposition frequency. As such, it is preferable for transposition to be monitored in the same cells as the event itself; in addition to overcoming some of the problems discussed above, this allows a screen at the level of single colonies. In the case of Tn7, it is possible to exploit the extreme site specificity of transposition, and design assays using the hotsite as a "trap" which can indicate whether the site is occupied. This chapter describes developement of one such assay, based on activation of a reporter gene engineered into the chromosomal hotsite.

RESULTS

6.4 Basis of the papilliation assay.

Given sporadic gene activation within a colony, if the gene concerned can confer a selective advantage on that cell and its descendants, then these will grow out as a papilla on the surface of the colony. With a suitable indicator medium, this change in gene expression can be readily visualised by staining, eg galK on MacConkey/ galactose agar, or lacZ on Xgal/lactose.

This could be utilised as an assay in the form of the construction described in figure 6.2. If the galK gene can be activated by transcription from within Tn7, then the $galK^+$ cells resulting from transposition (and their descendants) will have a growth advantage in the presence of galactose and limiting glucose, and so will grow out as



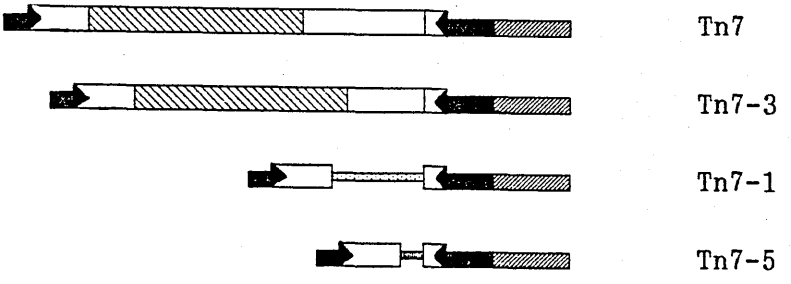
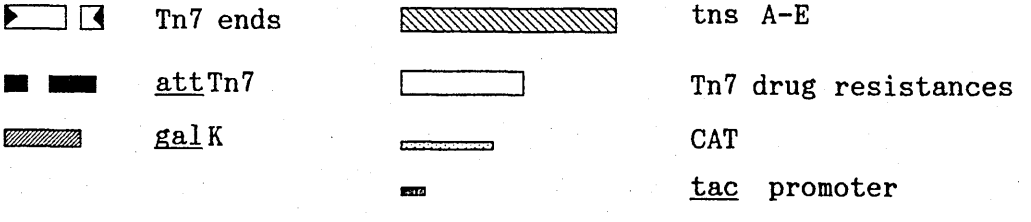


Figure 6.3 Construction of pLM50 and its derivatives
The 280 bp EcoRI-HindIII fragment of pKO1 was replaced by the EcoRI-HindIII minimal attTn7 fragment from pMR80, placing attTn7 upstream of a promoterless galK gene. The derivatives of Tn7 shown were transposed into the att site of pLM50 in vivo.
R=EcoRI H=HindIII



a red stained papilla; each one of these, then, represents an independent transposition event. In this way the frequency of transposition can readily be observed within a single colony, which allows a large number of strains to be assayed simultaneously. Any strains showing variant transposition levels can then be isolated for further study.

6.5 Testing in multicopy plasmids

To test the feasibility of this approach, ie whether transcription initiated within Tn7 can in fact extend through the left end into the flanking DNA, a test system was set up, using multicopy plasmids based on promoter probe vectors carrying galK as the reporter gene. In addition to the colour change seen on plates, galK expression can readily be assayed in cell lysates, providing quantitative data to support results from growth on indicator media.

The attTn7-galK plasmid pLM50, and a set of insertion derivatives, were constructed (figure 6.3). Clones of the transposition products were isolated; digestion of these with EcoRI showed that transposition retained the site- and orientation-specificity seen with other hotsite clones (data not shown).

galK expression in the transposition products was monitored both on plates and by assaying for galactokinase activity. The results are presented in table 6.1, and figure 6.4.

The insertion products were clearly redder than cells carrying the parent plasmid; however the colour was not as intense as that seen with galK under P_{lac}; this was borne out by the assays, with the insertion products showing only 5-10% of the P_{lac} level of kinase activity.

In order to try to increase the level of internal transcription initiation, a DNA fragment containing the strong tac promoter was cloned into a Tn7 derivative (Tn7-5). As there were no convenient restriction sites within

Table 6.1 GalK activities of insertion products of pLM50

Plasmid	IPTG[a]	GalK S.A.[b]	Bla S.A.[c]	GalK/Bla	ratio[d]
pLM50		2.70	1.72	1.57	1
pLM50::Tn7		9.24	2.30	4.02	2.56
pLM50::Tn7-3		12.1	1.03	11.7	7.45
pLM50::Tn7-1		5.51	1.28	4.30	2.74
pLM50::Tn7-5	-	3.46	0.70	4.94	3.15
pLM50::Tn7-5	+	3.72	0.47	7.90	5.03
pKO500		4.96	1.93	2.57	1.63
pKL500	-	1.60	2.55	0.63	0.40
pKL500	+	129	2.13	166	111

All values are the average of at least two assays.

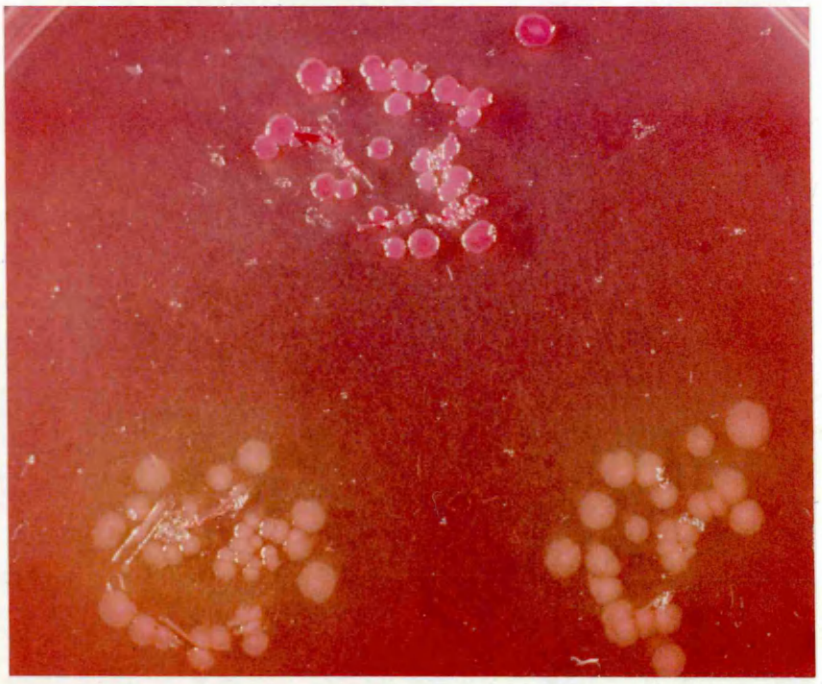
a. Cultures grown with and without IPTG for induction are indicated. Where no indication is made, the promoter concerned is not IPTG-inducible, and was not tested under inducing conditions.

b. GalK specific activity units are nmoles galactose phosphorylated / min / ml of culture at an A_{650} of 1.

c. B-lactamase activity units are defined in Materials and Methods, and are converted to specific activities by dividing by the A_{650} of the culture. The reasons for the variance in the values is that different subsets of the plasmids were used in repeat experiments; within an experiment, the rankings of the results were the same, although the absolute values varied.

d. The copy-number corrected GalK activity is divided by the copy-number corrected value for pLM50 with no insertion.

A



B

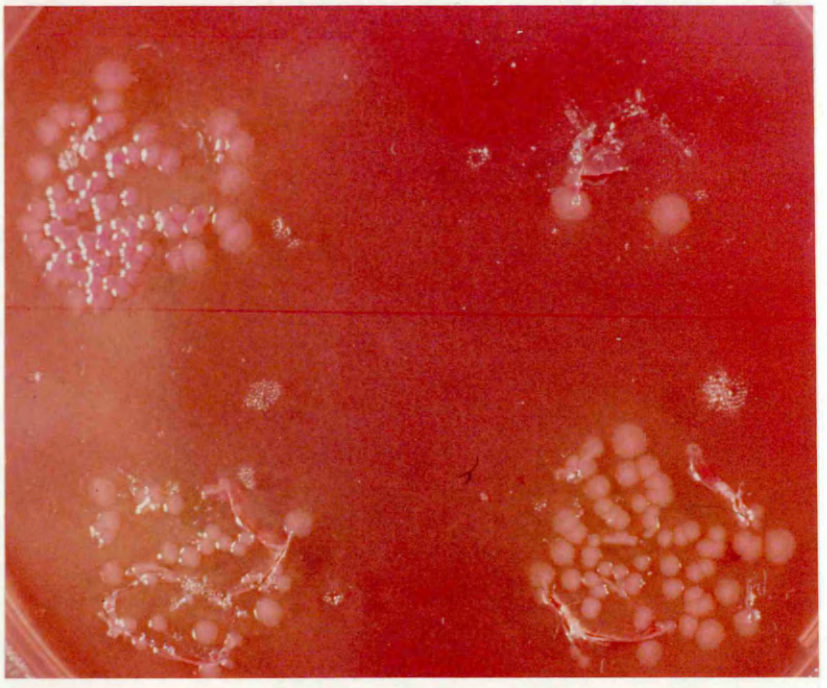


Figure 6.4 <u>gal</u>K phenotype of pLM50 and its derivatives
The indicated plasmids, in DS941, were grown on MacConkey /galactose agar at 30°C overnight.

A.

pKL500	
pKO500	pLM50

B.

::Tn7-1	::Tn7
::Tn7-5	::Tn7-3

this fragment, its orientation could not readily be determined. The resulting element was transposed into pLM50. None of the clones tested showed the significant increase in galK expression expected for this strong promoter in the presence of IPTG. This may be because the cloned promoter is in the wrong orientation in all cases, suggesting that the other orientation is selected against for some reason. Alternatively transcription from this promoter may be being terminated within the left end of the Tn7; the repetitive nature of these sequences may allow formation of secondary structures leading to transcript termination or instability. Indeed, it is not clear whether the increased galK expression seen in transposition products compared to pLM50 is true transcriptional readthrough, or an unrelated phenomenon; for example, structural changes following insertion of the transposon.

It was thought that the level of expression of galK in the insertion products might be sufficiently high to be detected when present in single copy in the chromosome, and so the required strain was constructed.

6.6 Construction of EM3 and EM5

The method used to generate a chromosomal hotsite-galK was forced integration of a selectable, replication-incompetent plasmid carrying the galK insertion (Flinn et al., 1989). It was necessary to provide further homology 3' of the gene. This was provided from a 1 kb hotsite clone, pEAL1 (Lichtenstein and Brenner, 1981); the construction is shown in figure 6.5.

The resulting plasmid, pLM55, carries a λdv origin sensitive to repression by λ repressor, and was exposed to repressor by introducing pEA305 (a cI overproducer). pLM55 will be unable to replicate, and chloramphenicol resistance can only be maintained by integration of the plasmid into the chromosome. The above order of addition allowed transformation and integration to occur in two

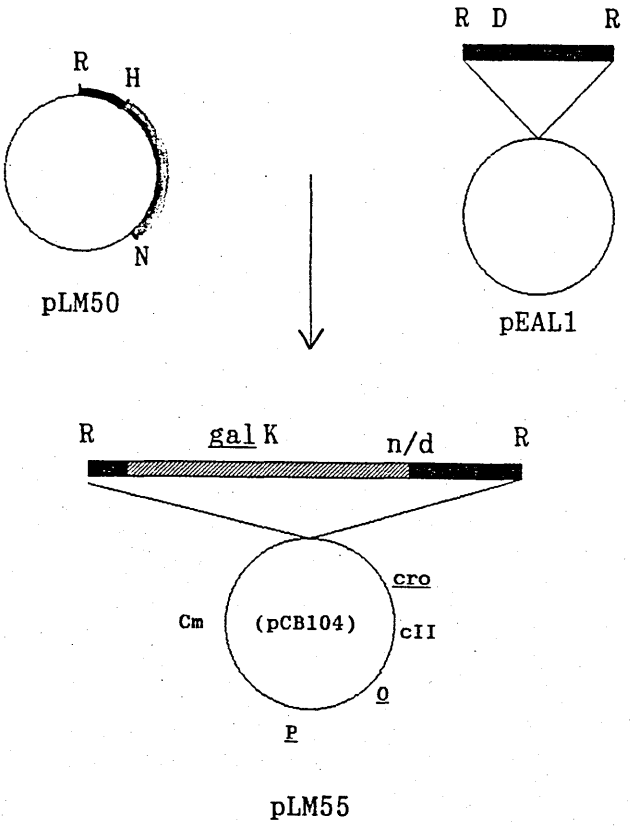
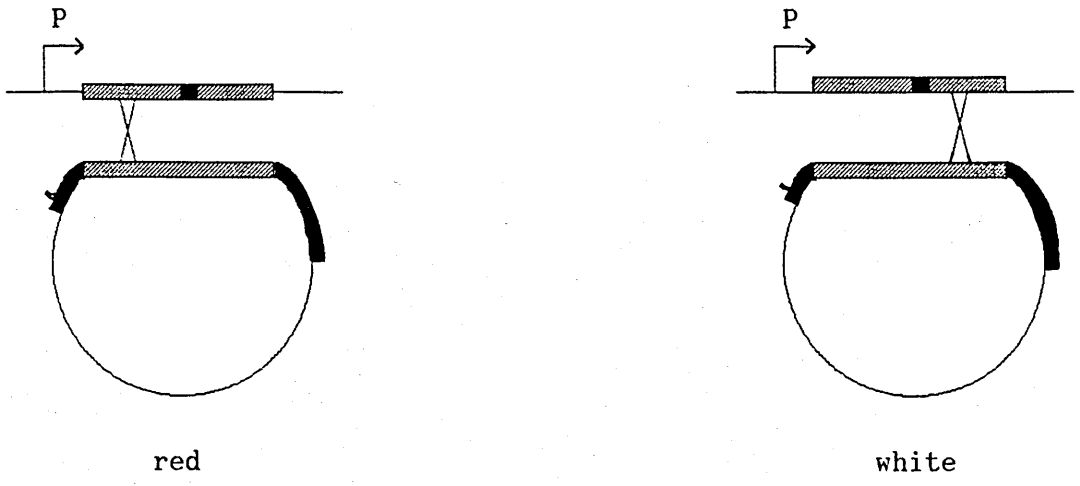


Figure 6.5 Construction of pLM55
A 1.9kb EcoRI-NdeI fragment from pLM50 contained attTn7 and the galK gene. After filling-in the NdeI end, the fragment was ligated to the 0.75 kb DraIII-EcoRI fragment from pEAL1 (which contained sequences from downstream of attTn7) in the presence of EcoRI. The resultant 2.65 kb EcoRI fragment was cloned into the EcoRI site of pCB104.
R=EcoRI H=HindIII N=NdeI D=DraIII
n/d=NdeI/DraI recombinant joint

cro, cII, O, P are required for replication of the lambda origin of pCB104

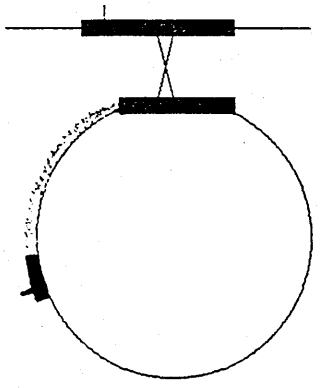
att Tn7 galK

A Insertion at galK



red white

B Insertion at attTn7



white

Figure 6.6 Proposed genetic structure of red and white integrant strains

A. Insertion of pLM55 at galK can give red or white colonies depending on whether the plasmid (wildtype) or the chromosomal (mutant) allele is downstream of the wildtype chromosomal gal promoter.

B. Insertion of pLM55 at attTn7 gives white colonies, as there is no promoter for galK expression.

galK
chromosomal galK mutation
attTn7 and point of insertion

stages.

ApCm^r colonies were isolated, and were screened for insertion by loss of the ability to transform cells to chloramphenicol resistance. Suitable strains were streaked onto MacConkey/galactose plates. Both red and white patches were seen; an interpretation of the result is shown in figure 6.6. In both types of white strain, the hotsite will be upstream of the wildtype galK gene, and so can be used in the assay.

However, Southern blotting indicated that there was some free plasmid present in these strains, which proved to be difficult to lose (data not shown); it is possible that plasmids were recombining in and out of the chromosome. Free plasmid would be able to survive if the repressor plasmid had mutated; this is possible, as pEA305 is highly unstable and deletes even in recombination-deficient strains (M. Burke, personal communication). As the presence of any free plasmid at all would destroy the papilliation screen, these strains were not tested further.

An alternative strategy was to use a lambda lysogen to repress the pLM55 origin. DS959, a lysogen of DS941, was used. Cm^r transformants were tested as above for loss of free plasmid. Again, both red and white clones were isolated, and white clones (EM5) tested for papilliation on introduction of Tn7.

6.7 Testing of insertion strains

EM5 cells were transformed with pEAL1::Tn7, and related plasmids which do not contain active Tn7, and the transformants were plated directly onto indicator medium. Firstly, there appeared to be some instability in the assay strain, especially following transformation, and red, pink or papilliated colonies appeared in the absence of Tn7. The instability of the galK genotype may again be due to plasmid excision and reintegration.

In the presence of Tn7, all the colonies appeared

pale pink, with white 'halos'. The colour change was not Tn7 transposition-dependent; the same result was seen for pEAL1 alone, for pEAL1::Tn7-1 (*tns*$^-$), and for pEAL1::Tn7ΔEcoRI (*tns*$^+$, but no left end). When the pink colonies were tested for their ability to grow on galactose as the sole carbon source, none grew, i.e they were all *gal*$^-$. It could be that the Tn7 was not transposition-competent (which is unlikely, as two different sources of Tn7 were tested), that EM5 is not able to support transposition, or that the level of expression of *gal*K is too low to confer a *gal*$^+$ phenotype. The third explanation is perhaps the most likely, given the GalK assays presented above.

EM5 shows approximately a tenfold deficit in the ability to take up plasmids containing active Tn7 (data not shown). The reason for this deficiency is not clear, but it demonstrates that there is some Tn7 transposition-related interaction in these cells, which might in itself be worthy of further study, or provide the basis for an *in vivo* assay.

The red colour of the colonies in the absence of transposition makes it difficult to detect any differential staining within the colony, particularly if *gal*K expression is low. Reasons for the observed colour change are discussed below (section 6.9), as are possible modifications to the assay to allow it to be used as a screen.

6.8 Alternative assays investigated

This section gives a brief summary of initial experiments on two other types of assay tested for use with Tn7, initially *in vivo*, but potentially also for *in vitro* work. Although the data are preliminary, they are presented here for the benefit of future workers, who may wish to develop these systems further for routine use.

6.8.1 Intramolecular transposition

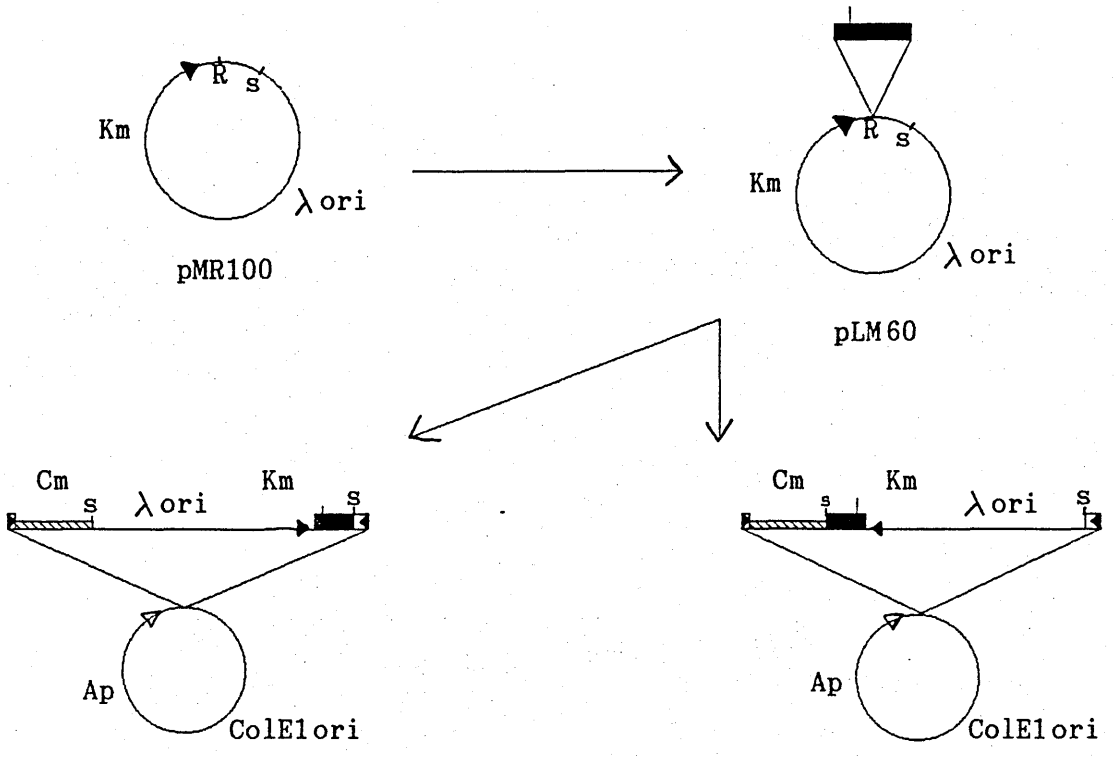
The constructions used to study the intramolecular transposition of Tn7, and the transposition products predicted, are shown in figure 6.7. Cells containing the plasmids grew poorly, forming small colonies which frequently failed to grow on subsequent plating; the poor viability was independent of the presence of tns genes.

Intramolecular transposition can be assayed by loss of drug resistance markers. Loss of both Ap^r and Cm^r was observed; however, it was not according to the predictions in the figure. Each plasmid gave rise to a variety of deletion products, and the deletion was not tns-dependent. The sizes of the product plasmids, and restriction maps where they could be obtained, could not be simply explained.

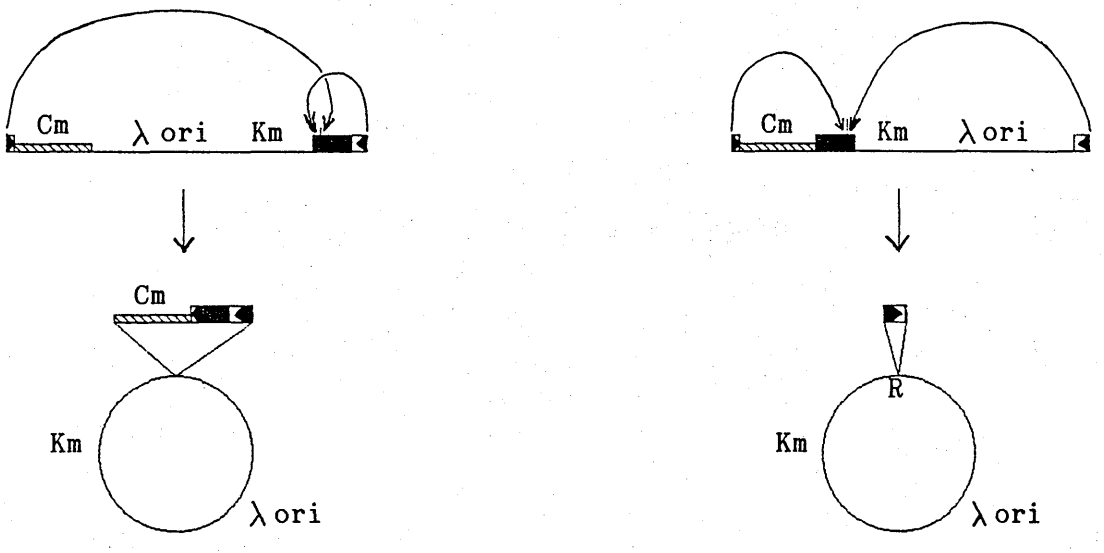
The plasmids were tested for the ability to act as targets for intermolecular Tn7 transposition, and for the ability of the att site mini-transposons to transpose. Both experiments gave positive results as assayed by transfer of drug resistances; on analysis of the products by gel electrophoresis, the sizes of the products could not readily be explained.

From these results, it appears that these plasmids, and some of their transposition products, are highly unstable, and undergo rearrangement in the presence or absence of Tn7 transposition functions. Various techniques were tried to increase their stability. One was to grow the strains on 1% glucose, to repress the lac and tac promoters, and so reduce the number of interacting domains of supercoiling (Wu et al.,1988) within the plasmids. However, this made little differenc to the stability. An sbcB recN strain, SP231, which had been found to reduce rearrangements of cosmid clones containing mammalian genes (Ishiura et al.,1989), was also investigated; however, our isolate of this strain was unable to maintain plasmids containing λ replication origins.

In their current state, these plasmids clearly are



TRANSPOSITION:



	Tn7 ends		cat gene
	attTn7 and insertion point	R	EcoRI
Km Cm Ap	resistances	s	SalI
▷	P_{lac}	▶	P_{tac}

Figure 6.7 Intramolecular transposition substrates
pLM60 was constructed by inserting the 1 kb EcoRI attTn7 fragment from pEAL1 into the EcoRI site of pMR100.
pLM601 and pLM602 contain pLM60, linearised at the polylinker SalI site, inserted into the SalI in Tn7-1 in pMR11; the two plasmids have pLM60 in the two possible orientations.
The transposition product generated will be determined by the orientational specificity of hotsite insertion; the configurations of origins and drug resistances relative to the att site in pLM601 and pLM602 will generate the transposition products shown.

not suitable for assaying transposition *in vivo*. Although the rearrangements occurring may be very interesting, the data are difficult to interpret due to the variety of interactions occurring, and intramolecular transposition was not studied further.

6.8.2 Induction of bacteriophage lambda prophage

Transposition, particularly by a conservative 'cut and paste'-type mechanism, creates cleavages in the DNA strands. DNA damage by other agents is known to induce expression of a set of genes which have functions in the repair of damage; this is known as the SOS response. The SOS genes are normally repressed by LexA; following DNA damage, cleavage of LexA (mediated by RecA) lifts the repression and the genes are expressed. Conditions which lead to SOS induction also cause induction of lytic developement in λ lysogens, because the phage repressor undergoes cleavage in the same way as LexA.

It was thought that the DNA damage caused by transposition, or by the effects of 'transposase', might also induce the SOS response. Roberts and Kleckner (1988) observed prophage induction that correlated with transposition of Tn10, and with other features of the SOS response. They suggested that the SOS response was being induced by degradation of the broken donor DNA molecule.

The ability of Tn7 transposition to induce the SOS response was tested by measuring induction of the λ prophage in DS959; the results are presented in table 6.2. In the first set of experiments, the effects of overexpression of *tns* genes was examined. DS959 is *lacI*^{q}, so should be able to repress the expression of the cloned gene to some extent, in the absence of IPTG. The ratios in the final two columns express the effect of the cloned gene relative to the parent vector. The most significant effect was for pMR205, encoding *tnsAB*; the induction is greatly increased by IPTG, indicating that the function responsible is under the control of the *tac* promoter. *tnsD*

Table 6.2 Induction of λ prophage by Tn7 and tns genes

Plasmid	tns gene(s)	Phage yield[a] -IPTG	Phage yield[a] +IPTG	IPTG induction	Ratio[b] -	Ratio[b] +
pMR78	---	9.5	13	1.37	1	1
pMR203	tnsE	22.5	17	0.76	2.4	1.3
pMR204	tnsA	6.5	19	2.92	0.7	1.5
pMR205	tnsA and B	4	66	16.5	0.4	5.1
pMR207	tnsB	20.5	38.5	1.88	2.2	3.0

a. Phage yields (in pfu/ml) are x 10^6, and are corrected for the spontaneous frequency of induction of the prophage in DS959, which gave 4 x 10^6 pfu/ml in the presence and absence of IPTG.
b. The phage yield for each plasmid is shown relative to that for the vector plasmid, under the relevant induction conditions.

Plasmid	Ends	tns	Phage yield
c. pEAL1	-	-	2.4 x 10^6
pEAL1::Tn7-1	L,R	-	2.0 x 10^6
pEAL1::Tn7△R1	R	+	1.5 x 10^6
pEAL1::Tn7	L,R	+	4.8 x 10^6
d. pEAL1	-	-	1.1 x 10^6
pEAL1::Tn7-1	L,R	-	3.0 x 10^5
pEAL1::Tn7	L,R	+	1.3 x 10^6

c. Cultures were grown to stationary phase.
d. Cultures were in log phase.

and E were also tested; no effect of the cloned gene, or of IPTG, was detected (data not shown).

The results could indicate a role in DNA cleavage for TnsA or B. However, it is not clear in these experiments whether the prophage induction is due to a DNA-metabolic activity *via* the SOS response, or is a more indirect effect due to the stress of overexpressing cloned genes.

The effect on prophage induction of plasmids carrying combinations of Tn7 ends and *tns* genes is shown in the lower part of table 6.2. There is no significant difference between plasmids with and without transposition activity. The values shown are averages from three experiments. The ranking of phage yield varied between experiments; there was not necessarily a correlation between transposition activity and prophage induction in a given experiment. As the aim of this part of the project was to design a rapid assay for actual transposition, this method of measuring transposition by SOS induction did not appear to be useful.

Induction of the SOS response was also tested in a strain in which *lacZ* is placed downstream of an SOS-inducible promoter in the chromosome (N3331). This strain was found to induce the SOS response under the stress of transformation with the *tns*-expressing plasmids; all colonies were blue on Xgal, and red on MacConkey/lactose agar. The vector plasmids alone caused the same effect. As the strain was found to be so sensitive, it was not used further.

6.9 DISCUSSION

The result of testing the *galK* papilliation assay strain was that no *galK* transposition products were detected, as assayed by the ability to utilise galactose; on MacConkey/galactose agar the colonies were uniformly pink, surrounded by an area of white. The reason for the colour change is not clear. It is our experience that the

results from MacConkey plates tend to be somewhat strain- and plasmid-dependent, and the utility of these indicator plates varies with the constructions in use. The white area around the colony is due to an alkaline pH, which is the expected result of the *gal*$^-$ cells utilising amino acids as their carbon source, and excreting ammonia-containing products. Presumably some characteristic of the combination of DS959, pLM55, and the tetracycline-resistant pEAL1 plasmids, is causing a change in the growth of the cells such that they create an acidic environment within the colony, leading to pink staining.

An alternative form of the papilliation assay is to include the promoterless reporter gene on the transposon, and look for activation by external transcription from the target gene. Some of the constructions described in chapter 7, with *lacZ* cloned into mini-Tn7, could be modified for this purpose. The reporter gene would need to be oriented such that it was transcribed from the left end of the element, as there is a transcriptional terminator in the right end (Gay et al., 1986), as well as a promoter; at least the latter is included in the region required in *cis* for transposition (Arciszewska et al., 1989). Such a construction would only assay coldsite transposition, given the orientation specificity of hotsite insertion.

The papilliation assays in use for other transposons are generally of this type, with the reporter gene within the transposon, eg Tn10 (Huisman and Kleckner, 1987) and IS903 (K.Derbyshire, personal communication). An exception is an IS1 and IS5 assay, which detects transposition by activation of a cryptic *bgl* operon by insertion upstream of the gene (Datta and Rosner, 1987). The internal reporter gene approach was considered when the Tn7 assay was first designed, and decided against for several reasons. The transcript produced by readthrough would be very long, including the target gene and the long Tn7 end as well as the reporter gene; there was concern that this

might reduce the stability of the transcript, and that this would vary between insertion sites. The level of reporter gene expression will fall under the same controls as the target gene, and so again will be variable for different insertion sites, and may depend on factors such as the growth phase of the cells, or nutritional status. Similarly, the state of the cell, and the resultant pattern of gene expression, might affect the availability of target sites, so altering the observed transposition frequency (Lodge and Berg, 1990). For these reasons, it was decided to use the hotsite, where the site of insertion is unique and well defined, so such position effects would not occur. The hotsite also had the advantage of high frequency of transposition, so increasing the probability of detecting an event.

A papilliation screen has been set up and is in use in Nancy Craig's laboratory (personal communication); it is probably similar to the coldsite assay discussed above. The two systems were developed independently; a priori the hotsite reporter is the more favourable option, but in reality has associated difficulties which were not apparent when the work was initiated.

The second assay considered was based on intramolecular transposition. Constructions which assay transposition within a single molecule are useful for two reasons. Firstly, they could be used as substrates in an in vitro system; as both of the sites involved are present on the same molecule, the number of interacting species is reduced, and the relative stoichiometries of the DNA sites are known. The current cell-free system for Tn10 utilises this type of substrate (Morisato and Kleckner, 1987; Benjamin and Kleckner, 1989). Secondly, the intramolecular transposition behaviour of an element can provide indirect evidence concerning the replicative or conservative nature of the transposition mechanism. Replicative intramolecular transposition can only utilise a target located outside the transposon, whereas 'cut and paste' transposition uses

a site within the transposon (Berg, 1989; Tomcsanyi et al., 1990).

The results indicated that there are interesting phenomena to be observed in such a system, but their interpretation is complicated by multiple interactions involving transposon and non-transposon DNA, transposition proteins, and cellular recombination systems. The instability of the products *in vivo* made their further analysis difficult; however, the intramolecular transposition substrates might be more useful in an *in vitro* system; if purified transposition proteins are used, then the rearrangements promoted by cellular recombination functions would not occur. The reasons for the instability are not known, but could include the presence of two replication origins or several promoters, or could be due to the recombinogenic nature of some sequence elements, or combination of sequence elements, included in the constructs.

The correlation between transposition frequency and the extent of SOS induction was described for Tn10 by Roberts and Kleckner (1988). A similar approach is taken here. No correlation was observed between phage yield and transposition activity, but this may be due to other variables in the growth of the cultures, which are causing stress to the cells. One such feature may be the density to which the culture is grown, or the nutritional state of the cells, which will affect the background induction frequency.

Once a rapid screen for changed transposition frequencies has been established, much data can be readily obtained, from mutagenesis experiments. As stated in the Introduction to this chapter, some of the targets for mutagenesis are the transposition genes, the transposon ends, and the host chromosome. For an initial screen for host mutants, a mutagen such as Tn5 would be favoured (Berg, 1989). To create mutations in the *tns* genes, and in the transposon termini, a less coarse method of

mutagenesis is required, such as chemical mutagenesis *in vivo* (Miller, 1972) or *in vitro* (Davis et al.,1980), mutator strains (eg Haniford et al.,1989), or degenerate oligonucleotides (Derbyshire et al., 1987), to saturate the sequence with point mutations. Once the essential regions of the Tns proteins have been mapped, further experiments could make use of site-directed mutagenesis, to ask specific questions about the functions of particular amino acids.

6.10 CONCLUSION

The assays of transposition frequency currently in use are not ideal for large scale *in vivo* experiments. This chapter presents a consideration of methods of assaying transposition *in vivo*, and the results of testing three approaches to this problem. The major assay described is based on detection of transposition by activation of a reporter gene, leading to formation of a papilla on the surface of the colony. The approach chosen was to engineer a reporter gene into the *att*Tn7 site, which would be activated by the insertion of a transposon.

A plasmid system to test this approach was designed; reporter gene activation was observed on insertion of Tn7 in test plasmids, but the level of expression was low. When the reporter construct was introduced into the *E.coli* chromosome, the strains produced were unstable. There was no differential reporter gene expression within the colonies; all the colonies became stained slightly pink, making it more difficult to observe any differential expression, and no *gal*$^+$ colonies were observed, indicating that there is an insufficient level of reporter gene expression. In order to make use of papilliation as an assay, then, the system needs to be modified, and the most likely alternative would be to place the reporter gene inside the transposon and screen for activation by external promoters.

An assay of intramolecular transposition was attempted, detecting transposition *in vivo* by loss of drug resistance markers. The test plasmids and their derivatives displayed extreme instability *in vivo*, and the dependence of these rearrangements on Tn7 transposition functions was not clear. However, the constructs may provide a viable approach to assaying transposition in a cell-free system, where such rearrangements would be less likely to occur, particularly if purified transposition functions were used.

Finally, induction of λ prophage by Tn7 transposition was observed, which was probably due to induction of the SOS response. The presence of TnsA and B caused slight induction, which might indicate that there is a DNA cleavage activity associated with one or both of the proteins. The level of SOS induction could not be correlated with transposition frequency, either by phage induction measurements, or using a test strain with an SOS promoter fused to *lacZ*, and so this approach to assaying transposition *in vivo* was not taken further.

CHAPTER 7

TRANSPOSITION MECHANISMS:

DOES Tn7 TRANSPOSE CONSERVATIVELY ?

7.1 INTRODUCTION

One important mechanistic question facing researchers in the field of transposition is that of how the information is transmitted from donor to target. Is the element simply cut from its original location and moved as a unit into the new site, or is replication of the element involved, generating two copies, one of which is then detected in a transposition event ? This is central to considerations of transposition models (discussed more fully in Chapter 1). If no replication is involved, simple 'cut-and-paste' models can be envisaged. However, if the element is shown to replicate during transposition, it is necessary to invoke more complex models. Clearly, when evaluating mechanistic proposals for any given transposon, one distinguishing feature will be the requirement or otherwise for replication of the transposon DNA.

In addition to its mechanistic interest, the question of replication during transposition is also central to our ideas about the evolutionary and population biology of the transposons and their hosts (eg Galas and Chandler, 1982). The extent and rate at which a transposon spreads through a population will depend on the way on which it increases its copy number. Tn7 is a medically important transposon, in both human and veterinary medicine, as the major response in wild populations to treatment of kidney and urinary tract infections with trimethoprim (Steen and Skold, 1985). It is therefore of interest to know how the element transposes, for considerations both of the its population biology in the field, and of the evolution of the bacteria and transposon since the modern high levels of antibiotic usage.

A third justification for wishing to determine whether an element replicates during transposition is to assist in the design and setting up of a cell-free system. With any transposon it is desirable to know in advance how much of the cellular machinery will be required in order

to allow a full transposition reaction *in vitro*. For Tn7, the requirements for transposition are already complex, in that five transposition functions are needed; it would be helpful to know whether a cell-free system must also be replication-competent.

As described in chapter 1, prokaryotic transposons can be classified according to whether or not the element is replicated during transposition. The Tn3 family move *via* an obligate intermediate (the cointegrate), containing the donor and target replicons joined by two copies of the transposon (Arthur and Sherratt, 1979). The structure of this intermediate requires that replication has occured. Demonstration that a cointegrate is formed can be taken as evidence for replicative transposition, providing that it can be shown that the cointegrate is a genuine intermediate, and not made by other cellular processes, such as homologous recombination or transposition from dimeric donors (Berg, 1983).

Phage Mu falls into the class of elements that are able to transpose both replicatively and conservatively. During lytic developement, phage growth is achieved by high levels of replicative transposition *via* unresolved cointegrates. This has also been demonstrated *in vitro*, where addition of a replication -competent extract to a intermediate will give rise to a cointegrate (Craigie and Mizuuchi, 1985).

Experimental demonstration of non-replicative transposition has been obtained in the case of initial infection by phage Mu, where a single transposition event occurs to create a lysogen. Ackroyd and Symons (1983) infected cell with phage containing a B^+/B^- heteroduplex, and found that both markers were retained in the progeny. Harshey (1984) showed that the methylation state of the infecting phage was maintained during transposition, and so no replication had occurred.

Tn10 has been shown genetically to transpose conservatively, by an experiment using heteroduplex

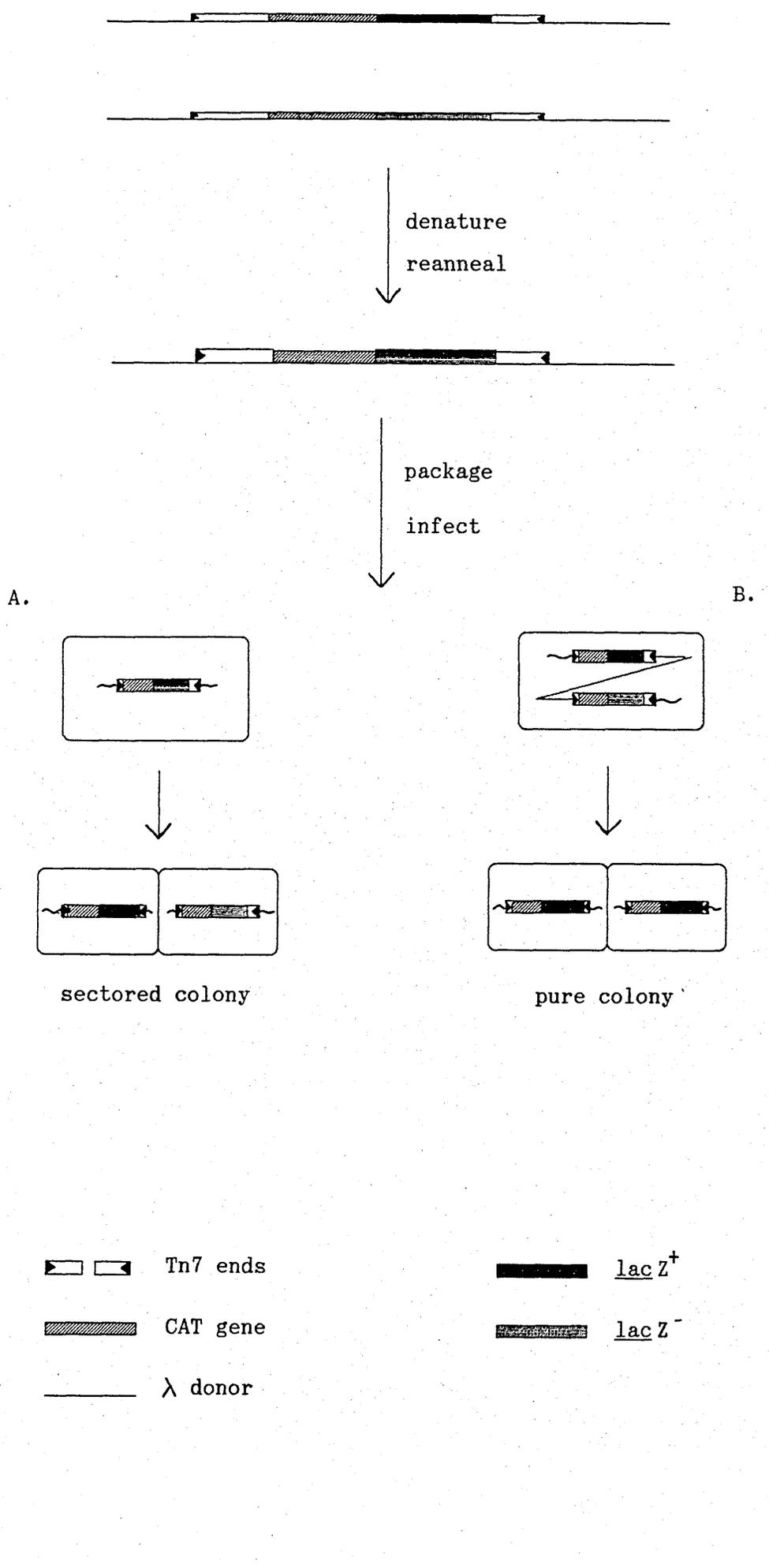

denature
reanneal
package
infect
A.
B.
sectored colony
pure colony
Tn7 ends
CAT gene
λ donor
lac Z+
lac Z-

Figure 7.1 The basis of the experiment
The two transposon donor phage contain Tn7-1 derivatives with *lacZ*$^+$ and *lacZ*$_{am}$ alleles. The donors are denatured and reannealed to generate a heteroduplex. On infection of the phage into cells, if transposition is conservative (A), then both alleles will be maintained, and will segregate at the next cell division, giving a sectored blue/white colony on Xgal plates. If the transposon undergoes semiconservative replication during transposition, the inserted element retains only one allele, and the resulting colony will be pure *lac*$^+$ or *lac*$^-$

transposons to follow the fate of each strand during transposition (Bender and Kleckner, 1986). However, for other elements the evidence concerning replication is largely indirect, such as the ability to mediate certain rearrangements (Berg, 1989).

The same is true for Tn7; there has been no direct demonstration that it transposes conservatively, although this has been the prejudice (see Discussion). In this chapter, an approach analogous to that of Bender and Kleckner (1986) has been taken to answer this question. The experiment follows the two strands of the element through the transposition process, by means of the slightly different genetic information they carry. From the segregation pattern of these markers in the transposition products, we can deduce whether the element has been replicated.

RESULTS

7.2 The basis of the experiment

The experiment is shown diagrammatically in figure 7.1. A heteroduplex transposon is generated, in which each strand of the element can be followed genetically through the reaction. The constructed element carries the *lacZ* gene, and a heteroduplex is formed between *lacZ*$^+$ and a *lacZ*$^-$ point mutant allele. Non-replicative transposition of this transposon will result in insertions which carry both markers; these will segregate at the first cell division, giving rise to *lacZ*$^+$/*lacZ*$^-$ sectored colonies (figure 7.1A), which can be scored on an indicator medium such as X-gal. If semiconservative replication of the element occurs before insertion (figure 7.1B), the product will carry only one type of *lacZ* information, and the colonies will be pure *lacZ*$^+$ or *lacZ*$^-$.

7.3 Construction of *lacZ* - Tn7-1

Both wildtype and an amber mutant allele of *lacZ* were

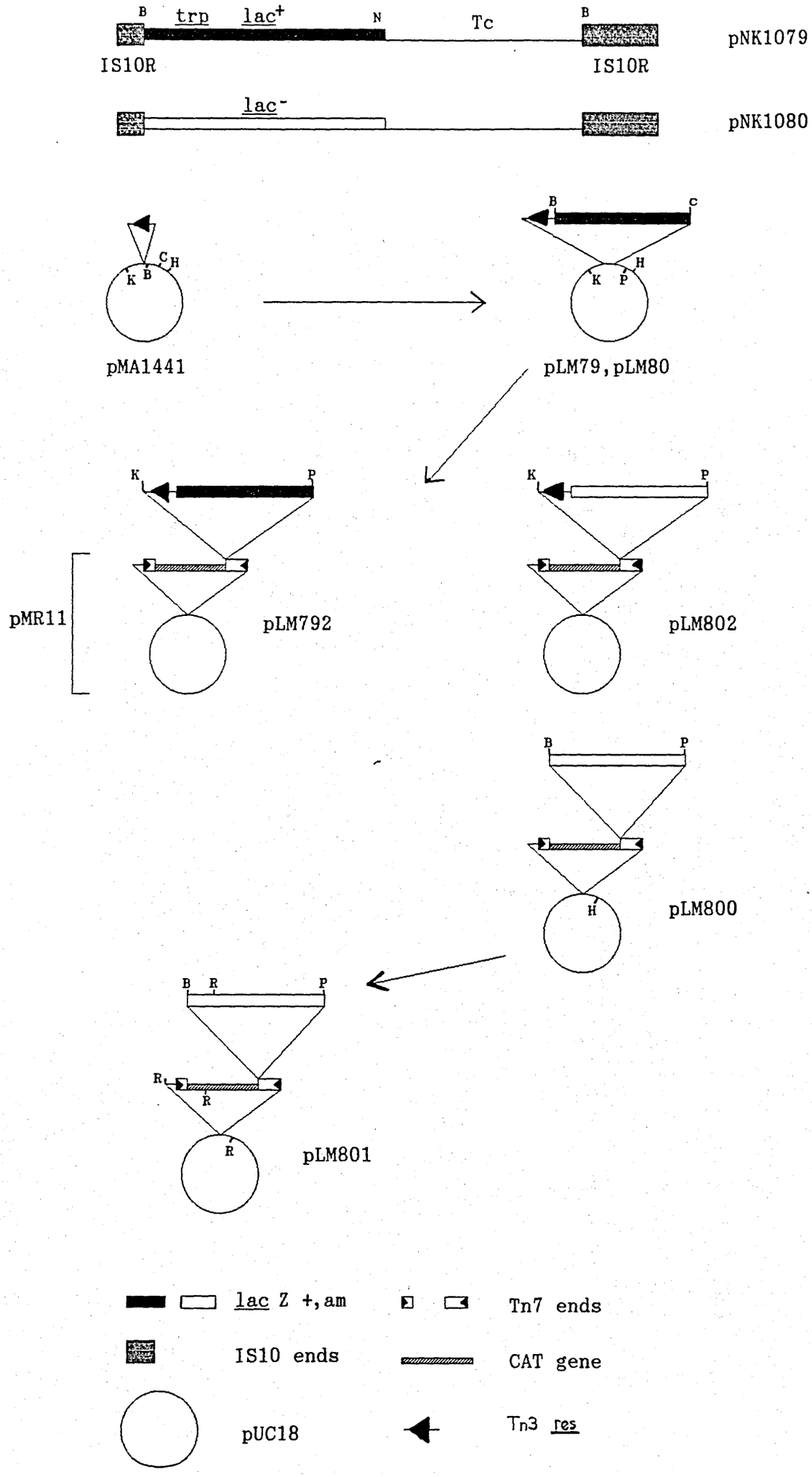

B
trp
lac+
N
Tc
B
pNK1079
IS10R
IS10R
lac-
pNK1080
B
C
K
B
C
H
K
P
H
pMA1441
pLM79, pLM80
K
P
K
P
pMR11
pLM792
pLM802
B
P
H
pLM800
B
R
P
R
R
pLM801
lac Z +, am
Tn7 ends
IS10 ends
CAT gene
pUC18
Tn3 res

Figure 7.2 Construction of Tn7-*lac*

lacZ was cloned as a 4.3 kb BamHI-NcoI fragment, which contained the 3' end of *trp*, into pMA1441 cut with BamHI and HincII; *lacZ* was expressed from the Tn3 transposase promoter in the cloned *res* site (pLM79, pLM80). KpnI-PstI *res-lac* fragments, and BamHI-PstI *lac* fragments, were both cloned into the SalI site of Tn7-1 in pMR11 (pLM792, pLM802, pLM800). The polylinker HindIII site in pLM800 was converted into an EcoRI site (pLM801).

B = BamHI	N = NdeI	K = KpnI
C = HincII	H = HindIII	R = EcoRI

obtained in trp-lac fusions from Nancy Kleckner. Both alleles of lacZ were cloned as described in figure 7.2. The cloned fragment carries the 3' end of the trpA gene, followed by the lacZ reading frame, lacking a promoter.

It was originally decided to use the Tn7-lac elements contained in pLM792 and pLM802, with lacZ under the res promoter. However, there was a concern that during their construction, some of our laboratory strains had been exposed to Tn1000. As resolution by Tn1000 resolvase would destroy cointegrates of a res-containing transposon, it was decided to use the element contained in pLM800, with lacZ expression presumably driven by P1 of Tn7. It is known that binding of TnsB to the transposon end can repress expression from P1; however, in the presence of pEAL1::Tn7△EcoRI (multi-copy plasmid carrying tnsA-E), expression of lacZ from pLM800 was sufficient to stain the cells blue (data not shown).

7.4 Strain and phage construction

The characteristics required of the host strain for this experiment, and the techniques used to obtain these, are described below.

a) lacZ$^-$, to allow the transposon phenotype to be seen.

b) supO, to avoid suppression of the amber allele.

c) lambda-free, to prevent interactions with the transposon vector.

d) recF, to reduce pseudo-transposition by homologous recombination.

e) mut$^-$, to prevent repair of the heteroduplex in the cells.

Strain DS947 has the first four of these characteristics. To make the cells mut$^-$, phage P1 was used to transduce a Tn5 insertion mutation in the mutL gene into this strain; the donor strain is RH302 (Bender and Kleckner, 1986). The mutation is marked with kanamycin

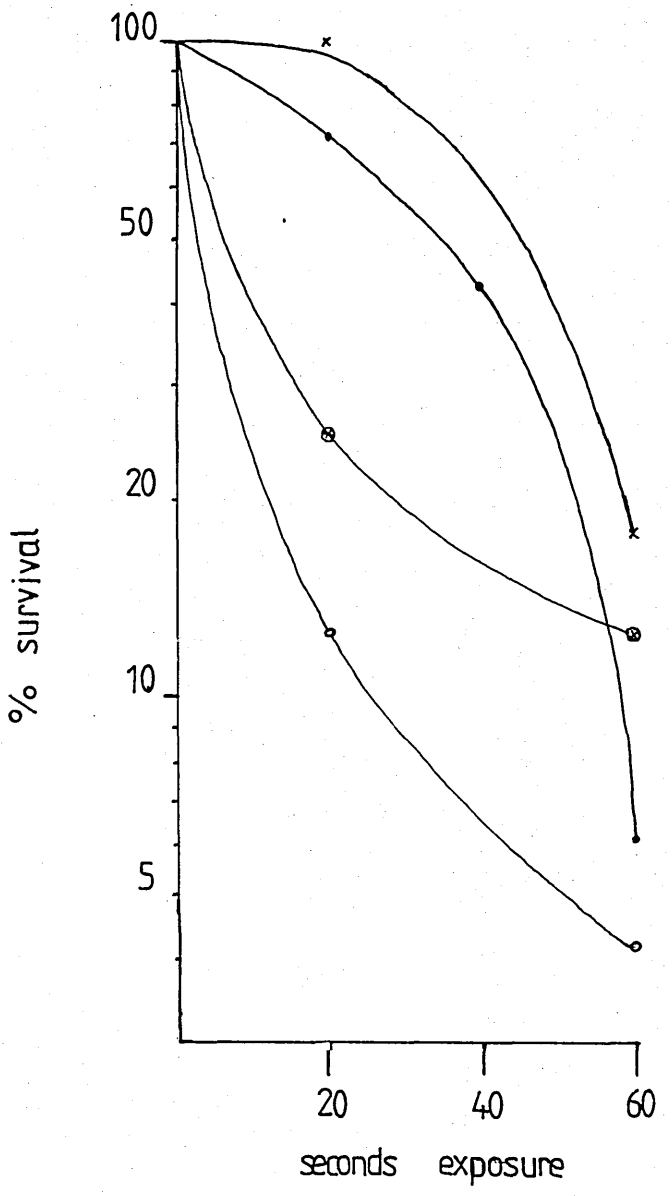


Figure 7.3 UV sensitivity of kanamycin-resistant transductants

Stationary cultures were serially diluted, spotted onto agar plates, and exposed to UV irradiation for the indicated time.

×	DS801	⊗	DS801K
•	DS947	o	DS947K

'K' indicates a Km^r, putative <u>mutL</u> derivative, of the strain concerned.

resistance and so can be readily selected. Isolates were checked for their mut status in two ways:

i) sensitivity to UV irradiation. The survival of kanamycin resistant transductants in DS801 (rec$^+$) and DS947 (recF), following UV irradiation, is shown in figure 7.3. The transductant strains were around 10-fold more sensitive to UV than the parent strains, or a control transductant strain carrying Tn5 at the xerB locus (data not shown), at this low level of cell killing.

ii) mutator phenotype at several loci. Mutation to resistance to rifampicin, naladixic acid, and valine were assayed, and the results are shown in table 7.1. The strains showed approximately a 100-fold increase in mutation rate at these loci; although many more independent cultures would need to be assayed to obtain good statistics, these data are consistent with previously reported mutation rates for this allele (Glickman and Radman, 1980). A Km^R strain with the described properties was chosen and named EM1.

In neither transformation nor conjugation is it clear that the entering DNA strands are retained in their original duplex. For this reason, phage lambda was chosen as a delivery vector for the transposon. It is then necessary to ensure that the phage does not survive in the cells, so that selection for transposon drug resistances will reveal only products of transposition. The phage must therefore be cI$^-$ to prevent maintenance by lysogeny, and must not be able to replicate in, and lyse, the assay strain. A suitable phage is λgt10.

The basis of exclusion of phage replication and cell lysis is the spi phenotype; a phage which carries wildtype red and gam genes cannot replicate on a P2 lysogen. The red and gam genes are carried on the 'stuffer' fragment of replacement vectors, eg EMBL4, allowing selection of recombinant phage by loss of these genes. They are also present on λgt10. It was therefore necessary to make the host strain, DS947 mutL, into a P2 lysogen.

Table 7.1 Mutation rates of kanamycin-resistant transductants

Strain	Naladixic acid	Rifampicin	Valine
DS947	$3x10^{-11}$	$7x10^{-11}$	$7x10^{-9}$
RH302	$2x10^{-9}$	$3x10^{-9}$	$3x10^{-8}$
DS947x	$1x10^{-9}$	$1x10^{-8}$	$5x10^{-9}$
DS947K1	$3x10^{-8}$	$3.5x10^{-6}$	$5.5x10^{-8}$
DS947K2	$1x10^{-7}$	$1.6x10^{-6}$	$1.3x10^{-8}$

Mutation rates are expressed as resistant colonies per viable cell.
In the top part of the table, the cultures were grown to stationary phase before assaying; in the lower part, the cultures were in log phase.

K1 and K2 are independently isolated kanamycin-resistant transductants. K1 was named EM1

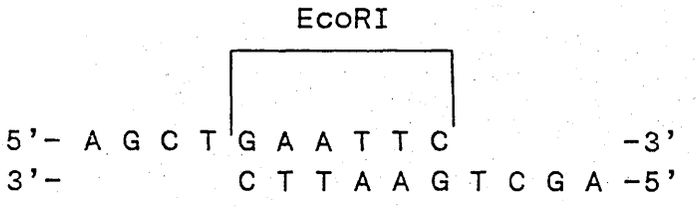


Figure 7.4 Sequence of the HindIII-EcoRI oligonucleotide
The oligonucleotide is self-complementary, and is shown in its double-stranded form.

P2 lysogens were isolated by spotting phage onto a lawn of EM1 cells and picking from the turbid regions of the plate. Cells isolated in this way were cross- streaked against P2, and against λ phage; strains were chosen which were unable to plate EMBL4, but were lysed by a phage stock of an EMBL4 replacement library which lacked the stuffer region. Their susceptibility to these phage was confirmed more quantitatively (table 7.2). The 10^4-fold exclusion of the parent phage was considered to be strong enough selection against phage persistence, given the expected transposition frequency of 1%. The same level of exclusion was seen for λgt10 (data not shown). The P2 lysogen was named EM2.

λgt10 has a unique EcoRI site, situated in the cI gene. Recombinant phage with insertions at the EcoRI site are therefore unable to form lysogens; this is desirable both for the transposition experiment, and to allow selection of recombinant phage. If the packaged ligation is plated on an hfl (high frequency of lysogeny) strain, any cI^+ non-recombinant phage will efficiently form lysogens, and the only clear plaques seen will be recombinants.

In order to make use of the EcoRI site to clone the lacZ transposons into the phage, it was necessary to create an EcoRI site at the right-hand end of the element in pLM800 (figure 7.2). This was achieved by means of a self-complementary oligonucleotide, which consisted of an EcoRI site plus the 5' extension of a HindIII site (figure 7.4). Digestion of pLM800 with HindIII, and religation in the presence of the annealed oligonucleotide and HindIII enzyme, generated a plasmid in which the polylinker HindIII site is replaced by an EcoRI site (pLM801). Partial digestion of this plasmid with EcoRI gives a 7.4 kb transposon fragment flanked by EcoRI sites, which can be cloned into the EcoRI site of λgt10.

A $lacZ^+$ derivative of pLM801 (referred to as p791) was required. The first approach taken was to transform

Table 7.2 Exclusion of EMBL4 from P2 lysogens

Strain	EMBL 4	EMBL 4 lib
EM 1	4×10^8	6×10^8
EM 2 (P2)	5×10^3	5×10^8

pLM801 into EM1 and screen for *lacZ*$^+$ on Xgal plates. Blue colonies were obtained, at a frequency of 1-20%, and their plasmids were checked for retention of all markers, and by restriction digestion. However, on retransformation into fresh DS947, all the plasmids gave white colonies on Xgal; presumably the original mutations were all chromosomal amber suppressors. Although more than 40 blue colonies were tested, none contained plasmids carrying a *lacZ*$^+$ gene.

The alternative approach was to construct p791 from pLM79 and pMR11 (as in figure 7.2), as intended when pLM800 was made. The HindIII site of pMR11 was converted into an EcoRI site using the oligonucleotide described above (pLM20). Unfortunately, time did not permit the construction of p791 to be completed, nor the construction of the transposon-bearing phage derivatives (referred to as λ79 and λ80).

7.5 Construction of control strains

In addition to transposition, Cm^R sectored colonies might be generated by homologous recombination with the chromosome. The number of sectored colonies produced by this means would give a measure of the number of heteroduplex transposons in the infecting population that survive packaging, infection, and transposition, and avoid repair before cell division. To promote events of this type, extensive homology for recombination was provided by making lambda lysogen derivatives of EM1 and EM2.

Wildtype lambda phage were spotted onto lawns of EM1 and EM2. Colonies were picked from turbid regions, and the strains isolated were tested for lysogeny; lysogens do not plate λ$^+$ and λ*cI*, but are lysed by λ*vir*. For EM1, lysogens were readily isolated and identified (EM11). In the case of EM2, the primary plates showed much fewer plaques than for EM2, due to the *spi* phenotype. Cells isolated from the turbid regions of the plate were assayed for susceptibility to the test phage as above; the results

were not as clear, due to the low level of lysis, but a strain was found which appeared to be additionally resistant to λ^+ but not to λvir (EM12). Before using this strain, however, it would be desirable to check for the presence of lambda, eg by testing for the ability to produce phage on induction, or for phage markers by complementation of defective phage. Interestingly, none of the P2 strains tested, including EM2, were able to plate λcI detectably; the reason for this is not known.

7.6 Preparation of repair-deficient packaging extracts

In addition to use of mutL strains, to reduce the amount of destruction of the heteroduplex by mismatch repair, the packaging extracts used in the experiment should also be repair-deficient. Extracts are made from standard packaging strains which have been made mutL by P1 transduction (Bender and Kleckner, 1986).

The defective prophage contained in the two packaging strains each make a subset of the structural proteins required for packaging, and so cannot produce phage; a mixture of extracts from the two strains is competent to package exogenous phage DNA. The repressors of the prophage contain a temperature-sensitive mutation, so expression of head and tail genes is induced by heating the cultures to 42^{o}C. Efficient induction is dependent on the maintenance of this temperature sensitivity, and so strains should be checked before use and isolates used which grow at 30^{o}C but not at 42^{o}C.

Extracts were prepared by standard procedures (Sambrook et al.,1989), and were assayed for packaging ability using EMBL4 phage DNA. The extracts produced around 10^5 pfu/ug DNA which is less than the efficiency of commercial packaging extracts (10^7/ug). The lower packaging activity is due to efficiency of induction. Although all small cultures used to inoculate the preparative cultures were tested before use for temperature sensitivity, by the time the large culture had

reached an A_{600} of around 0.3 (the stage at which the prophage were induced), a significant proportion of the cells were cI^+ revertants. Hence only a proportion of the cells will have been induced to synthesise head and tail proteins, leading to the lowered activity of the extracts. The high level of reversion is presumably due to the mutator activity of the strains. However, the activity of the extracts was sufficient for them to be used in the experiment.

7.7 DISCUSSION

Unfortunately there was insufficient time to complete the constructions required to carry out the heteroduplex experiment. The next stage of the work would be to construct p791, as described above. Both p791 and pLM801 would then be partially digested with EcoRI, and the transposon fragment ligated into EcoRI-cut λgt10. Recombinants can be selected on an hfl strain. The clear plaque-forming phage must then be checked for their lacZ status, which is necessary because pseudo-recombinants (cI mutants) can occur, and did so at a particularly high frequency when dephosphorylated vector was used (data not shown), presumably because phage which have managed to ligate have undergone mutation at the EcoRI site.

In addition to scoring for sectoring, there are several desirable control experiments. One is to show that any sectoring is not due to mutation; this is important given the observed increase in mutation rate in EM1. Transposition experiments as above, using homoduplex donors, will give an estimate of the level of $lacZ^+/lacZ_{am}$ interconversion.

Infection of the heteroduplex phage into EM11 will generate Cm^R colonies by homologous recombination. Ideally, 50% of the colonies would be sectored; however, this number might be reduced by inefficient denaturing, low efficiency of packaging of mismatched phage, or repair

of the mismatch in the packaging extracts or the recipient cell. Bender and Kleckner (1986) obtained around 16% sectored colonies. They suggest that one reason is mismatch repair, despite the use of repair-deficient strains and packaging extracts. Inefficiency of formation of heteroduplex is also a probable explanation. It should be possible to select against homoduplex phage after reannealing, if the two phage also carry restriction site polymorphisms; if the sites were engineered so that each phage had one of the sites, then in the presence of both enzymes the parent phage would be cut, and so would not package, but the heteroduplex phage would survive.

A concern with this experimental design is that cointegrates might be selected against. One way in which this could happen is if the lysis genes carried on the phage were to be expressed during the lifetime of a cointegrate, when phage sequences are maintained, so killing any cells which bore cointegrates. The homologous recombination experiment tests for this; if stable homologous insertions can be made, then we can expect cointegrates to survive. If not, the phage can be modified, for example by introducing $\underline{S}_{am}$.

A second potential for counterselecting cointegrates arises by insertion of $\underline{red}^+$ and $\underline{gam}^+$ into the chromosome of a P2 lysogen. The interaction between red/gam and P2 which acts to prevent lambda replication may also reduce the viability of a strain containing both functions, for example by causing a replication block. The fact that EM12 was constructed suggests that the double lysogen is viable (given the caveats on the verification of EM12's genotype discussed above), but its growth characteristics have not been investigated. The homologous recombination control in EM12 would indicate whether a P2 strain with an integrated lambda phage is viable.

Having obtained transposition products from the heteroduplex experiment, these can be further analysed genetically and physically, as described by Bender and

Kleckner (1986). The transposition products can be used to generate structures identical to transpositional cointegrates, in order to observe their behaviour. Cointegrates can be made by infecting a white (lacZ$^-$) transposition product with λ79 (lacZ$^+$) and selecting for lacZ$^+$ in the absence of transposition functions. Again, if the cointegrate can be formed in this way, we can be more confident that the donor exclusion system is not biassing the result against cointegrates.

There are alternative explanations for sectoring, which have been discussed by Bender and Kleckner. Transposition could generate cointegrates, which are then resolved with high efficiency within the first few cell generations. Resolution can give rise to sectored colonies, only if the recombination does not ocur at a fixed site in the transposon, so can occur in favour of either allele; this would not be the case with a Tn3-like resolution system. The model puts tight restrictions on the time at which resolution occurs; to explain the sectoring observed in the Tn10 experiment, for example, resolution would have to occur in the second or third generation. A second difficulty with the resolution scenario is that the number and distribution of sectored colonies was identical in the transposition and homologous recombination experiments. Randomly positioned resolution would show a bias in the proportion of retention of the lac alleles, depending on the position of the heterology within the transposon, and with respect to any preferred recombination sites.

Two other alternative explanations are more difficult to refute from the available data. The first is that the transposon replicates conservatively during transposition, generating a new mismatched transposon. Secondly, Howard Nash suggested that the resolution event itself may generate a large tract of heteroduplex, as proposed in models of homologous recombination, and over a shorter region in lambda integrative recombination (eg de

Massy et al.,1989). As the predictions from this model are identical to those based on conservative transposition, this argument cannot be ruled out.

The prediction for Tn7 is that transposition is non-replicative, for several reasons. Firstly, there is no evidence for a Tn3-like site-specific recombination system within Tn7. By analogy with Tn3, we might expect to find a site and enzyme related to res/resolvase if transposition involves cointegrate resolution; however no such homology has been found at the DNA level. There is a gene with homology to the Int family of site-specific recombinases in Tn7 (Sundstrom et al.,1988); the function of this gene is unclear, but it does not appear to be required for resolution of transpositional cointegrates, since transposition of mini-Tn7's, consisting of only the cis-acting sequences, produces only simple insertions (unpublished observations; Rogers et. al., 1986), indicating that any resolution system must reside completely within 200 bp of the left end or 560 bp of the right end. Artificially created cointegrate analogues are stable in recA strains, suggesting that Tn7 does not contain its own resolution system (Hodge, 1983); however this can also be interpreted as showing that resolution can only act co-transpositionally. Although these results suggest that Tn7 does not undergo replicative transposition via a cointegrate, they do not exclude other replicative pathways.

There is a correlation between conservative transposition and the presence of dam methylation sequences in the element's termini. IS10 (Roberts et. al., 1985) and IS50 (Yin et al., 1988) both contain dam sites and transpose conservatively; the Tn3 family transpose replicatively and do not contain dam methylation sites in their ends. In the case of IS10, transposition is coupled to DNA replication by the effects of dam methylation, which can be seen as an adaptation of non-replicating transposons, allowing them to exploit cellular replication

to increase their copy number.

The right end of Tn7 contains a dam methylation site, in the second 22 bp repeat (Lichtenstein and Brenner, 1982). By comparison with IS10 and IS50, it is tempting to suggest that Tn7 may also transpose conservatively. However, no differences in transpositional behaviour of Tn7 could be detected when compared in dam$^+$ and dam$^-$ strains (Ekaterinaki, 1987).

Nancy Craig's group in San Francisco have carried out in vitro transposition assays with Tn7, using substantially purified transposition functions (R.Bainton and N.Craig, personal communication). In this system, no cointegrates are observed, with all the transposition events being simple insertions, and the cleaved donor molecule being detected as a linear. However, it is not clear whether this system is replication-competent; it can be argued that it is not, and so this could be an artefactual result. At least one of the components is added in the form of a crude extract, however, which may be able to support at least repair synthesis.

The heteroduplex experiment to determine whether Tn7 transposes replicatively or conservatively, would be interesting and informative. The required strains have been constructed and tested, and most of the transposon constructs have been made. It should be relatively easy for a future worker to complete the experiment, and so find the answer to this basic mechanistic question.

CHAPTER 8

CONCLUDING REMARKS

The long-term aim of our work on Tn7 is to study the transposition reaction in a cell-free system. A wealth of information and insight has been gained from *in vitro* experiments on phage Mu; a similar depth of understanding of the transposition of elements like Tn10 and Tn7 should also be possible.

The approach we have taken is to overexpress and purify each of the transposition functions independently, and then to reconstitute a transposition system. This approach has the advantage that the initial cell-free system is very clean, so the difficulties inherent in working with crude cell extracts are removed. In addition, much can be learned, during the purification stages, about the properties, activities and requirements of each protein, so facilitating the design of an optimised *in vitro* system.

An alternative approach was followed by Nancy Craig in designing her Tn7 *in vitro* system, which was to set up the reaction with crude cell extracts ; the same route was taken by Mizuuchi in the early Mu experiments, and by Nancy Kleckner's group with Tn10. Once transposition activity has been established, a cleaner system can be evolved from this by fractionation of the crude extracts and assaying for maintenance of transposition competence. Transposition functions can be purified by the ability to complement extracts made from cells which are mutant in that function. As a route to purification, this has the advantage that the protein is known to be fully active; in contrast, protocols which follow purification by SDS-PAGE or by assay for a part of the function, might produce preparations which, though pure, are not transposition-competent.

In retrospect, in the case of Tn7, I would tend to favour the crude system as a primary approach. Tn7 transposition is notably complex. For transposition to an *att* site, four element-encoded functions would be required, and it may not be trivial to reconstitute all

the purified components such that they are all active. There are many other details of transposition about which we know very little; for example whether any host proteins are involved. There is a large number of experimental variables, which would make troubleshooting very complex, in contrast to simpler systems such as cointegrate resolution, where there are fewer components and unknown factors.

The final result of either approach is the same: an *in vitro* system and a set of purified transposition functions. The difficulty with the reconstitution approach (apart from purifying proteins for which there is no assay) is that the step from components to transposition reaction may be non-trivial; however, once a crude system has been established, much information can be generated, whilst purification work is in progress.

The preferred route to crude *in vitro* work would be to use extracts from a cell which overexpresses all the *tns* genes. This avoids the need to mix extracts; the overexpression is probably advantageous as the levels of transposition functions *in vivo* tend to be low. Such an overexpressing clone has been constructed in this laboratory; given the encouraging results from the Craig laboratory, the prospects for setting up a cell-free transposition reaction with highly enriched extracts, are very good.

The majority of the work presented in this thesis concerns the purification of TnsB and the characterisation of its DNA-binding activity. The purified protein can now be used in *in vitro* work by future researchers. Futher characterisation of the protein could include partial proteolysis to locate the active domains of the protein. An experiment which has been considered is to combine partial proteolysis with South-Western blotting, to locate a fragment with DNA-binding activity on SDS-PAGE; the peptide can be eluted from the gel and partially sequenced to determine its location within TnsB.

A useful experimental tool at this stage would be to have antisera raised against the various Tns proteins. Antibodies could be used to follow small amounts of protein through a purification or an experiment, or linked to a column to provide an affinity purification step. Antibodies could also be used for staining in electron-microscopic studies of the transposition functions and process, both *in vivo* and *in vitro*; for example, to determine which proteins are present in 'transpososome' complexes, as has been done for phage Mu functions by Brigitte Lavoie and George Chaconas.

In the work described in this thesis, the gel retardation system proved to be particularly informative, in both purification and characterisation. The main question remaining is the nature of the complexes seen on the gels. It should be possible to determine the stoichiometries of the major complexes, by using labelled protein, or by the method described in chapter 3, whereby gel slices containing the complexes are loaded into the wells of an SDS gel; in this experiment, antibodies to the protein would be useful, to detect the low levels of protein present. To complement and extend this work, footprinting of isolated complexes could be carried out.

The ligation experiments described in chapter 5 provide a powerful tool for analysing interactions between DNA sites and their binding proteins, and are worthy of further investigation. The experiments could be extended to examine the interaction between the transposon ends and the target site. For example, a ligation experiment containing TnsB, TnsD, a hotsite-containing linear fragment, and a supercoiled mini-Tn7 (eg pLM11) could be carried out to determine whether there was a linkage change in the hotsite ligation products, or an enhancement of catenation, due to the presence of the proteins. The substrates for these experiments are available, but they have not been done satisfactorily for want of a TnsD-containing extract of sufficient activity. Further

experiments could subsequently be done in which both the DNA fragments are linear, with compatible restriction termini; inferences concerning the three-dimensional arrangement of the strands extending from the synapse, can be made from the patterns of 'partner choice' in the ligation.

The *in vivo* experiments described here were initiated in order to fill some of the gaps in our knowledge about Tn7 transposition. Assaying transposition rapidly and accurately remains a problem. The papilliation assay in its current form does not provide a measure of transposition frequency, but it should be possible to modify the idea to generate a useful screen. As stated in chapter 6, much information could then be generated very rapidly by mutagenesis experiments.

In addition to its interesting transpositional properties, one fascinating feature of Tn7 biology is its complexity. The element transposes *via* a series of enzymatic steps with consequences that are indistinguishable from the transposition of an IS element, for example; Tn7 requires multiple protein functions, accounting for nearly 9 kb of genetic coding capacity, to carry out essentially the same reaction as the single IS element transposase. The reason for the complexity is not clear. However, it is apparent that there must be many interactions between DNA sites, between proteins, and between DNA and protein, in order to coordinate the activities required for transposition. It is possible that it is in this field, of communication and coordination between macromolecules, that Tn7 will feature prominently; its natural occurrence in *E.coli*, an organism of choice for genetic and molecular analysis, will facilitate its detailed characterisation, perhaps providing paradigms on which experiments in other systems can be based.

BIBLIOGRAPHY

Adler H.I., Fisher W.D., Cohen A., and Hardigree A.A. 1967. Miniature Escherichia coli cells deficient in DNA. Proc.Natl.Acad.Sci. 57: 321-326.

Adzuma K. and Mizuuchi K. 1988. Target immunity of Mu transposition reflects a differential distribution of MuB protein. Cell 53: 257-266.

Adzuma K. and Mizuuchi K. 1989. Interaction of proteins located at a distance along DNA: mechanism of target immunity in the Mu DNA strand-transfer reaction. Cell 57: 41-47.

Akroyd J., and Symonds N. 1983. Evidence for a conservative pathway of transposition of bacteriophage Mu. Nature 303: 84-86.

Almer A., Rudolph H., Hinnen A., and Horz W. 1986. Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. EMBO J. 5: 2689-2696.

Amann E., Brosius J., Ptashne M. 1983. Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in Escherichia coli. Gene 25: 167-179.

Arciszewska L., Drake D., and Craig N.L. 1989. Transposon Tn7: cis-acting sequences in transposition and transposition immunity. J.Mol.Biol. 207: 35-52.

Arthur A., and Sherratt D.J. 1979. Dissection of the transposition process: a transposon-encoded site-specific recombination system. Mol. Gen. Genet. 175: 267-274.

Balke V.L., and Gralla J.D. 1987. Changes in linking number of supercoiled DNA accompany growth transitions in Escherichia coli. J.Bacteriol. 169: 4499-4506.

Barth P.T. and Datta N. 1977. Two naturally occurring transposons indistinguishable from Tn7. J.Gen. Microbiol. 102: 129-134.

Barth P.T., Datta N., Hedges R.W., and Grinter N.J. 1976. Transposition of a deoxyribonucleic acid sequence encoding trimethoprim and streptomycin resistance from R483 to other replicons. J.Bacteriol. 125: 800-810.

Barth P.T. and Grinter N.J. 1977. Map of RP4 derived by insertion of transposon C. J.Mol.Biol. 113:455-474.

Barth P.T., Grinter N.J., and Bradley D.E. 1978. Conjugal transfer system of plasmid RP4: analysis by transposon 7 insertion. J.Bacteriol. 133: 45-52.

Barry G.F. 1986. Permanent insertion of foreign genes into the chromosome of soil bacteria. Bio/technology4:446-449.

Baumruker T., Sturm R., and Herr W. 1988. OBP100 binds remarkably degenerate octamer motifs through specific interactions with flanking sequences. Genes Devt. 2: 1400-1413.

Bednarz A.L. 1989. Dissection of the Tn3 res site. PhD. thesis, University of Glasgow .

Bender J., and Kleckner N. 1986. Genetic evidence that Tn10 transposes non-replicatively. Cell 45: 801-815.

Benjamin H.W., and Cozzarelli N.R. 1988. Isolation and characterization of the Tn3 resolvase synaptic intermediate. EMBO J. 7: 1897-1905.

Benjamin H.W., and Kleckner N. 1989. Intramolecular transposition by Tn10. Cell 59: 373-383.

Berg D.E. 1977. Insertion and excision of the transposable kanamycin resistance determinant Tn5. In: DNA insertion elements, plasmids and episomes, Eds. Bukhari A.I., Shapiro J.A., and Adhya S.L.; Cold Spring Harbor Press, New York.

Berg D.E. 1983. Structural requirement for IS50-mediated gene transposition. Proc.Natl.Acad.Sci. 80: 792-796.

Berg D.E. 1989. Transposon Tn5. In: Mobile DNA, Eds. Berg D.E., and Howe M.M. American Society for Microbiology, Washington, D.C.

Birnboim H.C., and Doly J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7: 1513-1523.

Bishop R., and Sherratt D.J. 1984. Transposon Tn1 intramolecular transposition. Mol.Gen.Genet. 196: 117-122.

Boocock M.R., Brown J.L., and Sherratt D.J. 1986. Structural and catalytic properties of specific complexes between Tn3 resolvase and the recombination site res. Biochem. Soc. Trans. 14: 214-216.

Bradford M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding. Anal. Biochem. 72: 248-254.

Brevet J., Faure F., and Borowski D. 1985. Tn7-encoded proteins. Mol.Gen.Genet. 210:258-264.

Cameron F.H., Groot Obbink D.J., Ackerman V.P., and Hall R.M. 1986. Nucleotide sequence of the AAD(2'') aminoglycoside adenylyl transferase determinant aadB. Evolutionary relationship of this region with those surrounding aadA in R358-1 and dhfrII in R388. Nucl. Acids Res. 14: 8625-8635.

Campbell A. 1981. Evolutionary significance of accessory DNA elements in bacteria. Ann.Rev.Microbiol. 35: 55-83.

Caparon M.G., and Scott J.R. 1989. Excision and insertion of the conjugative transposon Tn916 involves a novel recombination mechanism. Cell 59: 1027-1934.

Carey J. 1988. Gel retardation at low pH resolves trp repressor-DNA complexes for quantitative study. PNAS 85: 975-979.

Carthew R.W., Chodosh L.A., and Sharp P.A. 1985. An RNA polymerase II transcription factor binds to an upstream element in the Adenovirus major late promoter. Cell 43: 439-448.

Caruso M. and Shapiro J. 1982. Interactions of Tn7 and temperate phage F116L of Pseudomonas aeruginosa. Mol. Gen.Genet. 188: 292-298.

Chaconas G., Gloor G., and Miller J.L. 1985. Amplification and purification of the bacteriophage Mu-encoded B transposition protein. J.Biol.Chem. 260: 2662-2669.

Chao L., Vargas C., Spear B.B., and Cox E.C. 1983. Transposable elements as mutator genes in evolution. Nature 303: 633-635.

Connell N., Han Z., Moreno F., and Kolter R. 1987. An E.coli promoter induced by the cessation of growth. Molec.Microbiol. 1: 195-201.

Craig N.L. 1989. Transposon Tn7. In: Mobile DNA, Eds. Berg D.E., and Howe M.M. American Society for Microbiology, Washington, D.C.

Craig N.L., and Nash H.A. 1983. The mechanism of phage lambda site-specific recombination: collision versus sliding in att site juxtaposition. ICN-UCLA Symp. Mol. Cell. Biol. 10: 617-623.

Craigie R., Arndt-Jovin D.J., and Mizuuchi K. 1985. A defined system for the DNA strand-transfer reaction at the initiation of bacteriophage Mu transposition: protein and DNA substrate requirements. Proc. Natl. Acad. Sci. USA 82: 7570-7574.

Colloms S.D., Sykora P., Szatmari G, and Sherratt D.J. 1990. Recombination at ColE1 cer requires the Escherichia coli xerC gene product, a member of the lambda integrase family of site-specific recombinases. J. Bacteriol 172: 6973-6980.

Craigie R., Fujiwara T., and Bushman F. 1990. The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. Cell 62: 829-837.

Craigie R., and Mizuuchi K. 1985a. Cloning of the A gene of bacteriophage Mu and purification of its product, the Mu transposase. J.Biol.Chem. 260: 1832-1835.

Craigie R., and Mizuuchi K. 1985b. Mechanism of transposition of bacteriophage Mu: structure of a transposition intermediate. Cell 41: 867-876.

Craigie R., and Mizuuchi K. 1986. Role of topology in Mu transposition: mechanism of sensing the relative orientation of two DNA segments. Cell 45: 793-800.

Craigie R., and Mizuuchi K. 1987. Transposition of Mu DNA: joining of Mu to target DNA can be uncoupled from cleavage at the ends of Mu. Cell 51: 493-501.

Craigie R., Mizuuchi M., and Mizuuchi K. 1984. Site-specific recognition of the bacteriophage Mu ends by the Mu A protein. Cell 39: 387-394.

Darby N.J., and Creighton T.E. 1990. Folding proteins. Nature 344: 715-716.

Datta N., and Hedges R.W. 1972. Trimethoprim resistance conferred by W plasmids in Enterobacteriacae. J.Gen. Microbiol. 72: 349-355.

Datta N., Nugent M., and Richards H. 1981. Transposons encoding trimethoprim or gentamycin resistance in medically important bacteria. CSHSQB 45: 45-51.

Davis M.A., Simons R.W., and Kleckner N. 1985. Tn10 protects itself at two levels from fortuitous activation by external promoters. Cell 43: 379-387.

Derbyshire K.M., Hwang L., and Grindley N.D.F. 1987. Genetic analysis of the interaction of the insertion sequence IS903 transposase with its terminal inverted repeats. Proc.Natl.Acad.Sci. 84: 8049-8053.

Derbyshire K.M., Kramer M., and Grindley N.D.F. 1990. Role of instability in the cis action of the insertion sequence IS903 transposase. Proc.Natl.Acad.Sci. 87: 4048-4052.

Doolittle W.F., and Sapienza C. 1980. Selfish genes, the phenotype paradigm, and genome evolution. Nature 284: 601-603.

Dorman C.J., Barr G.C., NiBhriain N., Higgins C.F. 1988. DNA supercoiling and the anaerobic and growth phase regulation of tonB gene expression. J.Bacteriol. 170: 2816-2826.

Drew H., and Travers A. 1985. DNA bending and its relation to nucleosome positioning. J. Mol. Biol. 186: 773-790

Ekaterinaki N. 1987. Expression and function of Tn7 transposition proteins. PhD thesis, University of Glasgow.

Ely B. 1982. Transposition of Tn7 occurs at a single site on the Caulobacter crescentus chromosome. J.Bacteriol. 151: 1056-1058.

Engels W.R. 1989. P elements in Drosophila melanogaster. In: Mobile DNA, Eds. Berg D.E., and Howe M.M. American Society for Microbiology, Washington, D.C.

Evans L.R., and Brown N.L. 1987. Construction of hybrid Tn501/Tn21 transposases in vivo: identification of a region of transposase conferring specificity of recognition of the 38 bp terminal inverted repeats. EMBO J. 6: 2849-2853.

Falvey E., and Grindley N.D.F. 1987. Contacts between gamma-delta resolvase and the gamma-delta res site. EMBO J. 6: 815-821.

Fayet O., Ramond P., Prere M.F., and Chandler M. 1990. Functional similarity between retroviruses and the IS3 family of bacterial insertion sequences ? Molec. Microbiol. 4: 1771-1777.

Fedoroff N.V. 1989. Maize transposable elements. In: Mobile DNA, Eds. Berg D.E., and Howe M.M. American Society for Microbiology, Washington, D.C.

Feinberg and Vogelstein. 1983. A technique for radio-labelling DNA restriction fragments to high specific activity. Anal.Biochem 132: 6-13

Fling M., and Elwell L. 1980. Protein expression in Escherichia coli minicells containing recombinant plasmids specifying trimethoprim-resistant dihydro-folate reductases. J.Bacteriol. 141: 779-785.

Fling M., and Richards C. 1983. The nucleotide sequence of the trimethoprim-resistant dihydrofolate reductase gene harboured by Tn7. Nucl.Acids Res. 11: 5147-5158.

Fling M., Kopf J., and Richards C. 1985. Nucleotide sequence of the transposon Tn7 gene encoding an aminoglycoside-modifying enzyme, 3''(9)-O-nucleotidyltransferase. Nucl.Acids Res. 13: 7095-7106.

Flinn H., Burke M., Stirling C.J., and Sherratt D.J. 1989. Use of gene replacement to construct Escherichia coli strains carrying mutations in two genes required for stability of multicopy plasmids. J.Bacteriol. 171: 2241-2243.

Flores C.C. 1991. Investigation of the transposition functions of transposon Tn7. PhD thesis, University of London.

Flores C., Qadri M.I., and Lichtenstein C. 1990. DNA sequence analysis of five genes; tnsA, B, C, D and E, required for Tn7 transposition. Nucl.Acids Res. 18: 901-911.

Fried M., and Crothers D.M. 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucl. Acids Res. 9: 6505-6525

Fried M.G., and Crothers D.M. 1984. Equilibrium studies of the cyclic AMP protein-DNA interaction. Meths. Enzymol : 241-262.

Fried M.G., and Crothers D.M. 1984. Kinetics and mechanism in the reaction of gene regulatory proteins with DNA. Meths. Enzymol. : 263-282.

Galas D.J. and Chandler M. 1982. Structure and stability of Tn9-mediated cointegrates. J.Mol.Biol. 154: 245-272.

Galas D.J. and Chandler M. 1989. Bacterial insertion sequences. In: Mobile DNA, Eds. Berg D.E., and Howe M.M. American Society for Microbiology, Washington, D.C.

Galas D.J. and Schmitz A. 1978. DNAase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucl.Acids Res. 5: 3157-3170.

Gamas,P., Galas D., and Chandler M. 1985. DNA sequence at the ends of IS1 required for transposition. Nature 317: 458-460.

Garner M.M., and Revzin, A. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. Nucl. Acids Res. 9: 3047-3060

Gay N., Tybulewicz V., and Walker J. 1986. Insertion of transposon Tn7 into the Escherichia coli glmS transcriptional terminator. Biochem.J. 234: 111-117.

Gill R.E., Heffron F., and Falkow S. 1979. Identification of the protein encoded by the transposable element Tn3 which is required for its transposition. Nature 282: 797-801.

Glickman B.W., and Radman M. 1980. Escherichia coli mutator mutants deficient in methylation-instructed mismatch correction. Proc.Natl.Acad.Sci. 77: 1063-1067.

Gosti-Testu F., Norris V., and Brevet J. 1983. Restriction map of Tn7. Plasmid 10: 96-99.

Griffith J., Hochschild A., and Ptashne M. 1986. DNA loops induced by cooperative binding of lambda repressor. Nature 322: 750-752.

Grindley N.D.F., and Joyce C.M. 1980. Analysis of the structure and function of the kanamycin resistance transposon Tn903. CSHSQB 45: 125-133.

Grindley N.D.F., and Reed R. 1985. Transpositional recombination in prokaryotes. Ann.Rev.Biochem. 54: 863-896.

Grindley N.D.F., and Sherratt D.J. 1978. Sequence analysis at IS1 insertion sites: models for transposition. CSHSQB 43:1257-1262.

Gringauz E., Orle K.A., Waddell C.S. and Craig N.L. 1988. Recognition of Escherichia coli attTn7 by transposon Tn7: lack of specific sequence requirements at the point of Tn7 insertion. J.Bacteriol. 170: 2832-2840.

Hall R.M., and Vockler C. 1987. The region of the IncN plasmid R46 coding for resistance to B-lactam antibiotics, streptomycin/spectinomycin and sulphonamides is closely related to antibiotic resistance segments found in IncW plasmids and in Tn21-like transposons. Nucl. Acids Res. 15: 7491-7501.

Haniford D.B., Chelouche A.R., and Kleckner N. 1989. A specific class of IS10 transposase mutants are blocked for target site interactions and promote formation of an excised transposon fragment. Cell 59: 385-394.

Harshey R. 1984. Transposition without duplication of infecting Mu. Nature 311: 580-581.

Hartl D.L., Dukhuizen D.E., Mill R.D., Green L., DeFrammond J. 1983. Transposable element IS50 improves the growth rate of E.coli cells without transposition. Cell 35: 503-510.

Hassan D., and Brevet J. 1983. Absence of cis-acting transposition immunity with Tn7. Plasmid 10: 31-44.

Hauer B., and Shapiro J. 1984. Control of Tn7 transposition. Mol.Gen.Genet. 194: 149-158.

Hatfull G.F., Noble S.M., and Grindley N.D.F. 1987. The gamma-delta resolvase induces an unusual DNA structure at the recombinational crossover point. Cell 49: 103-110

Hedges R.W., and Jacobs A.C. 1974. Transposition of ampicillin resistance from RP4 to other replicons. Mol. Gen.Genet. 132:31-40.

Heery D.M.,. Gannon F., and Powell R. 1990. A simple method for subcloning DN fragments from gel slices. TIG 6: 173.

Higgins C.F., Dorman C.J., Stirling D.A., Waddell L., Booth I.R., May G., and Bremer E. 1988. A physiological role for DNA supercoiling in the osmotic regulation of gene expression in S.typhimurium and E.coli. Cell 52: 569-584.

Hochschild A., and Ptashne M. 1986. Cooperative binding of lambda repressors to sites separated by integral turns of the helix. Cell 44: 681-687.

Hodge D.M. 1983. Transposition of the transposon Tn7. PhD thesis, University of Sussex.

Hollingshead S. and Vapnek D. 1985. Nucleotide sequence analysis of a gene encoding a streptomycin/ spectinomycin adenyltransferase. Plasmid 13: 17-30.

Holmes D.S. and Quigley M. 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal.Biochem. 114: 193-197.

Hubner P., Haffter P., Iida S., and Arber W. 1989. Bent DNA is needed for recombinational enhancer activity in the site-specific recombination system Cin of bacteriophage P1. J. Mol. Biol. 205: 493-500

Huisman O., and Kleckner N. 1987. A new generalizable test for detection of mutations affecting Tn10 transposition. Genetics 116: 185-189.

Huisman O., Errada P.R., Signon L., and Kleckner N. 1989. Mutational analysis of IS10's outside end. EMBO J. 8: 2101-2109.

Ichikawa H., Ikeda K., Wishart W.L., and Ohtsubo E. 1987. Specific binding of transposase to terminal inverted repeats of transposable element Tn3. Proc. Natl.Acad.Sci. 84: 8220-8224.

Isberg R.R., Lazaar A.L., and Syvanen M. 1982. Regulation of Tn5 by the right repeat proteins: control at the level of the transposition reactions ? Cell 30: 883-892.

Jordan E., Saedler H., Starlinger P. 1968. O^C and strong polar mutations in the gal operon are insertions. Mol.Gen. Genet. 102: 353-363.

Johnson R.C., Bruist M.F., and Simon M.I. 1986. Host protein requirements for in vitro site-specific inversion. Cell 46: 531-539.

Katz R.A., Merkel G., Kulkosky J., Leis J., and Skalka A.M. 1990. The avian retroviral IN prootein is both necessary and sufficient for integrative recombination in vitro. Cell 63: 87-95.

Kahmann R., Rudt F., Koch C, and Mertens G. 1985. G inversion in bacteriophage Mu DNA is stimulated by a site within the invertase gene and a host factor. Cell 41: 771-780.

Kans J.A., and Casadaban M.J. 1989. Nucleotide sequences required for Tn3 transposition immunity. J.Bacteriol. 171: 1904-1914.

Kleckner N. 1981. Transposable elements in prokaryotes. Ann.Rev.Genet. 15: 341-404.

Kleckner N. 1989. Transposon Tn10. In: Mobile DNA, Eds. Berg D.E., and Howe M.M. American Society for Microbiology, Washington, D.C.

Kleckner N., and Ross D.G. 1980. A recA-dependent genetic switch generated by transposon Tn10. J.Mol.Biol. 144: 215-221.

Kleckner N., Chan R.K., Tye B-K., and Botstein D. 1975. Mutagenesis by insertion of a drug-resistance element carrying an inverted repetition. J.Mol.Biol. 97: 561-575.

Klug A. and Rhodes D. 1987. 'Zinc fingers': a novel protein motif for nucleic acid recognition. TIBS 12: 464-469.

Koch C., and Kahmann R. 1986. Purification and properties of the Esherichia coli host factor required for inversion of the G segment in bacteriophage Mu. J.Biol.Chem. 261: 15673-15678.

Kolb A., Spassky A., Chapon C., Blazy B., and Buc H. 1983. On the different binding affinities of CRP at the lac, gal, and malT ptomoter regions. Nucl. Acid Res. 11: 7833-7851.

Koo H.S., Wu H-M., and Crothers D.M. 1986. DNA bending at adenine-thymine tracts. Nature 320: 501-506.

Kotlarz D., Fritsch A., and Buc H. 1986. Variations of intramolecular ligation rates allow the detection of protein-induced bends in DNA. EMBO J. 5: 799-803.

Kraft C.A., Timbury M.C., and Platt D.J. 1986. Distribution and genetic location of Tn7 in trimethoprim-resistant Escherichia coli. J.Med. Microbiol. 22: 125-131.

Kramer H., Niemoller M., Amouyal M., Revet B., von Wilcken-Bergmann B., and Muller-Hill B. 1987. lac repressor forms loops with linear DNA carrying two suitably spaced lac operators. EMBO J. 6: 1481-1491.

Kramer H., Amouyal M., Nordheim A., Muller-Hill B. 1988. DNA supercoiling changes the spacing requirements of two lac operators for DNA loop formation with lac repressor. EMBO J. 7: 547-556.

Kubo K.M. and Craig N.L. 1990. Bacterial transposon Tn7 utilises two different classes of target sites. J. Bacteriol. 172: 2774-2778.

Kuhnke G., Theres C., Fritz H-J., and Ehring R. 1989. RNA polymerase and gal repressor bind simultaneously and with DNA bending to the control region of the Escherichia coli galactose operon. EMBO J. 8: 1247-1255.

Kuwabara M.D. and Sigman D.S. 1987. Footprinting DNA-protein complexes in situ following gel retardation assays using 1.10-phenanthroline-copper ion: Escherichia coli RNA polymerase-lac promoter complexes. Biochem. 26: 7234-7238.

Laemmli U.K. 1970. Cleavage and structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Laski F.A., Rio D.C., and Rubin G.M. 1986. Tissue specificity of Drosophila P-element transposition is regulated at the level of mRNA splicing. Cell 44: 7-19.

Lichtenstein C. and Brenner S. 1981. Site specific properties of Tn7 transposition into the E.coli chromosome. Mol.Gen.Genet. 183: 380-387.

Lichtenstein C. and Brenner S. 1982. Unique insertion site of Tn7 in the E.coli chromosome. Nature 297: 601-603.

Lin S-y., and Riggs A.D. 1975. The general affinity of lac repressor for E.coli DNA: implications for gene regulation in procaryotes and eucaryotes. Cell 4: 107-111.

Liu-Johnson H-N., Gartenberg M.R., and Crothers D.M. 1986. The DNA binding domain and bending angle of E.coli CAP protein. Cell 47: 995-1005.

Mahillon J., and Lereclus D. 1988. Structural and functonal analysis of Tn4430: identification of an integrase-like protein involved in the co-integrate-resolution process. EMBO J. 7: 1515-1526.

Marini J.C., Levene S.D., Crothers D.M., and Englund P.T. 1982. Bent helical structure in kinetoplast DNA. Proc. Natl.Acad.Sci. 79: 7664-7668.

Martin K., Huo L., and Schleif R.F. 1986. The DNA loop model for ararepression: AraC protein occupies the proposed loop sites in vivo and repression-negative mutations lie in these same sites. Proc.Natl.Acad.Sci. 83: 3654-3658.

Martin C., Grinsted J., and de la Cruz F. 1989. Effects of variation of inverted-repeat sequences on reactions mediated by the transposase of Tn21. J.Bacteriol. 171: 3996-4001.

Martinez E., and de la Cruz F. 1988. Transposon Tn21 encodes a RecA-independent site-specific integration system. Mol.Gen.Genet. 211: 320-325.

Martinez E., and de la Cruz F. 1990. Genetic elements involved in Tn21 site-specific integration, a novel mechanism for the dissemination of antibiotic resistance genes. EMBO J. 9: 1275-1281.

de Massy B., Dorgai L., and Weisberg R.A. 1989. Mutation of the phage lambda attachment site alter the directionality of resolution of Holliday structures. EMBO J. 8: 1591-1599.

Maxam A.M., and Gilbert W. 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA 74: 560-564.

McClelland M., Hanish J., Nelson M., and Patel Y. 1988. KGB: a single buffer for all restriction endonucleases. Nucl.Acids Res. 16:

McClintock B. 1950. The origin and behaviour of mutable loci in maize. Proc.Natl.Acad.Sci 36: 344-355.

McKenney K., Shimatake H., Court D., Schmeissner U., Brady C., and Rosenberg M. 1981. A system to study promoter and terminator signals recognised by the E.coli RNA polymerase. In Gene Amplification and Analysis, vol.III. Eds. Chirikjian J.C., and Papas T.S.:Elsevier/ North Holland, Amsterdam.

McKown R.L., Orle K.A., Chen T., and Craig N.L. 1988. Sequence requirements of Escherichia coli attTn7, a specific site of transposon Tn7 insertion. J.Bacteriol. 170: 352-358.

McKown R., Waddell C., Arciszewska L., and Craig N.L. 1987. Identification of a transposon Tn7-dependent DNA binding activity that recognises the ends of Tn7. Proc. Natl.Acad.Sci. 84: 7807-7811.

Miller J.H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York.

Miller J., McLachlan A.D., and Klug A. 1985. Repetitive zinc-binidng domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J. 4: 1609-1614.

Miskimins K.W., Roberts M.P., McClelland A., and Ruddle F.H. 1985. Use of a protein-blotting procedure and a specific DNA probe to identify nuclear proteins that recognize the promoter region of the transferrin receptor gene. Proc.Natl.Acad.Sci. 82: 6741-6744.

Mizuuchi K. 1983. In vitro transposition of bacteriophage Mu: a biochemical approach to a novel replicon. Cell 35: 785-794.

Mizuuchi K. 1984. Mechanism of transposition of bacteriophage Mu: polarity of the strand transfer reaction at the initiation of transposition. Cell 39: 395-404.

Mizuuchi K., and Craigie R. 1986. Mechanism of bacteriophage Mu transposition. Ann.Rev.Genet. 20: 385-429.

Mizuuchi M. and Mizuuchi K. 1989. Efficient Mu transposition requires interaction of transposase with a DNA sequence at the Mu operator: implications for regulation. Cell 58: 399-408.

Moore R.J., and Krishnapillai V. 1982. Tn7 and Tn501 insertions into Pseudomonas aeruginosa plasmid R91-5: mapping of two transfer regions. J.Bacteriol. 149: 276-283.

Morisato D., and Kleckner N. 1984. Transposase promotes double strand breaks and single strand joints at Tn10 termini in vivo. Cell 39: 181-190.

Morisato D., and Kleckner N. 1987. Tn10 transposition and circle formation in vitro. Cell 51: 101-111.

Morisato D., Way J.C., Kim H-J., and Kleckner N. 1983. Tn10 transposase acts preferentially on nearby ends in vivo. Cell 32: 799-807.

Mukherjee S., Erickson H., and Bastia D. 1988. Detection of DNA looping due to simultaneous interaction of a DNA-binding protein with two spatially separated sites on DNA. Proc.Natl.Acad.Sci. 85: 6287-6291.

Murphy E. 1989. Transposable elements in Gram-positive bacteria. In: Mobile DNA, Eds. Berg D.E., and Howe M.M. American Society for Microbiology, Washington, D.C.

Murphy E., Huwlyer L., and Bastos M.C.F. 1985. Transposon Tn554: complete nucleotide sequence and isolation of transposition-defective and antibiotic-sensitive mutants. EMBO J. 4: 3357-3365.

Murray N.E. 1983. Phage lambda and molecular cloning. In: Lambda II Eds. Hendrix R.W., Roberts J.W., Stahl F.F., and Weisberg R.A. Cold Spring Harbor Press, New York.

New J.H., Eggleston A.K., and Fennewald M. 1988. Binding of the Tn3 transposase to the inverted repeats of Tn3. J.Mol.Biol. 201: 589-599.

Nomark S., Bergstrom S., Edlund T., Grundstrom T., Jaurin B., Lindberg F.P., and Olsson O. 1983. Overlapping genes. Ann.Rev.Genet 17: 499-525.

Ohtsubo E., Zenilman M., Ohtsubo H., McCormick M., Machida C., Machida Y. 1980. Mechanism of insertion and cointegration mediated by IS1 and Tn3. CSHSQB 45:283-295.

Orgel L.E., and Crick F.H.C. 1980. Selfish DNA: the ultimate parasite. Nature 284: 604-607.

Orle K.A. and Craig N.L. (in press). Identification of transposition proteins encoded by the bacterial transposon Tn7.

Ouartsi A., Borowski D., and Brevet J. 1985. Genetic analysis of Tn7 transposition. Mol.Gen.Genet 198: 221-227.

Pato M.L. 1989. Bacteriophage Mu. In: Mobile DNA, Eds. Berg D.E., and Howe M.M. American Society for Microbiology, Washington, D.C.

Pedersen S., Bloch P.L., Reeh S., and Neidhart F.C. 1978. Patterns of protein synthesis in E.coli: a catalogue of the amount of 140 individual proteins at different growth rates. Cell 14: 179-190.

Phillips S.E.V., Manfield I., Parsons I., Davidson B.E., Rafferty J.B., Somers W.S., Margarita D., Cohen G.N., Saint-Girons I., and Stockley P.G. 1989. Cooperative tandem binding of met repressor of Escherichia coli. Nature 341: 711-715.

Potter S.S. 1982 DNA sequence of a foldback transposable element in Drosophila. Nature 297:201-204.

Prentki P., Chandler M., and Galas D.J. Escherichia coli integration host factor bends the DNA at the ends of IS1 and in an insertion hotspot with multiple IHF binding sites. EMBO J. 6: 2479-2487.

Ptashne M. 1986. Gene regulation by proteins acting nearby and at a distance. Nature 322: 697-701

Pulkkinen L., Huovinen P., Vuorio E., and Toivanen P. 1984. Characterization of trimethoprim resistance by use of probes specific for transposon Tn7. Antimicrob. Agents Chemother. 26: 82-86.

Qadri M.I., Flores C.C., Davis A.J., and Lichtenstein C.P. 1989. Genetic analysis of attTn7, the transposon Tn7 attachment site in Escherichia coli, using a novel M13-based transduction assay. J.Mol.Biol. 207: 85-98.

Reznikoff W.S. 1982. Tn5 transposition and its regulation. Cell 31: 307-308.

Richet E., Abcarian P., and Nash H.A. 1988. Synapsis of attachment sites during lambda integrative recombination involves capture of a naked DNA by a protein-DNA complex. Cell 52: 9-17

Roberts D., and Kleckner N. 1988. Tn10 transposition promotes RecA-dependent induction of lambda prophage. Proc.Natl.Acad.Sci. 85: 6037-6041.

Roberts D., Hoopes B.C., McClure R.W., and Kleckner N. 1985. IS10 transposition is regulated by dam methylation. Cell 43: 117-130.

Robertson C.A., and Nash H.A. 1988. Bending of the bacteriophage lambda attachment site by Esherichia coli integration host factor. J.Biol.Chem. 263: 3554-3557.

Robinson M.K., Bennett P.M., Grinsted J., and Richmond M.H. 1977. Inhibition of TnA translocation by TnA. J. Bacteriol. 129: 407-414.

Rogers M.S. 1986. Transpositional properties and organisation of Tn7. PhD thesis, University of Glasgow.

Rogers M., Ekaterinaki N., Nimmo E., and Sherratt D. 1986. Analysis of Tn7 transposition. Mol.Gen.Genet. 205: 550-556.

Roth M.J., Tanese N., and Goff S.P. 1988. Gene product of Moloney Murine Leukemia Virus required for proviral integration is a DNA-binding protein. J.Mol.Biol 203: 131-139.

Rothstein S.J., Jorgensen R.A., Postle K., and Reznikoff W.S. 1980. The inverted repeats of Tn5 are functionally different. Cell 19: 795-805.

Rowland S-J., and Dyke K.G.H. 1990. Tn552, a novel transposable element from Staphylococcus aureus. Molec. Microbiol 4: 961-976.

Salvo J.J., and Grindley N.D.F. 1988. The gamma-delta resolvase bends the res site into a recombinogenic complex. EMBO J. 7: 3609-3616.

Sambrook J., Fritsch E.F., and Maniatis T. 1989. Molecular cloning: a laboratory manual (second edition). Cold Spring Harbor Laboratory Press, New York.

Satchwell S.C., Drew, H.R., and Travers A.A. 1986. Sequence periodicities in chicken nucleosome core DNA. J.Mol.Biol. 191: 659-675.

Scott J.R., Kirchman P.A., and Caparon M.G. 1988. An intermediate in transposition of the conjugative transposon Tn916. Proc.Natl.Acad.Sci. 85: 4809-4813.

Sekine Y. and Ohtsubo E. 1989. Frameshifting is required for production of the transposase encoded by insertion sequence 1. Proc.Natl.Acad.Sci. 86: 4609-4613.

Shapiro J.A. 1979. Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proc.Natl.Acad.Sci. 76: 1933-1937.

Sherratt D.J. 1989. Tn3 and related transposable elements; site-specific recombination and transposition. In: Mobile DNA, Eds. Berg D.E., and Howe M.M. American Society for Microbiology, Washington, D.C.

Simmons M.J., and Bucholz L.M. 1985. Transposase titration in Drosophila melanogaster: a model of cytotype in the P-M system of hybrid dysgenesis. Proc.Natl.Acad.Sci. 82: 8119-8123.

Simons R.W., and Kleckner N. 1983. Translational control of IS10 transposition. Cell 34: 683-691.

Simons R.W., Hoopes B., McClure W., and Kleckner N. Three promoters near the ends of IS10: p-IN, p-OUT, and p-III. Cell 34: 673-682.

Simonsen C.C., Chen E.Y., and Levinson A.D. 1983. Identification of the type I trimethoprim-resistant dihydrofolate reductase specified by the Escherichia coli R-plasmid R483: comparison with procaryotic and eucaryotic dihydrofolate reductases. J.Bacteriol. 155: 1001-1008.

Slatter R. 1987. The regulation and properties of Tns transposase. PhD thesis, University of Glasgow.

Smith G., and Jones P. 1984. Effects of deletions in transposon Tn7 on its frequency of transposition. J. Bacteriol. 157: 962-964.

Smith P.K., Krohn R.I., Hermanson G.T., Mallia A.K., Gartner F,.H., Provenzano M.D., Fujimoto E.K., Goeke N.M., Olson B.J., and Klenk D.C. 1985. Measuerement of protein using bicinchoninic acid. Anal.Biochem. 150: 76-85.

Staacke D., Walter B., Kisters-Woike B., von Wilcken-Bergmann B., and Muller-Hill B. 1990. How Trp repressor binds to its operator. EMBO J. 9: 1963-1967.

Stark W.M., Sherratt D.J., and Boocock M.R. 1989. Site-specific recombination by Tn3 resolvase: topological changes in the forward and reverse reactions. Cell 58: 779-790.

Steen R., and Skold O. 1985. Plasmid-borne or chromosomally mediated resistance by Tn7 is the most common response to ubiquitous use if trimethoprim. Antimicrob.agents Chemother. 27: 933-937.

Stirling C.J., Szatmari G., Stewart G., Smith M.C.M., and Sherratt D.J. 1988. The arginine repressor is essential for plasmid-stabilizing site-specific recombination at the ColE1 cer locus. EMBO J. 7: 4389-4395.

Stirling C.J., Colloms S.D., Collins J.F., Szatmari G., and Sherratt D.J. 1989. xerB, an Escherichia coli gene required for plasmid stability ColE1 site-specific recombination, is identical to pepA, encoding aminopeptidase A, a protein with substantial similarity to bovine lens leucine aminopeptidase, and generates the longest paper titles in the history of molecular biology. EMBO J. 8: 1623-1627.

Stokes H.W., and Hall R.M. 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. Mol.Microbiol. 3: 1669-1683.

Strauss F., and Varshavsky A. 1984. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell 37: 889-901.

Straney D.C., and Crothers D.M. 1985. Intermediates in transcription initiation from the E.coli lacUV5 promoter. Cell 43: 449-459.

Straney S.B., and Crothers D.M. 1987. Lac repressor is a transient gene activating protein. Cell 51: 699-707.

Summers D.K. and Sherratt D.J. 1984. Multimerisation of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerisation and stability. Cell 36: 1097-1103

Sundstrom L., and Skold O. 1990. The dhfr trimethoprim resistance gene of Tn7 can be found at specific sites in other genetic surroundings. Antimicrob. agents Chemother. 34: 642-650.

Sundstrom L., Radstrom P., Swedberg G., and Skold O. 1988. Site-specific recombination promotes linkage between trimethoprim and sulphonamide resistance genes. Sequence characterization of dhfrV and SulI and a recombination active locus of Tn21. Mol.Gen.Genet. 213: 191-201.

Surette M.G., and Chaconas G. 1989. A protein factor which reduces the negative supercoiling requirement in the Mu strand transfer reaction is Escherichia coli Integration Host Factor. J. Biol. Chem. 264: 3028-3034.

Taylor A.L. 1963. A bacteriophage-induced mutation in E.coli. Proc.Natl.Acad.Sci. 50: 1043-1051.

Thomasian S.K., and Voll M.J. 1989. Specificity of transposition of Tn7 in Vibrio parahaemolyticus. Plasmid 22: 82-85.

Thomson J.A., Hendson M., and Magnes R.M. 1981. Mutagenesis by insertion of drug resistance transposon Tn7 into a Vibrio species. J.Bacteriol. 148: 374-378.

Teitze E., Brevet J., and Tschape H. 1987. Relationships among the streprothricin resistance transposons Tn1825 and Tn1826 and the trimethoprim resistance transposon Tn7. Plasmid 18: 246-249.

Tietze E., Prager R., and Tschape H. 1982. Characterisation of Tn1822 (Tc) and Tn1824 (TpSm) and the light they throw on the natural spread of resistance genes. Plasmid 8: 253-260.

Travers A.A., and Klug A. 1987. The bending of DNA in nucleosomes and its wider implications. Phil. Trans. R. Soc. Lond. B

Tsudu M., Minegishi K-I., and Iino T. 1989. Toluene transposons Tn4651 and Tn4653 are class II transposons. J.Bacteriol. 171: 1386-1393.

Tullius T.D. 1987. Chemical 'snapshots' of DNA: using the hydroxyl radical to study the structure of DNA and DNA-protein complexes. TIBS 12: 297-300.

Tullius T.D., and Dombroski B.A. 1986. Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to lambda repressor and Cro protein. Proc.Natl.Acad.Sci. 83: 5469-5473.

Tullius T.D., Dombroski B.A., Churchill M.E.A., and Kam L. 1987. Hydroxyl radical footprinting: a high-resolution method for mapping protein-DNA contacts. Meths.Enzymol. 155:537-558.

Turner P., Barber C., and Daniels M. 1984. Behaviour of the transposons Tn5 and Tn7 in Xanthomonas campestris pv. campestris.

Waddell C.S., and Craig N.L. 1988. Tn7 transposition: two transposition pathways directed by 5 Tn7-encoded genes. Genes Dev. 2: 137-149.

Waddell C.S., and Craig N.L. 1989. Tn7 transposition: recognition of the attTn7 target sequence. Proc.Natl. Acad.Sci. 86: 3958-3962.

Way J.C., and Kleckner N. 1984. Essential sites at transposon Tn19 termini. Proc.Natl.Acad.Sci. 81: 3452-3456.

Weinert T.A., Derbyshire K.M., Hughson F.M., and Grindley N.D.F. 1984. Replicative and conservative transpositional recombination of insertion sequences. CSHSQB 49: 251-260.

Wiater L.A., and Grindley N.D.F. 1988. Gamma-delta transposase and integration host facto bind cooperatively at both ends of gamma-delta. EMBO J. 7: 1907-1911.

Wiater L.A., and Grindley N.D.F. 1990a. Integration host factor oncreases the transpositional immunity conferred by gamma-delta ends. J.Bacteriol. 172: 4951-4958.

Wiater L.A., and Grindley N.D.F. 1990b. Uncoupling of transpositional immunity from gamma-delta transposition by a mutation at the end of gamma-delta. J.Bacteriol. 172: 4959-4963.

Wu H.-M., and Crothers D.M. 1984. The locus of sequence-directed and protein-induced DNA bending. Nature 308: 509-513

Wu H.-Y., Shyy S., Wang J.C., and Liu L.F. 1988. Transcription generates positively and negatively supercoiled domains in the template. Cell 53: 433-440.

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103-119.

Zerbib D., Jakowec M., Prentki P., Galas D.J., and Chandler M. 1987. Expression of proteins essential for IS1 transposition: specific binding of InsA to the ends of IS1. EMBO J. 6: 3163-3169.

Zerbib D., Polard P., Escoubas J.M., Galas D., and Chandler M. 1990a. The regulatory role of the IS1-encoded InsA protein in transposition. Molec.Microbiol. 4: 471-477.

Zerbib D., Prentki P., Freund E., Galas D.J., and Chandler M. 1990b. Functional organisation of the ends of IS1: specific binding site for an IS1-encoded protein. Molec.Microbiol. 4: 1477-1486.

Zwieb C., Kim J., and Adhya S. 1989. DNA bending by negative regulatory proteins Gal and Lac repressors. Genes Dev. 3: 606-611.