https://theses.gla.ac.uk/

Theses Digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

EXPRESSION AND FUNCTION OF TN7 TRANSPOSITION PROTEINS

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

by

Nelly Ekaterinaki

Institute of Genetics
University of Glasgow
Church St
Glasgow

August 1987

ProQuest Number: 10997344

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10997344

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

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.

Contents

Abbreviations

Chemicals

APS - ammonium persulphate
ATP - adenosine triphosphate
CIP - alkaline phosphatase
CsCl - caesium chloride
DMF - dimethylformamide
DMS - dimethylsulphate
DNA - deoxyribonucleic acid
DTT - dithiothreitol
EDTA - ethylenediaminetetra-acetic acid (disodium salt)
EtBr - ethidium bromide
EtOH - ethanol
IPTG - isopropylthio-B-D-galactoside
NaAc - sodium acetate
RNA - ribonucleic acid
SDS - sodium dodecylsulphate
Tris - tris (hydroxymethyl) amino ethane
X-gal - 5-bromo-4-chloro-3-indolyl-B-galactoside

Antibiotics

Ap - Ampicillin
Cm - Chloramphenicol
Kan - Kanamycin
Nal - Nalidixic acid
Rif - Rifampicin
Sp - Spectinomycin
Sm - Streptomycin
Tc - Tetracycline
Tp - Trimethoprim

Phenotypes

X^r - resistance to X
X^s - sensitive to X
ori - origin of replication
res - resolution site

Measurements

bp - base pairs
kb - kilo base pairs
D - dalton
kD - kilodalton (10^3 dalton)
A - amps
mA - milliamps
V - Volts
mV - millivolts
Ci - curies
mCi - millicuries
uCi - microcuries
^oC - degrees centigrade
g - grammes
mg - milligrammes
ug - microgrammes
ng - nanogrammes
l - litres
ml - millilitres
ul - microlitres
cm - centimetres
mm - millimetres
nm - nanometres
M - molar
mM - millimolar
pH - acidity defined as $[-\log_{10}(\text{Molar concentration } H^+ \text{ ions})]$
min - minutes
sec - seconds

Miscellaneous

D&M - Davis and Mingioli
UV - ultraviolet light
Tn - transposon
LMP - low melting point
% - percentage

Acknowledgements

I thank all the members of the Genetics Institute in Glasgow, too many to mention individually, for their support and help of any sort. My particular thanks go to Colin and Mark for their willingness to share my scientific anxieties with me and offer of invaluable suggestions. Grateful thanks to Gill, George, Mary and Ros for their presence in the lab made me forget that it was always pouring with rain outside. Many thanks to Amy for her kindness and for alleviating the tension of my last days of writing by helping me with the diagrams. Thanks to Martin for introducing me to the protein world and for his gift of resolvase. No acknowledgement would be complete without referring to the ladies in the prep room for keeping my bacteria satisfactorily fed, providing clean glassware, and being so warm and kind.

Very special thanks to my mother for her love, support, and understanding.

Finally, I could never express adequately my gratitude to one particular person who has especially inspired me over the past three years: My supervisor, Prof. Dave Sherratt, for his inexhaustible supply of advice, support, and encouragement of all kinds. I also thank Dave for proofreading this thesis and fueling so many fascinating discussions that broadened my insight of the ways of DNA within the cell and kept my mind continually alert.

Summary

Transcriptional and translational fusion vectors were used to identify and characterise transcriptional control signals and translated reading frames in the transposition region of the transposon Tn7. Two promoters, one strong, P_1, and one much weaker, P_2, were detected and their precise levels of transcription were measured. Five proteins, corresponding to the tnsA, tnsB, tnsC, tnsD, and tnsE genes on the complementation map, were also identified. All but one of these proteins were expressed in low amounts. Transcription and translation of all five genes appeared to proceed in the same direction. The control of transcription was studied by placing fragments expressing Tn7 transposition functions *in trans* with the major Tn7 promoter, P_1, and examining the effects on transcription. It was found that the presence of the tnsB gene product *in trans* with P_1, resulted in a considerable decrease in P_1 activity.

The binding properties of the tnsB and tnsD gene products were examined employing a gel electrophoresis protein-DNA binding assay. In the absence of purified Tn7 proteins, whole-cell crude extracts from overproducing strains were used. It was found that tnsB, a protein required for both modes of Tn7 transposition, recognises and binds specifically to sequences in the right end of the element. Preliminary results suggested that tnsD, the "hot site" specific protein, binds to the chromosomal attachment site ("hot site"). The implications of these events on the mechanism of transposition were discussed.

The role of DNA adenine methylase (dam protein) in Tn7 transposition was investigated. The existence of a GATC (dam sensitive site) in a critical region in the element (within the repeats at the right end terminus) and our belief that Tn7 transposes via a conservative mechanism, prompted this investigation. Tn7 transposition was monitored in unmethylated, hemimethylated and fully methylated environments. No significant difference in the transposition levels in any of these three backgrounds was observed. This seemed to suggest that Tn7 transposition is not regulated by dam methylation, however, alternative interpretations were discussed.

CHAPTER 1

INTRODUCTION

1.1 Historical Background

Transposable genetic elements are unique, non permuted DNA sequences that can insert into genomes at sites that lack substantial DNA homology with the element by mechanisms that are independent of a functional host homologous recombination system. Their existence was first suggested by McClintock (McClintock, 1950, 1951) on the basis of her analysis of patterns of variegation in maize kernels and leaves. They have since been identified in bacteria, yeast, plants, and invertebrates (e.g. Drosophila, Trypanosoma), while the proviral forms of the vertebrate RNA tumour viruses share a number of common features with some eukaryotic transposable elements, possibly suggestive of a close evolutionary relationship between them (Temin, 1980). See Calos and Miller, (1980); Doring, (1985); Kleckner, (1981); Shapiro, (1983); for reviews.

Apart from their basic ability to transpose, transposable elements have been shown to exhibit a variety of other properties, such as the ability to fuse unrelated DNA molecules, to promote deletions and inversions nearby, and to control gene expression by their possession of transcriptional regulatory signals. All these properties make transposable elements major contributors to genome plasticity and generation of diversity by providing substrates for evolution to occur.

The work presented in this thesis concerns the prokaryotic transposable element Tn7, and this introduction will be confined to prokaryotic transposable elements.

1.2 Bacterial Transposable Elements

Bacterial transposable elements are divisible into two classes depending on whether they transpose predominantly via a replicative mechanism (Tn3 family; Arthur and Sherratt, 1979; bacteriophage Mu; Chaconas et al, 1981) or via a conservative transposition mechanism (IS10, Tn10; Kleckner et al, 1984; IS50, Tn5; Berg, 1977). A few elements (notably IS1, Biel and Berg,

1984; Tn903, Grindley and Joyce, 1981) appear to employ both mechanisms with similar frequencies (see section 1.7 for detailed description of mechanisms of transposition).

Upon insertion, transposons (with the exception of the Staphylococcus aureus transposon Tn554; Murphy and Lofdahl, 1984) generate 3-13bp direct duplications of the target site. These direct duplications result from replication across staggered nicks in the target replicon (Grindley and Sherratt, 1978), and are not required for subsequent transposition (Kleckner, 1981). Short inverted repeats (9-40bp) occur at the ends of the element which are thought to constitute at least a part of the recognition sequence for enzymes which separate the element from the host during the transposition process. Mu, D108, and Tn554 carry asymmetrical termini (Kahmann and Kamp, 1979; Murphy and Lofdahl, 1984; Groenen et al, 1985), though some DNA sequences are conserved at their ends.

Conventionally, transposable elements in bacteria have been divided into three classes on the basis of their genetic organisation, mechanistic properties and DNA sequence homologies (Kleckner, 1981). Tn7, a transposon found in several enteric species, and Tn554, do not easily fall into any of the classes studied so far. The next sections will describe the classification and properties of the prokaryotic transposons.

1.2.1 Class I Elements: Insertion Sequences and Composite Transposons

Insertion sequences are short (0.77-1.75kb, Kleckner, 1981) elements encoding no determinants unrelated to their own transposition functions (Iida et al, 1983). Typically, each insertion sequence element encodes for at least one protein which probably acts at the ends of the element (Foster et al, 1981; Grindley and Joyce, 1980; Isberg and Syvanen, 1981).

Composite elements are made up of an accessory determinant, usually an antibiotic resistance gene, flanked by two insertion

sequences in either direct (Tn9, Machattie and Jackowski, 1977) or inverted (Tn5, Berg et al, 1982a, Tn10, Kleckner, 1981) orientation. In composite elements, all the functions and sites required for transposition are contained within the insertion sequence elements. Both flanking insertion sequence elements are structurally intact, but one is often functionally deficient (IS10L in Tn10, Foster et al, 1981).

1.2.2 Class II Elements: The Tn3 Family

Class II elements are large (>5kb) and encode their own transposition determinants. Most elements of this class carry accessory determinants which are not separated from the transposition functions but form an integral part of the main transposon body (Kleckner, 1981; Grindley and Reed, 1985). All class II elements generate 5bp duplications of the target site upon insertion and have nearly perfect 38-40bp inverted repeats at their ends. The strong similarity of the repeats between different elements of this class is suggestive of a common evolutionary origin.

Class II elements are subdivided into two groups, represented by Tn3 and Tn501, according to their organisation and ability to cross complement each other.

In the Tn3 subgroup, transcription of the transposition genes is divergent and initiates within one region, namely res (Chou et al, 1979a, 1979b; Gill et al, 1979; Heffron et al, 1979). In the Tn501 subgroup, transcription of the tnpA and tnpR genes is unidirectional (Diver et al, 1983; Grinstead et al, 1982).

Intermolecular transposition of class II elements proceeds via an obligatory cointegrate intermediate (Arthur and Sherratt, 1979; Muster and Shapiro, 1980; McCormick et al, 1981) that is resolved by an element-encoded site-specific recombination system (Arthur and Sherratt, 1979; Kitts et al, 1981, 1982). In contrast, intramolecular transposition of class II transposons occurs in the absence of a functional transposon-encoded site-specific

recombination system to give adjacent deletions and/or replicative inversions (Bishop and Sherratt, 1984).

1.2.3 Class III Elements: The Transposing Bacteriophages

Class III elements are represented by the closely related phages Mu and D108. These elements apart from their "phage characteristics" differ from class I and II elements in that their ends are not inverted repeats although some homology exists at either end (Kahmann and Kamp, 1979). Mu transposition generates 5bp duplications of the target site and is generally replicative, apart from the initial step after infection which is conservative (Liebart *et al*, 1982; Akroyd and Symonds, 1983; Harshey, 1984). Cointegrates constitute the major products of transposition whereas simple inserts are rarely observed (Chaconas *et al*, 1981; Howe and Shumm, 1980). *In vitro* studies with Mu have shown that both simple inserts and cointegrates arise from the same transposition intermediate (Craigie and Mizuuchi, 1985).

1.3 Insertional Specificity of Transposable Elements

Different transposons show different degrees of target specificity upon insertion, suggesting that transposon- encoded functions are responsible for the recognition event. At one end of the spectrum are elements like IS50 and IS10, which show little specificity for insertion over large regions of the genome, although fine mapping analysis of target sites has revealed a local sequence specificity (Engler and Van Bree, 1981; Foster, 1977; Halling and Kleckner, 1982; Kleckner, 1981; Noel and Ames, 1978). At the other end are elements like IS4, Tn7, and Tn554 which show a high insertion preference. IS4 has only been found in three sites in the chromosome of *E.coli* (Klaer *et al*, 1980; Klaer and Starlinger, 1980; Pfeifer *et al*, 1977). Tn554 has a single insertion site in the chromosome of *S.aureus* (Krolewski *et al*, 1981). Tn7 shows an extremely high insertion preference for one specific site in the chromosome of *E.coli* and other genera, but it inserts into other sites when this specific

site is occupied or deleted (N. Craig, pers. comm; C. Lichtenstein, pers. comm; Rogers, 1986). It also inserts at many sites in conjugative plasmids (Barth and Grinter, 1977; Barth and Datta, 1977; Barth et al, 1978; Ely, 1982; Lichtenstein and Brenner, 1981, 1982; C. Lichtenstein, pers. comm; Thomson et al, 1981; Turner et al, 1984). Between those extremes, are elements like Mu, Tn3, and Tn9, which exhibit only a regional specificity for insertion. They insert preferentially into A+T rich regions and/or regions which show homology with their ends (Miller et al, 1980; Tu and Cohen, 1980).

1.4 Regulation of Transposition

The acquisition of a large number of copies of an element within a genome could prove deleterious for the organism, since each transposition event can induce disadvantageous mutations and cause chromosomal rearrangements by homologous recombination. However, uncontrollable proliferation of transposons is not observed and regulatory processes exist which limit the rate of transposition to beteween 10^{-4} and 10^{-8} events per cell per generation (Kleckner, 1981). Limitation in the number of target sites within the genome and/or precise excision of an element are not included in the regulatory processes which keep the transposition level low. The availability of unoccupied potential insertion sites within a genome shows that the transposition frequency is not limited by the saturation of potential target sites (Nymann et al, 1981). Also, for most elements, the observed rates of excision are orders of magnitude less than the rates of transposition (Kleckner, 1981).

Prokaryotic transposable elements employ a diversity of mechanisms aimed to control their transposition rates. For most elements, reduced transposition levels are attained by regulation at the level of transcription and/or translation. In Tn3, the tnpR gene product represses transcription of both the tnpA and tnpR genes (Chou et al, 1979a, 1979b; Gill et al, 1979; Kitts et al, 1981). Mu regulation occurs by repressing transcription of the MuA and MuB genes (Howe and Bade, 1975; Toussaint and

Resibois, 1983). Tn5 encodes a truncated form of the transposase protein that is translated from the same open reading frame as transposase and controls transposition either by complexing with the transposase or by competing for the transposon ends (Biek and Roth, 1980; Isberg *et al*, 1982; Jonhson *et al*, 1982; Reznikoff, 1982; Rothstein and Reznikoff, 1981).

One regulatory feature mainly confined to the Class II elements is transposition immunity; this is a transposon-encoded mechanism that acts only *in cis* to limit (by a factor of 100) transposition of a second transposon copy into a replicon that already possesses a copy or even a single end of that transposon, although transposition to other replicons in a cell is unaffected (Robinson *et al*, 1977, Sherratt *et al*, 1980; Wallace *et al*, 1980). Transposition immunity has been observed in Tn7, though it does not belong to the class II elements. However, the effect is less pronounced (N. Craig, pers. comm; Hassan and Brevet, 1983; Hauer and Shapiro, 1984).

Host- encoded functions and in particular the DNA adenine methylase (*dam* protein) are also involved in the control of transposition. Tn5, Tn10, and Tn903 exhibit a higher transposition frequency in unmethylated environments (Kleckner *et al*, 1984; Sternberg, 1985). In the case of IS10, increased transposition frequency is associated with hemi-methylation of DNA, indicating that transposition occurs only during host DNA replication and is therefore determined by the methylation state of the DNA.

In class I elements, regulation of transposition is attained by preferentially *cis*-acting transposases (Grinley and Joyce, 1981; Isberg and Syvanen, 1981).

Mu and D108 seem to use the stoichiometric activity of the *A* protein (transposase) to control their transposition (Pato and Reich, 1982, 1984). Continual synthesis of the Mu*A* protein is needed for continual transposition.

Finally, a novel form of regulatory mechanism has been proposed for Tn21 (Hyde and Tu, 1985). A modulator function (M) acts to increase transposition frequency and cointegrate formation, and provides the first example of an element-specified positive effector of transposase expression.

1.5 Consequences of Transposition

Transposable elements can insert themselves into new locations at frequencies between 10^{-4} and 10^{-8} per cell per generation (Kleckner, 1981). Insertion, in all but one case, (Tn554; Murphy and Lofdahl, 1984), is accompanied by a short (3-12bp) duplication of the target DNA sequence, so that the inserted element is flanked by a direct repeat of host DNA. Apart from transposition into new locations, transposable elements can mediate a series of other rearrangements, both intramolecular and intermolecular (Figure 1.1). An intermolecular transposition event can generate either a simple insertion or a cointegrate structure (Figure 1.1). Simple insertions arise from the integration of the element into a new site in the target replicon. Cointegrates are fusions between the donor and target replicon separated by direct repeats of the transposon. If stable, cointegrates can have important consequences. For example, in the case of the F plasmid, integration into the E.coli chromosome can result in subsequent transfer of chromosomal genes to a new host (Cullum and Broda, 1979).

Transposition into a coding sequence results in the inactivation of the target gene (Lieb, 1981). In the case of many insertion elements (Iida et al, 1983), Tn3 (Heffron et al, 1975a, Rubens et al, 1976), and Tn5 (Berg et al, 1980), it has been observed that transposition exerts strong polar effects, diminishing or abolishing the expression of genes downstream of the insertion site. However, in other instances, adjacent genes may be trancriptionally activated by insertion of an insertion sequence element or transposon as a result of transcriptional readthrough from a promoter carried on the element (IS3, Charlier et al, 1982; Tn3, Heffron et al, 1979).

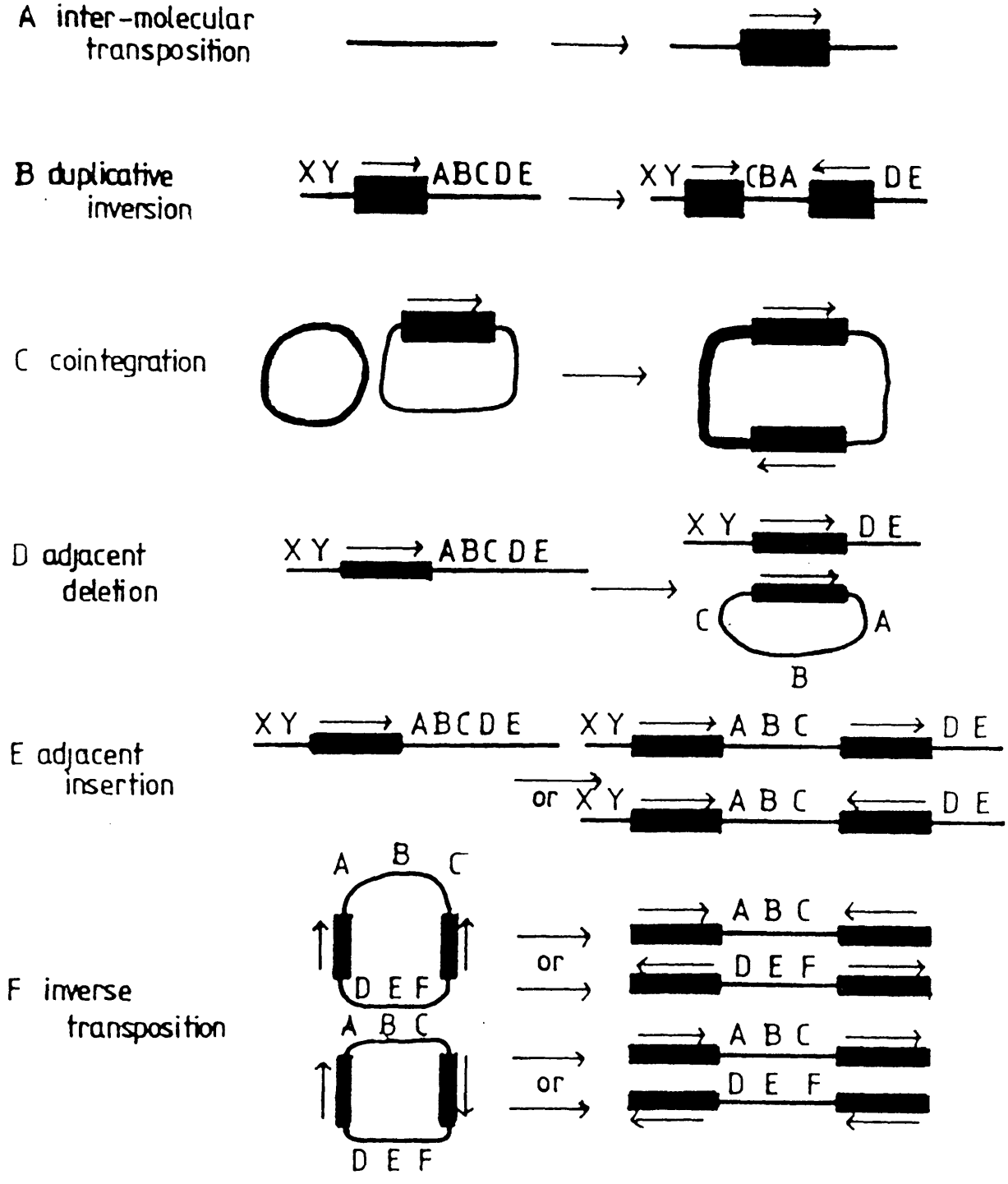


Figure 1.1 Illustration of transposon mediated DNA rearrangements. Letters represent genes external to the transposable element. Thick boxes represent the transposon. (Kleckner, 1981)

From the summary of transposon-mediated events described above, it becomes apparent that transposable elements are of essential importance in genome evolution acting both as portable mutator agents and sources of genetic rearrangements.

1.6 Persistence of Transposable Elements

As a result of the additional determinants they carry, some transposable elements can confer a selective advantage to the organism which propagates them. For example, Tn5 (Km^r) confers a selective advantage on cells challenged by the onset of carbon source limitation in the absence of kanamycin (Biel and Hartl, 1983). This selective advantage is due to the Tn5 transposase and not to a beneficial mutation resulting from the insertion of Tn5 or to the Kan gene product itself. Campbell (1981) has attributed the maintenance of transposons to their ability to reorganise unrelated DNA sequences either directly through transposition or indirectly by providing a substrate for the host-encoded recombination system. In this regard, transposons survive because they act as sources of genetic rearrangements.

However, it would appear that the deleterious effects of some elements, due to increased mutation rate, would outweigh any possible benefits which they might confer upon their host. The persistence of transposable elements and their ability to "over-replicate" their host can be explained by the "selfish DNA" concept. The "selfish DNA" hypothesis suggests that once a unique segment of DNA acquires the ability to duplicate itself independently of the duplication of the rest of the genome, as a consequence of transposition to new sites, then such an entity will be maintained purely because of the difficulty in simultaneously eliminating all the copies of that element (Doolittle and Sapienza, 1980; Doolittle, 1982; Orgel and Crick, 1980). This hypothesis can apply to elements which transpose either replicatively or conservatively, since plasmid copy number effects can account for the increase in copies of elements transposing via a conservative mechanism.

The overall situation probably consists of components from all these sources.

1.7 Models for Transposition

A number of models has been proposed to explain the mechanism of transposition. They fall into two main classes, conservative and replicative (Figure 1.2) depending on whether replication of the element occurs during transposition. Replicative models are subdivided into symmetric and assymmetric, depending on whether both ends of the element are treated similarly or not (for review, see Grindley and Reed, 1985).

The "cut and paste" conservative model explains transposition which occurs by non-replicative simple insertion. It was proposed by Berg (1977) to account for the transposition of Tn5 and is applicable to all transposons that do not form cointegrates or their intramolecular analogs at any stage of the transposition process. According to this model, double strand breaks occur at the ends of the element and at the target site which are brought together by the transposase protein. The transposon is ligated to the recipient molecule and the remaining linear donor molecule gets degraded. This model can explain the transposition of Tn5 and Tn10, but is inadequate for elements like ISI, Mu and Tn903 that can transpose either replicatively or conservatively, though most probaly via a branched pathway.

The symmetrical replicative model was developed to account for the transposition of Tn3 which proceeds via obligate cointegrate intermediates. In this model, the donor duplex is nicked on opposite strands at opposite ends of the element, and the target duplex is nicked on both strands in a staggered way. The attachment of the free ends of the element to the target DNA leads to the formation of two replication forks and replication occurs through both of these to replicate the entire element. The replicated DNA is ligated to the flanking sequences to generate a cointegrate which is subsequently resolved by site-specific or homologous recombination to the transposition end

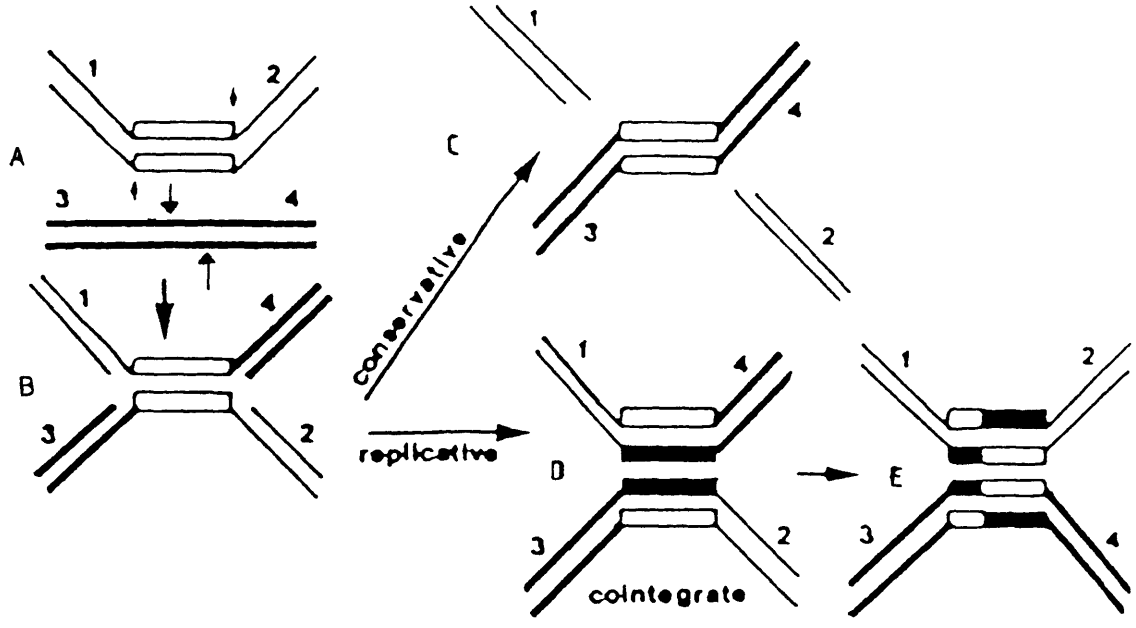


Figure 1.2 Transposition model accounting for simple insertion and cointegrate formation

a) Single strand cleavage at transposon ends and staggered double strand cleavage at the target.
b) Strand exchange transfers the ends of the transposon to the target forming two replication forks. Then either
c) Breakage and transfer of the second transposon strand and repair of staggered breaks produces a conservative transposition event. The donor molecule is probably degraded. Alternatively,
d) Replication across the element and sealing of replicated DNA gives a cointegrate type intermediate, which is (e) resolved via an element encoded site-specific recombination system to produce a simple insert.

Parental transposon DNA is shown as open bars; Newly synthesised DNA is shown as solid bars.

products (Arthur and Sherratt, 1979; Shapiro, 1979). Ohtsubo et al (1980b) modified this model to allow for non-replicative simple insertions without cointegrate formation be the outcome of this pathway. In vitro studies with Mu (Craigie and Mizuuchi, 1985; Mizuuchi, 1983), provide direct evidence that supports the modified symmetrical transposition model which also seems to agree with all of the present data. The variable ratio of simple inserts to cointegrates observed for different transposons could be explained by the variety of efficiencies with which replisomes are formed at the transposon termini (Grindley and Reed, 1985). The absence of simple inserts during transposition of class II elements could be due to the reversal of the polarity of the strand transfer.

The asymmetrical replicative model has been proposed to explain the existence of both simple inserts and cointegrates as final transposition products (Galas and Chandler, 1981; Harshey and Bukhari, 1981). In this model, the donor duplex is nicked at one end of the element and ligated to one strand of the target duplex that has been cut in a staggered fashion. The replication fork that is thus generated proceeds through the element until it reaches and replicates its opposite ends. A second cleavage on the originally nicked parental strand will result in a replicative simple insertion, whereas a second nick on the nascent strand at the equivalent sequence to the initial will form a cointegrate. The accumulating evidence against the asymmetrical model has made clear that it cannot interpret all the current data (Grindley and Reed, 1985). Under this model, simple inserts and cointegrates occur via a replicative pathway. It has been shown for Tn10 and Mu (Bender and Kleckner, 1986; Craigie and Mizuuchi, 1985) that no extensive replication of the element accompanies the formation of simple inserts. Additionally, results with artificially constructed giant transposons do not (Harayama et al, 1984; Rosner and Guyer, 1980) support the variation of transposition frequency with transposon length (Chandler et al, 1982; Morisato et al, 1983), suggested by this model. Finally, according to the model, simple inserts could be either intermolecular or intramolecular. Intramolecular

inserts are rarely, if at all, observed.

1.8 Enzymology of Transposition

The transposition process is complex and dependent on the presence of a number of element-specified and host-specified proteins. An early step of the transposition reaction is thought to involve appropriate cuts at the transposon ends and target site, coupled with DNA strand transfers. Most transposons code for a single protein, commonly called a "transposase", that is essential for this early stage of the transposition reaction. The role of transposases might be confined to the recognition of the ends of the element, with a requirement for host proteins to complete the strand transfer reaction. Alternatively, transposases may have the activity to carry out the complete strand transfer reaction without the participation of host proteins. Tn10 transposase has been found to promote double strand breaks and single strand joints at Tn10 termini *in vivo* (Morisato and Kleckner, 1984). Craigie *et al* (1984), using the DNAase I protection method have shown that the MuA protein (Mu transposase) recognises and binds site-specifically to the ends of Mu DNA. Recently, Zerbib *et al* (in press) have reported the specific binding of the InsA protein (IS1 transposase) to the ends of the element with the use of the gel retardation technique. The *E.coli* integration host factor (IHF) binds specifically to the ends of IS1, though the functional purpose of this event remains to be elucidated. Another *E.coli* host factor, the HU protein has been found to be required for efficient formation of the intermediate in Mu transposition *in vitro* (Craigie *et al*, 1985).

1.9 The Transposon Tn7

Tn7 is an unusually large (14kb) transposable element which confers resistance to the antibiotics trimethoprim, streptomycin and spectinomycin. It was first isolated from a trimethoprim resistant *E.coli* strain on the plasmid R483 and was initially known as TnC (Hedges *et al*, 1972). It has since been found to be

capable of translocating these resistances to the E.coli chromosome and a variety of plasmids harboured by different bacterial strains (Barth et al, 1976).

The most interesting aspect of Tn7 is its transpositional properties. Within the sensitivity of P1 transduction, independent insertions into the E.coli chromosome map at the same site (Barth et al, 1976). Insertions into plasmids are not specific, but they occur in a single orientation relative to the restriction maps of the plasmid and transposon (Hodge, 1983). This latter property was found when the transposon was used to generate a map of the transfer regions of RP4 by insertional mutagenesis (Barth and Datta, 1977; Barth and Grinter, 1977; Barth et al, 1978). It was noticed that all 49 insertions into RP4 occured in a single orientation relative to the restriction maps of Tn7 and RP4. Only one insertion was found in the opposite orientation (Moore and Krishnapillae, 1982). This orientation specificity of insertion implies that the transposon recognises some feature of the plasmid, apart from replication (RP4 has a bidirectional origin), that is orientation specific. Tn7 also inserts into the chromosome of many other genera at one specific site which is usually referred to as the "hot site" for Tn7 (Ely, 1982; Thomson et al, 1981; Turner et al, 1984) as opposed to a "cold site" that is any other site to which Tn7 can transpose.

The E.coli chromosomal "hot site" has been cloned, sequenced and identified as the transcriptional terminator of the glmS gene (Gay et al, 1986; Lichtenstein and Brenner, 1981). A 969bp fragment carrying the "hot site" functions as an effective site for insertion of Tn7 when cloned in multicopy plasmids. The orientation of insertion of Tn7 into this site is determined by the sequence of the site and not by any plasmid determinants. Inversion of the "hot site" within a plasmid results in inversion of Tn7 insertions (Lichtenstein and Brenner, 1981). Further sequencing studies have revealed that the site of integration between different bacterial strains is highly conserved over an extensive region, which is larger than the 70bps required to

define the "hot site" (C. Lichtenstein, pers. comm.).

Tn7 generates 5bp duplications of the target site upon insertion and sequencing of the ends has revealed characteristic terminal inverted repeats (28bp) (Lichtenstein and Brenner, 1982). The same workers have also found four tandem repeats of a 22bp consensus sequence in the right end which, are also repeated three times in the left end but are separated by non homologous sequences (Figure 1.3). This asymmetry between the ends may account for the orientational specificity of Tn7 transposition. There is an observed difference in activity between the ends of Tn7. The right end is capable of promoting transposition immunity, while the left end is not (N. Craig, pers. comm.). Also, two right ends in inverted orientation can transpose when complemented, while two left ends cannot (N. Craig, pers. comm.).

Conventionally, the 5kb left end of the transposon carries the antibiotic resistance genes. Two genes have been defined within this region confering resistance to Tp and Sm/Sp (Hodge, 1983; Smith and Jones, 1984). The dihydrofolate reductase (DHFR) gene has been sequenced (Fling and Richards, 1983; Simonsen et al, 1983), and the 18kD polypeptide encoded by it has been observed in minicell and maxicell protein preparations (Brevet et al, 1985; Fling and Elwell, 1980). Comparison of the Tn7 restriction map derived from the sequence places this gene between 2.32kb and 2.8kb and indicates that it is transcribed from left to right. Maxicell analysis of the proteins produced by Tn7 indicates that the adenyltransferase gene determining resistance to Sm/Sp encodes a protein of 32kD (Brevet et al, 1985). This gene is probably, at least partially, contained within the BstEII fragment between 3.9kb and 5.8kb because deletion of this region results in a $Sm^{S}Sp^{S}$ phenotype.

Transposition immunity, which is thought to be a property of Class II elements is also observed in Tn7, though the strength of the effect is unclear (N. Craig, pers. comm; Hassan and Brevet, 1983; Hauer and Shapiro, 1984, M. Rogers, pers. comm.).

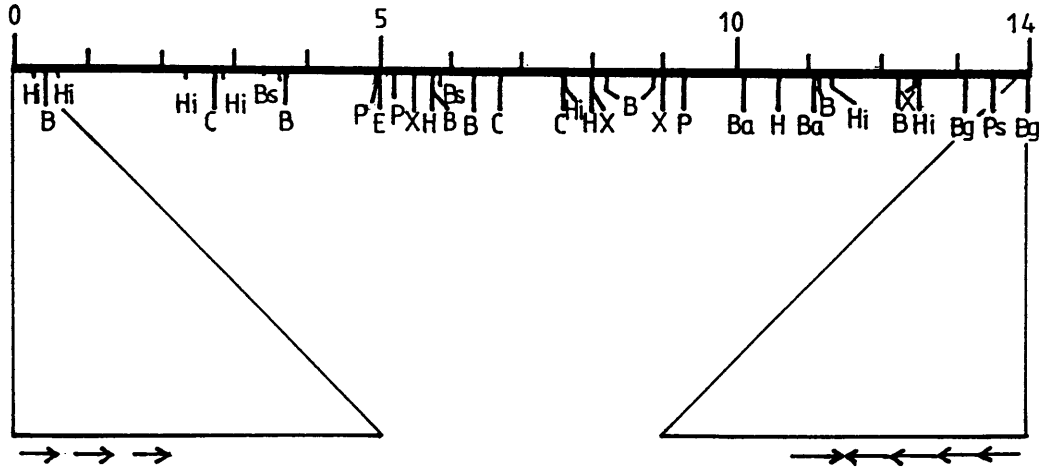


Figure 1.3 Restriction map of Tn7 (Gosti-Testu et al, 1983) and current view of the organisation of the ends. The left end contains three copies of a 22bp consensus sequence separated by unrelated sequences. The right end contains a further copy in direct repeat and four contiguous copies in inverted repeat.

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

Work by Rogers (1986), provides evidence against the presence of a site-specific recombination system in Tn7 similar to that observed in class II elements (Arthur and Sherratt, 1979; Gill *et al*, 1978, 1979; Heffron *et al*, 1977). If such a system exists in Tn7, the internal resolution site must be located within the 703bp end sequences (168bp of the left end and 535bp of the right end). There is no detectable homology between these regions of Tn7 and the *res* site of Tn3. Also, replicon fusions mediated by homologous recombination are stable in *recA*$^-$ strains (Rogers, 1986) implying that any resolvase system in Tn7 must be very inefficient under the conditions studied. The only remaining possibility is that such a system resides in one or both ends and acts during the transposition process.

When this work was started, all information about the essential regions needed for Tn7 transposition came from studies using deletion mutants. These mutants had been generated *in vitro* using convenient restriction sites (Smith and Jones, 1984) or *in vivo* using the genetic instability caused by Tn9 insertion to generate adjacent deletions (Hauer and Shapiro, 1984; Hodge, 1983). The deletion mutants were assayed for their ability to transpose both independently and in the presence of other deletions of Tn7 or fragments of Tn7 cloned into multicopy vectors. Hodge (1983), and Hauer and Shapiro (1984), examined transposition to plasmids and to the "hot site". Smith and Jones (1984), only looked at transposition to plasmids. These studies defined at least three *trans*-acting functions that are involved in Tn7 transposition, either to the "hot site" or to plasmids or to both. One of these functions was encoded entirely within the 2.2kb *Hin*dIII fragment between 5.8kb and 8.0kb and was only required for transposition to plasmids (Hauer and Shapiro, 1984; Smith and Jones, 1984). The other two functions were encoded between the *Bgl*II site at 14.0kb and the *Sst*I site at 10.9kb and between 13.1kb and 7.9kb (figure 1.4).

The data described above could only loosely localise the proposed transposition functions. More work was needed before a precise map of the transposition genes could be constructed.

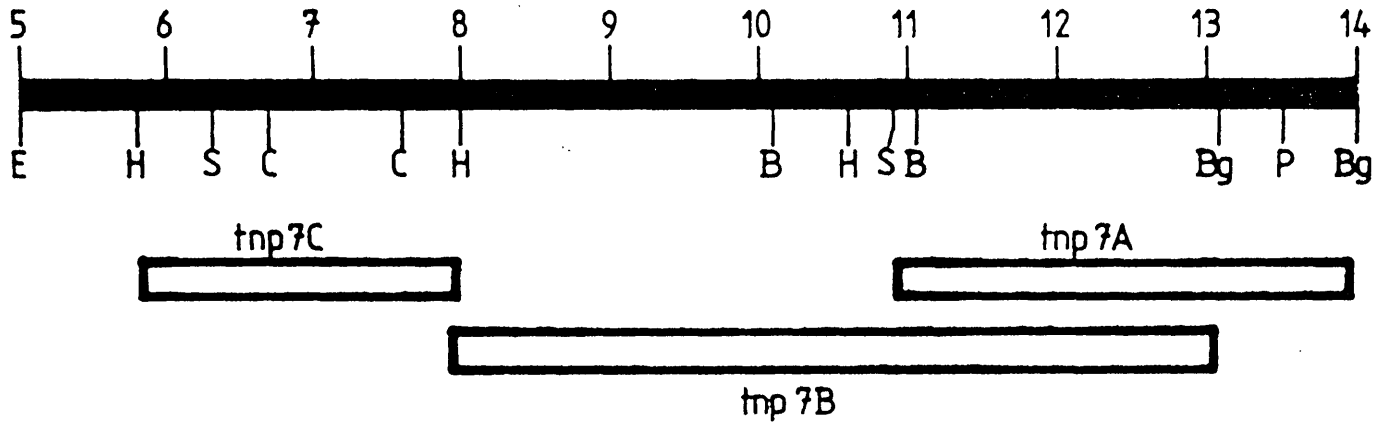


Figure 1.4 Map of Tn7 transposition functions.
Modified from Hauer and Shapiro, (1984). Each open box represents a complementation group.

Abbreviations: B, BamHI; Bg, BglII; C, ClaI; E, EcoRI; H, HindIII; P, PstI; S, SstI

During the course of this thesis, two more papers describing work on Tn7 have been published. Data by Ouartsi et al (1985), argued against the presence of a function at the right end of Tn7. The results presented in chapter 3 together with other published and unpublished data confirm both its presence and requirement for transposition (Hauer and Shapiro, 1984; Rogers et al, 1986; C. Lichtenstein, pers. comm; C. Wadell and N. Craig, pers. comm.). Ouartsi et al (1985), have published a maxicell analysis of the proteins produced by Tn7. They identify six polypeptides including the resistance genes. Four of these are encoded in the region required for transposition. The complementation groups tnsB, tnsC and tnsE (Figure 1.5; Rogers et al, 1986) can be related directly to the polypeptides p85a, p54 and p85b respectively in their nomenclature, but they did not observe a polypeptide which correlates with tnsA (Figure 1.6). Also, complementation data by Rogers et al, (1986) indicates that tnsD extends beyond the HindIII site at 8.0kb. Brevet et al (1985), found that a 40kD polypeptide, p40, was translated from a fragment of Tn7 with this site as one of its boundaries.

At the time this work was initiated, Mark Rogers, in our group in Glasgow, was investigating the cis and trans-acting functions required for Tn7 transposition using complementation studies. The transposition of a mini-Tn7, Tn7-1, deleted for all but 703bp of Tn7 end sequences (cis -acting functions required for transposition) was complemented with fragments of Tn7 cloned onto multicopy plasmids. These studies led to the construction of the map illustrated in figure 1.5. Five diffusible functions, tnsA, tnsB, tnsC, tnsD, and tnsE have been identified that are required for Tn7 transposition. Three of these, tnsA, tnsB and tnsC are essential for all types of insertions. A fourth, tnsD, is required for transposition into the "hot site" but not transposition into plasmids lacking this site. The fifth, tnsE is required for transposition to "cold sites" but is dispensable for insertion to the "hot site".

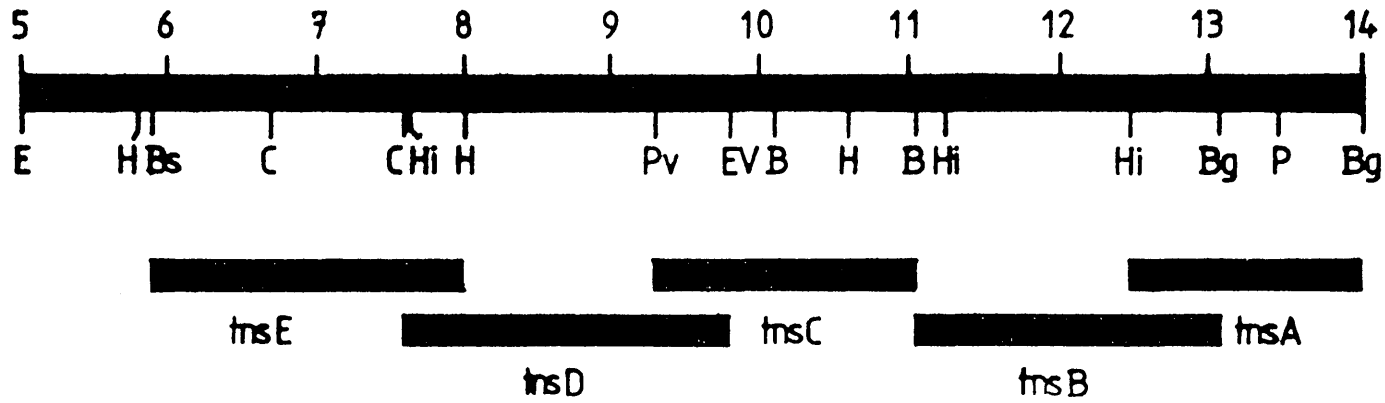


Figure 1.5 Organisation of Tn7 transposition functions

The thick line represents the right 9kb of Tn7. The boxes named tnsA-tnsE indicate the minimum sequences identified which encode these functions.
Rogers et al, 1986

Abbreviations: B, BamHI; Bg, BglII; Bc, BclI; C, ClaI; E, EcoRI;

Ev, EcoRV; H, HindIII; Hi, HincII; P, PstI; Pv, PvuII

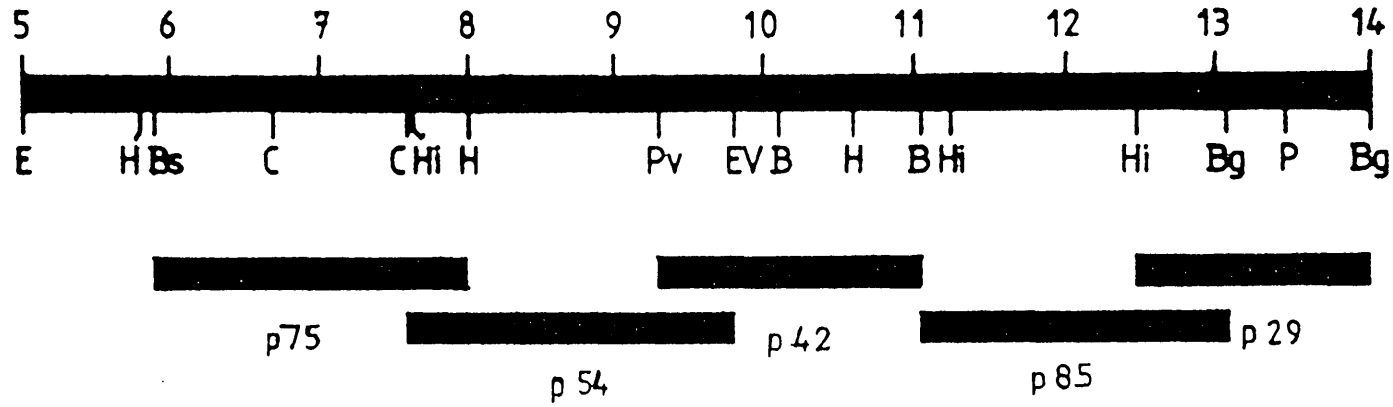


Figure 1.6 Polypeptides encoded by the transposition specific functions in Tn7

The thick line represents the 9kb right part of Tn7. The boxes indicate the genetic organisation of the Tn7 transposition functions, as shown in Figure 1.5. The apparent molecular weight of the polypeptides encoded by these regions is given below each box (in kD), as estimated by minicell analysis (Rogers, 1986).

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

In this thesis, transcriptional and translational fusions have been used for the detection and characterisation of genes involved in Tn7 transposition and its control. The levels of transcription and translation of all proteins identified have been assessed and the effect of *trans* -acting factors on their expression has been examined. The binding of Tn7 transpositon proteins to the right end of the transposon and to the "hot site" has been investigated using the gel retardation technique. Finally, the role of host factors and in particular of the DNA adenine methylase (*dam* protein) in transposition has been studied by monitoring Tn7 transposition in both unmethylated and hemimethylated hosts.

CHAPTER 2

MATERIALS AND METHODS

Table 2.1 Bacterial Strains

Strain	Characteristics	Reference/Source
CB51	CSH26 but *dam*-3	Chris Boyd
CGSC4294	Hfr PO68 *thi*1 *rel*A1 *spo*T1	Bachman (KL14)
CGSC4312	Hfr PO3 *met*B1 *rel*A1 *spo*T1 lambda$^-$	Bachman (KL227)
CSH26	*ara thi* del{*lacpro*}	Miller, 1972
DS801	*pro*A7 *str*31 *thr*1 *leu*6 *tsx*33 *mtl*2 *his*4 *arg*E3 *lac*Y1	Dave Sherratt
DS883	*rec*A$^+$ deriv. of Hfr PO45 *srl*::Tn10 *thr*300 *ilv*318	David Leach
DS902	DS801 but *rec*A13 *arg*E$^+$	Howard-Flanders (1966) (AB2463)
DS941	DS801 but *lac*I^q *lac*ZdelM15 *lac*Y$^+$ *rec*F143	Dave Sherratt
DS947	DS801 but del{*lacpro*} sup^o *rec*F143	Dave Sherratt
DS949	*thr leu thi* *pro arg*E *rec*A *rps*L *lac*	Dave Sherratt
JM101	*lac pro sup*E *thi* F' *tra*D36 *pro*AB *lac*I^q *lac*ZdelM15	Messing, 1983
ZMR1	*rif*^{r} *nal*^{r} *sp*^{r} *his* *trp rec*A	Mark Rogers
ZNE1	CSH26 *srl*::Tn10	Chapter 5
ZNE2	CB51 *srl*::Tn10	Chapter 5
ZNE3	ZNE2 but 2APur^r	Chapter 5
ZNE8	CSH26 but *rec*A$^-$	Chapter 5
ZNE9	CB51 but *rec*A$^-$	Chapter 5
ZNE15	DS801 but *rif*^{r}	Chapter 5
ZNE16	ZNE8 but *rif*^{r}	Chapter 5
ZNE18	ZNE9 but *rif*^{r}	Chapter 5
ZNE19	CGSC4294::Tn7	Chapter 5
ZNE20	CGSC4312::Tn7	Chapter 5
ZNE21	DS902 but *rif*^{r}	Chapter 5
ZNE22	DS949 but *rif*^{r}	Chapter 5

2.1 Bacterial Strains. The bacterial strains used were all derivatives of Escherichia coli K-12 and are listed in Table 2.1. Genotype and phenotype symbols are those recommended by Bachman et al (1976) and Novick et al (1976).

2.2 Plasmids. The plasmids used and constructed in this study are listed in Table 2.2 and their nomenclature follows that of Novick et al (1976).

2.3 Bacteriophages. P1 phage used in this work was a gift from Dr Millie Masters.

2.4 Chemicals.

CHEMICALS	SOURCE
General chemicals	B.D.H., Hopkins and Williams, Kochlight Laborotories, May and Baker
Media	Difco, Oxoid
General biochemicals	Sigma, Pharmacia
Agarose	BRL
X-gal, IPTG	BRL
Radiochemicals	NEN
10 X core buffer	BRL
Antibiotics	Sigma
Restriction enzymes	BRL, Boerhinger Mannheim

2.5 Culture Media.

L-Broth: 10g tryptone, 5g yeast extract, 5g NaCl, 1g glucose, 20mg thymine, made up to 1 litre in distilled water and adjusted to pH 7.0 with NaOH.

L-Agar: As L-broth without glucose and the addition of 15g/l agar.

Table 2.2 Plasmids

Plasmid	Marker	Compatibility	Comment	Source/reference
R388	Tp^r	IncW	-	Datta and Hedges (1972)
pACYC184	Cm^r, Tc^r	P15A	-	Chang and Cohen (1978)
pEA305	Ap^r	pMB1	lambda cI overproducer	Amann *et al* (1983)
pZEN300	Tp^r	IncW	Tn7 "hot site" cloned into R388	Elaine Nimmo
pGLW8	Ap^r	pMB1	P_{tac}	Fiona Stuart
pKO500	Ap^r	pMB1	Promoter probe vector	McKenney *et al* (1981)
pKL500	Ap^r	pMB1	Terminator probe vector	McKenney *et al* (1981)
pMA1441	Ap^r	pMB1	Tn3 *res* site cloned into pUC8	Martin Boocock
pMA6114	Ap^r	pMB1	Tn3 *tnpR* gene cloned into pKK223-3	Martin Boocock
pPE14	Cm^r	p15A	*recA* gene cloned into pACYC184	Emmerson
pZMR64	Ap^r	pMB1	*EcoRV* (9.8kb)-*ClaI* (7.6kb) fragment from Tn7 cloned into pGLW8	Mark Rogers
pZMR80	Ap^r	pMB1	Tn7 "hot site" cloned into pUC8	Mark Rogers
pZMR80::Tn7	Ap^r Tp^r	pMB1	Tn7 insertion into "hot site" in pZMR80	Mark Rogers
pZMR88	Ap^r	pMB1	RE[2]-*PstI* (13.5kb) fragment from Tn7 cloned into pUC8	Mark Rogers
pZMR100	Km^r	λori	P_{tac}	Mark Rogers
pZMR112	Km^r	λori	RE-*BamHI* (11.05kb) with *BclI* deletion (11.1kb-12.2kb) fragment from Tn7 cloned into pZMR100	Mark Rogers
pZMR113	Km^r	λori	*BamHI* (11.05kb)-*ClaI* (7.6kb) with *XbaI* deletion (8.05kb-9.0kb) fragment from Tn7 cloned into pZMR100	Mark Rogers
pZMR117	Km^r	λori	*BglII* (13.1kb)-*BamHI* (11.05kb) fragment from Tn7 cloned into pZMR100	Mark Rogers

Notes to Table 2.2

1. Fragments are described in the direction that transcription or translation would run though them.
2. RE; includes sequences up to the right end of Tn7.

Table 2.2 cont.

Plasmid	Marker	Compatibility	Comment	Source/reference
pZMR118	Km^r	λori	BamHI (10.1kb)-ClaI (7.6kb) fragment from Tn7 cloned into pZMR100	Mark Rogers
pZNE1	Tc^r	λori	Tn7 inserted into cloned "hot site"	Chapter 5
pZNE51	Ap^r	pMB1	RE-PstI (13.5) fragment from Tn7 cloned into pKO500	Chapter 3
pZNE62	Ap^r	pMB1	BglII (13.1kb)-BamHI (11.05kb) fragment from Tn7 cloned into pKL500	Chapter 3
pZNE64	Ap^r	pMB1	pZNE62 with BglII-XbaI (13.1kb-12.4kb) deletion	Chapter 3
pZNE65	Ap^r	pMB1	RE-PstI (13.5kb) fragment from Tn7 cloned into pKO500	Chapter 3
pZNE66	Ap^r	pMB1	pZNE83 with BglII-XbaI deletion (13.1kb-12.4kb)	Chapter 3
pZNE71	Ap^r	pMB1	BglII (14.0kb-13.1kb) fragment from Tn7 cloned into pNM482	Chapter 3
pZNE77	Cm^r Tc^r	P15A	HindIII (5.8kb-8.0kb) fragment from Tn7 cloned into pACYC184	Chapter 3
pZNE78	Ap^r	pMB1	RE-PstI (13.5kb) fragment from Tn7 cloned into pNM481	Chapter 3
pZNE81	Ap^r	pMB1	BglII (14kb-13.1kb) fragment from Tn7 cloned into pKO500	Chapter 3
pZNE83	Ap^r	pMB1	BglII (13.1kb)-BamHI (11.05kb) fragment from Tn7 cloned into pKO500	Chapter 3
pZNE89	Ap^r	pMB1	RE-BamHI (11.05kb) fragment from Tn7 cloned into pKO500	Chapter 3
pZNE90	Ap^r	pMB1	RE-HindIII (10.6kb) fragment from Tn7 cloned into pNM482	Chapter 3
pZNE91	Ap^r	pMB1	RE-PvuII (9.3kb) fragment from Tn7 cloned into pNM480	Chapter 3
pZNE93	Ap^r	pMB1	RE-ClaI (7.6kb) fragment from Tn7 cloned into pNM482	Chapter 3
pZNE200	Ap^r	pMB1	RE-BalI (13.85) fragment from Tn7 cloned into pUC8	Chapter 4
pNM480	Ap^r	pMB1	-	Minton, 1984
pNM481	Ap^r	pMB1	-	Minton, 1984
pNM482	Ap^r	pMB1	-	Minton, 1984

Iso-Sensitest Broth: 23.4g iso-sensitest broth made up to 1 litre with distilled water.

Iso-Sensitest Agar: 31.4g iso-sensitest agar made up to 1 litre with distilled water.

Minimal Agar: 7g K_2HPO_4, 2g KH_2PO_4, 4g NH_4SO_4, 0.25g trisodium citrate, 0.1g $MgSO_4.7H_2O$, 17.5g agar, made up to 1 litre in distilled water.

Davis-Mingioli (D&M) Salts (X4): 28g K_2HPO_4, 8g KH_2PO_4, 4g $(NH_4)_2SO_4$, 1g sodium citrate, 0.4g $MgSO_4.7H_2O$, made up to 1 litre with distilled water.

D&M Minimal Medium: 25ml D&M salts, 5ml 20% casamino acids, 250ul 20% glucose, 0.5ml vitamin B1 (1mg/ml), made up to 100ml with distilled water.

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

Supplements: When required, supplements were added to minimal media at the following concentrations:

glucose 2mg/ml
amino acids 40ug/ml
thiamine (vitamin B1) 20ug/ml
casamino acids 1%
thymine 50ug/ml

2.6 Sterilisation. All growth media were sterilised by autoclaving at 120°C for 15 minutes; supplements, gelatin solution and buffer solutions at 108°C for 10 minutes, and $CaCl_2$ at 114°C for 10 minutes.

2.7 Buffer Solutions and Other Solutions.

Electrophoresis

10 X E Buffer: 48.4g Tris-HCl, 16.4g NaAc, 3.6g $Na_2EDTA.2H_2O$, made up to 1 litre in distilled water, pH adjusted to 8.2 with acetic acid.

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

10 X Laemmli Gel Buffer: 144g glycine, 30g Tris-HCl made up to 1 litre in distilled water. When diluted, 10ml 10% SDS (w/v) was added for every 1 litre of 1 X concentration.

Single Colony Gel Loading Buffer: 2% Ficoll, 1% SDS, 0.01% bromophenol blue, 0.01% Orange G in buffer E.

Horizontal Agarose and Polyacrylamide Gel Loading Buffer: 25% (w/v) sucrose, 0.06% bromophenol blue, 10mM Tris-HCl (pH 8.0).

Vertical Agarose Gel Loading Buffer: 1% Ficoll, 0.5% SDS, 0.06% bromophenol blue, 0.06% Orange G in 1 X TBE.

Restriction and Ligation Buffers:

10 X Low salt: 100mM Tris-HCl pH 7.5, 100mM $MgSO_4$, 10mM DTT. Stored at 4^oC.

10 X Medium salt: 500mM NaCl, 100mM Tris-HCl pH 7.5, 100mM $MgSO_4$, 10mM DTT. Stored at 4^oC.

10 X High salt: 1M NaCl, 500mM Tris-HCl pH 8.0, 100mM $MgCl_2$, 10mM DTT. Stored at 4^oC.

10 X SmaI Buffer: 200mM KCl, 100mM Tris-HCl pH 8.0, 100mM $MgCl_2$, 10mM DTT. Stored at 4^oC.

10 X Ligation Buffer: 660mM Tris-HCl pH 7.6, 66mM $MgCl_2$, 100mM DTT. Stored at -20°C.

4mM ATP: Dissolve 60mg of ATP in 0.8ml distilled water. Adjust to pH 7.5 with 0.1M NaOH, made up to 1ml with distilled water; Stored at -20°C in aliquots and thawed once only.

TE Buffer: 10mM Tris-HCl, 1mM EDTA; pH 8.0.

All these buffers were stored for long term at -20°C.

Protein Sample Buffer: 10% glycerol, 0.01% bromophenol blue, 5% b-mercaptoethanol, 3% SDS, 0.625M Tris-HCl pH 8.0. Stored at room temperature.

Phenol All phenol used in the purification of DNA contained 0.1% 8-hydroxyquinoline, and was buffered against 0.5M Tris-HCl pH 8.0.

Birnboim-Doly I: 50mM glucose, 25mM Tris-HCl pH 8.0, 10mM EDTA; add lysozyme to 1mg/ml immediately before use.

Birnboim-Doly II: 0.2M NaOH, 1% SDS; made fresh.

Birnboim-Doly III: 5M KAc pH 4.8; mix equal volumes of 3M CH_3COOK and 2M CH_3COOH, pH should be 4.8.

STET Buffer: 8% sucrose, 5% Triton X-100, 50mM EDTA, 50mM Tris-HCl pH 8.0

Bradford Assay Solutions

Bradford reagent: Bradford reagent was made up by dissolving 100mg Coomassie Brilliant Blue G-250 in 50ml 95%x (W/V) ethanol, adding 100ml 85% (W/V) phosphoric acid to the mixture, and making up the volume to 1lt with distilled water. This was then filtered through Whatman No1 paper and

stored in a dark bottle.
Standard: 1mg/ml BSA in 150mM NaCl.

b-galactosidase Assay Solutions

Z buffer is per litre: 60mM $Na_2HPO_4.7H_2O$
40mM $NaH_2PO_4.H_2O$
10mM KCl
1mM $MgSO_4.7H_2O$
50mM b-mercaptoethanol

pH was adjusted to 7.0.

Galactokinase Assay Solutions

1. 5mM DTT, 16mM NaF.
2. 8mM $MgCl_2$, 200mM Tris HCl, pH 7.9, 3.2mM ATP.
3. 100mM $Na_2EDTA.2H_2O$, 100mM DTT, 50mM Tris HCl, pH 8.0.
4. 240ul 4mM galactose + 10ul D(1-^{14}C)galactose.

2.8 Antibiotics. The antibiotic concentrations used throughout for both liquid and plate selections were as follows:

Name	Source of resistance	Selective conc^{n}	Stock sol^{n}
Ampicillin (Ap)	plasmid	50ug/ml	5mg/ml (water)
Streptomycin (Sm)	chromosomal	300ug/ml	30mg/ml (water)
Streptomycin	Tn7	5ug/ml	0.5mg/ml (water)
Rifampicin (Rif)	chromosomal	50ug/ml	5mg/ml (methanol)
Tetracycline (Tc)	plasmid	10ug/ml	1mg/ml (10mM HCl)
Chloramphenicol (Cm)	plasmid	25ug/ml	2.5mg/ml (ethanol)
Kanamycin (Kan)	plasmid	50ug/ml	5mg/ml (water)
Trimethoprim (Tp)	plasmid/Tn7	50ug/ml	5mg/ml (50% water/ 50% ethanol)
Spectinomycin (Sp)	Tn7	25ug/ml	2.5mg/ml (water)
Naladixic acid (Nal)	chromosome	20ug/ml	2mg/ml (1M NaOH)

All stock solutions were stored at 4°C and when required added to molten agar pre-cooled to 55°C.

2.9 Indicators. X-gal (5-bromo-4-chloro-3-indolyl-b-galactoside) was used in conjuction with the host strains DS941 and DS947 and the pNM vectors providing a screen for plasmids with inserts in the polylinker region. Clones containing inserts were blue; clones lacking inserts were white. X-gal was stored at a concentration of 20mg/ml in dimethylformamide (DMF) at -20°C and added to L-agar plates to a final concentration of 20ug/ml. McConkey galactose agar was used to test for galactokinase expression (see sections 3.2.3 and 3.3.2).

2.10 Growth Conditions. Liquid culture for transformation, DNA preparations or conjugations were routinely grown in L-broth, or when Tp selection was required in isosensitest broth, at 37°C with vigorous shaking. Growth on plates was on L-agar, minimal medium plus supplements, or on isosensitest agar for selection on Tp. Antibiotics were used as required. Plates contained 25ml of agar solution and were incubated at 37°C overnight unless otherwise stated. All dilutions were carried out in D&M salts.

Bacterial strains were stored in 50% L-broth, 20% glycerol and 1% peptone at -20°C. Inocula from these stocks were applied to L-agar plates and grown, prior to replating on selective plates.

2.11 Determination of Viable Cell Counts. Viable cells were determined by spreading a measured aliquot of a diluted bacterial culture onto agar plates and counting the resulting colonies after a period of incubation. For growing cultures of ZNE15, ZNE16, ZNE18, ZNE19, and ZNE20 the number of viable cells was found to be proportional to the $O.D_{650}$.

2.12 Plasmid DNA Preparation. Two methods were used to obtain DNA from cells.

Birnboim and Doly (1979) DNA preparation: 200 ml cultures of stationary phase cells were harvested by centrifugation (12000g,

5mins at 4°C). The pellet was resuspended in 4ml of Birnboim-Doly I solution and incubated on ice for 5mins. 8ml of Birnboim-Doly II solution were added and the solution left on ice for 5mins before 6ml of cold Birnboim-Doly III solution was added, gently mixed and left on ice for a further 5mins. The cell debris and most chromosomal DNA was removed by centrifugation (32000g, 5mins at 4°C) and the plasmid DNA precipitated by addition of an equal volume of isopropanol followed by centrifugation at 39200g for 15mins. This DNA was further purified by banding on a CsCl/EtBr gradient. The DNA was resuspended in 2.09ml of water and added to 270ul of a 15mg/ml EtBr 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 20°C. Two bands were visible, a lower supercoiled plasmid band and an upper chromosomal and relaxed plasmid DNA band. The lower band was removed using a 1ml syringe and the EtBr removed by repeated butanol extractions (using water saturated butanol). The salts were removed by dialysis in 2 X 500ml 1 X TE. The DNA was then ready for use.

Mini DNA preparation using the method of Holmes and Quigley (1981): 1.5ml of an overnight culture containing the plasmid of interest was harvested by centrifugation in a 1.5ml eppendorf tube and resuspended in 350ul of STET buffer. 25ul of STET buffer containing lysozyme at a concentration of 10mg/ml was added and the tube vortexed briefly. This solution was boiled for 40secs and centrifuged in an eppendorf microfuge for 15mins at 4°C. The pellet was discarded using a toothpick and 40ul of 3M NaAc and 400ul of cold isopropanol was added, followed by microcentrifugation for 7mins which precipitated the nucleic acid. The pellet was washed twice in 70% ethanol and dried briefly in a vacuum drier before being resuspended in 20ul - 50ul 1 X TE. This DNA was suitable for digestion and other <u>in vitro</u> manipulations.

2.13 Ethanol Precipitation of DNA. The DNA solution was made 0.3M in NaOAc and 2 volumes of cold ethanol added. After mixing, the DNA was precipitated by cooling on ice for up to 1 hour and pelleted by centrifugation (27000g, 15mins, $4^{o}C$ for large volumes or 12000g, 15mins, $4^{o}C$ for small volumes in eppendorf centrifuge tubes). The pellet was usually washed in 70% ethanol and dried briefly in a vacuum drier.

2.14 Restriction of DNA. Restrictions were usually performed in a total volume of 20ul containing between 0.25ug and 1ug of DNA, 2ul of 10X restriction buffer and 1 unit/ug DNA of enzyme, the volume being made up with distilled water. For larger scale restrictions the volumes were scaled up accordingly. The reactions were allowed to proceed for 1 to 2 hours at the appropriate temperature. Reactions were stopped either by the addition of gel loading buffer or by rapid heating to $70^{o}C$ for 5 min followed by rapid cooling on ice.

2.15 End-Labelling of DNA. Restriction fragments which produce a using the Klenow fragment of DNA polymerase I. Only one labelled (alpha-^{32}P) triphosphate needs to be incorporated adjacent to the recessed 3' hydroxyl group for end labelling, (the labelled nucleotide used depends on the sequence of the restriction enzyme site). 2-5ug of plasmid DNA was appropriately digested in restriction buffer. 1unit of Klenow polymerase per ug of DNA and 3-5uCi (specific activity 3000Ci/mM) alpha labelled dNTP in 1ul (alpha-^{32}PdATP for 3' end-labelling EcoRI sites), was added to the digest, and incubated for 30mins at room temperature. The end labelled fragments were located by autoradiography on a 5% polyacrylamide gel.

2.16 Ligation of DNA Fragments. The restriction fragments to be ligated were mixed such that the insert was in 3-fold excess over the vector (10-fold excess for "blunt" end ligations) and made up to 20ul by the addition of 2ul 10 X ligation buffer, 2ul 4mM ATP and distilled water. T4 DNA ligase was added (0.01 units/ug DNA for "sticky" end ligation and 1 unit/ug DNA for "blunt" end

ligation) and the solution ligated for 1 hour at room temperature (overnight at 16°C for "blunt" end ligations). Aliquots of the ligation mix were used to transform competent cells.

2.17 Calf Intestinal Phosphatase (CIP) Treatment. To increase cloning efficiency, CIP was used to remove the 5'-terminal phosphate groups from the linearised vector to prevent recircularisation of the vector. This technique is particularly useful if there is no direct selection for the insert. CIP works in high, medium or low salt buffers and was used at a concentration of 1-2 units/ug DNA. It was added directly to the restriction digest for the final 15mins of the incubation and was heat inactivated in the manner described for restriction enzymes.

2.18 Conversion of Fragments with Protruding 5' Ends to Blunt Ends. Protruding 5' ends are filled in using the DNA polymerising activity of the Klenow fragment of E.coli DNA polymerase I. The reaction was set up as follows: restriction fragment (up to 1ug of DNA in 10ul)

2mM solution of all four dNTPs	1ul
10X nick translation buffer	2.5ul
water	to make up volume up to 25ul

To this reaction mix was added 2 units of Klenow fragment of DNA polymerase I. This was mixed and incubated for 30mins at 22°C. Subsequently the Klenow was heat inactivated at 70°C for 5mins. The blunt end fragment was then ready for use.

2.19 Quantitation of DNA. DNA concentration was determined spectrophotometrically. Readings (or preferably a scan) were taken at 260nm and 280nm. An OD_{260} of 1 is equivalent to 50ug/ml for double-stranded DNA. OD_{260}/OD_{280} ratio of pure DNA is 1.8. To obtain accurate quantitation pure samples of DNA were used (lacking e.g. phenol, agarose, protein, or other nucleic acids).

2.20 Transformation with Plasmid DNA. Plasmids were introduced to different strains by genetic transformation. An overnight culture of the recipient was diluted 1 in 100 into 20ml L-broth and was grown to a density of approximately 10^8 cells/ml (about 90mins - 2 hrs). The cells were harvested (12000g, 5mins, 4°C) and resuspended in 10ml of cold 50mM $CaCl_2$. The cells were pelleted again, resuspended in 1ml of cold 50mM $CaCl_2$ and kept on ice for at least 15mins before use. 200ul aliquots of the competent cells were added to the plasmid DNA, mixed gently and left on ice for up to 1 hr. The cells were heat shocked (2mins, 42°C) and returned to the ice for a further 15mins. An equal volume of L-broth was added and the cells incubated at 37°C for 1 hour to allow expression of the plasmid resistance genes. The cells were plated out on the appropriate selections. For transformation to ampicillin resistance, no expression time was necessary.

The presence of the new plasmid in the transformed colonies was confirmed by single colony gel electrophoresis.

2.21 Preparation of P1 Lysates. The donor strain was grown to mid-log phase in L-broth, the bacteria were then pelleted at 10K (1200g) and resuspended in a small volume of L-broth plus $CaCl_2$. Sufficient phage to produce an m.o.i. of 0.001 were added to the bacteria and allowed to absorb for 20mins at 37°C. 500ul of infected cells were added to 3mls of top agar and this was poured over a thin wet L-agar plate. These were then incubated in a sandwich box containing a small beaker of water for 6-8 hrs only. At this time substantial lysis of the bacteria had occured. The lysate was scraped off the plate with a glass spreader and placed in a centrifuge tube. A few drops of chloroform were added and the the lysate was vortexed vigorously for 30secs. This was allowed to stand for 30mins at room temperature and then the debris was pelleted by spinning at 10K (1200g) for 10mins. The resulting supernatant contains the P1 phage particles and was subsequently titred to determine the number of phage.

2.22 P1 Transduction. The recipient strain was grown to mid-log phase (2 x 10^8 cells/ml) in L broth plus 10mM $CaCl_2$ and the cells pelleted at 10K (1200g) for 10mins. Sufficient phage to give an m.o.i. of 0.01-0.1 were added to 200ul aliquots of bacteria in phage buffer. This mixture was incubated at 37^{o}C for 20mins, then the cells were pelleted and washed twice in 1M Na citrate. These cells were then plated in 3ml of soft agar on selective plates and incubated at 37^{o}C for 48hrs.

2.23 Single Colony Gel Analysis. This technique enables the plasmid content of a colony to be observed without the purification of plasmid DNA. A single colony was patched out (1cm square) and grown overnight. The patch was scraped off the plate using a toothpick and suspended in 150 - 250ul of single colony gel buffer. The cells were left to lyse at room temperature and centrifuged in a microfuge (12000g, 4^{o}C) for at least 15mins. 30ul of the supernatant were loaded onto an agarose gel which did not contain ethidium bromide.

2.24 Gel Electrophoresis. Both vertical and horizontal agarose gels were used. Vertical gels were used mainly for single colony analysis while horizontal agarose gels were used for restriction analysis of plasmids and for southern analysis. Unless otherwise stated 1% agarose gels were used.

Horizontal gels. Two types were commonly run.

(1) 100ml gels - 100ml of molten agarose was poured into a 11 X 19cm perspex gel former with a 13 space teflon well former. After the gel had set, the comb was removed and the gel placed in a horizontal gel tank, submerged in E buffer and loaded with 20 - 30ul of sample. Gels were usually run overnight at 20V and stained in ethidium bromide (0.5ug/ml) for 30mins. The stained DNA was photographed on a 254nm wavelength UV transilluminator. Some gels were made up with ethidium bromide to a concentration of 0.5ug/ml.

(2) 200ml gels - These gels were made by pouring 200ml of molten agarose into a 16.5 X 23cm gel former with a 20 space well former. The gels were run in E buffer in a gel tank with a buffering capacity of 3 litres overnight at 20V. The gels were stained with ethidium bromide (0.5ug/ml) and visualised on a 254nm wavelength transilluminator.

Vertical gels: The gel kits held two 16 X 15cm glass plates separated by 3mm spacers. After sealing the edges with molten agarose the agarose was precooled to 55°C and poured between the glass plates. The comb (10 or 15 teeth) was inserted and the gel allowed to set. The top and bottom wells were filled with E buffer and the comb removed. Samples were loaded and the gel run, generally at 5 V/cm for about 4 hours prior to staining in 0.5ug/ml ethidium bromide. The gel was photographed on a 254nm wavelength transilluminator.

Vertical gels were used mainly for single colony analysis and the interpretation of the DNA band seen followed that of Dugaiczyk et al (1975). The fastest migrating and generally most abundant band was the supercoiled plasmid monomer. Behind this ran the open cicular plasmid band often comigrating with the supercoiled plasmid dimer. Open circular dimers and other higher forms ran higher up the gel. Plasmid linears could sometimes be detected running between the supercoiled monomer and the open circle monomer. Sheared fragments of chromosomal DNA ran as a thick band toward the top of the gel. Large plasmids like R388 run above the chromosomal band.

Low melting point agarose gels. These gels were used to separate fragments which were to be purified from the agarose and used in ligation reactions. The agarose was dissolved in 100ml 1 X E buffer at 100°C, precooled to 37°C and poured as a 100ml gel described above.

Photographing of gels. Ethidium bromide stained gels were viewed on a 254nm UV transilluminator and photographed using a Polaroid CU-5 camera loaded with type 67 land film or using a Pentax 35mm

SLR loaded with Ilford HP5 film. Both cameras were fitted with a Kodak Wratten filter No.9.

Sizing of restriction fragments. The size of linear restriction fragments was estimated from graphs of the $\log_{10}$ molecular size plotted against the distance migrated in the gel according to;

$$\log M = C \times 1/D$$ (Helling *et al*, 1974)

M = Molecular size in base pairs.
D = Distance migrated
C = Arbitary constant

Molecular weight standards were obtained by restriction of Lambda cI857Sam7 (Philippsen *et al*, 1978; Haggerty and Scheif, 1976) or the pUC plasmids (Yanish-Perron *et al*, 1985).

2.25 Extraction of DNA from Low Melting Point Agarose Gels. After staining, the gel was placed on a long wave transilluminator (300nm - 360nm) and the band of interest excised. The agarose slice was added to 3 volumes of E buffer and heated to 65°C. The solution was cooled to 37°C and an equal volume of room temperature phenol added, mixed throughly by vortexing and the phases separated by centrifugation (12000g, 2mins). The supernatant was recovered, phenol extracted 3 times, chloroform extracted twice and ether extracted once. The DNA was precipitated and dried. This DNA was used for ligation.

2.26 UV Sensitivity Test. This technique was used to rapidly confirm the recombination status of strains. Stationary phase cultures of the strain to be tested and control strains were spotted onto L-agar plates at 10^{-2}, 10^{-4} and 10^{-6} dilutions. These plates were exposed to UV radiation (17.5ergs/sec/M^2) for 30 and 60 seconds and incubated overnight in the dark. *rec*$^+$ strains usually grow after 60 seconds exposure while *rec*A$^-$ strains show reduced growth after 30 seconds and do not grow after 60 seconds exposure. *rec*F$^-$ strains show an intermediate phenotype. Periodically all strains were checked for their

recombination status.

2.27 Protein Gel Electrophoresis: The electrophoresis of proteins followed the procedure of Laemmli (1970). Unless specifically stated all SDS-polyacrylamide gels were 10% running gel with a 4% stacking gel. The gel plates were separated by 0.8mm spacers and the gel was pre-sealed using molten 0.6% agarose in H_2O. The gels were poured according to the table below;

	Running Gel	Stacking Gel
Acrylamide/Bis (30 ; 0.8)	10ml (10%)	2.0ml (4%)
Tris-HCl	11.25ml (1M; pH 8.8)	1.9ml (1M; pH 6.8)
10% SDS	0.3ml	0.15ml
TEMED	0.01ml	0.01ml
APS (made fresh; 100mg/ml)	0.15ml	0.15ml
dH_2O	8.3ml (10%)	10.9ml (4%)

The running gel was poured leaving about 1.5-2.0cm between the top of the running gel and the bottom of the comb. Isopropanol was layered on top of the running gel to create a sharp interphase. Once the running gel had set, the stacking gel was poured after removing the isopropanol and washing away any residue with distilled water. The comb was pushed into the stacking gel firmly. After the gel had set the comb was removed, and running buffer poured into the gel tank covering the wells. Residue acrylamide was washed out of the slots using running buffer prior to loading the samples. The gel was run overnight at about 6mA until the bromophenol blue dye had reached the bottom of the gel.

The gel was fixed in 30% methanol/10% glacial acetic acid for 1 hour before being soaked in EN^{3}HANCE for a further hour. The EN^{3}HANCE was removed and the gel soaked in water for a further 30 minutes before being dried in a Biorad vacuum drier at 80^{o}C for at least 1 hour. The dried gel was placed against Blue sensitive film (Kodak Xomat S) and kept in the dark at -70^{o}C for a variable time (up to 1 month) to allow exposure of the film.

2.28 Conjugation and Transposition Assays

Bacterial matings in culture. Overnight cultures of donor and recipient cells (usually in minimal media) were diluted 1:100 in L broth and grown with shaking to mid log phase (to a density of about 2-3 X 10^8 cells/ml). Donor and recipient cells were then mixed in a ratio of 1:10, in a total volume of 5ml. Aeration during mating was achieved by having a large surface to volume ratio using a 100ml flask on a 33rpm rotor at 37°C (vigorous agitation would disrupt mating pairs). 200ul aliquots of the mating mixture and several dilutions (usually 10^{-1}, 10^{-2}) were then spread onto minimal selective plates (containing selection against both the donor and the recipient). Aliquots of both the donor and the recipient were also spread onto the selective plates as controls. The plates were incubated at 37°C for 24-72 hrs.

Interrupted liquid matings. These matings were carried out to determine the time of entry of particular markers on the E.coli K-12 chromosome, as Hfr strains transfer markers in a linear fashion from a fixed point of origin. The donor and recipient strains were grown until they reached the exponential phase of growth (2 X 10^8 cells/ml) and then mixed in the ratio of 1Hfr:10F$^-$. This mating mixture was incubated in a 37°C water bath. Samples were removed at different time points and the mating pairs were disrupted by vortexing, diluting and then plating out on selective media.

To assay transposition frequencies the following methods were employed:

Streak mating. When a conjugative plasmid had to be mated into a number of different strains, this method was used. The donor was spread onto half an agar plate containing antibiotics or lacking supplements that would both counter select the donor strain and select for the plasmid. Streaks of the recipient were made, starting on the clean half of the plate and crossing into the area containing the donor. Only at the boundary between

these strains should colonies grow. These were picked and purified.

Plate matings. This technique was employed to measure the transposition frequency between a non mobilisable plasmid containing Tn7 and a conjugative plasmid. Donor and recipient cells were grown overnight and 0.5ml of donor was mixed with 1.5ml of recipient ensuring an excess of the recipient. These cells were concentrated 10-fold and spread on a well-dried agar plate. The plate was incubated inverted for 3 hours at 37°C. The cells were washed off the plate using D&M salts and concentrated 10 fold. This was the 0 dilution. Serial dilutions were made and plated out on the relevant selective media which selected for recipients with the plasmid or for recipients which had received the conjugative plasmid and the transposon. The proportion of recipients receiving both plasmid and transposon compared to those receiving only the conjugative plasmid was termed the transposition frequency. The standard errors of transposition assays were less than 1/5 of the frequency of transposition unless otherwise stated.

2.29 Bradford Protein Assay for Total Protein Estimation. 50ul of culture lysed by the toluene/ethanol method was added to 2.5ml Bradford reagent, mixed well, and incubated at room temperature for 10mins. The OD_{595} was measured. A standard curve was also constructed using a range of BSA concentrations(0-50ug/ml).

Ref: Bradford, 1976.

2.30 Determination of the Content of b-galactosidase. Overnight cultures of the strains to be assayed were grown under selection in minimal medium supplemented with D&M salts, B1, $MgSO_4$, and glucose. Each of the cultures was diluted in fresh medium of the exact type used in the overnight. 4 drops of the overnight were added to 5ml of fresh media. The cultures were grown at 37°C with shaking until the OD_{650} was in the range of 0.2-0.6, and then cooled on ice. After 20mins the bacterial density was

measured and recorded. Immediately aliquots of the cultures were added to the assay medium (Z buffer). The final volume was always one ml. If high levels of b-galactosidase were being assayed, 0.1ml of culture were added to 0.9ml of Z buffer. If low levels were being determined, 0.5ml of culture were added to 0.5ml of Z buffer. 2 drops of chloroform and 1 drop of 0.1% SDS solution were added to each ml of assay mix. The tubes were then vortexed for 10secs, and placed in a water bath at 28°C for 5mins. The reaction was started by adding 0.2ml of ONPG (4mg/ml) to each tube and shaking for a few seconds and the time of the reaction was recorded with a stop clock. The reactions were stopped by adding 0.5ml of a 1M Na_2CO_3 after sufficient yellow color had developed. The optical density at both 420nm and 550nm for each tube was recorded.

b-galactosidase specific activity was calculated by using the following equation:

$$\text{Units} = 1000 \times \frac{OD_{420} - 1.75 \times OD_{550}}{t \times v \times OD_{600}}$$

t=time of reaction in minutes, and v=volume of culture used in the assay in ml.

Ref: Miller, 1972

2.31 Determination of the Content of Galactokinase. Overnight cultures of the strains to be assayed were grown under selection in L-broth without glucose. These were subcultured 1 in 20 in the same medium and grown at 37°C with shaking till their OD_{650} was in the range 0.2-0.35. The OD_{650} of each culture was recorded. 1ml aliquots of each culture were taken, 10ul of mix 3 containing CTAB (mixed alkyltrimethyl-ammonium bromide) at a concentration of 5mg/ml, added to each, and the mixture was incubated at 30° C for 10 mins. Assay reaction systems were made up in small microfuge tubes as follows:

5.0ul of mix 1
+ 12.5ul of mix 2
+ 2.5ul of mix 4

(See section for the composition of the solutions).
The reactions were started by the addition of 5ul. of cell lysate. (Blank systems were set up by adding 5ul. of L broth rather than cell lysate and additional systems to act as unwashed, "total count" systems were also set up, using cell lysates taken at random). The reactions were incubated at 32°C for 30 mins. and stopped by placing the tubes on ice. The whole reaction mixture was spotted onto a 2.3 cm diameter Whatman DE81 filter paper and the filters carrying the unwashed reaction systems dried. The other filters were added twice (in 1 litre of distilled water for 10 mins. each time) and then dried. Lastly, the filters were counted in 5ml of "Biofluor" scintillant in a scintillation couter. Each filter was counted three times for two minutes each time.
Galactokinase activity was calculated using the following equation:

$$\text{Number of units of galactokinase activity} = \frac{\text{cpm from test system} - \text{cpm from blank system}}{\text{average cpm from 2 unwashed filters} \times \text{incubation time (mins)} \times OD_{650}} \times 10400$$

cpm= counts per minute

Ref.: McKenney et al, 1981

2.32 Determination of the Content of b-lactamase. The content of b-lactamase was determined by measuring the rate of hydrolysis of nitrocefin (O'Callaghan et al, 1972). 200ul of a bacterial culture with OD_{650} in the range 0.2-0.6 was centrifuged and resuspended in 0.5ml of 0.1M Tris HCl, pH 8.0. To the suspension, 20ul of a lysis medium consisting of 100mM EDTA and 100mM DTT in 50mM Tris HCl, pH 8.0 and a small drop of toluene were added. The suspension was kept at 30°C for 30mins and then cooled to 0°C. 20ul of the toluenised cells were incubated for 10mins with 1ul of 100uM nitrocefin in 0.1M Tris HCl, pH 8.0 at

30°C. The reaction was terminated by the addition of 20ul of 20% SDS and the increase in OD_{482} was determined within 10mins. The difference of molar extinction coefficients of the native and hydrolysed nitrocefin is 1.59 X 10^4 at 482nm (O' Callaghan et al, 1972). The unit of activity is defined as nanomoles of nitrocefin hydrolysed per min per ml of cells at OD_{650} of 1.0. In the standard assay, a bacterial culture with one unit per ml of activity gives a change in OD_{482} of 0.00071 in 10mins.

Ref: O' Callaghan et al, 1972.

2.33 DNA-Binding Gel Electrophoresis Assay.

Protein-DNA complexes were resolved on low ionic strength polyacrylamide gels as described (Fried and Crothers, 1981; Strauss and Varshavsky, 1984) with a few modifications. Protein samples were incubated with approximately 1ng of end-labelled double-stranded DNA fragment in the presence of 1-2ug of "carrier" DNA (see chapter 4) in a final volume of 10-20ul. Incubations were carried out at 37°C for 5-10mins in 10% glycerol, 10mM Tris/HCl pH 8.2, 0.1mM EDTA and 40mM NaCl. Samples were layered onto low ionic strength 5% polyacrylamide gels. Gels were preelectrophoresed for 90mins at 150V in buffer consisting of 10mM Tris/HCl pH 8.2 and 0.1mM EDTA. Buffer was recirculated between compart ments. Gels were electrophoresed at 150V at room temperature for 2.5hrs. They were then transferred to Whatman 3MM, dried, and autoradiographed.

2.34 Preparation of Crude Extracts. Strains pZMR117/DS941, pZMR100/DS941, pZMR64/DS941, pGLW8/DS941, pMA6114/JM101 and JM101 were grown in L-broth containing the appropriate antibiotics at 37°C to an $O.D_{600}$ of 0.5. An equal volume of L-broth with antibiotics and 1mM IPTG was then added to the cultures for induction of expression of the tnsB and tnsD gene products. After about 90mins ($O.D_{600}$ may begin to drop), cells were harvested by centrifugation at 4000g at 4°C for 15mins and washed by resuspension in extraction buffer (50mM Tris/Hcl pH 8.2; 50mM NaCl; 5mM EDTA; 0.4mM DTT; 1mM benzamidine; 1.2mM

phenylmethylsulfonylfluoride), repelleted as above and weighed. To each gram of cell was added 1.5ml of extraction buffer. Cells were then burst open by sonication, and spun at 45700g for 30mins at 4°C. The supernatant (crude extracts) was stored at -20°C in 50% glycerol. The protein concentration of the crude extracts was determined by the method of Bradford, (1976).

CHAPTER 3

ANALYSIS OF TN7 GENE EXPRESSION USING TRANSCRIPTIONAL AND TRANSLATIONAL FUSIONS

3.1 INTRODUCTION

When this work was initiated, little information was available on the genetic organisation of the transposition functions of Tn7. All such information came from the complementation of deletion mutants of the transposon using other deletions within Tn7 or plasmids carrying cloned Tn7 fragments. The major results from these studies were the following:

a) All functions required for transposition were entirely encoded within a 7.9kb fragment starting from the BstEII site at 6.1kb to the BglII site at 14kb (Grinter, 1983; Hauer and Shapiro, 1984; Hodge, 1983; Ouartsi *et al*, 1985).

b) Three complementation groups were defined, each of which contained at least one function required for transposition. All three were necessary for "cold site" transposition, but only two were essential for transposition to the "hot site" (Figure 3.1; Hauer and Shapiro, 1984; Smith and Jones, 1984).

These data were incomplete and could only loosely localise the proposed transposition functions. They indicated, however, that Tn7 transposition is complex and unlike any of the transposons studied so far.

For this thesis, a totally independent strategy aimed at identifying transcriptional control signals and translated reading frames in Tn7 was undertaken. Restriction fragments from the 7.9kb right part of the element (transposition region) were cloned into vectors specially designed to detect transcriptional and translational units (see sections 3.2 and 3.3). This approach resulted in the identification of genes whose transcriptional and translational efficiencies were measured (see sections 3.2 and 3.3). The control of transcription in Tn7 was also studied.

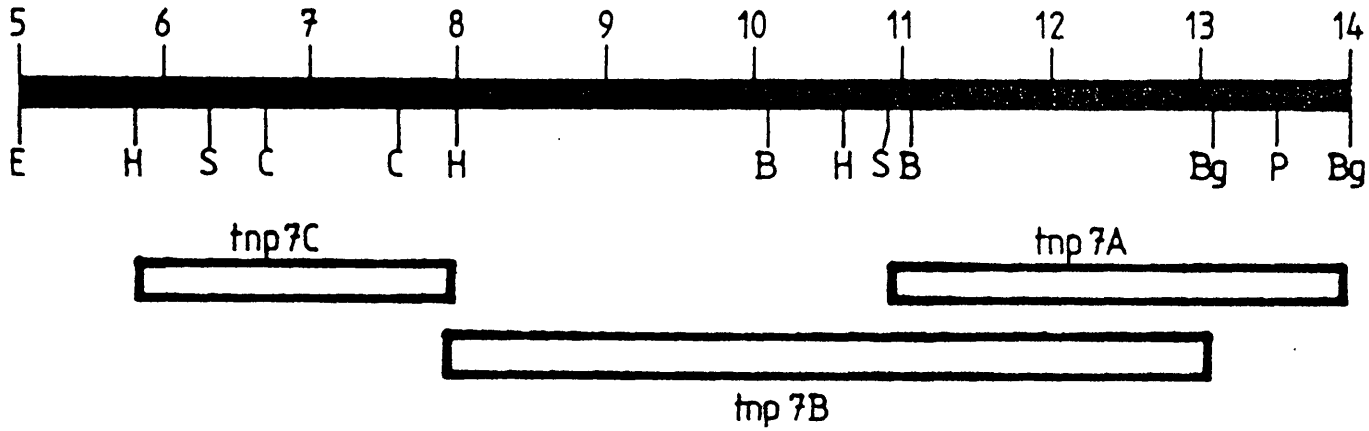


Figure 3.1 Map of Tn7 transposition functions.
Modified from Hauer and Shapiro, (1984). Each open box represents a complementation group.

Abbreviations: B, BamHI; Bg, BglII; C, ClaI; E, EcoRI; H, HindIII; P, PstI; S, SstI

RESULTS AND DISCUSSION

3.2 Transcriptional Fusion Vectors

Transcriptional fusion vectors provide systems to isolate, compare and characterise essentially any promoter or terminator signal recognised by the E.coli RNA polymerase. A promoter probe fusion vector carries a gene with an assayable and selectable gene function, such as b-galactosidase (lacZ), galactokinase (galK), chloramphenicol acetyltransferase (cat), or tetracycline resistance (tetA) (Casadaban and Cohen, 1980; Close and Rodriguez, 1982; McKenney et al, 1981; Munson et al, 1984). Expression of this gene is prevented by the removal of its promoter. Foreign DNA segments cloned into the polylinker sites located adjacent to the structural portion of the indicator gene can be screened for promoter activity according to their ability to express the gene function. A terminator probe fusion vector is a derivative of a promoter probe vector which has its promoter intact, and hence expresses the gene function encoded by the indicator gene. Cloning of any terminator-bearing DNA segments into the polylinker region which exists between the promoter and the indicator gene results in partial or total disruption of gene expression. The level of the reduction in gene expression gives a measure of the terminator strength.

The plasmid vector system has several limitations most of which are due to the high copy number. A cloned regulatory signal when present in a large number of copies, may not respond normally in experiments designed to study the in vivo regulation of the site. If, for example, the signal is affected by various ancillary factors, these may become limiting to the cell. In addition, the cell may respond abnormally to the presence of many copies of the regulatory site, again complicating regulatory studies. A different problem arises from potential variations in plasmid copy number. In the latter case, however, copy number assays provide a means of making possib le valid

comparisons among different constructs. These and other potential problems created by the plasmid vector system can be alleviated by integrating the fusions into the chromosome.

In this study, galK fusion vectors have been used for the detection and study of transcriptional control signals in Tn7. The reason for deciding to work with these vectors was the easy, swift and sensitive assay that can be used to measure galK expression (see Materials and Methods).

3.2.1 GalK Gene Fusions

The galK fusion vectors provide a system to study promoter and terminator signals recognised by the E.coli RNA polymerase. In these vectors, the promoter-distal gene of the galactose operon, galK, including the 168bp region normally preceding the galK coding sequence which includes the galK ribosome binding site, is inserted into a pBR322 derivative deleted for the entire Tc^r region. Any of the unique cloning sites upstream galK can be used to insert fragments with potential transcriptional regulatory signals. When fused in the proper orientation, the inserted signal will control galK expression. Because many DNA segments that carry transcriptional regulatory sites also contain translational start sites, translational stop codons in all three reading frames are constructed beyond the cloning sites, to minimize the potential effects of upstream translation that might start in the inserted DNA and traverse all or part of the leader region preceding galK, thereby exerting different effects on galK translational efficiency (Figure 3.2).

3.2.2 The Transcriptional Fusion Vectors pKO500 and pKL500 (McKenney et al, 1981)

In this work, pKO500 and pKL500 were the vectors used for the detection of promoter and terminator signals, respectively. pKO500 (Figure 3.2) is a pKL500 derivative deleted for a 300bp fragment upstream the polylinker that carries the wild type lac promoter.

3.2.3 Selection for and against GalK

The galK gene provides a readily selectable genetic marker that can be used for the isolation of DNA fragments that carry transcription regulatory sites. Galactokinase expression is detected on McConkey galactose agar, a medium containing phenol red as a pH indicator. GalK$^-$ bacteria that can grow on this medium form white colonies. If the bacteria can ferment galactose (galK$^+$) the acid by-products of the reaction turn the pH indicator, and thus the colony, red. The degree of redness is proportional to the amount of galK expression. To ensure accuracy, the reaction on McConkey galactose plates was read soon after growth was complete, since color distinctions on McConkey agar fade with time.

The galK gene in pKO500 is only transcribed at a very low level by some uncharacterised promoter on the plasmid and hence, a galK$^-$ cell carrying this plasmid will not grow on minimal media with galactose as the only carbon source and will grow on McConkey galactose indicator plates as a white colony. However, when a DNA fragment carrying a promoter signal is inserted into pKO500, galK is transcribed and the plasmid complements the galK$^-$ host, resulting on growth on minimal galactose media and growth as red colonies on McConkey galactose plates.

3.2.4 Detection of Transcriptional Initiation in Tn7

Initially, a preliminary experiment was undertaken to search for promoters in the 7.9kb transposition region of Tn7. Shotgun restriction fragments from pZMR80::Tn7, a plasmid containing the wild-type transposon (Table 2.2), were cloned into pKO500 (Figure 3.2) and recombinants exhibiting galK expression were kept for further analysis. Cloning of BamHI, BglII, BamHI-BglII, Sau3A and HindIII restriction fragments from pZMR80::Tn7 into the BamHI and HindIII polylinker sites of pKO500 gave rise to two clones expressing galactokinase, as judged by the color test described in section 3.2.3. One of these clones, pZNE81, carried the 0.9kb

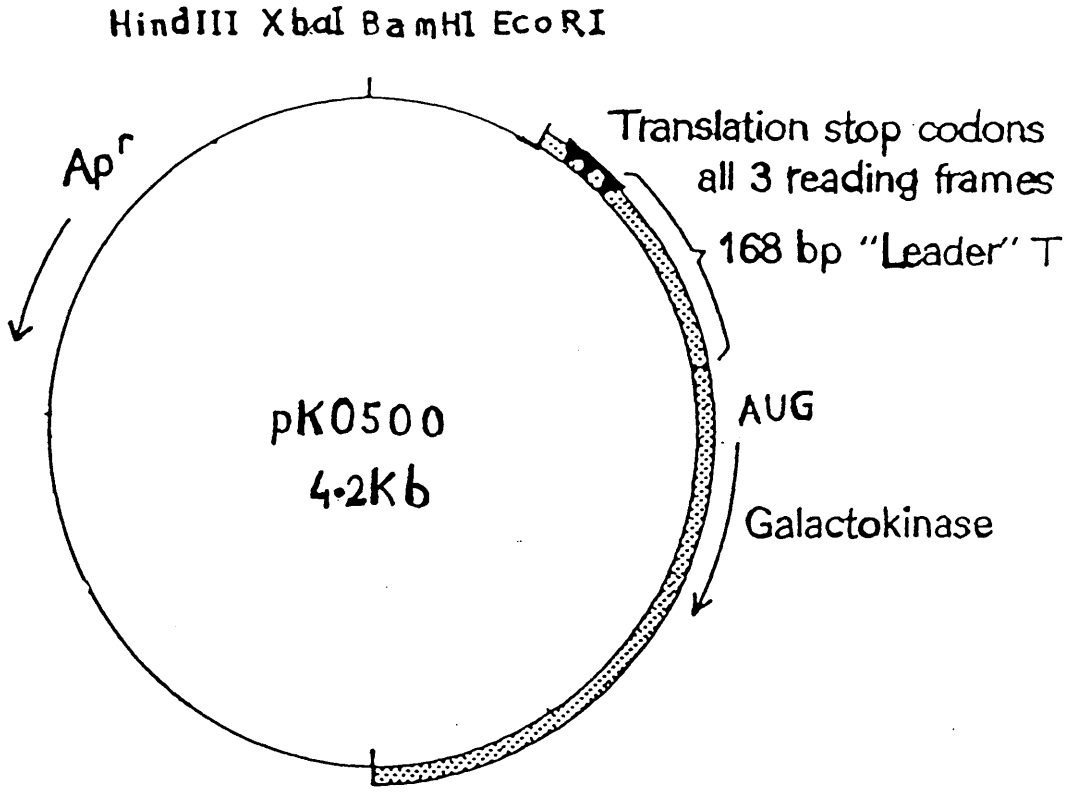


Figure 3.2 The transcriptional fusion vector pKO500. The dotted area represents the galactokinase gene (galK) and 168bp of "Leader" galT gene sequences preceding the AUG codon of galK. Translation stop codons in all three reading frames prevent any translation originating in the inserted DNA fragment from reaching the ribosome binding site and AUG codon of galK.
McKenney et al, 1981

BglII Tn7 fragment and the other, pZNE83, carried the 2.05kb BglII-BamHI Tn7 fragment (Figure 3.3).

During the course of these experiments, Mark Rogers in our group in Glasgow, was using complementation tests to define the functions required for transposition of Tn7 to "hot" and "cold" sites. At least three functions participating in transposition were indicated by his preliminary results. Completion of the complementation map coincided with the completion of the fusion analysis.

In order to better understand the organisation of the transcriptional units in Tn7, a second, more specific approach was undertaken: A series of nested fragments from the 7.9kb right part of the element was fused to galK in pKO500 (Figure 3.3). In choosing the endpoints of these fragments the preliminary results from the complementation analysis were taken into account. Figure 3.3 gives an illustration of all the Tn7 sequences fused to galK. Galactokinase assays (see Materials and Methods) were performed on all clones constructed (Table 3.1) to measure the levels of galactokinase expression, and hence promoter activity. Table 3.2 correlates the Tn7 function fused to galK in each of these constructs with its presumptive gene product as deduced from the preliminary data from the complementation tests.

The major conclusions from the results presented in Table 3.1 are discussed below:

a) A strong promoter is contained within the 0.5kb BglII (14kb)-PstI (13.5kb) Tn7 fragment, reading leftwards into the element. This promoter was named P_1 and appears to be 0.5 times the strength of P_{lac} (pZNE65, pZNE81; see Table 3.1 and Figure 3.3).

b) A weaker promoter is contained within the 0.7kb BglII (13.1kb)-XbaI (12.4kb) Tn7 fragment, initiating trascription from right to left. This promoter was named P_2 and appears to be 0.04 times the strength of P_{lac} (pZNE83; Table 3.2). P_2 was first

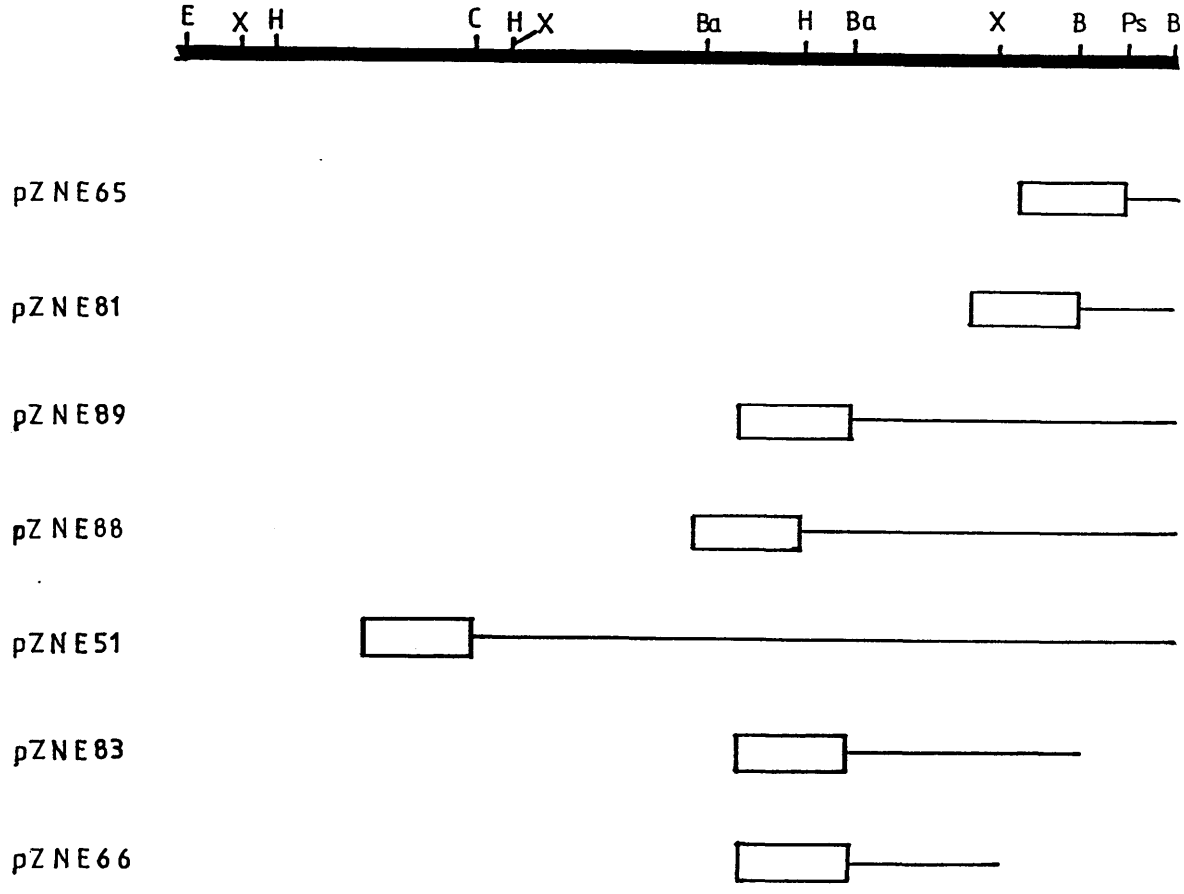


Figure 3.3 Illustration of the Tn7 sequences fused to galK in pKO500. The right 9kb of Tn7 is presented as a thick solid line. The thin solid line represents the region of Tn7 present in the plasmid named alongside. Open boxes represent the galK gene.

Abbreviations: Ba, BamHI; B, BglII; C, ClaI; H, HindIII; E, EcoRI; P, PstI; X, XbaI

Table 3.2 Properties of Plasmids Used in Table 3.1

Plasmid	Tn7 Protein fused to galK	Tn7 promoter
pZNE65	tnsA	P_1
pZNE81	tnsA	P_1
pZNE89	tnsB	P_1 and P_2
pZNE83	tnsB	P_2
pZNE66	tnsB	-
pZNE88	tnsC	P_1 and P_2
pZNE51	tnsE	P_1 and P_2
pKO500	-	-
pKL500	-	P_{lac}

Table 3.1 Detection of Transcriptional Initiation in Tn7

Plasmid[4]	galK/ml ($\times 10^2$)	b-lact/ml ($\times 10^2$)	galK(S.A)[1]	b-lact(S.A)[2]	galK[3] / b-lact ($\times 10^2$)
pZNE65	87	246	352	998	35
pZNE81	57	146	268	690	39
pZNE89	24	249	102	1045	10
pZNE83	12	459	42	1584	3
pZNE66	0	343	0	1320	0
pZNE88	4	320	14	1229	1
pZNE51	2	163	6	698	1
pKO500	3	362	10	1320	1
pKL500	109	148	476	648	73

1. GalK specific activity (S.A) units are expressed as nanomoles of galactose phosphorylated per min per ml of cells at $O.D_{650}$ of 1.0, as described in Materials and Methods. GalK/ml units are derived from galK specific activity units by multiplying by the $O.D_{650}$ of the culture.

2. b-lactamase (b-lact) specific activity units (S.A) are expressed as nanomoles of nitrocefin hydrolyzed per min per ml of cells at $O.D_{650}$ of 1.0, as in Materials and Methods. b-lact/ml units are derived from b-lact specific activity units by multiplying by the $O.D_{650}$ of the culture.

3. Copy number corrected galK activity gives identical results whether calculated from specific activity units or units/ml.

4. All plasmids were assayed in strain DS941 (galK$^-$, Table 2.1). Each assay was repeated at least three times.

detected within the 2.05kb BglII (13.1kb)-BamHI (11.05kb) fragment (pZNE83) and was further localised within the 0.7kb BglII-XbaI region by constructing pZNE66 (Figure 3.3 and Table 3.1).

c) The level of transcription at the BamHI site (11.05kb) in pZNE89 appears to be too high to be explained as transcription coming solely from P_2. This transcription level, however, cannot be due to a promoter contained within the 1.45kb BglII (13.1kb)-XbaI (12.4kb) fragment, since no detectable levels of transcription were found at the BamHI site (11.05kb) in pZNE66. It appears, therefore, that the level of transcription at the BamHI site in pZNE89 is due to transcription coming from both P_1 and P_2. Comparison of the levels of transcription at the BamHI site at 11.05kb in pZNE89, pZNE83 and pZNE65 (or pZNE81) implies the existence of a factor in the 3.95kb BglII-BamHI fragment reducing P_1 activity by approximately 5-fold. This factor could either be a Tn7 function encoded within the 3.95kb BglII (14kb)-BamHI (11.05kb) region or a terminator existing in the 0.7kb BglII-XbaI region. These two possibilities are discussed in sections 3.2.5 and 3.2.6.

d) The levels of transcription at the HindIII site (10.6kb) and ClaI site (7.6kb) are similar to the levels obtained for pKO500 indicating the presence of one or more terminators within the 3.45kb BamHI (11.05kb)-ClaI (7.6kb) region.

3.2.5 Effects of *trans*-Acting Functions on P_1 Activity

It is known that one of the major functions of the transposition complex is to interact with the ends of the element and that such interaction is important for initiation of transposition. Figure 3.4 shows that the -35 region of P_1 is contained within the innermost direct repeat of the right end of Tn7, a likely candidate for binding of transposition proteins. To investigate whether the observed reduction in P_1 activity at the BamHI site in pZNE89 was due to repression by one or more Tn7 proteins

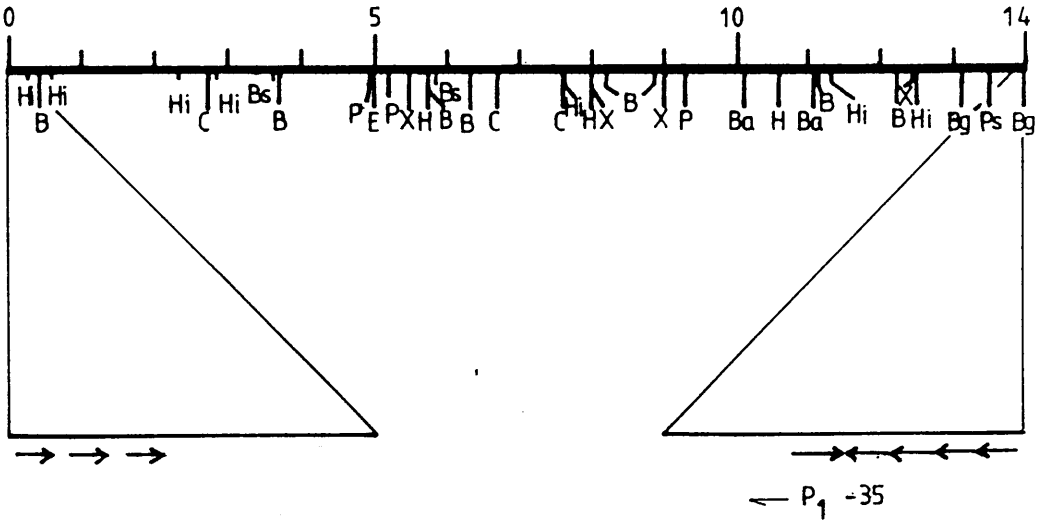


Figure 3.4 Organisation of the termini of Tn7 indicating that the -35 region of P_1 is contained within the innermost direct repeat in the right end of the element.

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

encoded within the 3.95kb BglII-BamHI region, the following experiment was carried out:

A series of plasmids (pZMR112, pZMR113, pZMR117, pZMR118 and pZNE77; Figure 3.5) containing each at least one function required for transposition, as deduced from the genetic analysis, was used to complement pZNE65, thus assessing the effects on P_1 activity. Each of these plasmids is capable of fully complementing transposition of Tn7 mutants deleted for the transposon sequences carried on the plasmid. Table 3.4 correlates the Tn7 fragments cloned in these constructs with their presumptive gene functions. To simplify the nomenclature only one function is assigned to each construct. A strong promoter, P_{tac}, is used to transcribe the Tn7 functions in pZMR112, pZMR117, pZMR113 and pZMR118, while the tnsE function in pZNE77 is driven by the anti-Tc promoter in pACYC184 (Table 2.2).

Results in Table 3.3 show that the presence in trans of pZMR117 with pZNE65 results in an approximate 5-fold reduction in P_1 activity. This repression of P_1 transcription was believed to occur by binding of the Tn7 gene products encoded within the 2.05kb BglII-BamHI fragment either directly at P_1 or the repeats at the ends of the element. This latter hypothesis was further investigated in chapter 4.

None of the other Tn7 functions used to complement P_1 transcription appeared to have any significant effect on its activity. It should be noted, however, that it is difficult to study the effect of the Tn7 product (or products) encoded by pZMR112 on P_1 activity, since P_1 is already present in this construct. In order to examine the real effect of the gene product encoded in the 0.9kb BglII fragment, a pZMR112 derivative should be made, deleted for P_1.

3.2.6 Detection of Transcriptional Termination in Tn7

In section 3.2.4 it was discussed that the observed reduction in the level of transcription at the BamHI site at 11.05kb could be

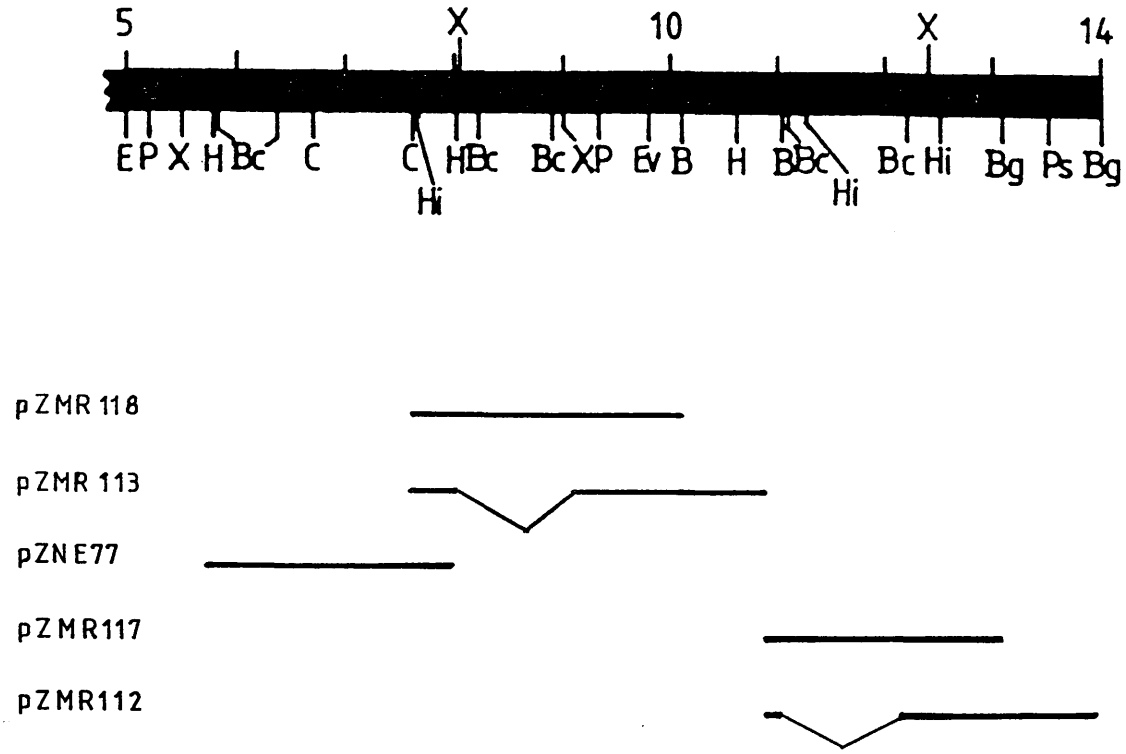


Figure 3.5 Illustration of the Tn7 sequences used to complement pZNE65 in Table 3.3
The right 9kb of Tn7 is presented as a thick solid line. The thin lines represent the regions of Tn7 present in the plasmid named alongside. A ⟨ represents a deletion of Tn7 sequences.

Abbreviations: B, BamHI; Bc, BclI; Bg, BglII; C, ClaI; E, EcoRI; Ev, EcoRV; H, HindIII; Hi, HincII; P, PvuII; Ps, PstI; X, XbaI

Table 3.4 Properties of Plasmids Complementing pZNE65 in Table 3.3

Plasmid	Tn7 protein	Promoter
pZMR100	-	P_{tac}
pZMR112	tnsA	P_{tac} and P_1
pZMR117	tnsB	P_{tac} and P_2
pZMR113	tnsC	P_{tac}
pZMR118	tnsD	P_{tac}
pZNE77	tnsE	-

Table 3.3 Effects of *trans* Acting Functions on P_1 Activity

Plasmids[4]	galK/ml (X 10^2)	b-lact/ml (X 10^2)	galK(S.A)[1]	b-lact(S.A)[2]	galK[3] / b-lact (X 10^2)
pZNE65-pZMR100$^-$	243	292	860	1037	83
pZNE65-pZMR100$^+$	180	169	811	761	107
pZNE65-pZMR112$^-$	234	198	852	720	118
pZNE65-pZMR112$^+$	220	153	844	585	144
pZNE65-pZMR117$^-$	107	207	386	747	52
pZNE65-pZMR117$^+$	20	194	72	693	10
pZNE65-pZMR113$^-$	172	109	832	528	158
pZNE65-pZMR113$^+$	212	174	877	721	122
pZNE65-pZMR118$^-$	253	195	859	664	129
pZNE65-pZMR118$^+$	215	165	807	619	130
pZNE65-pACYC184	103	437	408	1726	24
PZNE65-pZNE77	83	331	390	1560	25

1. GalK specific activity (S.A) units are expressed as nanomoles of galactose phosphorylated per min per ml of cells at $O.D_{650}$ of 1.0, as in Materials and Methods. GalK/ml units are derived from galK specific activity units by multiplying by the $O.D_{650}$ of the culture.

2. b-lactamase (b-lact) specific activity (S.A) units are expressed as nanomoles of nitrocefin hydrolyzed per min per ml of cells at $O.D_{650}$ of 1.0, as in Materials and Methods. b-lact/ml units are derived from b-lact specific activity units by multiplying by the $O.D_{650}$ of the culture.

3. Copy number corrected galK activity gives identical results whether calculated from specific activity units or units/ml.

4. All plasmids were assayed in strain DS941 (galK$^-$ lacI^q, Table 2.1). Each assay was repeated at least three times.
+/- indicates the presence or absence of IPTG during the growing of the cultures.

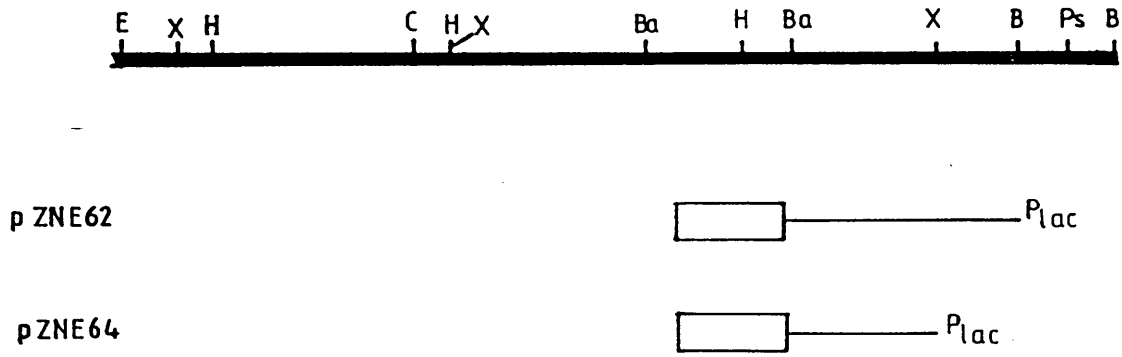


Figure 3.6 Illustration of the Tn7 sequences fused to galK in pKL500. The right 9kb of Tn7 is presented as a thick solid line. The thin solid line represents the region of Tn7 present in the plasmids named alongside. Open boxes represent the galK gene. The wild-type lac promoter (P_{lac}) is present on the vector to direct transcription of galK.

Abbreviations: B, BglII; Ba, BamHI; C, ClaI; E, EcoRI; H, HindIII; P, PstI; X, XbaI

Table 3.5 Detection of Transcriptional Termination Within Tn7

Plasmid[4]	galK/ml (X 10^2)	b-lact/ml (X 10^2)	galK(S.A)[1]	b-lact(S.A)[2]	galK[3] / b-lact (X 10^2)
pZNE62	156	223	421	601	70
pZNE64	160	230	479	690	69
pKL500	109	148	476	648	73

1. GalK specific activity (S.A) units are expressed as nanomoles of galactose phosphorylated per min per ml of cells at $O.D_{650}$ of 1.0, as in Materials and Methods. GalK/ml units are derived from galK specific activity units by multiplying by the $O.D_{650}$ of the culture.

2. b-lactamase (b-lact) specific activity units (S.A) are expressed as nanomoles of nitrocefin hydrolyzed per min per ml of cells at $O.D_{650}$ of 1.0, as described in Materials and Methods. b-lact/ml units are derived from b-lact specific activity units by multiplying by the $O.D_{650}$ of the culture.

3. Copy number corrected galK activity gives identical results whether calculated from specific activity units or units/ml.

4. All plasmids were assayed in strain DS941 (galK$^-$, Table 2.1). Each assay was repeated at least three times.

attributed to the presence of a weak terminator in the 2.05kb BglII (13.1kb) - BamHI (11.05kb) region. If present, this terminator would partially terminate P_1 transcripts and would therefore reside upstream of P_2, in the 0.7kb BglII (13.1kb) - XbaI (12.4kb) fragment. To investigate this possibility, the 1.35kb XbaI-BamHI fragment was cloned into the terminator probe vector pKL500 (Figure 3.6 and Table 3.5).

Results presented in Table 3.5 suggest that there is no significant reduction in the level of galK expression in pZNE64 compared to pZNE62. This together with the observed 5-fold reduction in the level of P_1 transcription in the presence of the the tnsB gene product in trans with P_1 (section 3.2.5) point out to the fact that binding of the tnsB protein at P_1 represses its activity.

3.3 Translational Fusion Vectors

Translational fusion vectors provide a means to identify, isolate and characterise transcriptional and translational control signals by fusing them to the 5' end of a gene encoding an easily assayable and/or selectable function, such as b-galactosidase (lacZ; Bassford et al, 1978). In the parent vector this gene is not expressed since no promoter, ribosome binding site and translation initiation codon exist to programme its messenger synthesis. Cloning of a fragment containing a promoter, ribosome binding site and translation initiation codon in the correct orientation and reading frame relative to the indicator gene turns on its expression. These fusion vehicles are also useful for obtaining a qualitative estimate of a given protein's level of expression. Because cloning of DNA segments carrying different transcriptional regulatory signals can result in the alteration of the vector's copy number, any quantitative measurements of activity require the control of copy number variations.

3.3.1 LacZ Gene Fusions

Since the amino-terminal end of the b-galactosidase protein is not essential for enzymatic activity (Muller-Hill and Kania, 1974), translational lacZ fusions that replace the corresponding 5' coding segment of the lacZ gene with other DNA sequences, but which still have functional b-galactosidase activity, can also be constructed (Bassford et al, 1978; Heidecker and Muller-Hill, 1977; Muller-Hill and Kania, 1974). At least the first 27 amino acid codons of lacZ can be removed, and the remainder of the gene can be joined to a DNA segment that encodes the regulatory signals and amino-terminal region of another gene, yielding a hybrid protein that retains the enzymatically active b-galactosidase remnant and has the same specific activity as the intact enzyme (Sarthy et al, 1979). The length of the amino-terminal segment fused to lacZ can vary from as few as six amino acid codons (Chou et al, 1979a; Chou et al, 1979b; Cohen et al, 1979) to a segment almost as large as an entire gene (Muller-Hill and Kania, 1974).

3.3.2 The translational Fusion Vectors pNM480, pNM481, and pNM482 (Minton, 1984)

pNM480, pNM481, and pNM482 are pBR322 derivatives in which a part of the lac operon replaces the Tc^r determinant, but leaves the Ap^r gene intact. The segment of lacZ cloned into these vectors is deleted for the first 8 amino acid codons and is missing the sites necessary for transcription and translation initiation. Therefore, all information needed to specify b-galactosidase is contained on the vector. Any DNA insert that correctly aligns translational and transcriptional start signals with lacZ will yield a hybrid gene that confers a $lacZ^+$ phenotype. The cloning sites in the three vectors differ with respect to their translational phasing relative to the 8th codon of the lacZ gene, allowing the cloning of restriction fragments to produce fusions of the proteins of interest in all three reading frames relative to b-galactosidase (Figure 3.7).

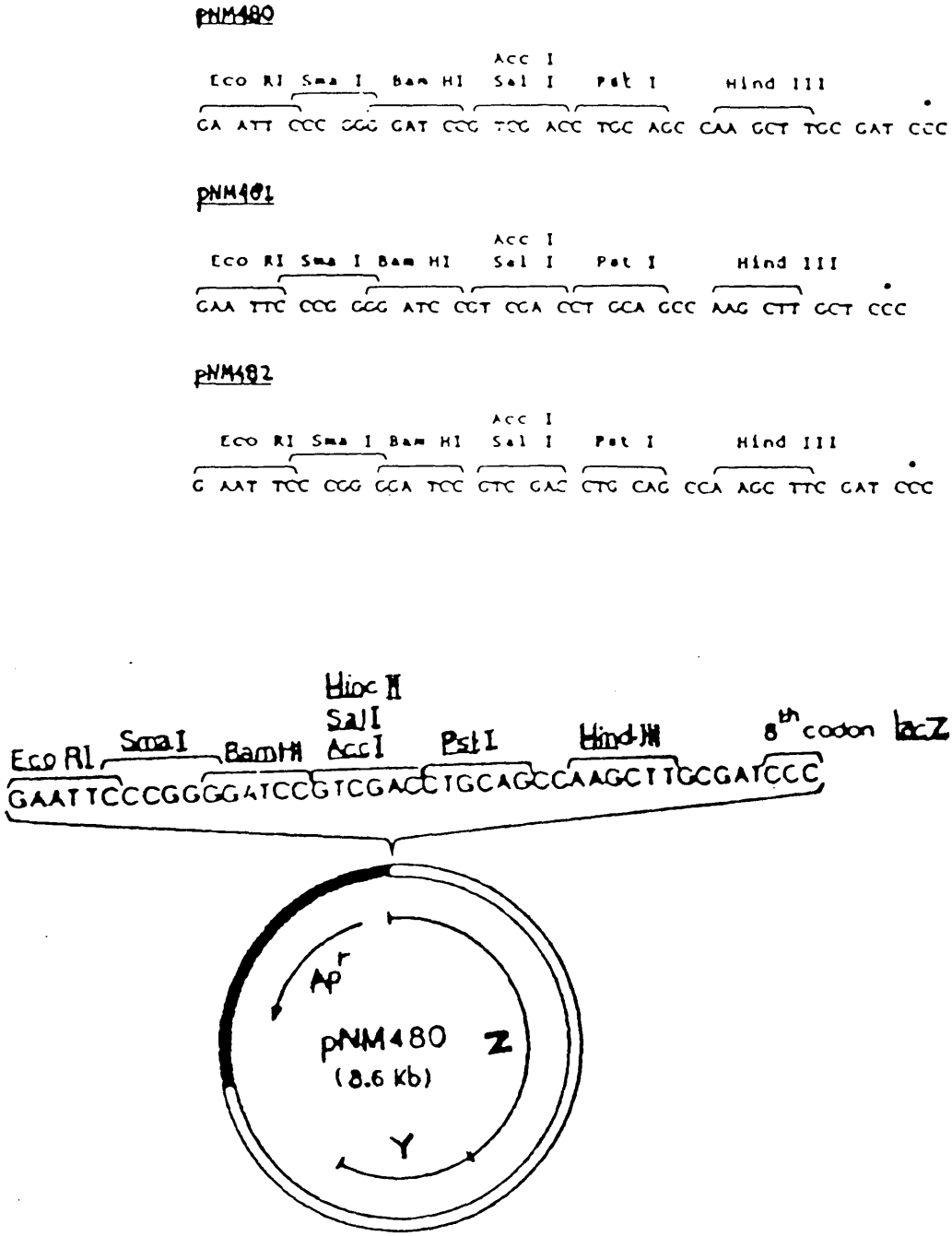


Figure 3.7 The translational fusion vector pNM480 and the translational phasing of the restriction sites of the three vectors (pNM480, pNM481, and pNM482) relative to the 8th codon of lacZ (indicated by an asterisk).
Minton, 1984

3.3.3 Selection for and against LacZ

A blue/white test was used for the selection of recombinants that expressed b-galactosidase on X-gal plates. On agar containing X-gal, lacZ$^+$ bacteria form blue colonies. The translational fusion vectors, described in section 3.2.1, are white on X-gal (lacZ$^-$), in a lacZ$^-$ host, and hence clones harboring transcriptional and translational start signals, were selected on their ability to form blue colonies on X-gal containing medium.

3.3.4 Detection of Translational Start Signals in Tn7

Section 3.2.4 described the study of the transcriptional levels at different sites within the 7.9kb transposition region of Tn7. To be able to correlate the results derived from this transcriptional analysis with the distribution of proteins in the 7.9kb right part of the transposon, a search for translated reading frames within this region was undertaken. The same series of nested fragments fused to galK in pKO500 (Figure 3.3) was fused to lacZ in the translational fusion vectors described in section 3.3.2, with the intent of selecting recombinants that expressed b-galactosidase. To ensure cloning of each restriction fragment in the proper reading frame relative to lacZ, all three pNM vectors were employed for the cloning. The blue/white test (section 3.3.3) was applied for the screening of clones that produced b-galactosidase, the activity of which was measured by lacZ assays (see Materials and Methods). Because genetic tests implied the existence of one or more functions in the EcoRV (9.8kb)-ClaI (7.6kb) region, the 4.4kb RE-ClaI fragment was also cloned into the three vectors. The paragraphs that follow analyse the results obtained from the translational fusion study (Table 3.6, Figure 3.8). Interpretation of these data was based on two assumptions: a) The specific activity of the lacZ hybrid protein equals the lacZ specific activity (Sarthy et al, 1979). b) The level of lacZ expression from the fusion protein is irrespective of the length of the foreign protein fused to lacZ.

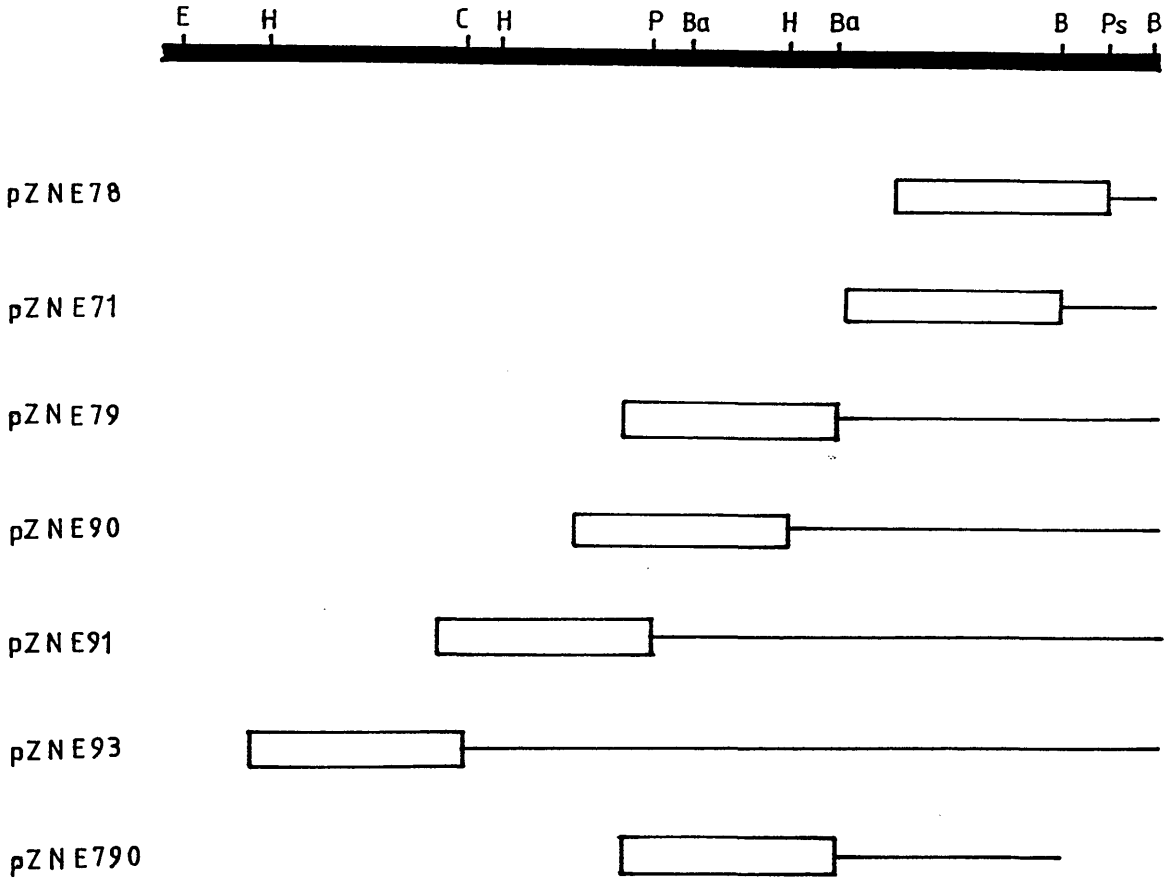


Figure 3.8 Illustration of the Tn7 sequences fused to *lacZ* in pNM480, pNM481, and pNM482. The right 9kb of Tn7 is presented as a thick solid line. The thin solid line represents the region of Tn7 present in the plasmid named alongside. Open boxes represent the *lacZ* gene.

Abbreviations: Ba, *Bam*HI; B, *Bgl*II; C, *Cla*I; E, *Eco*RI; P, *Pvu*II; Ps, *Pst*I

Table 3.7 Properties of Plasmids Used in Table 3.6

Plasmid	Parent Plasmid	Tn7 protein fused to *lacZ*	Tn7 promoter
pZNE78	pNM481	*tnsA*	P_1
pZNE71	pNM482	*tnsA*	P_1
pZNE79	pNM480	*tnsB*	P_1 and P_2
pZNE790	pNM480	*tnsB*	P_2
pZNE90	pNM482	*tnsC*	P_1 and P_2
pZNE91	pNM480	*tnsD*	P_1 and P_2
pZNE93	pNM482	*tnsE*	P_1 and P_2

Table 3.6 Detection of Translated Reading Frames in Tn7

Plasmid[4]	lacZ/ml	b-lact/ml	lacZ(S.A)[1]	b-lact(S.A)[2]	lacZ[3] ------(X 10^3) b-lact
pNM480	0	1332	0	3600	0
pZNE78	7	342	23	1200	19
pZNE71	12	544	20	900	22
pZNE79	82	851	310	3200	97
pZNE790	23	634	173	4800	36
pZNE90	4	344	26	2100	12
pZNE91	3	837	17	5300	3
pZNE93	6	586	33	3200	10

1. LacZ specific activity (S.A) units are expressed as nanomoles of o-nitrophenol per min per ml of cells at $O.D_{600}$ of 1.0, as in Materials and Methods (Miller, 1972).

2. b-lactamase (b-lact) specific activity units (S.A) are expressed as nanomoles of nitrocefin hydrolyzed per min per ml of cells at $O.D_{650}$ of 1.0, as in Materials and Methods. b-lact/ml units are derived from b-lact specific activity units by multiplying by the $O.D_{650}$ of the culture.

3. Copy number corrected lacZ activity gives identical results whether calculated from specific activity units or units/ml.

4. All plasmids were assayed in strain DS947 (dellac, Table 2.1). Each assay was repeated at least three times.

The level of translation at the PstI site (13.5kb) is similar to the translational level at the BglII site (13.1kb). This could either mean that there is one protein crossing both these sites or that there are two proteins, one of which crosses the PstI site but ends before the BglII site and the other beginning after the PstI site and crossing the BglII site. Complementation data exclude the second possibility. The Tn7 protein fused to lacZ in pZNE78 and pZNE71 was named tnsA.

The difference between the levels of translation at the BamHI (11.05kb), HindIII (10.6kb), PvuII (9.3kb) and ClaI (7.6kb) sites suggests that a different protein is crossing each of these sites. These proteins were named tnsB, tnsC, tnsD, and tnsE, respectively, and their presence was supported by the final evidence obtained from the complementation analysis (Figure 1.5).

3.4 Conclusions

This chapter described the use of transcriptional and translational fusion vectors in the study of the expression and activity of Tn7 transposition proteins. The approach presented here, revealed the existence of two promoters, P_1 and P_2, and five proteins tnsA, tnsB, tnsC, tnsD, and tnsE in the 7.9kb right part of Tn7. P_1 has also been identified by S1 mapping (Gay et al, 1986) , and appears to be the major Tn7 promoter initiating transcription of all five genes. P_2 seems to partially transcribe tnsB. Provision of the tnsB gene product in trans with P_1 followed by no complete open reading frame results in a 5-fold repression of transcription by this promoter. This repression of transcription could be explained by the binding of the tnsB gene product to the direct repeat carrying P_1. Binding of transposition proteins to the ends of transposons is one of the prime events in the transposition reaction. This possibility was examined in detail in chapter 4.

Results presented in this chapter showed that all but one (tnsB) of the Tn7 transposition proteins are expressed in very low

amounts. This is not surprising considering the detrimental effects of the elevated transposition frequencies on the host, an immediate consequence of an increase in the levels of transposases (Casadaban et al, 1982).

From the levels of transcription and translation of the five genes detected in this study, it is possible to draw conclusions concerning the translational efficiency of each protein. It is known that a single copy of a strong promoter, say, P_{lac}, makes approximately 30 transcripts per cell per generation (Reznikoff and McClure, 1986). Considering that P_1 is 0.5 times the strength of P_{lac}, it appears that this promoter would make approximately 15 transcripts per cell per generation if it were present on a single copy in the chromosome. It is also known that a wild-type *lacZ*$^+$ strain has about 1000 units of b-galactosidase per unit absorbance of culture (Miller, 1972). This corresponds to approximately 12,000 molecules of *lacZ* monomer per cell (Watson, 1975). Unpublished data from our laboratory demonstrate that a plasmid with a copy number similar to that of the pNM vectors, carrying the wild-type *lacZ* gene under the control of P_{lac}, has approximately 10,000 units of b-galactosidase per unit absorbance of culture, corresponding to 120,000 molecules of *lacZ* monomer per cell (D. Sherratt, pers. comm.). Taken together the latter three figures suggest that the gene product (or products) made by P_1 would be present in 60,000 copies per cell, if it were translated at the same rate as b-galactosidase. Results in Table 3.6, however, show that there are only 240 molecules of the *tnsA* gene product per cell (20 units of b-galactosidase per unit absorbance of culture), indicating that *tnsA* is translated 250 times less efficiently than *lacZ*. Doing similar calculations it is possible to estimate the translational efficiency of all five Tn7 gene products as compared to b-galactosidase.

The b-galactosidase fusion proteins constructed in this study can be used by future workers to isolate biologically active Tn7 transposition proteins. Hybrid proteins exhibiting b-galactosidase activity can be purified using a one-step method

based on affinity chromatography (Ulmann, 1984). Antisera raised against the hybrid molecule will cross react with sequences derived from the gene to which *lac* is fused, thus providing a means to identify and purify the target gene product (Shuman *et al*, 1980).

CHAPTER 4

DNA BINDING PROPERTIES OF THE TN7 *tnsB* AND *tnsD* GENE PRODUCTS

4.1 INTRODUCTION

The termini of Tn7 consist of 28bp nearly perfect inverted repeats. Immediately inside the right end are four copies of a directly repeated consensus consisting of the internal 22bp of the repeat. Inside the left terminal 28bp there are three repeated copies of the same consensus (yet inverted with respect to the sequence in the right terminus), though in this case the repeats are interspersed with other sequences (Lichtenstein and Brenner, 1982; Figure 1.3). It is almost certain that one or several of the Tn7 transposition proteins (independently or in concert) bind to these repeats (independently or to the protein-DNA complex) and such binding is required for initiation of transposition.

Studies with transcriptional fusions (chapter 3) have shown that provision of the tnsB gene product in trans with the 500bp right end-PstI fragment of Tn7 (repeats and P_1 promoter intact yet lacking any complete open reading frame) results in a considerable reduction in P_1 promoter activity (as measured by the levels of galactokinase expression). These data provide a strong indication for the binding of the tnsB gene product in the region which contains P_1. Figure 4.1 shows that the -35 region of P_1 is contained within the innermost direct repeat in the right end of Tn7.

The tnsD gene product is a good candidate for binding to the "hot site". It is only required for transposition to "hot sites" and is likely to mediate the recognition and integration of the transposon at that site.

This chapter presents an in vitro approach that has been taken to examine the binding of the tnsB and tnsD gene products to the right end of Tn7 and the "hot site", respectively. It is based on a gel electrophoresis technique for detecting complexes of DNA with purified DNA-binding proteins (Fried and Crothers, 1981; Garner and Revzin, 1981).

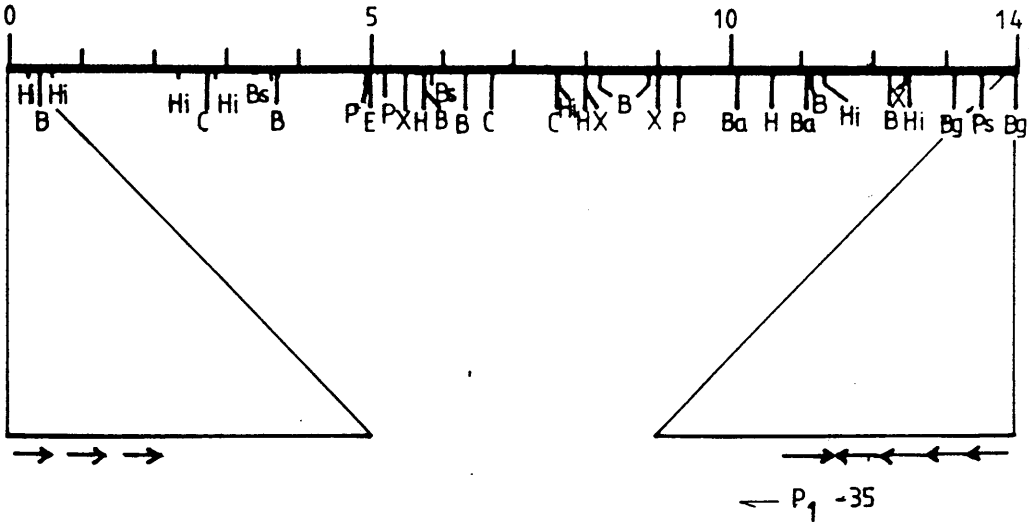


Figure 4.1 The -35 region of the major Tn7 promoter, P_1, is contained within the innermost direct repeat in the right end of the element.

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

4.2 The Use of Gel Electrophoresis to Detect and Study Protein-DNA Interactions

Proteins can bind at specific sites in DNA and thereby mediate either catalytic or regulatory effects. Multiple protein binding sites exist and are necessary to obtain full repression in the gal operon (Fritz et al, 1983), the araBAD operon (Dunn et al, 1984), the nrd operon (Tuggle and Fuchs, 1986), and the deo operon (Dandanell and Hammer, 1985).

It has been known for several years that tightly bound protein-nucleic acid complexes are stable during electrophoresis under non-denaturing conditions, as found, for example, in RNA polymerase-DNA complexes. Intensive application of non-denaturing gel electrophoresis to regulatory protein-DNA complexes began in 1981 with work by Garner and Revzin, (1981) on the E.coli CAP protein and by Fried and Crothers, (1981) on the lac repressor. Such work has been based on the technique of low-ionic strength gel electrophoresis developed originally for the analysis of oligo- and mononucleosomes (Varshavsky et al, 1976) and has demonstrated that the DNA complexes could be detected and characterised quantitatively by the gel method.

The gel method has several advantages over the standard filter binding technique for detecting protein-DNA complexes. The nitrocellulose filter assay (Jones and Berg, 1966) depends on the fact that proteins and protein-DNA complexes are retained on the filter, and therefore leads to a direct measure of the complexes without distinguishing between different complexes which may coexist in the solution. It is also possible that DNA-protein interactions can be perturbed when the protein binds to the filter. These serious limitations are eliminated by the use of native gels. In the gel electrophoresis method the products of the protein-DNA interaction are separated into discrete species that migrate with different mobilities, rather than grouped together in the "filter bound" category. The low-salt conditions in the gel, and a potential "cage effect" created by the gel matrix, stabilise even weak complexes so that they can be

quantified, enabling determination of binding equilibria and kinetics, and thermodynamic parameters of the reaction (Fried and Crothers, 1981; Fried and Crothers, 1984; Garner and Revzin, 1981). Recent work by Liu-Johnson et al, (1986) has shown that it is possible to dissect the binding affinity into contributions by individual base pairs in the binding site.

Work by Wu and Crothers, (1984) has demonstrated that the location of a bend in a DNA sequence can be found by comparing the electrophoretic mobilities of a series of circularly permutated variants of the same sequence. Isomers with the bend near the end have a much larger end-to-end distance, and hence greater gel mobility, than that observed for molecules in which the bend is near the middle of the chain. In their study the above workers have shown that CAP-lac promoter DNA complexes migrate at different rates depending on whether the CAP binding site is in the center or at the end of the DNA fragment; in contrast, lac repressor-operator complexes move at the same rate no matter where the operator is located on the fragment. The authors concluded that the binding of CAP causes DNA to bend.

Craigie et al, (1984) have used the non-denaturing polyacrylamide gel electrophoresis assay to determine binding sites for MuA protein on Mu DNA using purified protein and phage DNA digested with restriction enzymes; MuA binding fragments were determined after dissociation of the complex bands followed by electrophoresis in a second dimension. Recently, Kramer et al, (1987) have used the same technique to demonstrate that one lac repressor tetramer may bind to two lac operators on one DNA fragment, thus causing the intervening DNA to form a loop. Exploiting the potential of the method, the same workers have shown that by varying the concentration of the components of the reaction different configurations of the intervening DNA are observed.

The gel binding technique is now widely used to separate "free" and "bound" DNA in "interference" experiments which reveal sites where protein and DNA are in close proximity (Hendrickson and

Schleif, 1985). Perhaps, the most wide-spread current use of the technique is in the assay of complex nuclear protein mixtures for their ability to produce a discrete protein-DNA gel band due to binding to a specific DNA sequence (Carthew et al, 1985).

4.3 The Band-Competition Assay as a General Method for Detecting Specific DNA-Binding Proteins in Crude Extracts

Although many of the current assays for DNA-binding proteins can be used to detect specific DNA-protein interactions in crude extracts, these assays are often not sufficiently sensitive or selective, particularly for DNA-binding proteins with a moderate degree of nucleotide sequence specificity. The band-competition assay as a method for detecting specific DNA-binding proteins in crude extracts (Strauss and Varshavsky, 1984) is based upon an electrophoretic method for detecting complexes of DNA with purified DNA-binding proteins (Garner and Revzin, 1981; Fried and Crothers, 1981). The essence of the method is simply to mix a DNA fragment containing a presumptive binding site, with a crude extract containing the protein of interest, including any necessary divalent ions and cofactors. After completion of the binding reaction, the sample is applied to a low ionic strength polyacrylamide gel and electrophoresed for the shortest period of time which gives resolution of components. The resulting electrophoretic pattern will display a diminution or disappearance of the DNA band and possibly the appearance of a lower mobility band containing the DNA-protein complex (Garner and Revzin, 1986). To reduce the effect of proteins that bind non-specifically to DNA, an amount of "carrier" DNA is added to the reaction mixture. Care must be taken in the choice of "carrier" DNA to avoid introducing adventitious sites with high affinity for the protein under study.

Carthew et al, (1985) have employed the "band-shift" assay using Hela cell crude extacts to show that an RNA polymerase II transcription factor binds specifically to an upstream element in the adenovirus major late promoter. Tolias and Dubow, (1985,

1986) have employed the same technique using ner-containing crude protein extracts, to demonstrate specific binding of the bacteriophage Mu and D108 regulatory DNA-binding proteins ner to DNA restriction fragments containing ner binding sites. Zerbib et al, (in press) have shown that cell extracts from InsA overproducers (presumptive IS1 transposase) display a DNA binding activity specific for the ends of IS1.

RESULTS AND DISCUSSION

4.4 Application of the Band-Competition Assay Using Whole-Cell Crude Extracts to the Tn3 resolvase/res System

At the time it was decided to employ the band-shift assay using whole-cell crude extracts from strains overproducing Tn7 proteins, there was no evidence in the literature reporting the application of this method to the study of E.coli systems. It was thought, therefore, to undertake a preliminary "trial" experiment that would test whether it was possible to apply this method to E.coli crude extacts. Martin Boocock, in our group in Glasgow, was succesfully employing the band-competition assay to study the binding of Tn3 resolvase (tnpR gene product) at the res site, using purified tnpR protein.

The binding profile of the tnpR gene product at its binding site res was examined using crude extracts from a strain overproducing Tn3 resolvase, in the hope of obtaining an acceptable degree of correlation with the binding profile obtained in the in vitro experiment using purified protein (M. Boocock, pers. comm.).

pMA6114/JM101 (Tables 2.1 and 2.2) was the IPTG inducible system overproducing Tn3 resolvase, used to compare the binding pattern of the tnpR gene product at res with the one obtained using purified resolvase. Figure 4.2 describes in detail the experiment performed and shows that it is possible to apply the band-competition assay to E.coli systems.

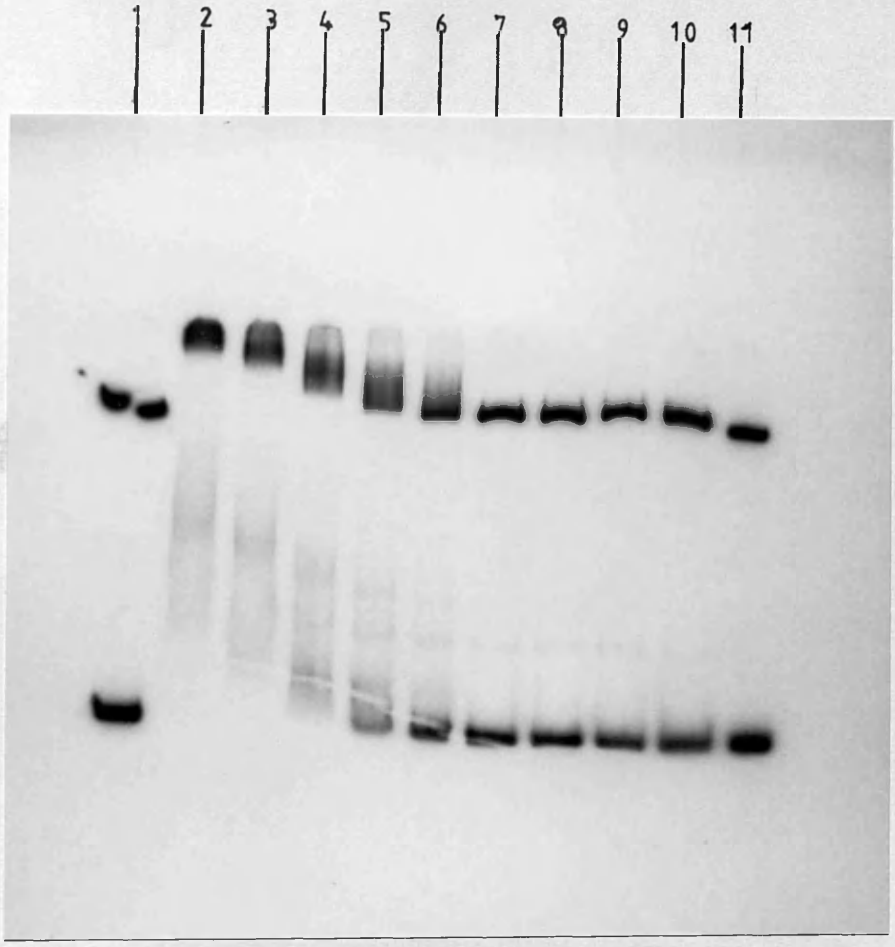


Figure 4.2 Application of the band-competition assay using whole-cell crude extracts to the res/resolvase system

pMA1441 was cleaved with EcoRI. This restriction excises a 238bp fragment containing the Tn3 res site from the vector DNA (pUC8; 2.7kb). The whole digest was end-labelled and used in the binding assay. 1ng of this labelled DNA was mixed under the conditions of the assay described in Materials and Methods with increasing amounts of MA6114/JM101 crude extract (lanes 2-6) and purified resolvase (lanes 7-11). 2ug of carrier DNA (supercoiled pUC8) were included in each assay.

1 no extract
2 5.66 ug pMA6114/JM101 extract
3 2.83 ug pMA6114/JM101 extract
4 707.1 ng pMA6114/JM101 extract
5 353.6 ng pMA6114/JM101 extract
6 176.8 ng pMA6114/JM101 extract
7 39.2 ng purified resolvase
8 19.6 ng purified resolvase
9 10.2 ng purified resolvase
10 5.0 ng purified resolvase
11 2.5 ng purified resolvase

pMA6114/JM101 extract = tnpR extract

4.5 Use of the Band-Competition Assay to Study the Binding Profile of the Tn7 tnsB Gene Product

In order to characterise the DNA-binding properties of the tnsB gene product, the IPTG inducible pZMR117/DS941 system was used. pZMR117 (Figure 4.3, Tables 2.1 and 2.2) is a lambda dv plasmid carrying the tnsB gene (2.05kb fragment, from BglII at 13.1kb to BamHI at 11.05kb) cloned under the control of P_{tac} and is capable of fully complementing transposition of tnsB$^-$ Tn7 deletion mutants (Rogers et al, 1986). Sequencing information reveals that the tnsB gene product in this construct is lacking seven amino acids from its N-terminus (C. Lichtenstein, pers. comm.). This system allowed the tnsB protein to be expressed in sufficient quantities, as observed in minicells (Figure 4.4, Rogers, 1986), but did not allow any detectable amounts of tnsB to be seen in whole cell extracts.

pZMR100/DS941 (Table 2.2) was the tnsB-free system that acted as a negative control in experiments assessing the binding activity of the tnsB protein. pZMR100 is the parent plasmid of pZMR117 and contains no insert cloned in the polylinker region that follows P_{tac}.

4.5.1 Source of Tn7 Right End DNA

Plasmid pZNE200 (Figure 4.5, Table 2.2) was used as a source of Tn7 right end DNA in the tnsB-right end binding assays. This construct carries a 205bp fragment from the right end of Tn7 (BalI site at 13.85kb to end of the element) cloned in the pUC8 polylinker, and was made by deleting the SmaI-BalI region from pZMR88 (Table 2.2). Restriction of pZNE200 with EcoRI and HindIII excises a 239bp insert containing the 205bp fragment of Tn7 DNA flanked by 34bp of foreign DNA sequences (Figure 4.5).

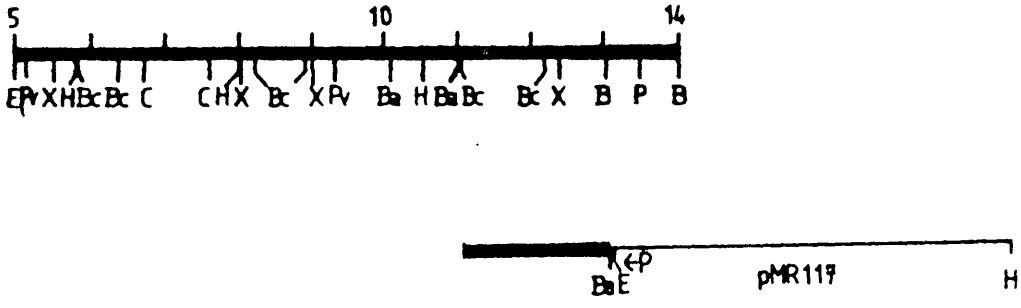


Figure 4.3 Representation of the plasmid pZMR117. The 2.05kb BglII (13.1kb)-BamHI (11.05kb) Tn7 fragment, containing most of the tnsB coding region (see section 4.5 for details) is transcribed from the tac promoter. Thick lines represent regions of Tn7. The thin line represents pZMR100, the vector into which the Tn7 sequences are cloned.
Rogers et al, 1986

Abbreviations: B, BglII; Ba, BamHI; Bc, BclI; C, ClaI; E, EcoRI; H, HindIII; P, PstI; Pv, PvuII; X, XbaI; P ,tac promoter.

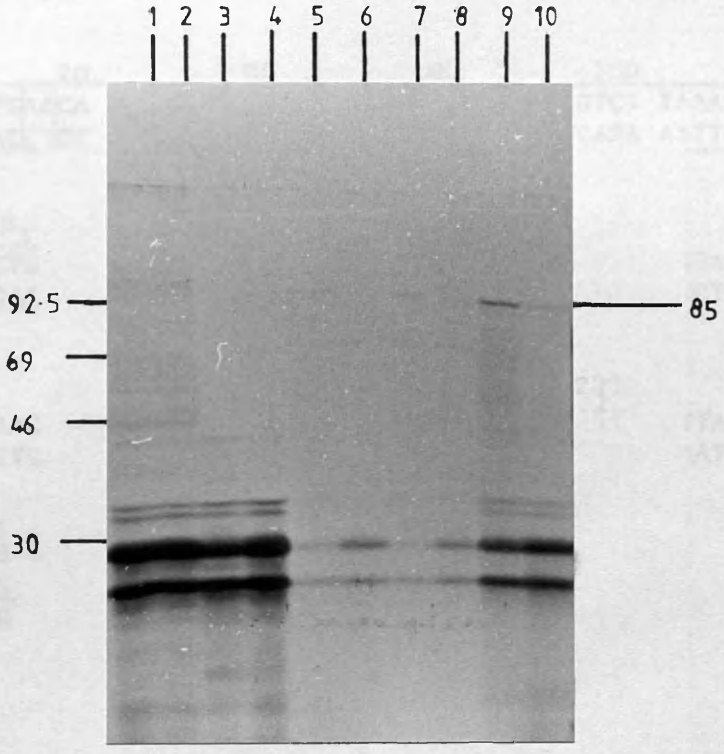


Figure 4.4 Polypeptides produced by pZMR117 in E.coli minicells

1 pZMR100 (+)
2 pZMR100 (-)
3-8 not relevant to this work
9 pZMR117 (+)
10 pZMR117 (-)

pZMR100 is the parent of pZMR117 (negative control)
+/- in parentheses indicates the presence or absence of IPTG during the labelling of the minicells.
The size of bands is given in kD.

The photograph was a gift kindly offered by Mark Rogers.

```
HindIII  10          20          30          40          50          60
AAGCTTGGCT GCAGGTCGAT GAGTTCCCGG ACTTCTTGTG TGGGCGGACA ATAAAGTCTT
TTCGAACCGA CGTCCAGCTA CTCAAGGGCC TGAAGAACAC ACCCGCCTGT TATTTCAGAA

         70          80          90          100         110         120
AAACTGAACA AAATAGATCT AAACTATGAC AATAAAGTCT TAAACTAGAC AGAATAGTTG
TTTGACTTGT TTTATCTAGA TTTGATACTG TTATTTCAGA ATTTGATCTG TCTTATCAAC

- 35 P1 130          140  -10 P1 150          160         170         180
TAAACTGAAA TCAGTCCAGT TATGCTGTGA AAAAGCATAC TGGACTTTTG TTATGGCTAA
ATTTGACTTT AGTCAGGTCA ATACGACACT TTTTCGTATG ACCTGAAAAC AATACCGATT

         190         200         210         220         230         240
AGCAAACTCT TCATTTTCTG AAGTGCAAAT TGCCCGTCGT ATTAAAGAGG GGCGTGGGGG
TCGTTTGAGA AGTAAAAGAC TTCACGTTTA ACGGGCAGCA TAATTTCTCC CCGCACCCCC

EcoRI
 245
AATTC
TTAAG
```

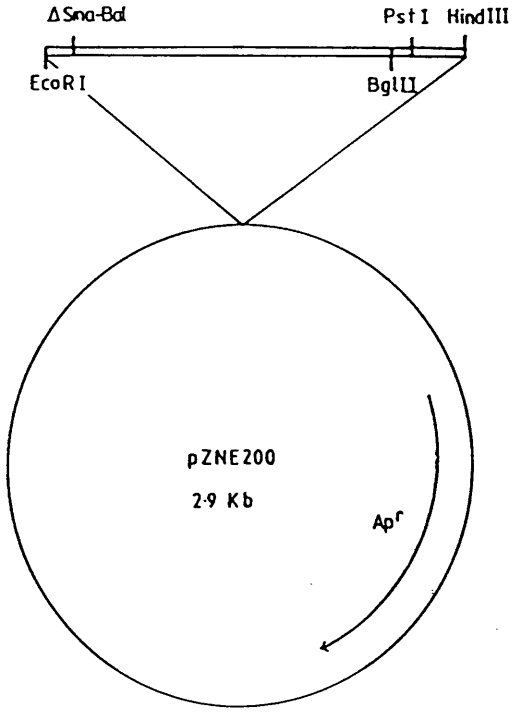


Figure 4.5 Representation of the plasmid pZNE200. The 239bp insert containing sequences from the right end of Tn7 is represented by the horizontal line, while the circle represents the vector (pUC8) sequences. The Ap^r gene and the positions of closely spaced restriction sites are not drawn to scale. The sequence of the 239bp fragment is also illustrated. The four direct repeats occuring in the right end of the element are enclosed within boxes. The position of P_1 is indicated in the innermost direct repeat in the right end of Tn7.

4.5.2 DNA-Binding Properties of the tnsB Gene Product

DNA-binding assays using tnsB (and tnsB-free) crude extracts and ^{32}P-end-labelled pZNE200 DNA were carried out as described in Materials and Methods. In order to examine the binding of tnsB to the right end of Tn7, pZNE200 was digested with EcoRI-HindIII and ^{32}P-end-labelled. This digestion gave rise to two restriction fragments: the 2.7kb pUC8 vector fragment and the 239bp fragment containing the presumptive tnsB binding sites. It was thought that the presence in the binding reaction of the 2.7kb fragment together with the 239bp fragment could act as an internal control for assessing the binding of proteins that bind non-specifically to DNA. Whole-cell crude extracts were made of pZMR117/DS941 (tnsB-containing extract) and pZMR100/DS941 (tnsB-free extract), as described in Materials and Methods. The protein concentration of the extracts was determined by the method of Bradford, (1976) as in Materials and Methods.

Figures 4.6 and 4.7 represent autoradiographs showing the retardation pattern obtained for binding of the tnsB gene product to the right end of Tn7 and are analysed in detail in the following paragraphs.

In the absence of extract, the end-labelled fragments migrated as discrete bands (Figure 4.6, lane 5; Figure 4.7, lane 5), while in the presence of extract most of the DNA failed to enter the gel (data not shown), the expected result of adding a vast excess of DNA-binding proteins to a small amount of naked DNA. A sufficiently large excess of "carrier" DNA (1-2ug of supercoiled or linear pUC8 or lambda DNA; the precise amounts and type of DNA used in each experiment are indicated in the legends of figures 4.6 and 4.7) was included in each assay to sequester proteins in the extract that would otherwise bind non-specifically to the labelled fragments. These conditions resulted in a substantial reduction of non-specific binding, though the observed retardation of the 2.7kb band (as seen in Figures 4.6 and 4.7) indicated that some of the proteins present in the extracts were

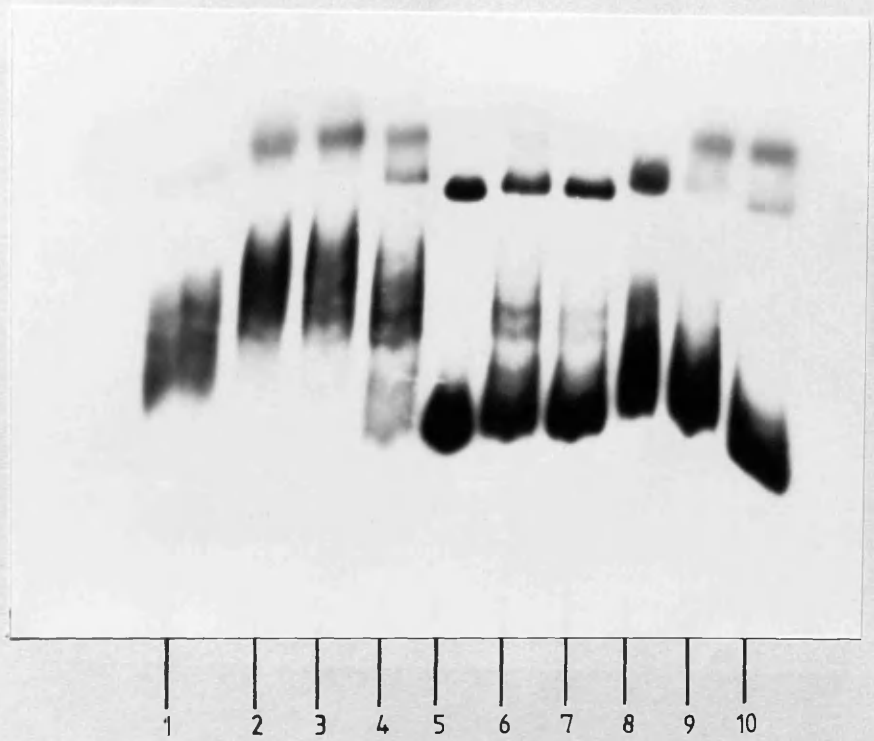


Figure 4.6 Detection of DNA-binding activity in tnsB crude extracts by the band-competition assay

pZNE200 was cleaved with EcoRI and HindIII and ^{32}P-labelled. 1ng of this end-labelled DNA was mixed under the conditions of the assay, (see Materials and Methods), with decreasing amounts of tnsB and tnsB-free crude extract. 2ug of carrier DNA were included in each binding reaction (supercoiled lambda DNA).

1 21 ug tnsB extract
2 10.5 ug tnsB extract
3 5.25 ug tnsB extract
4 2.63 ug tnsB extract
5 no extract
6 0.66 ug tnsB extract
7 0.33 ug tnsB extract
8 2.63 ug tnsB-free extract
9 0.66 ug tnsB-free extract
10 0.33 ug tnsB-free extract

tnsB extract = pZMR117/DS941 extract
tnsB-free extract = pZMR100/DS941 extract

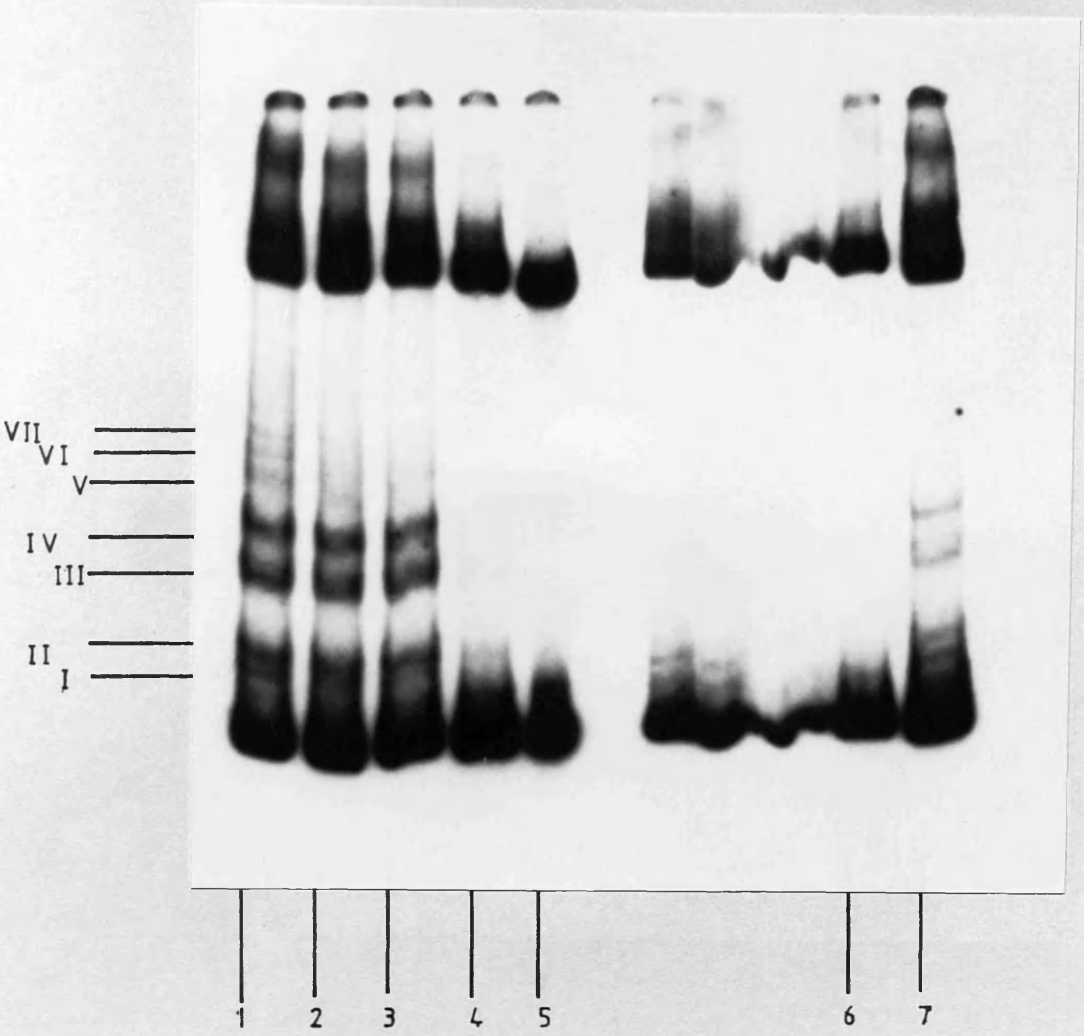


Figure 4.7 Detection of specific DNA-binding activity in tnsB crude extracts by the band-competition assay

pZNE200 was digested with EcoRI and HindIII and ^{32}P-end labelled. 1ng of this labelled DNA was mixed under the assay conditions with decreasing amounts of tnsB and tnsB-free crude extract. 2 ug of carrier DNA (EcoRI-HindIII cleaved lambda DNA) were included in each assay. Various amounts of unlabelled 239bp DNA fragment were added to some of the reactions to compete for tnsB binding. The 239bp EcoRI-HindIII fragment was gel purified as described in Materials and Methods.

1 1.32 ug tnsB extract
2 0.66 ug tnsB extract
3 1.32 ug tnsB extract + 40ng of linear unlabelled pUC8 DNA
4 1.32 ug tnsB extract + 4 ng of unlabelled 239bp fragment
5 no extract
6 1.32 ug tnsB extract + 10 ng of unlabelled 239bp fragment
7 1.32 ug tnsB extract + 2 ng of unlabelled 239bp fragment

tnsB extract = pZMR117/DS941

binding to this fragment.

When a constant amount (1ng; DNA concentration was measured as in Materials and Methods) of ^{32}P-labelled DNA fragments was incubated with an increasing amount of tnsB-containing whole-cell crude extract, in the presence of "carrier" DNA, seven distinct bands (I, II, III, IV, V, VI, VII) of lower mobility were observed (Figure 4.7, lane 1). The relative intensity of these bands and the number of complexes observed was a function of protein concentration (Figure 4.7, lanes 1-2; Figure 4.6, lanes 1-4 and 6-7). The exact protein concentrations used in each binding reaction are indicated in the legends of Figures 4.6 and 4.7.

Considering the amount of total labelled DNA used in each assay (1ng), and the ratio of Tn7 right end DNA/total DNA in pZNE200 (1/10), the amount of labelled DNA carrying the presumptive tnsB binding sites (239bp fragment) included in each reaction was estimated to be about 0.1ng. Assuming that the molecular weight of 1bp is 650D, the molecular weight of the 239bp fragment was calculated to be approximately 156kD. Taken together the last two figures, the number of moles and molecules of the 239bp fragment per assay were estimated as $6.4X10^{-16}$ and $3.9X10^{8}$, respectively.

The fact that all seven complexes were seen at a lambda DNA/right end DNA weight ratio of 20,000 (0.1ng of the 239bp fragment and 2ug of lambda DNA) suggested that the protein in the complexes had a considerably higher affinity for right end DNA than for lambda DNA.

When the same amount of ^{32}P-labelled DNA (1ng) was incubated with tnsB-free crude extracts (pZMR100/DS941), the 239bp fragment migrated with the mobility of naked DNA (Figure 4.6, lanes 8-10), at identical protein concentrations that gave the four to seven distinct bands in the case of the tnsB extract.

Figures 4.6 and 4.7 also show that the range of protein added to

the reaction which still showed specific DNA-binding activity extended from 2.63 ug to 330 ng of tnsB crude extract per reaction. The molecular weight of the tnsB gene product is approximately 85kD (Ouartsi et al, 1985; Rogers, 1986). On the assumption that tnsB represents 0.1% of the total protein in the extract, the moles and molecules of the tnsB gene product for the extract concentration that gave the seven discrete complexes (1.32 ug; Figure 4.7, lane 1), were estimated as $1.6X10^{-14}$ and $9.6X10^{9}$, respectively. Assuming that binding occurs at a 1:1 stoichiometry and that there are four binding sites for tnsB per DNA molecule (the four repeats in the right end of Tn7), it is estimated that tnsB was present in a 6-fold excess in the binding reaction. However, only 10-20% of the DNA appeared to be in a bound form, as estimated by looking at Figures 4.6 and 4.7. This is not an unreasonable situation, considering that a large proportion of the protein is expected to be inactive in the extract and that proteases which may exist in the extract are likely to inhibit the binding to a certain extent.

Proteinase K treatment of binding reactions containing tnsB protein extract and ^{32}P-labelled pZNE200 DNA abolished retardation and converted all the fragment in bands I, II, III, IV, V, VI, and VII to a form that comigrated with unbound DNA (Figure 4.8, lane 2). This strongly suggested that bands I, II, III, IV, V, VI, and VII represented discrete protein-DNA complexes.

4.5.3 Specificity of tnsB binding

To determine which complexes represented specific protein-DNA interactions and obtain an estimate of the relative affinity and specificity of tnsB for the right end, a series of competition experiments was performed. The 239bp EcoRI-HindIII fragment from pZNE200 was gel purified (see Materials and Methods) and various amounts of this unlabelled DNA were added to the extract together with a constant amount of end-labelled (1ng) and "carrier" (2ug) DNA. Also, a constant amount of tnsB extract (1.32ug) was used for the competition studies. Products were resolved by gel

Figure 4.8 Proteinase K treatment of the tnsB-DNA binding reaction

pZNE200 was cleaved with EcoRI and HindIII and ^{32}P-end labelled. 1ng of this labelled DNA was mixed under the conditions of the assay, (see Materials and Methods), with 1.32 ug of tnsB whole-cell crude extract. In addition to the carrier DNA (2 ug of supercoiled lambda DNA), proteinase K (1mg/ml) was included in the binding reaction.

1 no extract
2 1.32 ug tnsB extract + proteinase K

electrophoresis. When a small amount of unlabelled fragment (1-10 ng) was added to the reaction mixture, the intensity of bands I, II, III, IV, V, VI, and VII was dramatically decreased (Figure 4.7, lanes 4,6,7). An additional competition experiment was performed by incubating 1.32ug of tnsB extract with 1ng of end-labelled DNA and 40ng of unlabelled linear pUC8 DNA (negative control), in the presence of 2ug of "carrier" DNA (Figure 4.7, lane 3). Results from both of these experiments showed that while unlabelled right end fragment effectively competed with tnsB binding to the right end (Figure 4.7, lanes 4,6,7), unlabelled pUC8 DNA failed to compete for binding even at a competitor to labelled fragment ratio of 400:1 (Figure 4.7, lane 3).

Based on the number of tnsB molecules present in the binding reaction which gave the seven discrete complexes (9.6×10^9), and the number of cells contained in each assay (3×10^7; extracts were made on a 1lt culture of an $O.D_{650}$ of 0.8 which contained 3×10^8 cells/ml and were concentrated 100-fold before use), the abundance of active tnsB was estimated to be 320 molecules per cell. The actual abundance of tnsB in the cell is probably higher, since the binding assay does not detect tnsB molecules that bind the "carrier" DNA. In any case, tnsB appears to be a rare protein, present at a low concentration.

4.6 Source of "Hot Site" DNA

Plasmid pZMR80 (Figure 4.9, Table 2.2) was used as a source of "hot site" DNA. This plasmid contains a 290bp fragment carrying the reduced Tn7 "hot site" (Figure 4.9) cloned into the EcoRI-HindIII sites of the pUC8 polylinker.

4.6.1 DNA Binding Properties of the tnsD Gene Product

In order to study the DNA-binding properties of the tnsD gene product, the IPTG-inducible pZMR64/DS941 system was used. pZMR64 (Figure 4.10, Table 2.2) is a pMB1 based plasmid carrying the tnsD gene under the control of P_{tac} and is capable of fully

```
        10          20          30          40          50          60
GAATTCCCCG GATCAAAGGC ACCGACGTTG ACCAGCCGCG TAACCTGGCA AAATCGGTTA
CTTAAGGGGC CTAGTTTCCG TGGCTGCAAC TGGTCGGCGC ATTGGACCGT TTTAGCCAAT

        70          80          90          100         110         120
CGGTTGAGTA ATAAATGGAT GCCCTGCGTA AGCGGGGCAT TTTTCTTCCT GTTATGTTTT
GCCAACTCAT TATTTACCTA CGGGACGCAT TCGCCCCGTA AAAAGAAGGA CAATACAAAA

        130         140         150         160         170         180
TAATCAAACA TCCTGCCAAC TCCATGTGAC AAACCGTCAT CTTCGGCTAC TTTTTCTCTG
ATTAGTTTGT AGGACGGTTG AGGTACACTG TTTGGCAGTA GAAGCCGATG AAAAAGAGAC

        190         200         210         220         230         240
TCACAGAATG AAAATTTTTC TGTCATCTCT TCGTTATTAA TGTTTGTAAT TGACTGAATA
AGTGTCTTAC TTTTAAAAAG ACAGTAGAGA AGCAATAATT ACAAACATTA ACTGACTTAT

        250         260         270         280
TCAACGCTTA TTTGGGGATC CGTCGACCTG CAGCCAAGCT T
AGTTGCGAAT AAACCCCTAG GCAGCTGGAC GTCGGTTCGA A
```

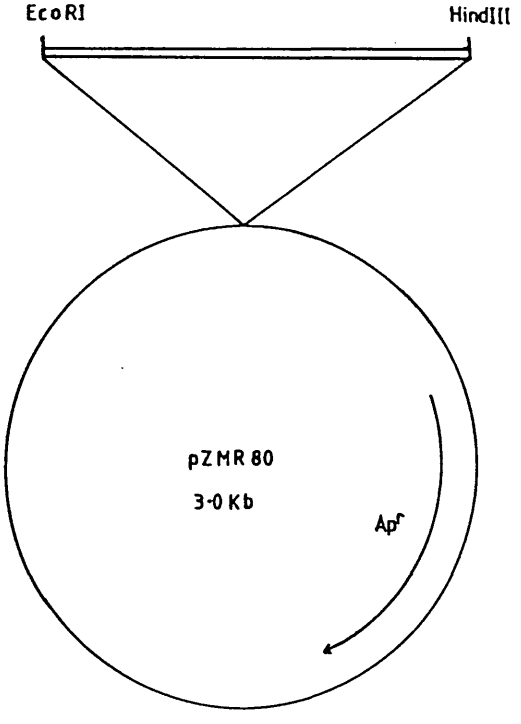


Figure 4.9 Representation of the plasmid pZMR80. The 290bp insert containing the Tn7 "hot site" is represented by the horizontal line. The circle represents the vector (pUC8) sequences. The Ap^r gene is not drawn to scale. The sequence of the 290bp insert is illustrated on the top of the page.

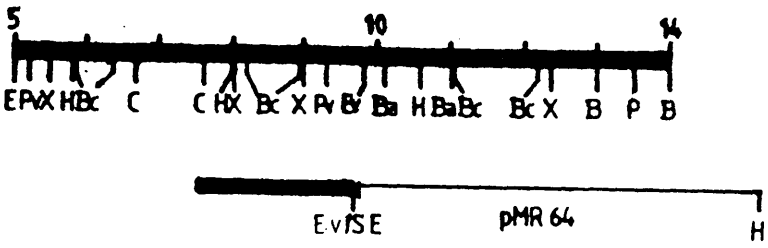


Figure 4.10 Representation of the plasmid pZMR64. The 2.2kb EcoRV (9.8kb)-ClaI (7.6kb) Tn7 fragment, containing the tnsD coding region, is transcribed from the tac promoter. Thick lines represent regions of Tn7. The thin line represents pGLW8, the vector into which the Tn7 sequences are cloned.
Rogers et al, 1986

Abbreviations: B, BglII; Ba, BamHI; Bc, BclI; C, ClaI; E, EcoRI; Ev, EcoRV; H, HindIII; P, PstI; Pv, PvuII; X, XbaI

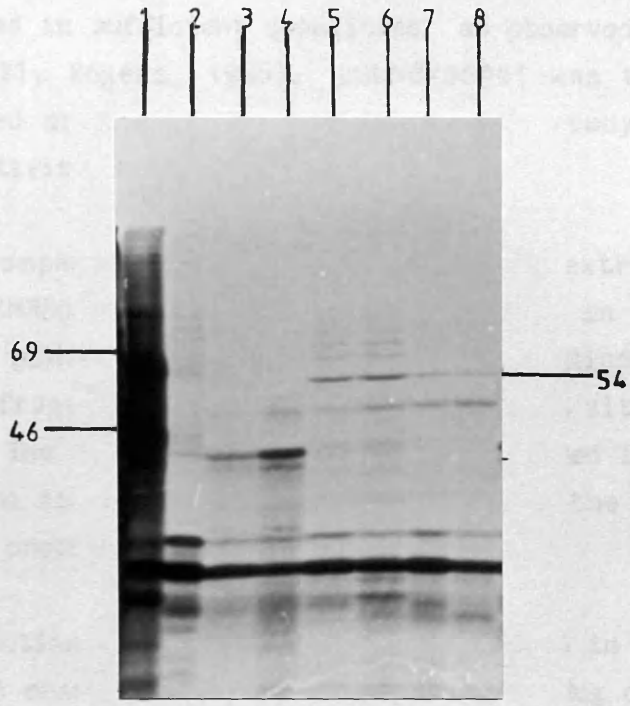


Figure 4.11 Polypeptides produced by pZMR64 in E.coli minicells

1 Molecular weight markers
2 pGLW8 (+)
3-6 not relevant to this work
7 pZMR64 (-)
8 pZMR64 (+)

pGLW8 is the parent of pZMR64 (negative control)
+/- in parentheses indicates the presence or absence of IPTG during the labelling of the minicells.
The size of bands is given in kD.

The photograph was a gift kindly offered by Mark Rogers.

complementing transposition of tnsD⁻ Tn7 deletion mutants (Rogers, et al, 1986). Recent sequencing data that was made available to us reveals that the entire tnsD gene product is encoded by the Tn7 fragment cloned in this plasmid (C. Lichtenstein, pers. comm.). This system allowed tnsD protein to be expressed in sufficient quantities, as observed in minicells (Figure 4.11, Rogers, 1986). pGLW8/DS941 was the tnsD-free system used as a negative control in the study of the tnsD binding activity.

The band-competition assay using tnsD crude extracts and ^{32}P-labelled pZMR80 DNA was employed as described in Materials and Methods. pZMR80 was digested with EcoRI-HindIII and both resulting fragments (2.7kb pUC8 and 290bp "hot site") were end labelled. The 2.7kb pUC8 fragment was included in the binding reaction to act as an internal control for the non-specific binding of proteins contained in the extract.

Binding reactions were carried out as described in Materials and Methods. A constant amount of "carrier" DNA (2ug of supercoiled pUC8 DNA) was included in each assay to reduce the effect of proteins that would bind non-specifically to DNA.

When a constant amount (4ng) of ^{32}P-labelled DNA was incubated with an increasing amount of tnsD whole-cell crude extract, in the presence of 2ug of "carrier" DNA (supercoiled pUC8 DNA), one band (I) of lower electrophoretic mobility was observed (Figure 4.12, lanes 3-5). Considering the amount of total labelled DNA used in each assay (4ng) and the ratio of "hot site" DNA/total DNA in pZMR80 (1/9), the amount of labelled DNA carrying the presumptive tnsD binding sites (290bp fragment) included in each reaction was estimated to be about 0.4ng. Assuming that the molecular weight of 1bp is 650D, the molecular weight of the 290bp fragment was calculated to be approximately 189kD. Taken together the latter two figures, the number of moles and molecules of the 290bp fragment per assay, were estimated as $2.1X10^{-15}$ and $1.3X10^{9}$, respectively.

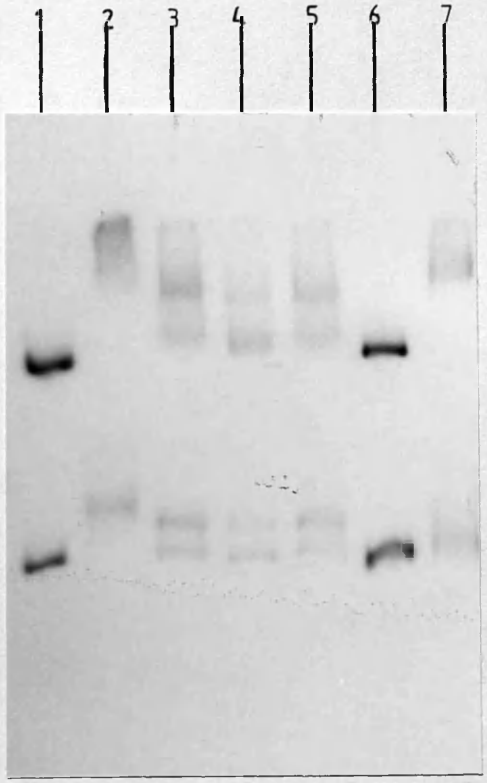


Figure 4.12 Detection of DNA-binding activity in tnsD crude extracts by the band-competition assay

pZMR80 was digested with EcoRI and HindIII and ^{32}P-end labelled. 4ng of this labelled DNA were mixed under the conditions of the assay (see Materials and Methods) with decreasing amounts of tnsD and tnsD-free crude extract. 1ug of carrier DNA (supercoiled lambda DNA) was included in each assay.

1 no extract
2 5 ug tnsD extract
3 2.5 ug tnsD extract
4 1.25 ug tnsD extract
5 0.63 ug tnsD extract
6 0.63 ug tnsD extract + proteinase K
7 2.5 ug tnsD-free extract

tnsD extract = pZMR80/DS941
tnsD-free extract = pGLW8/DS941

When the same amount of ^{32}P-labelled DNA (4ng) was incubated with tnsD-free crude extracts (pGLW8/DS941), the 290bp fragment migrated with the mobility of naked DNA, at an identical protein concentration that gave the retarded form in the case of the tnsD extract (Figure 4.12, lane 7).

The molecular weight of the tnsD gene product is approximately 54kD (Rogers, 1986). On the assumption that tnsD represents 0.05% of the total protein in the extract, the moles and molecules of the tnsD gene product for the extract concentration that gave the retarded complex (0.63ug, Figure 4.13; lane 5), were estimated as $5.8X10^{-15}$ and $3.5X10^{9}$, respectively. Assuming that binding occurs at a 1:1 stoichiometry and that there is one binding site for tnsD per DNA molecule, it is estimated that tnsD was present in a 3-fold excess in the binding reaction. However, only 50-60% of the DNA appeared to be in a bound form (Figure 4.12, lanes 3-5).

Based on the number of molecules of tnsD added (1.3×10^{9}), and the number of cells included in each binding reaction (3×10^{7}), the abundance of active tnsD was estimated to be 43 molecules per cell. The actual abundance of tnsD in the cell is likely to be higher, since not all molecules in the cell are likely to be detected by the binding assay.

Proteinase K treatment of the binding reactions containing tnsD protein extract abolished retardation and converted all the fragment in band I to a form that comigrated with unbound DNA (Figure 4.12, lane 6).

4.6.2 Specificity of tnsD Binding

Had time permitted competition experiments, similar to those described in section 4.5.3 for tnsB, would have been carried out using unlabelled "hot site" DNA, to examine whether the binding of tnsD observed in Figure 4.12 is specific for the "hot site".

4.7 Conclusions

One of the main functions of the transposition complex is the recognition of the ends of the element. Results presented in chapter 3 provided strong evidence supporting the interaction of the Tn7 tnsB gene product, one of the proteins involved in both modes of Tn7 transposition (Rogers et al, 1986), with the right end of Tn7. In this chapter, the DNA binding properties of the tnsB gene product were examined. In addition, a preliminary study of the binding profile of the Tn7 tnsD gene product, the "hot site" specific protein (Rogers et al, 1986), was undertaken. It was thought that the transposon "hot spot" for insertion in the E.coli chromosome is a likely candidate for binding of this protein.

Results presented in this chapter described the specific binding of the tnsB protein to the right end of Tn7 and the binding of the tnsD protein to the Tn7 chromosomal attachment site ("hot site"). A modification of the gel electrophoresis DNA binding assay described by Fried and Crothers, (1981); Garner and Revzin, (1981); and Strauss and Varshavsky, (1984) was used. In this assay a highly radioactive fragment is mixed with an excess of unlabelled non-specific DNA ("carrier" DNA) and is incubated with a crude extract. Proteins that form specific complexes with the labelled fragment retard the mobility of the fragment when electrophoresed through low ionic strength polyacrylamide gels.

Much of the binding of tnsB to the right end of Tn7 was dissociated by competition with a 4 to 10-fold molar excess of specific unlabelled fragment, but was resistant to dissociation when competed with unlabelled heterologous DNA.

Evidence presented in chapter 3 suggested that direct binding of the tnsB gene product in the region encompassing P_1 results in a considerable repression of transcription from P_1. This repression is most probably due to the binding of the tnsB protein at the direct repeat containing the -35 region of P_1 (Figure 4.1).

The tnsB protein binds specifically and with high affinity to sequences contained within the 205bp fragment from the right end of Tn7. Such binding results in a considerable increase in the apparent size of the complexes (Figures 4.6 and 4.7) which is suggestive of a major change in the conformation of DNA upon binding. Similar increases resulting from bending of the DNA have been observed upon specific binding of the E.coli integration host factor (IHF protein) to the IS1 termini (Gamas and Chandler, 1987) and of the InsA protein of IS1 to the ends of the element (Zerbib et al, in press).

To confirm that the retardation observed is due to binding of the tnsB gene product and not to binding of another protein present in the extracts that is activated by tnsB, retardation assays analogous to those performed by Zerbib et al, (in press) using ^{35}S-methionine labelled cell extracts (with or without the tnsB gene product) and unlabelled DNA fragments can be performed.

Using the "band-shift" assay described in this chapter, it is possible to examine the time required for the specific protein-DNA association to be established. All binding reactions performed to demonstrate specific binding of the tnsB protein to the right end, were incubated for 5-7mins. However, it is likely that the protein-DNA complexes are formed very rapidly after the addition of the extract to the DNA substrate. A time course starting, say, 30secs after the addition of the extract, and stopping at different time points, can provide us with an estimate for the time of formation of the various complexes. In such an experiment, it is expected that the non-specific extract (tnsB-free) will not show any binding activity, even after long incubation periods. In a similar manner, it is possible to determine the length of time required to dissociate the specific binding in each complex after addition of a 10-fold molar excess of unlabelled DNA substrate. Again, dissociation is expected to occur immediately after the addition of the unlabelled fragment.

Circularly permu ted variants of the same sequences used to detect specific DNA-binding of the Tn7 tnsB and tnsD proteins,

can be used to determine the location of the bends within these regions. It is known that isomers with the bend at different positions display different electrophoretic mobilities on low ionic strength polyacrylamide gels, with the slower migrating forms possessing the bend in the middle of the fragment (Wu and Crothers, 1984).

The binding pattern obtained for the tnsB gene product (Figures 4.6 and 4.7) indicates that there might be more than one binding site for tnsB within the 239bp right end fragment. This is supported by the sequence organisation of the fragment which, as seen in Figure 4.5, possesses four stretches of homology and hence, leaves enough space for speculations concerning the nature of the seven bound forms (complexes I-VII) observed. Each of these forms could arise from occupation by tnsB of one or more binding sites within the 239bp fragment. For example, the four major complexes (I-IV; as deduced from the intensity of the bands in Figure 4.7, lane 1) could represent binding of one tnsB molecule to the inverted repeat (I) and binding of two protein molecules to the inverted repeat and each of the three direct repeats (II-IV). Similarly, the three minor complexes (V-VII) could represent binding of two protein molecules to pairs of the direct repeats.

To delineate and locate more precisely the regions of tnsB binding within the right end of Tn7 and the regions of tnsD binding within the Tn7 "hot site", protein-DNA complexes can be analysed by a modified DNAase I footprinting technique which displays DNA sequences protected from DNase I digestion by crude bacterial (or cell-free) extracts (Galas and Schmitz, 1978; Tolias and Dubow, 1986). Footprint experiments can be performed on the 239bp right end fragment using tnsB whole-cell extract and on the 290bp "hot site" fragment using tnsD extracts.

Another method for examining binding to Tn7 ends would be the use the *in vivo* photofootprinting technique (Becker and Wang, 1984). This technique has succesfully been used in our laboratory for Tn3 res/resolvase interactions *in vivo* (Brown, 1986) and should

have no difficulty in applying it to examine "footprints" of Tn7 proteins to Tn7 termini separately and together in the same molecule. Genetic dissection can be used to look at derivatives of the ends deleted for parts of the repeats.

DNAase footprint analyses have shown that RNA polymerase and the E.coli integration host factor (IHF protein) bind specifically to the ends of IS1 (Gamas and Chandler, 1987; Machida et al, 1984). It is also known that IHF binds at the ends of most transposable elements, though such binding does not seem to be required for transposition (N. Grindley, pers. comm; Gamas and Chandler, 1987). Tn7 possesses one good consensus IHF binding sites at either end and it is therefore likely that IHF is binding to one or both of them.

Results presented in this chapter suggest that the specific association and dissociation of the tnsB protein from its binding sites may transiently expose the major Tn7 promoter (P_1) and thus allow some transcription to occur from P_1 in preference to the weaker Tn7 promoter P_2. This would have the dual effect of regulating P_2 transcription as well as allowing more efficient expression from P_1 and thereby tnsA, tnsB, tnsC, tnsD, and tnsE proteins. Elevated expression of the latter functions could result in increased transposition frequencies, and thus the transient dissociation of the tnsB gene product from its binding sites could coincide with bursts in transposition.

The effect (if any) of having the left hand terminal repeat present in cis in the tnsB binding reaction can also be examined. It is almost certain that the two terminal repeats interact before transposition occurs. Also, it is worth examining the binding of the tnsD gene product to the termini of Tn7, since it is possible that tnsD brings the transposon ends together by binding to each of them before recognising the "hot site".

The modified DNase I footprinting technique and the in vivo photofootprinting technique can provide us with information concerning the specific binding sites of the tnsB gene product

within the right end of Tn7. These techniques on their own, however, cannot relate each of the seven protein-DNA complexes observed in Figure 4.7 with their binding sites. DMS treatment of the binding reactions before resolving them by gel electrophoresis may allow us to isolate specific DNA-protein complexes, and examine which purine residues in the DNA are protected from methylation. Direct comparison between the sequencing profiles of the 205bp Tn7 right end- containing fragment and that of DMS treated complexes would then assign each of the seven complexes to a site within the right end of Tn7.

Finally, it should be noted, that the band-competition assay employed here to detect specific binding of the tnsB and tnsD gene products to the right end and "hot site" respectively, can be used by future workers to monitor the purification of active forms of these two proteins.

CHAPTER 5

THE ROLE OF DNA ADENINE METHYLASE IN TN7 TRANSPOSITION

5.1 INTRODUCTION

DNA adenine methylase, the product of the E.coli dam gene, is the enzyme responsible for the postreplicative methylation of adenine at GATC sequences in the DNA of E.coli (Lacks and Greenberg, 1977; Marinus and Morris, 1973). Dam methylation is involved in the correction of replication errors that generate basepair mismatches so as to preserve the information on the parental (methylated) strand (Pukkila et al, 1983) and is known to play an important role in several E.coli systems either by coordinating cellular processes related to replication of the bacterial chromosome or by affecting gene regulation.

Methylation influences trpR promoter activity (Marinus, 1985) and is required for expression of the mom gene of phage Mu (Plasterk et al, 1983), the cre gene of phage P1 and for cleavage at the specific packaging site (pac) during encapsidation of DNA by phage P1 (Sternberg, 1985). Absence of methylation at the IS10 transposase gene promoter, pIN, increases its activity 10 to 30 times (Roberts et al, 1985), while precise excision of Tn10, a transposase-independent illegitimate event, is enhanced in dam$^-$ mutants (Lundblad and Kleckner, 1984). The unusually high density of methylation sites within oriC, has been implicated to be involved in regulating initiation of chromosomal replication (Smith et al, 1985). Methylation also influences expression of the dnaA and sulA genes, which are important for initiation of chromosomal replication and bacterial cell division, respectively (Braun and Wright, 1986; Mizusawa et al, 1983).

The physiological role of DNA methylation in promoter activity is unclear since all measurable dam sites in non-replicating DNA are methylated in vivo (Marinus, 1984). Work by Roberts et al (1985) on the IS10 transposase gene promoter, pIN, suggests that promoter activity is enhanced on hemimethylated templates which exist at the replication fork.

As a host factor, DNA adenine methylase has been shown to modulate the activity of several non-replicative transposable

elements. The effects of dam$^-$ mutations on transposition have been studied using the dam3 allele. Transposition frequencies of Tn5/IS50 and Tn903/IS903 are increased in dam3 by 5 to 20-fold, while transposition frequencies of Tn3 and Tn9 are unaltered (Roberts et al, 1985). Tn10-promoted events are increased in dam3, though the magnitude of the effect varies for different events. Tn10 transposition is increased 10 to 20-fold, IS10 transposition is increased 100-fold, and Tn10-promoted deletions and "inside-out" transpositions 200-fold or more. All effects of dam$^-$ mutations on IS10 transposition are mediated through the two GATC sites in IS10, one of which overlaps the -10 region of the transposase promoter of IS10, pIN, and the other is 7-10bp from the inside end of IS10, within the region where transposase presumably binds (Roberts et al, 1985). In the cases of Tn5/IS50 and Tn903/IS903, the increases are probably due to increased promoter function precisely analogous to that responsible for the increase in Tn10 transposition. Like Tn10, both transposons have GATC methylation sites located in the -10 regions of the transposase gene promoters (Grindley and Joyce, 1980), and both promoters are more active in vivo in dam$^-$ strains. Moreover, IS50 has two dam methylation sites at its inner terminus, indicating that IS50 transposition may be regulated by dam methylation at both of the levels observed for IS10. However, not all transposable elements are affected by dam methylation. IS1 and Tn3, which undergo replicative transposition at least 10% of the time, contain no dam sites in critical regions and transpose at equal frequencies in dam$^+$ and dam$^-$ hosts (Heffron et al, 1979; Johnsrud, 1979; Ohtsubo and Ohtsubo, 1978).

Roberts et al, (1985) have looked at the effect of hemimethylated DNA on IS10 transposition, since there is no evidence that wild type cells ever contain fully unmethylated DNA analogous to that found in dam$^-$ mutants (Marinus and Morris, 1973; Marinus, 1984; Razin et al, 1980). On the contrary, hemimethylated DNA is generated every time DNA synthesis crosses a methylation site during passage of a replication fork or a repair tract, or after transfer and complementary strand synthesis of single stranded Hfr DNA during conjugation. This hemimethylated DNA exists only

transiently, 2-6mins out of 40-60mins of a generation time (Lyons and Schendel, 1984; Szyf et al, 1982). Their work has shown that one hemimethylated species of IS10 (promoter and terminus methylated only on the template strand) is much more active than the other (pIN-transposase gene methylated only on the non-template strand) and is estimated to be at least 1000 times more active than a fully methylated element. This observation, supported by the fact that the promoter and inner terminus of IS10 function coordinately in a methylation dependent manner, led them to suggest that the natural occurence of hemimethylated DNA in vivo is used to regulate transposition of IS10. They proposed specifically, that IS10 is activated to transpose following passage of a replication fork across the element. Figure 5.1 illustrates how activation of one hemimethylated element immediately after passage of a replication fork might be particularly advantageous for elements that move by non-replicative excision from the donor location followed by insertion at a new target site without religation of the ends of the donor molecule.

In this chapter, an analogous approach to that followed by Roberts et al, (1985) for Tn10 has been taken to examine whether dam methylation is a host factor participating in the transposition of Tn7. The in vivo transposition of Tn7 has been studied in unmethylated and hemimethylated hosts and a direct comparison with the transposition levels in isogenic fully methylated strains has been made. Also, the transposition frequencies of the two hemimethylated species of Tn7 have been compared. The main reasons for undertaking this series of experiments have been the following: a) The existence of a dam methylation site (GATC) at the right end of Tn7, in a region that is believed to act as a substrate for binding of transposition proteins (see chapter 4). This GATC site comprises part of the BglII site at 14.0kb and is contained within the second direct repeat proximal to the end of the element, (figure 5.2) b) Tn7 appears to transpose conservatively (no evidence for cointegrate intermediates, Rogers, 1986), and therefore it is possible that it uses methylation to regulate its transposition, as has been

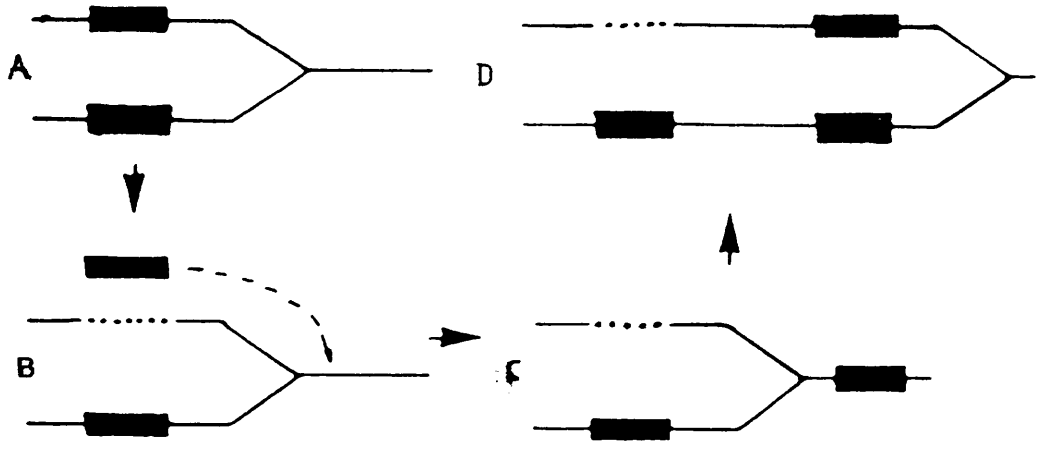


Figure 5.1 Possible strategy for increasing the copy number of a conservative element

A) The replication fork crosses the element activating transposition.
B) The element transposes ahead of the replication fork (or into the newly synthesised sister strand).
C) The donor chromosome is lost due to double strand breaks during transposition.
D) Passage of the replication fork across the newly inserted element gives a product that would be indistinguishable from a replicative transposition event.

Solid boxes represent the transposon.

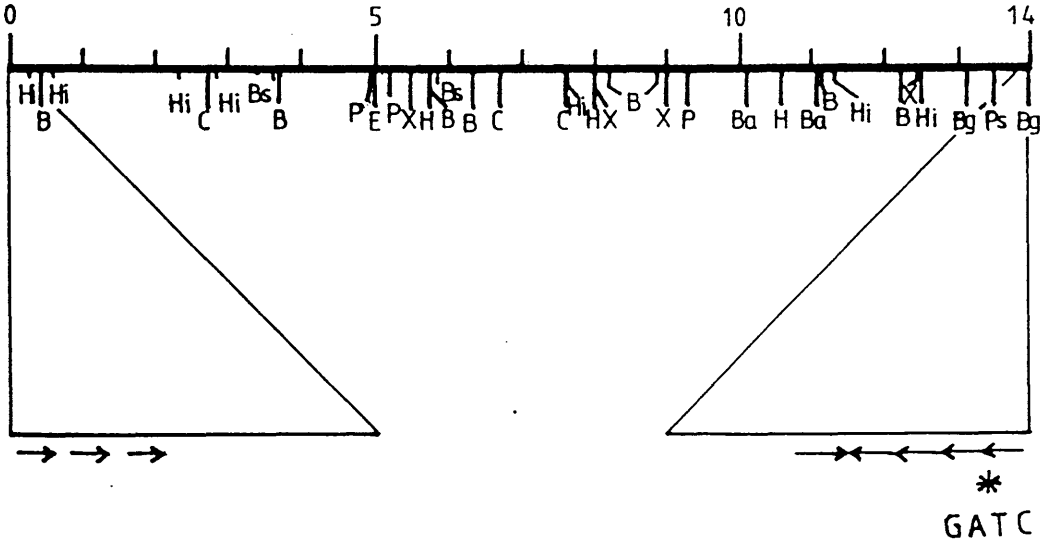


Figure 5.2 Location of the dam sensitive site (GATC) in the right end of Tn7.

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

observed for other non replicative-elements (IS10/Tn10, IS50/Tn5; Roberts *et al*, 1985).

The effects of *dam*$^-$ mutations on Tn7 transposition have been studied using the *dam*3 allele.

RESULTS AND DISCUSSION

5.2 Transposition of Tn7 in a *dam*$^-$ Background

Initially, an experiment was designed that could give some preliminary evidence as to whether changes in the state of methylation in the only one *dam* methylation site (GATC) within Tn7 affect transposition. A series of mating-out transposition assays, (Tables 5.1 and 5.2) was performed, as described in Materials and Methods, in CSH26 and its isogenic *dam*$^-$ derivative CB51. pUC8::Tn7 and pZMR80::Tn7 were transformed into CSH26(pZEN300), CSH26(R388), CB51(pZEN300), and CB51(R388). R388 and pZEN300 were mated out into ZMR1 (Table 2.1) in a standard transposition assay.

It is important to note, at this point, that the above described experiment examines an artificially created situation, since bacterial DNA is rarely, if ever, fully unmethylated *in vivo* (Marinus and Morris, 1973; Marinus, 1984; Razin *et al*, 1980). Therefore, any data obtained should be interpreted cautiously with respect to their importance in determining the role of *dam* methylation in Tn7 transposition. Also, the *in vivo* activity of fully methylated Tn7 is difficult to measure directly, because in wild-type cells the hemimethylated and fully methylated forms of Tn7 exist alternatively, and therefore the activity of Tn7 in a *dam*$^+$ strain is the total of contributions from both fully methylated and hemimethylated elements.

Results presented in Figure 5.2 show that there is a 5 to 9-fold increase in Tn7 transposition in a *dam*$^-$ environment. Though this effect was found to be consistently repeatable between different experiments, it was not immediately attributed to an elevated

Table 5.1 Transposition of Tn7 in a *dam*⁻ Host.

Donor plasmid	Recipient plasmid	Host	Transposition Frequency
pUC8::Tn7	R388	CSH26	6.1×10^{-7}
pUC8::Tn7	R388	CB51	3.9×10^{-6}
pZMR80::Tn7	R388	CSH26	2.2×10^{-6}
pZMR80::Tn7	R388	CB51	2.0×10^{-5}
pUC8::Tn7	PZEN300	CSH26	1.7×10^{-4}
pUC8::Tn7	PZEN300	CB51	1.3×10^{-3}
pZMR80::Tn7	PZEN300	CSH26	4.3×10^{-4}
pZMR80::Tn7	PZEN300	CB51	1.3×10^{-3}

A. Each frequency represents the average of at least three experiments. The standard errors of these frequencies never exceeded 1/10 of the transposition frequencies.

B. Transposition frequencies were measured as described in Materials and Methods.

Table 5.2 Properties of the Plasmids Used in Table 5.1.

Plasmid	Hot site	Tn7
R388	-	-
PZEN300	+	-
pUC8::Tn7	-	+
pZMR80::Tn7	+	+

activity of the element in unmethylated hosts. Physiological artifacts, such as differential generation time of dam$^+$ and dam$^-$ cells, could account for this phenomenal increase. To investigate the relevance of this result to the transposition of Tn7, an experiment was set up that closely mimicked the in vivo conditions: Tn7 transposition was examined in hemimethylated and fully methylated backgrounds (section 5.3).

5.2.1 Considerations in Using the Conjugative Transposition Assay

The conjugative or mate-out transposition assay (see Materials and Methods) was employed in this chapter for the study of Tn7 transposition in dam$^+$ and dam$^-$ environments. This assay involves growing up cultures of E.coli containing a recipient conjugative plasmid (R388 or its "hot site" derivative pZEN300) and the donor plasmid which should be non-mobilisable and non-conjugative. A mating is performed which samples the conjugative plasmid population. The proportion of these plasmids which also carry the transposon (as evidenced by their acquisition of the transposon markers) is termed transposition frequency. Though this figure does not reflect an absolute measure of transposition, it allows meaningful comparisons between assays to be made.

Rogers, (1986) examined the following two possible sources of error in this assay, before studying the transposition frequencies:

a) The growth rate of cells containing a transposon insertion in R388 or pZEN300 was determined under conditions similar to those used for measuring the transposition frequencies. It was found that the presence of Tn7 in a plasmid had no effect on the growth rate of the host.

b) The efficiencies of mating of R388 (and pZEN300) and R388::Tn7 (and pZEN300::Tn7) were compared. It was found that there was no significant difference between the efficiencies of transfer in both cases.

5.3 Transposition of Tn7 in a *dam*$^+$/*dam*$^-$ Background

Transposition of hemimethylated Tn7 was examined *in vivo* by measuring transposition from newly transferred Hfr DNA (Figure 5.3). During conjugal transfer, a single DNA strand moves from the donor cell to the recipient cell, where the complementary strand is synthesised. If the donor is *dam*$^+$ and the recipient is *dam*$^-$, the newly replicated DNA will be hemimethylated and will remain so until it gets either degraded or integrated into the host genome by transposition or homologous recombination. The frequency of transposition of hemimethylated Tn7 can be measured as the frequency of Tp resistant exconjugants obtained in such an experiment using a *dam*$^-$/*recA*$^-$ recipient strain where integration by homologous recombination is prevented. Tn7 can occur in two hemimethylated forms, one methylated on the presumptive transposase binding site template strand and the other on the non template strand (Figure 5.3). Transposition of the two hemimethylated forms of Tn7 was compared by using two donor strains carrying Tn7 in opposite orientations at the "hot site". These frequencies were then compared with those obtained in an isogenic *dam*$^+$/*recA*$^-$ strain to examine the effects of full methylation on the activity of hemimethylated transposon DNA.

It should be stressed, however, that that this experiment can only provide a minimum estimate of the intrinsic difference in activity of the fully methylated and hemimethylated forms of Tn7, since much or all of the transposition observed in the *dam*$^+$ strain could result from transposition of the hemimethylated elements prior to full methylation.

The efficiency of transfer of Tp^r was measured separately, each time the experiment was repeated, by measuring the transfer of this marker from both donors into a *recA*$^+$ recipient. Possible orientation differences unrelated to methylation can be detected by comparing the transposition frequencies of the two donors into the *dam*$^+$/*recA*$^-$ recipient. Because the two donor Hfr strains are *recA*$^+$, additional experiments were performed, to test for the existence and subsequent transfer of F' from the donors into the

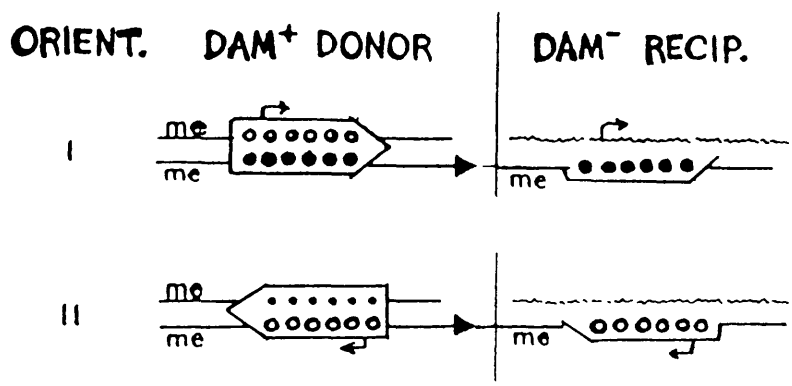


Figure 5.3 Two hemimethylated forms of Tn7 (I and II) are generated *in vivo* during conjugal transfer. From a *dam*$^+$, Tn7-containing Hfr donor, (ZNE19 or ZNE20) a single methylated strand is transferred, 5' end first, (indicated by large arrow), and the complementary strand is synthesised in the recipient (*dam*$^-$; ZNE16 or ZNE18), thus generating hemimethylated Tn7 DNA. The two Hfr donors used in this experiment contained Tn7 insertions in opposite orientations at the "hot site". The two strands of Tn7 are drawn as closed and open circles for P_1 template strand and non-template strand, respectively.
The diagram was taken from Roberts *et al*, 1985

recipients during the matings. Such F' existence could interfere with the interpretation of the results (see section 5.8).

5.4 Construction of Strains

Two isogenic recA$^-$ strains, namely ZNE16 and ZNE18, which differ in their dam phenotypes, were constructed to act as recipients in the Hfr crosses. The following three sections discuss the principles upon which construction of the recA$^-$/dam$^-$ strain was based.

5.4.1 Secondary Phenotypes of dam$^-$ Strains

Dam mutant strains differ from wild type in several other respects for which the term "secondary phenotypes" is used to distinguish them from the primary dam$^-$ phenotype which is loss of DNA adenine methylase (the dam3 mutation results in a 5-fold reduction in the number of 6-methyl-adenine residues in the DNA of E.coli or phage lambda, Marinus and Morris, 1975). These secondary phenotypes include: increased sensitivity to 2-aminopurine (Glickman et al, 1978); slight sensitivity to ultra-violet light (Marinus and Morris, 1974, 1975); inviability of mutants containing dam3 and recA, recB, recC, polA or lex-1 mutations (Marinus and Morris, 1975); high spontaneous induction of wild-type lambda prophage (Marinus and Morris, 1975); high spontaneous mutability (Glickman et al, 1978; Marinus and Morris, 1974); increased mass per cell (Marinus and Morris, 1975).

The sensitivity of the dam$^-$ strains to the base analog 2-aminopurine was used in the construction of the dam$^-$/recA$^-$ strain NE18 (see section 5.5).

5.4.2 Interpretation of the Sensitivity of a dam Mutant Strain to the Base Analog 2APur

Figure 5.4 illustrates the sensitivity of dam mutant strains to 2APur. Relatively closely spaced, newly incorporated 2APur residues (or other mismatched bases) will result in overlapping

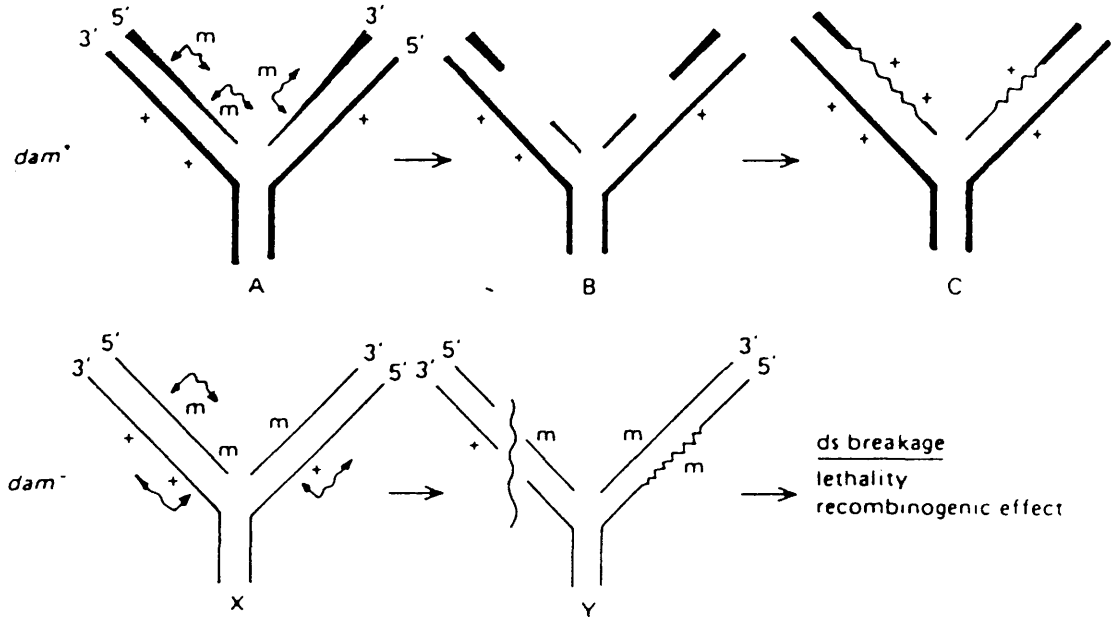


Figure 5.4 Interpretation of the sensitivity of a dam mutant strain to the base analog 2-aminopurine (after Glickman and Radman, 1980)

The thickness of the DNA strands indicates their degree of methylation. A replicational error or a base analog-induced mismatch is symbolised by m. Arrows perpendicular to the DNA strands symbolise endonucleolytic attack at mismatched sites. Wavy lines parallel to the DNA strand indicate the 5'→3' direction of the excision repair tract. The wavy portion of the single-stranded DNA indicates an excision repair tract.

Wild-type situation: (A) Endonuclease/exonuclease-mediated mismatch removal occurs specifically from the undermethylated, newly synthesised strands. (B) Exonucleolytic degradation and resynthesis leads to the intact, mutation-free molecule shown as (C).

dam⁻ situation: (X) Endonuclease/exonuclease can act on both the newly synthesised and the parental DNA strands due to the absence of methylation. (Y) The left arm depicts the production of a double-strand (ds) break due to overlapping excision tracts extending over several thousand base pairs. The right arm illustrates a situation where mismatch repair results in error fixation rather than error avoidance, thus illustrating the mutator phenotype of dam mutants. The creation of double strand breaks accounts for the sensitivity to 2-aminopurine.

excision repair tracts if excision occurs in both the parental and newly synthesised strands as postulated for the dam$^-$ mutants. Consequently, mutants deficient in mismatch correction (e.g. in endonucleolytic cleavage or exonucleolytic strand degradation) are expected to restore resistance to 2APur in a dam$^-$ mutant while maintaining the mutator properties associated with the dam$^-$ phenotype. Such 2APur resistant revertants of dam$^-$ strains which, when separated from the dam mutation, are mutators themselves, have been isolated and mapped to the locations of the three known mutator loci mutS, mutL, and mutH (Glickman and Radman, 1980). MutH, mutS, mutL, and dam$^-$ mutations result in the same mutational spectrum (increased transition and frameshift mutagenesis, Cox, 1976; Glickman, 1979). This is interpreted to mean that mutH, mutL, mutS, and dam are involved in the same error-avoidance mechanism.

5.4.3 Restored Viability of dam/recA Mutants by Addition of a mutH, mutL, or mutS Mutation.

The lethality of a dam$^-$/recA$^-$ double mutant (Marinus and Morris, 1975) might reflect a requirement for the recA gene to repair double-strand breaks, as speculated in Figure 5.4. Thus, an additional deficiency in mismatch correction should prevent the occurence of the double-strand breaks and alleviate the need for the recA gene in the maintenance of viability in a dam$^-$ background.

5.4.4 Construction of ZNE16 (dam$^+$/recA$^-$)

ZNE16 is the recA$^-$ derivative of CSH26 and its construction has been made as follows:

a) CSH26 was made srl$^-$Tc^r by P1 transduction using a P1 lysate made on JC10241 (srl::Tn10, recA$^+$, see Table 2.1). Transductants were selected on minimal media containing tetracycline (to select for the aquisition of Tn10) and CSH26 amino acid requirements. The inability to metabolise sorbitol was checked by growth on minimal media containing sorbitol (instead of glucose) as a

carbon source. It was found, that none of the Tc^r colonies that grew on MM + glucose plates could grow on MM + sorbitol plates. Therefore, all Tc^r colonies were transductants carrying Tn10 inserted in the sorbitol locus, and had not arisen by transposition of Tn10 from the P1 infecting lysate to the chromosome of CSH26. One of these transductants was named ZNE1 (CSH26 srl::Tn10) and was selected to act as a recipient in the following step of the construction.

b) ZNE1 was made recA$^-$ by P1 tranduction, using a P1 lysate made on DS902 containing pPE14, a Cm^r pACYC184 based plasmid carrying the recA gene. From previous experience, the presence of a recA harboring plasmid in a recA$^-$ strain facilitates the infection by P1 of this strain, and gives higher titer lysates. Transductants were selected for their ability to metabolise sorbitol as a carbon source on minimal media containing CSH26 amino acid requirements. The recA status of these isolates was then determined by UV sensitivity, as described in Materials and Methods. Several isolates, (66% of the total number), were recA$^+$ probably due to homogenotisation, however, a few were obtained that were recA$^-$. Finally, the recA$^-$ transductants were tested for their sensitivity to tetracycline (confirmation of loss of Tn10) and were all found to be sensitive. This showed that Tn10 had not hopped into any other sites in the chromosome of ZNE1 during its residence in the sorbitol gene. The presence of another transposon in the recipient strain could possibly distort the Tn7 transposition frequency. One of these recA$^-$ transductants was selected and named ZNE8.

c) Because ZNE8 carries no additional markers that can be used to counter select against it, a rif^r spontaneous mutant of this strain was obtained, by growing ZNE8 on plates containing 50ug/ml rifampicin. The resulting strain was named ZNE16 and was tested for the maintenance of the recA$^-$ status and all the phenotypic markers of CSH26.

5.4.5 Construction of ZNE18 (*dam*$^-$/*recA*$^-$)

ZNE18 is the isogenic *dam*$^-$ derivative of ZNE16. Its construction was based on CB51 and was made as follows:

a) CB51 (*dam*$^-$ derivative of CSH26) was transduced to a srl$^-$Tc^r phenotype by P1 transduction using a P1 lysate made on JC10241 (*recA*$^+$, srl::Tn10). Transductants were selected on minimal media containing tetracycline and CSH26 amino acid requirements. All isolates obtained were tested for their inability to metabolise sorbitol, as described previously for ZNE16, and were found unable to metabolise it. One of these transductants was was named ZNE3 (CB51 srl::Tn10), and was selected for further work.

b) ZNE3 is *dam*$^-$ and therefore sensitive to 2-aminopurine. To allow the viability of the final strain, which is required to be *dam*$^-$/*recA*$^-$, 2-aminopurine revertants of ZNE3 were selected. 2-aminopurine was used at 500ug/ml and four colonies were obtained after plating different concentrations of ZNE3 on minimal media containing ZNE3 requirements and 2-aminopurine. These four isolates were subsequently tested for their maintenance of the *dam* phenotype. pZNE89, a plasmid with known *Bcl*I sites ("methylation" sensitive enzyme), was transformed in all four isolates, and the restriction patterns after digestion with *Bcl*I were compared. Three strains were found to be *dam*$^+$ (pZNE89 DNA remained uncut) probably due to reversion of the *dam*3 allele to wild-type (isolation of *dam*$^+$ revertants has been reported before, McGraw and Marinus, 1980) and the fourth was found to be *dam*$^-$. This isolate, namely ZNE4, had an unmapped second site mutation which suppressed the 2-aminopurine sensitivity of ZNE3 and hence permitted the viability of a *dam*$^-$/*recA*$^-$ double mutant.

c) ZNE4 was made *recA*$^-$ by P1 transduction using a P1 lysate made on DS902/pPE14 (as described earlier for ZNE16). Transductants were selected for their ability to grow on sorbitol, and the procedure followed to characterise them was identical to that described for ZNE16. Here again, 2/3 of the obtained isolates were *recA*$^+$ due to homogenotization. One of these *recA*$^-$

transductants was named ZNE9.

d) ZNE9 was made rifampicin resistant, as described for ZNE2. The resulting strain was named ZNE18, and maintained all the phenotypic characteristics of ZNE9.

5.4.6 Construction of ZNE19 and ZNE20 (Hfr::Tn7)

ZNE19 and ZNE20, are the two Hfr donors carrying Tn7 on opposite orientations, at the "hot site", which have been used to examine the effect of hemimethylation on Tn7 transposition. They are CGSC 4294 and CGSC 4312 derivatives, respectively. Their construction were made as follows:

a) The pMB1 based, cI (lambda repressor) overproducing plasmid pEA305 (Ap^r, Amann et al, 1983), was used to select for the transposition of Tn7 from the lambda dv Tn7-harboring plasmid pZNE1 into the chromosome of these two strains. Introduction of pEA305 in a cell that already contains pZNE1 (lambda origin of replication) inhibits any further replication of this plasmid, as a consequence of the presence of the lambda repressor. Therefore, all Tp^r/Ap^r transformants should have Tn7 inserted into the chromosome.

CGSC 4294 and CGSC 4312 were sequentially transformed with pZNE1, (Tc^r) and pEA305 (Ap^r, see Table 2.2). Indeed, all Ap^r transformants checked were Tc^s, but Tp^r and Sm^r/Sp^r, indicating that Tn7 had inserted into the chromosome of these strains.

b) One Ap^r transformant, for each strain, was selected and grown in L-broth - Ap for several generations. The cultures were then plated on L-agar -Ap, and colonies that grew on such plates, were patched on -Ap and + Ap L-agar, to screen for loss of pEA305. It was found that pEA305 was very stable. Ap^s colonies could be isolated after growing overnight cultures of the initial Ap^r transformant for one week, under no selection, subculturing every day.

5.5 Entry Time for Tp^r

To define the optimum time for entry of Tp^r, and thus Tn7 from each of the two Hfr donors (ZNE16 and ZNE18) into the recipient cell, interrupted matings (see Materials and Methods) were performed in ZNE15 (Table 2.1), a $\underline{dam}^+/\underline{rec}^+$ recipient. Figure 5.5 shows that the optimum time for transfer of Tn7 is 60mins.

5.6 Do Tp^r Exconjugants Represent Solely Transposition Events?

One important consideration that had been taken into account in interpreting the data from the Hfr crosses, was whether a part or all of the Tp^r colonies were due to F' formed in the donors and transferred into the recipients during the cross. Although both Hfr donors appeared to be stable according to their pedigrees, it did not seem unreasonable to expect formation of F', even at a low level, in a $\underline{rec}A^+$ background. If this turned out to be the case, the transposition frequency, measured as the number of Tp^r exconjugants, would not necessarily totally represent transposition events. Single colony DNA made on a number of Tp^r exconjugants seemed to be clear of F'. This, however, was not considered as conclusive evidence supporting their absence, since low copy number or unstable F' cannot be detected by this method. To finally determine whether formation and subsequent transfer of F' from the donors was distorting the transposition frequencies, two possibilities were examined:

a) The existence and transfer of F' carrying recombination genes would allow homologous recombination to occur in the $\underline{rec}A^-$ background of the recipients. Hence, the number of Tp^r exconjugants measured would represent the total of transposition and recombination events.

b) Occurence and transfer of F' carrying Tn7 would interfere with the results, since the Tp^r exconjugants measured could represent colonies that carry Tn7 on either the F' or the chromosome.

To investigate the first possibility, the $\underline{rec}A$ status of all

Figure 5.5 Time of entry for Tp^r

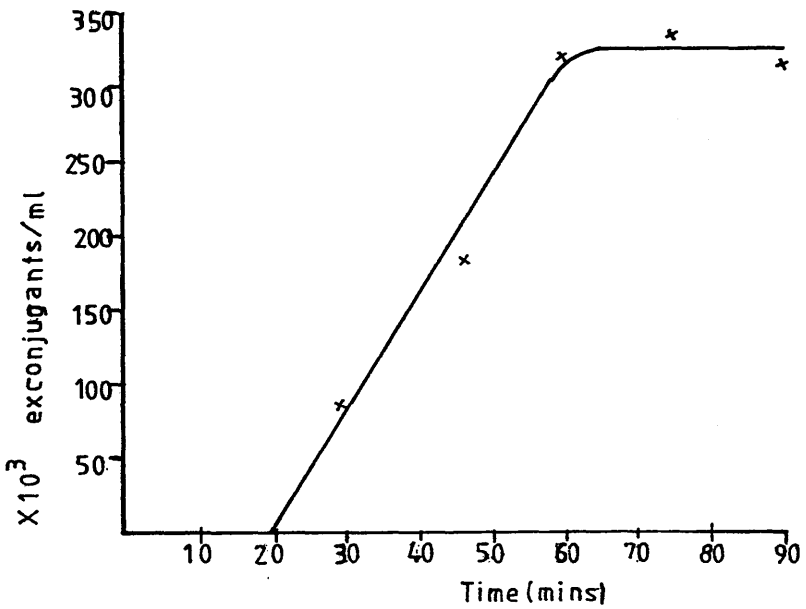
Data for the cross ZNE19 X ZNE15

Time (mins)	No of Tp^r exconjugants/ml (X 10^3)	
30	28	
45	58	
60	100	Plotted on graph 1
75	102	
90	96	

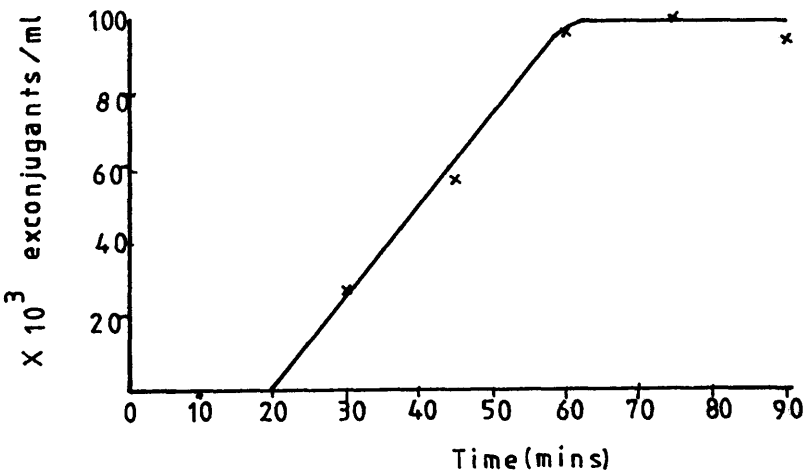
Data for the cross ZNE20 X ZNE15

Time (mins)	No of Tp^r exconjugants/ml (X 10^3)	
30	90	
45	185	
60	325	Plotted on graph 2
75	340	
90	320	

Each cross has been repeated twice.


350
300
250
200
150
100
50
X10³ exconjugants/ml
10
20
30
40
50
60
70
80
90
Time (mins)

1


100
80
60
40
20
X 10³ exconjugants/ml
0
10
20
30
40
50
60
70
80
90
Time (mins)

2

Table 5.3 Transposition of Hemimethylated Tn7 after Hfr Transfer.

Hfr Donor	F^- Recipient	Frequency of Tp^r transfers/donor	Transposition frequency
ZNE19 (orientation I)	ZNE18 (*dam*$^-$*rec*$^-$)	3.0×10^{-7}	9.4×10^{-5}
ZNE19	ZNE16 (*dam*$^+$*rec*$^-$)	6.6×10^{-7}	2.0×10^{-4}
ZNE20 (orientation II)	ZNE18	1.6×10^{-5}	8.0×10^{-4}
ZNE20	ZNE16	1.2×10^{-5}	6.0×10^{-4}
ZNE19	ZNE15 (*dam*$^+$*rec*$^+$)	3.2×10^{-3}	$<1.0 \times 10^{-8}$
ZNE20	ZNE15	2.0×10^{-2}	$<1.0 \times 10^{-8}$

A. The frequency of Tp^r transfer/donor and the transposition frequency were measured as described in Materials and Methods.

B. Each frequency represents the average of at least three experiments. The standard error of these frequencies never exceeded 1/10 of the frequency.

exconjugants from crosses into ZNE16 and ZNE18 was tested by UV sensitivity (see Materials and Methods). Every colony was found to be *recA*$^-$, which suggested that there were no F' containing recombination genes in the conjugated cells. However, this experiment could not detect unstable F' that got lost soon after the transfer and therefore it could not conclusively eliminate the hypothesis tested.

To examine the second possibility, the following experiment was performed:

Two strains, auxotrophic for markers that flank the site of insertion of Tn7 into the chromosome (82mins) of the two Hfr donors were selected (Figure 5.6). ZNE21 is *xyl*$^-$ (80mins) and was used to test for transfer of *xyl*-carrying F' into ZNE19. ZNE22 is *argE*$^-$ (89mins) and was used to test for transfer of *argE*-carrying F' into ZNE20 (Table 5.4). Results in Table 5.4 indicated that there were no F' carrying either *xyl* or *argE* in the recipient cells after Hfr transfer, though unstable F' remained unaccounted for.

Finally, it should be noted that another way of testing whether the Tp^r exconjugants measured represented transposition events, would be by applying the Southern blot analysis to a number of the exconjugants.

5.8 Conclusions

Data shown in Table 5.3 suggest that Tn7 transposes at similar frequencies in hemimethylated and fully methylated backgrounds. Additionally, there is no apparent difference in the transposition levels of the two hemimethylated species of the element. The observed slight increase in transposition of the hemimethylated form II (ZNE20 transfer) of the transposon is a methylation-independent effect that is attributed to the elevated efficiency of Hfr transfer from this donor strain. This latter conclusion is drawn by comparing the efficiencies of Tp^r transfer from ZNE19 and ZNE20 into ZNE16 and/or ZNE15 (Table 5.3), and

Table 5.4 Transfer of *argE* and *xyl* into NE21 and NE22.

Donor	Recipient	Frequency of *xyl* transfers/donor
ZNE19	ZNE21 (*xyl*$^-$/*rec*$^-$)	$<1.0 \times 10^{-8}$
ZNE19	ZNE15 (*xyl*$^-$/*rec*$^+$)	1.8×10^{-4}

Donor	Recipient	Frequency of *argE* transfers/donor
ZNE20	ZNE22 (*argE*$^-$/*rec*$^-$)	$<1.0 \times 10^{-8}$
ZNE20	ZNE15 (*argE*$^-$/*rec*$^+$)	1.9×10^{-3}

A. The frequencies of *xyl* and *argE* transfers/donor were measured as described in Materials and Methods.

B. Each frequency represents the mean of three expreriments.

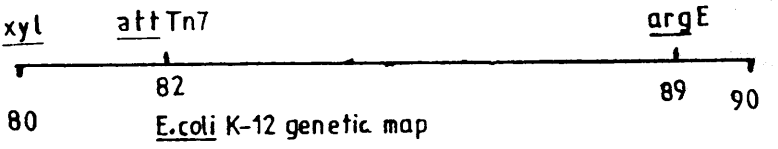


Figure 5.6 The region of the E.coli K-12 chromosome spanning the site of insertion of Tn7

the efficiencies of *xyl* and *arg*E transfer from ZNE19 and ZNE20, respectively, into ZNE15 (Table 5.4). Taken together, the data in Tables 5.3 and 5.4 suggest that ZNE20 transfers Hfr DNA 2-11 times more efficiently than ZNE19.

Results presented in Table 5.1 show that there is a slight increase (3 to 9-fold) in Tn7 transposition in a *dam*$^-$ background in all cases examined. Though this increase is consistently reproducible between different experiments, it is believed to be associated with physiological artifacts (e.g. difference in generation time between the *dam*$^+$ and *dam*$^-$ cells) and not with the methylation state of DNA. This interpretation is supported by the observation that hemimethylation, which represents the *in vivo* occurence of DNA, does not seem to have any effects on Tn7 transposition, as discussed in the previous paragraph.

The experiments described in this chapter do not differentiate between "hot" and "cold" site transposition. It is conceivable that all events measured represented, say, "cold" site transpositions and that transposition to the "hot" site can only occur from a fully methylated substrate. If this were the case, one of the two modes of Tn7 transposition would be regulated by *dam* methylation. To investigate this hypothesis, Southern blot analysis can be applied to a number of the exconjugants. Alternatively, the same experiment performed here with wild-type Tn7 could be repeated with *tnsD*$^-$ or *tnsE*$^-$ transposons (deficient for "hot" and "cold" site transposition respectively), thus providing a means of directly assessing "hot" and "cold" site transposition.

It has been shown that *dam* methylation may modulate the activity of a number of non-replicative prokaryotic transposable elements. Transposition of the conservative elements Tn5/IS50, and Tn10/IS10 is elevated in *dam*$^-$ mutants (Roberts *et al*, 1985), while transposition of replicative elements like Tn3 and IS1 is not affected by *dam* methylation (Heffron *et al*, 1979; Johnsrud, 1979; Ohtsubo and Ohtsubo, 1978). Roberts *et al*, (1985) have shown in a very elegant manner, that IS10 transposition is

coupled to normal chromosomal DNA replication. According to their hypothesis, transposition is low during most of the cell cycle since DNA is fully methylated. Immediately after passage of the replication fork, there will briefly be two hemimethylated copies of the element, which are methylated on opposite strands. Transposase will not be made in all cells (Raleigh and Kleckner, 1986), but in the cells in which it is synthesised, the more highly activated element will transpose. Methylation will then occur, reducing expression and transposition to a low level.

Methylation does not seem to be a general mechanism employed by non-replicative elements to regulate their transposition. Tn7, appears to transpose conservatively (absence of evidence for cointegrate formation, replicative deletions and inversions) and independently of the methylation state of DNA. Some indication of the control of Tn7 transposition was obtained and discussed in chapters one and two. The number of proteins required for transposition of the element (more than that of any transposon studied to date) and the fact that a 8kb region is used in transposition, indicate that Tn7 transposition is complex, occuring at different levels and involving a number of host factors.

In conclusion, the results described in this chapter provide no evidence for involvement of the DNA adenine methylase in the regulation of Tn7 transposition. Transposition was found to occur at similar frequencies in unmethylated, hemimethylated and fully methylated hosts. Furthermore, no difference between the transposition frequencies of the two hemimethylated forms of Tn7 was detected. It is important to stress, however, that data presented in this study do not necessarily rule out the possibility of involvement of *dam* methylation in the transposition of Tn7. It is conceivable that methylation plays an important role in transposition of the element that is impossible to be identified under the experimental conditions studied here. For example, it is possible to underestimate the transposition frequency in a *dam*$^-$ host if transposition is also dependent on another host factor that is limiting in that cell.

CHAPTER 6

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The first aim of the work presented in this thesis was to identify and characterise the functions involved in Tn7 transposition and its control. Five proteins were detected within the 7.9kb right part of the element and their levels of transcription and translation were measured. The locations of these proteins were in total agreement with the position of the five genes on the complementation map (Rogers et al, 1986). Two promoters were identified, one of which, P_1, appeared to be the major Tn7 promoter, initiating transcription of all five genes. All but one of the Tn7 transposition proteins appeared to be present in low amounts within the cell. This did not seem surprising, since large quantities of transposition proteins would result in elevated transposition frequencies and thus increased detrimental risks for the host. The second promoter, P_2, detected by the transcriptional fusion analysis, seemed to be present to allow efficient transcription of the tnsB gene, the most efficiently translated Tn7 transposition gene.

The major result of the fusion analysis was the finding that the presence of tnsB, one of the proteins involved in both modes of Tn7 transposition, in trans with P_1, resulted in a significant decrease in promoter activity. This provided us with a first notion concerning the regulation of Tn7 transposition and assigned a function to the tnsB gene product. Since the -35 region of P_1 is contained within one of the repeats at the right end of the element, it was reasoned that the observed repression of transcription was due to the binding of the tnsB protein to the repeats in the right end terminus, thereby indirectly blocking transcription. The latter hypothesis was tested in chapter 4.

Evidence provided in chapter 4 demonstrated the specific binding of the tnsB gene product to the right end of Tn7. The fragment used to test the binding properties of tnsB contained all four repeats at the right end terminus. The complexity of the binding profile obtained suggested that, most likely, all of these sites were occupied by tnsB molecules. However, the system used in this study to detect the binding could not provide us with

information concerning the number and the specific binding sites for tnsB in the right end terminus. To clarify the nature of the binding reaction and physically locate the binding sites, footprinting experiments need to be performed. Future workers, can capitalise on previous work carried out in this lab with res/resolvase (Brown, 1986) in which in vitro "footprinting" (methylation protection on linear substrates and UV photofootprinting on circular supercoiled substrates), was used to examine specific protein-DNA interactions.

Since both ends of the element are most probably brought together by the transposition complex during initiation of transposition, it is believed that the left terminus of Tn7 is also recognised by the tnsB gene product. It would be interesting to see whether tnsB exhibits a higher affinity for either of the ends of the element. If this turns out to be the case, it is expected that such increased binding affinity will be displayed for the right end, since all functions essential for transposition are carried in the right part of Tn7.

Preliminary information suggested that none of the other proteins involved in Tn7 transposition bind the right end of the element. These experiments need to be repeated since other reasons, such as inability to maintain active forms of the proteins in the extracts, might have resulted in the failure to observe binding. In addition, the assay which was used in this work to detect binding relied on the occurence of conformational changes of DNA upon binding. Binding of proteins to DNA, however, does not necessarily result in bending of the DNA chain. DNA footprinting analysis can provide information on protein-DNA interactions that do not induce conformational changes in the DNA, and therefore cannot be detected by their decreased electrophoretic mobility on non-denaturing gels.

The binding of the tnsD gene product to the Tn7 chromosomal attachement site was also discussed in chapter 4. This binding event may be one of the prime steps in the "hot site" transposition reaction. tnsD might be acting by recognising

the "hot site" and by bringing together the transposition complex, (consisting of the ends and the three proteins required for both modes of transposition), with this site. This seems to argue indirectly that a similar role can be assigned to tnsE, the "cold site" specific protein. tnsE could be a DNA-binding protein exhibiting no site specificity, since no sequence homology between "cold sites" has been detected so far (Rogers, 1986).

No function has yet been assigned to the tnsA and tnsC gene products. It is likely that both of these proteins are participating in the initiation of transposition by interacting either directly with the ends of the element or with the tnsB-DNA complex. Sequencing data that has recently been made available to us (C. Lichtenstein, pers. comm.) reveals the existence of a good consensus ATP binding site within the tnsC coding region. This immediately seems to implies the involvement of ATP in the transposition reaction and it should be kept in mind when an in vitro system is about to be set up.

The most efficient way of examining the molecular consequences, DNA intermediates and enzyme requirements for transposition is the development of an in vitro assay system. Such a system has the additional advantage of being able to freeze the transposition process at various stages, thus enabling the analysis of the DNA structure of intermediates of the reaction.

Tn7 is unique among transposable elements in many respects, and most importantly for utilising an 8kb region for transposition. The five genes required for both modes of transposition occupy most of this region according to the complementation map and the sequencing data (Rogers et al, 1986; C. Lichtenstein, pers. comm.). The reason for employing four genes for each mode of transposition is still unclear. Data provided in this study, suggest that tnsD is involved in recognition of the target site during "hot site" transposition and it is possible that tnsE has an analogous role upon "cold site" transposition. tnsB seems to be acting by repressing P_1 activity and thereby keeping

expression of the rest of the transposition proteins in low amounts.

Probably the most intriguing property of Tn7 is its ability to insert to a preferred site in the chromosome of E.coli at an unusually high frequency. This preferred attachment site is also conserved in other genera (C. Lichtenstein, pers. comm.) and is responsible for the stable maintenance of Tn7 among bacterial populations. This property allows the element to transpose at a high frequency without running the risk of causing detrimental mutations on its host. Being inserted in a bacterial chromosome means that a transposable element can be stably inherited, whereas a transposon inserted in a plasmid can be lost from a strain if the plasmid is unstable.

Another interesting aspect of the transposition of Tn7 is its orientational specificity upon insertion into plasmids. Almost all insertions occur in a specific orientation relative to the restriction map of the transposon and the plasmid. The mechanism by which Tn7 is able to sense the orientation of the recipient plasmid remains a mystery. It is unlikely to be the replication origin of the plasmid, since insertions in a specific orientation occur in plasmids with bidirectional origins of replication. Also, it cannot be a specific sequence in the recipient plasmid because this sequence would have to be complex enough to be present only once in large conjugative plasmids and simple enough to be present in small plasmids.

In chapter 5 the role of DNA adenine methylase in Tn7 transposition was examined. The experimental approach used provided no evidence supporting the involvement of dam methylation in Tn7 transposition. This together with our view that Tn7 uses a conservative mechanism to transpose seems to indicate that methylation does not constitute a general mechanism for regulating transposition of non-replicative elements. However, the data presented in this study should be interpreted with caution since no distinction between the two modes of Tn7 transposition was made.

Some preliminary experiments were undertaken to examine the role of the E.coli integration host factor (IHF) in Tn7 transposition. In vivo transposition of Tn7 to "hot" and "cold" sites was monitored in isogenic IHF^+ and IHF^- hosts. The preliminary data obtained appeared to suggest that "cold site" transposition was approximately 100-fold reduced in the absence of IHF, while "hot site" transposition seemed unaffected. On the basis of these data alone no valid conclusions concerning the role of IHF in transposition can be made. It is however a result to be kept in mind in the preparation of an in vitro transposition system.

A lot more work is needed before we are in a position of fully understanding the mechanism of Tn7 transposition. Data elucidating the different steps taking place in the transposition process are expected to be mainly derived from in vitro studies using purified components of the transposition reaction. The unravelling of the transposition process will throw light on other processes taking place in E.coli, such as protein-DNA interactions, protein-protein interactions and site-specific recombination.

BIBLIOGRAPHY

Akroyd JE, Symonds N (1983) Evidence for a conservative pathway of transposition of bacteriophage Mu. Nature **303**:84-86

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**:167-178

Arthur A, Sherratt DJ (1979) Dissection of the transposition process: A transposon-encoded site-specific recombination system. Mol. Gen. Genet. **175**:267-274

Bachmann BJ, Low KB, Taylor A (1976) Recalibrated linkage map of Escherichia coli K-12. Bacteriol. Rev. **40**:116-167

Barth PT, Datta N, Hedges RW, Grinter NJ (1976) Transposition of a deoxyribonucleic acid sequence encoding trimethoprim and streptomycin resistance from R483 to other replicons. J. Bacteriol. **125**:800-810

Barth PT, Datta N (1977) Two naturally occurring transposons indistinguishable from Tn7. J. Gen. Mircobiol. **102**:129-134

Barth PT, Grinter NJ (1977) Map of RP4 derived by insertion of transposon C. J. Mol. Biol. **113**:455-474

Barth PT, Grinter NJ, Bradley DE (1978) Conjugal transfer system of plasmid RP4: Analysis by transposon 7 insertion. J. Bacteriol. **133**:43-52

Bassford P, Beckwith J, Berman M, Brinkman E, Casadaban M, Guarente L, Saint-Girons I, Sarthy A, Schwartz M, Shuman H, Silhavy, T (1978) Genetic fusions of the lac operon: a new approach to the study of biological processes, p.245-261. In The Operon, Eds. Miller J, Reznikoff W, Cold Spring Harbor, New York, pp245-261

Becker MM, Wang JC (1984) Use of light for footprinting DNA in vivo. Nature **309**:682-687

Bender J, Kleckner N (1986) Genetic evidence that Tn10 transposes by a nonreplicative mechanism. Cell **45**:801-815

Berg DE, Davies J, Allet B, Rochaix J-D (1975) Transposition of R factor genes to bacteriophage lambda. Proc. Natl. Acad. Sci. USA **72**:3628-3632

Berg DE (1977) Insertion and excision of the transposable kanamycin resistance determinant Tn5. In DNA Insertion Elements, Plasmids and Episomes, Eds. AI Bukhari, JA Shapiro, SL Adhya, Cold Spring Harbor press pp205-212

Berg DE, Weiss A, Crossland L (1980) The polarity of Tn5 insertion mutations in Eschrichia coli. J. Bacteriol. **142**:439 -446

Berg DE, Johnsrud L, McDivitt L, Ramaohadran R, Hirschel BJ (1982a) The inverted repeats of Tn5 are transposable elements. Proc. Natl. Acad. Sci. USA **79**:2632-2635

Berg DE (1983) Structural requirement for IS50-mediated gene transpositions. Proc. Natl. Acad. Sci. USA **80**:792-796

Berg DE, Lodge J, Sasakawa C, Nag DK, Phadnis SH, Weston-Hafer K, Carle GF (1984) Transposon 5: Specific sequence recognition and conservative transposition. Cold Spring Harbor Symp. Quant. Biol. **49**:215-226

Biek D, Roth JR (1980) Regulation of Tn5 transposition in Salmonella typhimurium. Proc. Natl. Acad. Sci. USA **77**:6047 -6051

Biel SW, Hartl DL (1983) Evolution of transposons: Natural selection for Tn5 in E.coli K-12. Genetics **103**:581-592

Biel SW, Berg DE (1984) Mechanism of IS1 transposition in E.coli; choice between simple insertion and cointegration. Genetics **108**:319-330

Birmboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acid Res. **7**:1513-1523

Bishop R, Sherratt DJ (1984) Transposon Tn1 intra-molecular transposition. Mol. Gen. Genet. **196**: 117-122

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Anal. Biochem. **72**: 248-254

Braun R, Wright A (1986) DNA methylation differentially enhances the expression of one of the two E. coli dnaA promoters. Mol. Gen. Genet. **202**: 246-250

Brevet J, Faure F, Borowski D (1985) Tn7-encoded proteins. Mol. Gen. Genet. **201**:258-264

Brown JL (1986) Properties and action of Tn3 resolvase. D.Phil. University of Glasgow

Bukhari AI (1976) Bacteriophage Mu as a transposition element. Ann. Rev. Genet. **10**:389-412

Calos M, Miller J (1980) Transposable elements. Cell **20**: 579-595

Campbell A (1981) Evolutionary significance of accessory DNA elements in bacteria. Ann. Rev. Microbiol. **35**:55-83

Casadaban MJ, Cohen SN (1980) Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. J. Mol. Biol. **138**: 179-207

Casadaban MJ, Chou J, Cohen SN (1982) Overproduction of the Tn3 transposition protein and its role in DNA transposition. Cell **28**:345-354

Cathew RW, Chodosh LA, Sharp PA (1985) An RNA polymerase II

transcription factor binds to an upstream element in the adenovirus major late promoter. Cell 43:439-446

Chaconas G, Harshey RM, Sarvetrich J, Bukhari AI (1981) Predominant endproducts of prohage Mu DNA transposition during the lytic cycle are replicon fusions. J. Mol. Biol. 150:341-359

Chandler M, Clerget M, Galas DJ (1982) The transposition frequency of IS1-flanked transposons is a function of their size. J. Mol. Biol. 154:229-243

Chang ACY, Cohen SN (1978) Construction and characterisation of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol.141:1141-1156

Charlier D, Piette J, Glansdorff N (1982) IS3 can function as a mobile promoter in Escherichia coli. Nucl. Acid Res. 10: 5935-5948

Chou J, Casadaban M, Lemaux PG, Cohen SN (1979a) Identification of a self regulated repressor of translocation of the Tn3 element. Proc. Natl. Acad. Sci. USA 76:4020-4024

Chou J, Lemaux PG, Casadaban M, Cohen SN (1979b) Transposition protein of Tn3: Identification and characterisation of an essential repressor-controlled gene product. Nature 282:801-806

Close TJ, Rodriguez RL (1982) Construction and characterization of the chloramphenicol-resistance gene cartridge: A new approach to the transcriptional mapping of extrachromosomal elements. Gene 20: 305-316

Cohen SN, Casadaban M, Chou J, Tu C (1979) Studies of the specificity and control of transposition of the Tn3 element. Cold Spring Harbor Symp. Quant. Biol. 43:1247-1255

Cox EC (1976) Bacterial mutator genes and the control of spontaneous mutation. Annu. Rev. Genet. **10**:135-156

Craigie R, Mizuuchi M, Mizuuchi K (1984) Site specific recognition of the bacteriophage Mu ends by the Mu A protein. Cell **39**:387-394

Craigie R, Mizuuchi K (1985) Mechanism of transposition of bacteriophage Mu: Structure of the transposition intermediate. Cell **41**:867-876

Craigie R, Arndt-Jovin DJ, 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

Craigie R, Mizuuchi K (1986) Role of DNA topology in Mu transposition: Mechanism of sensing the relative orientation of two DNA segments. Cell **45**:793-800

Cullum J, Broda P (1979) Chromosome transfer and Hfr formation by F in *recA*$^+$ and *recA*$^-$ strains of *Escherichia coli* K-12. Plasmid **2**:358-365

Dandanell G, Hammer K (1985) Two operator sites separated by 599bp are required for *deoR* repression of the *deo* operon of *Escherichia coli*. EMBO J. **4**:3333-3338

Datta N, Hedges RW (1972) Trimethoprim resistance conferred by W plasmids in Enterobacteriacae. J. Gen. Micobiol. **72**:349-355

Datta N, Nugent M, Richards H (1980) Transposons encoding trimethoprim or gentamycin resistance in medically important bacteria. Cold Spring Harbor Symp. Quant. Biol. **45**:45-51

Diver WP, Grinstead J, Fritzinger DC, Brown NL, Altenbuchner J (1983) DNA sequence of and complementation by the *tnpR* genes of Tn21, Tn501 and Tn1721. Mol. Gen. Genet. **191**:189-193

Doolittle WF, Sapienza C (1980) Selfish genes, the phenotype paradigm and genome evolution. Nature 284:601-603

Doolittle WF (1982) Selfish DNA after fourteen months. In Genome Evolution, Eds. GA Dover RB Flavell Academic Press London pp3-28

Doring HP (1985) Plant Transposable Elements. Bioassays 3:164-166

Dunn TM, Hahn S, Ogden S, Schleif R (1984) An operator at -280 base pairs that is required for repression of araBAD operon promoter: Addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc. Natl. Acad. Sci. USA 81:5017-5020

Ely B (1982) Transposition of Tn7 occurs at a single site on the Caulobacter crescentus chromosome. J. Bacteriol. 151:1056-1058

Engler JA, van Bree MP (1981) The nucleotide sequence and protein coding capability of the transposable element IS5. Gene 14:155-163

Fling M, and Elwell L (1980) Protein expression in E.coli minicells containing recombinant plasmids specifying trimethoprim resistant dihydrofolate reductases. J. Bacteriol. 141:779-785

Fling M, Richards C (1983) The nucleotide sequence of the trimethoprim resistance gene harbored by Tn7. Nucl. Acid Res. 11:5147-5158

Foster TJ (1977) Insertion of the tetracycline resistance translocatable unit Tn10 in the lac operon of Escherichia coli K-12. Mol. Gen. Genet. 154:305-310

Foster TJ, Davis MA, Takeshita K, Roberts DE, Kleckner N (1981)

Genetic organisation of transposon Tn10. Cell 23:201-207

Fried M, Crothers DM (1981) Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucl. Acid Res. 9:6505-6525

Fried M, Crothers DM (1984) Equilibrium studies of the cAMP receptor-operator interaction. J. Mol. Biol. 172: 241-262

Fritz HJ, Bicknase H, Gleumes B, Heibach C, Rosahl S, Ehring R (1983) Characterisation of two mutations in the E.coli galE gene inactivating the second galactose operator and comparative studies of repressor binding. EMBO J. 2:2129-2135

Galas DJ, Chandler M (1981) On the molecular mechanisms of transposition. Proc. Natl. Acad. Sci. USA 78:4858-4862

Galas DJ, Chandler MA (1982) Structure and stability of Tn9-mediated cointegrates: Evidence for two pathways of transposition. J. Mol. Biol. 154:245-272

Galas DJ, Schmitz A (1978) DNAase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucl. Acid Res. 5:3157-3170

Gamas P, Chandler MG (1987) Escherichia coli integration host factor binds specifically to the ends of the insertion sequence IS1 and to its major insertion hot-spot in pBR322. J. Mol. Biol. 195: 261-272

Garner M, Revzin A (1986) The use of gel electrophoresis to detect and study nucleic acid-protein interactions. TIBS 11:395-396

Garner MM, 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. Acid Res. 9:3047-3059

Gay N, Tybulewicz VLJ, Walker JE (1986) Insertion of transposon Tn7 into E.coli. Biochem. J. 234:111-117

Gill R, Heffron F, Dougan G, Falkow S (1978) Analysis of sequences transposed by complementation of 2 classes of transposition deficient mutants of transposition element Tn3. J. Bacteriol. 136:742-756

Gill RE, Heffron F, 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, Radman M (1980) Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction. Proc. Natl. Acad. Sci. USA 77:1063-1067

Glickman BW, Vanden Elsen P, Radman M (1978) Increased mutagenesis in dam⁻ mutants of E.coli, a role of 6-methyladenine residues in mutation avoidance. Mol. Gen. Genet. 163:307-321

Gosti-Testu F, Norris V, Brevet J (1983) Restriction map of Tn7. Plasmid 10:96-99

Grindley NDF, Sherratt DJ (1978) Sequence analysis at IS1 insertion sites: Models for transposition. Cold Spring Harbor Symp. Quant. Biol. 43:1257-1262

Grindley N, Joyce C (1980) Genetic and DNA sequence analysis of the kanamycin resistance transposon Tn903. Proc. Natl. Acad. Sci. USA 77:7176-7180

Grindley N, Joyce C (1981) Analysis of the structure and function of the kanamycin resistance transposon Tn903. Cold Spring Harbor Symp. Quant. Biol. 45:125-133

Grindley NDF, Reed RR (1985) Transpositional recombination in prokaryotes. Ann. Rev. Biochem. 54:863-896

Grinstead J, de la Cruz F, Altenbuchner J, Schmitt R (1982) Complementation of transposition of tnpA mutants of Tn3, Tn21, Tn501 and Tn1721. Plasmid 8:276-286

Grinter NJ (1983) A broad-host-range cloning vector transposable to various replicons. Gene 21:133-143

Groenen MAM, Timmers E, van de Putte P (1985) DNA sequences at the end of the genome of bacteriophage Mu essential for transposition. Proc. Natl. Acad. Sci. USA 82:2087-2091

Halling SM, Kleckner N (1982) A symmetrical six-base pair target site sequence determines Tn10 insertion specificity. Cell 28:155-163

Harayama S, Oguchi T, Iino T (1984) Does Tn10 transpose via a cointegrate molecule? Mol. Gen. Genet. 194:444-450

Harshey R, Bukhari A (1981) A mechanism of DNA transposition. Proc. Natl. Acad. Sci. USA 78:1090-1094

Harshey R (1984) Nonreplicative DNA transposition: Integration of infecting bacteriophage Mu. Cold Spring Harbor Symp. Quant. Biol. 49:273-278

Hassan D, Brevet J (1983) Absence of cis-acting transposition immunity with Tn7. Plasmid 10:31-44

Hauer B, Shapiro J (1984) Control of Tn7 transposition. Mol. Gen. Genet. 194:149-158

Hedges R, Datta N, Flemming P (1972) R factors conferring resistance to trimethoprim but not to sulphonamides. J. Gen. Mircobiol. 73:573-575

Heffron F, Rubens C, Falkow S (1975a) The translocation of a DNA sequence which mediates ampicillin resistance: molecular nature and specificity of insertion. Proc. Natl. Acad. Sci.

USA **72**:3623-3627

Heffron F, Bedinger P, Champoux JJ, Falkow S (1977) Deletions affecting the transposition of an antibiotic resistance gene. Proc. Natl. Acad. Sci. USA **74**:702-706

Heffron F, McCarthy BJ, Ohtsubo H, Ohtsubo E (1979) DNA sequence analysis of the transposon Tn3. Three genes and three sites involved in the transposition of Tn3. Cell **18**:1153-1164

Heidecker G, Muller-Hill B (1977) Synthetic multifunctional proteins. Mol. Gen. Genet. **155**: 301-307

Hendrickson W, Schleif R (1985) A dimer of AraC protein contacts three adjacent major groove regions fo the araI site. Proc. Natl. Acad. Sci. USA 82:3129-3133

Hodge DM (1983) Transposition of the transposon Tn7. D.Phil. University of Sussex

Holmes DS, Quigley M (1981) A rapid boiling method for the preparation of bacterial plasmids. Analytical Biochem. **114**: 193-197

Howe M, Bade E (1975) Molecular biology of bacteriophage Mu-1. Science **190**:624-632

Howe M, Schumm JW (1980) Transposition of bacteriophage Mu: Properties of lambda phages containing both ends of Mu. Cold Spring Harbor Symp. Quant. Biol. **45**:337-346

Hyde DR, Tu CD (1985) tnpM: A novel regulatory gene that enhances Tn21 transposition and suppresses cointegrate resolution. Cell **42**:629-638

Iida S, Meyer J, Arber W (1983) Prokaryotic IS elements. In Mobile Genetic Elements, Eds: Shapiro JA, Academic Press, New York, pp.159-221

Isberg RR, Syvanen M (1981) Replicon fusions promoted by the inverted repeats of Tn5. J. Mol. Biol. **150**:15-32

Isberg RR, Lazaar AM, Syvanen M (1982) Regulation of Tn5 by the right repeat protein: Control at the level of the transposition reaction. Cell **30**:883-892

Johnson RC, Yin JCP, Reznikoff WS (1982) Control of Tn5 transposition in Escherichia coli is mediated by protein from the right repeat. Cell **30**:873-882

Johnsrud L (1979) DNA sequence of the transposable element IS1. Mol. Gen. Genet. **169**: 213-218

Jones OW, Berg P (1966) Studies on the binding of RNA polymerase to polynucleotides. J. Mol. Biol. **22**: 199-209

Kahmann R, Kamp D (1979) Nucleotide sequence of the attachment sites of the bacteriophage Mu DNA. Nature **280**:247-250

Kitts PA, Symington L, Burke M, Reed RR, Sherratt DJ (1981) Transposon-encoded site-specific recombination. Proc. Natl. Acad. Sci. USA **79**:46-50

Kitts PA, Lamond A, Sherratt D (1982) Inter-replicon transposition of Tn1/3 occurs in two genetically distinguishable steps. Nature **295**:626-628

Klaer R, Starlinger P (1980) IS4 at its chromosomal site in E.coli K12. Mol. Gen. Genet. **178**:285-291

Klaer R, Pfeifer D, Starliger P (1980) IS4 is still at its chromosomal site after transposition to galT. Mol. Gen. Genet. **178**:281-284

Kleckner N (1981) Transposable genetic elements in prokaryotes. Ann. Rev. Genet. **15**:341-404

Kleckner N (1983) Transposon Tn10. In Mobile Genetic Elements Eds. JA Shapiro, Academic press pp 261-299

Kleckner N, Morisato D, Roberts D, Bender J (1984) Mechanism and regulation of Tn10 transposition. Cold Spring Harbor Symp. Quant. Biol. **49**:235-244

Kramer H, Niemoller M, Amouyal M, Revet B, von Wilcken-Bergamann B, Muller-Hill B (1987) lac repressor forms loops with linear DNA carrying two suitably spaced lac operators. EMBO J. **6**:1481-1491

Krolewski JJ, Murphy E, Novick RP, Rush MG (1981) Site-specificity of the chromosomal insertion of Staphylococcus aureus transposon Tn554. J. Mol. Biol. **152**:19-33

Lacks S, Greenberg B (1977) Complementary specificity of restriction endonuclease of Diplococcus pneumoniae with respect to DNA methylation. J. Mol. Biol.**114**:153-168

Laemmli UK (1970) Cleavage and structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685

Lichtenstein C, Brenner S (1981) Site specific properties of Tn7 transposition into the E.coli chromosome. Mol. Gen. Genet. **183**:380-387

Lichtenstein C, Brenner S (1982) Unique insertion site of Tn7 In the E.coli chromosome. Nature **297**:601-603

Lieb M (1981) A fine structure map of spontaneous and induced mutations in the lambda repressor gene, including insertions of IS elements. Mol. Gen. Genet. **184**: 364-371

Liebart JC, Ghelardini P, Paolozzi P (1982) Conservative integration of bacteriophage Mu DNA into pBR322 plasmid. Proc. Natl. Acad. Sci. USA **79**:4362-4366

Liu-Johnson HN, Gartenberg MR, Crothers D (1986) The DNA binding domain and bending angle of the E.coli CAP protein. Cell 47: 995-1005

Lundblad V, Kleckner N (1985) Mismatch repair mutations of Escherichia coli K12 enhance transposon excision. Genetics 109:3-19

Lyons SM, Schendel PF (1984) Kinetics of methylation in E.coli K-12. J. Bacteriol. 159:421-423

Machida C, Machida Y, Ohtsubo E (1984) Both inverted repeat sequences located at the ends of IS1 provide promoter functions. J. Mol. Biol. 177:247-267

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory; New York

Marinus MG (1985) DNA methylation influences trpR promoter activity in Escherichia coli K-12. Mol.Gen.Genet.200:185-186

Marinus MG (1984) Methylation of prokaryotic DNA. In DNA Methylation: Biochemistry and Bioligical Significance, Razin AD, Cedar H, Riggs AD, eds. New York: Springer-Verlag, pp. 81-109

Marinus MG, Morris R (1973) Isolation of deoxyribonucleic acid methylase mutants of Escherichia coli K-12. J.Bacteriol. 114:1143-1150

Marinus MG, Morris R (1975) Pleiotropic effects of a DNA adenine methylation mutation (dam-3) in Escherichia coli K12. Mutation Res. 28:15-26

McClintock B (1950) The origin and behaviour of mutable loci in maize. Proc. Natl. Acad. Sci. USA 36:344-355

McClintock B (1951) Chromosome organisation and gene expression.

Cold Spring Harbor Symp. Quant. Biol. **16**:13-47

McCormick M, Wishart W, Ohtsubo H, Heffron F, Ohtsubo E (1981) Plasmid cointegrates and their resolution mediated by transposon Tn3 mutants. Gene **15**:103-118

McGraw BR, Marinus MG (1980) Isolation and characterization of *dam*$^+$ revertants and suppressor mutations that modify secondary phenotypes of *dam*-3 strains of *Escherichia coli* K-12. Mol. Gen. Genet. **178**:309-315

McHattie LA, Jackowski JB (1977) Physical structure and deletion of the chloramphenicol resistance element in Tn9 in phage lambda. In DNA Insertion Elements, Plasmids and Episomes, Eds: Bukhari A, Shapiro JA, Adhya S, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp.219-228

McKenney K, Shimatake H, Court D, Schmeissner U, Brady C, Rosenberg M (1981) A system to study promoter and terminator signals recognised by the *E.coli* RNA polymerase. In Gene Amplification and Analysis, Vol.II, Eds. Chirikjian JC, Papas TS, Elsevier/North Holland, Amsterdam, pp.383-415

Messing J (1983) New M13 vectors for cloning. Methods Enzymol. **101**: 20-78

Miller JH (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York.

Miller JH, Calos MP, Galas D, Hofer M, Buchel DE, Muller-Hill B (1980) Genetic analysis of transposition in the *lac* region of *E. coli*. J. Mol. Biol. **144**: 1-18

Minton N (1984) Improved plasmid vectors for the isolation of translational *lac* gene fusions. Gene **31**:1-4

Mizusawa S, Court D, Gottesman S (1983) Transcription of the *sulA* gene and repression of *lexA*. J. Mol. Biol. **171**:337-343

Mizuuchi K (1983) In vitro transposition of bacteriophage Mu: A biochemical approach to a novel replication reaction. Cell 35:785-794

Moore RJ, Krishnapillae V (1982) Tn7 and Tn501 insertions into Pseudomonas aeruginosa plasmid R91-5: Mapping of two transfer regions. J. Bacteriol. 149:276-283

Morisato D, Way JC, Kim HT, Kleckner N (1983) Tn10 transposition acts preferentially on nearby transposon ends in vivo. Cell 32:799-807

Morisato D, Kleckner N (1984) Transposase promotes double strand breaks and single strand joints at Tn10 termini in vivo. Cell 39:181-190

Muller-Hill B, Kania J (1974) Lac repressor can be fused to b-galactosidase. Nature 249:561-562

Munson LM, Stormo GD, Niece RL, Reznikoff WS (1984) lacZ translation initiation mutations. J. Mol. Biol. 177:663-683

Murphy E, Lofdahl S (1984) Transposition of Tn554 does not generate a target duplication. Nature 307:292-294

Muster CJ, Shapiro JA (1980) Recombination involving transposable elements: On replicon fusions. Cold Spring Harbor Symp. Quant. Biol. 45:239-244

Noel KD, Ames GF (1978) Evidence for a common mechanism for the insertion of the Tn10 transposon and for the generation of Tn10-stimulated deletions. Mol. Gen. Genet. 166:217-223

Novick RP, Clones RC, Cohen SN, Curtiss R, Datta N, Falkow S (1976) Uniform nomenclature for bacterial plasmids: A proposal. Bacteriol. Rev. 40:168-189

Nyman K, Nakmura K, Ohtsubo H, Ohtsubo E (1981) Distribution of

the insertion sequence IS1 in gram-negative bacteria. Nature **289**:609-612

O'Callaghan C, Morris A, Kirby SM, Shingler AH (1972) Novel method for the detection of b-lactamases by using a chromogenic cephalosporin substrate. Antimicrobial Agents and Chemotherapy **1**: 283-288

Ohtsubo E, Zenilman M, Ohtsubo H, McCormick M, Machida C, Machida Y (1980b) Mechanism of insertion and cointegration mediated by IS1 and Tn3. Cold Spring Harbor Symp. Quant. Biol. **45**:283-295

Ohtsubo H, Ohtsubo E (1978) Nucleotide sequence of the insertion element IS1. Proc. Natl. Acad. Sci. USA **75**:615-619

Orgel LE, Crick FHC (1980) Selfish DNA: The ultimate parasite. Nature **284**:604-607

Ouartsi A, Borowski D, Brevet J (1985) Genetic analysis of Tn7 transposition. Mol. Gen. Genet. **198**:221-227

Pato ML, Reich C (1982) Instability of transposase activity: Evidence from bacteriophage Mu replication. Cell **29**:219-225

Pato ML, Reich C (1984) Stoicheiometric use of the transposase of bacteriophage Mu. Cell **36**:197-202

Pfeifer D, Habermann P, Kubai-Maroni D (1977) Specific sites for integration of insertion elements within the transferase gene of the *gal* operon of *E.coli* K12. In DNA Insertion Elements, Plasmids and Episomes. Eds. A Bukhari, JA Shapiro, S Adhya, Cold Spring Harbor press, pp31-36

Plasterk RHA, Vrieling H, Van de Putte P (1983) Transcription initiation of Mu *mom* depends on methylation of the promoter region and a phage-coded transactivator. Nature **301**:344-347

Pukkila PJ, Peterson J, Herman G, Modrich P, Meselson M (1983)

Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in Escherichia coli. Genetics 104:571-582

Raleigh EA, Kleckner N (1986) Quantitation of insertion sequence IS10 transposase gene expression by a method generally applicable to any rarely expressed gene. Proc. Natl. Acad. Sci. USA 83:1787-1791

Razin A, Urieli S, Pollack Y, Gruenbaum Y, Glaser G (1980) Studies on the biological role of DNA methylation; IV. Mode of methylation of DNA in E.coli cells. Nucl. Acid Res. 8:1783-1782

Reznikoff WS (1982) Tn5 transposition and its regulation. Cell 31:307-308

Reznikoff WS, McClure WR (1986) E.coli promoters. In Maximizing Gene Expression, Eds. W Reznikoff, L Gold, Butterworths

Roberts DE, Hoopes BC, McClure WR, Kleckner N (1985) IS10 transposition is regulated by adenine methylation. Cell 43: 117-130

Robinson MK, Bennett PM, Richmond MH (1977) Inhibition of TnA transposition by TnA. J. Bacteriol. 129:407-414

Rogers MS, Ekaterinaki N, Nimmo E, Sherratt DJ (1986) Analysis of Tn7 transposition. Mol. Gen. Genet. 205:550-556

Rogers MS (1986) Transpositional properties and organisation of Tn7. D.Phil. University of Glasgow

Rosner JL, Guyer MG (1980) Transposition of insertion sequences IS1 lambda bio IS1 from a bacteriophage lambda derivative carrying the IS1 cat IS1 transposon Tn9. Mol. Gen. Genet. 178:111-120

Rothstein SJ, Reznikoff WS (1981) The functional differences in the inverted repeats of Tn5 are caused by single base pair nonhomology. Cell **23**:191-199

Rubens C, Heffron F, Falkow S (1976) Transposition of a plasmid DNA sequence that mediates ampicillin resistance: independence from host rec functions and orientation of insertion. J. Bacteriol. **128**:425-434

Sarthy A, Fowler A, Zabin I, Beckwith J (1979) Use of gene fusions to determine a partial signal sequence of alkaline phosphatase. J. Bacteriol. **139**:932-939

Shapiro JA (1979) Molecular model for the transposition and replication of bacteriophage Mu and other transposable element. Proc. Natl. Acad. Sci. USA **76**:1933-1937

Shapiro JA (1983) Mobile Genetic Elements. Academic press

Sherratt D, Arthur A, Burke M (1980) Transposon-specified site-specfic recombination systems. Cold Spring Harbor Symp. Quant. Biol. **45**:275-281

Simonsen C, Chen E, Levenson A (1983) Identification of the type I trimethoprim-resistant dihydrofolate reductase specified by the E.coli R-plasmid R483: Comparison with prokaryotic and Eukaryotic dihydrofolate reductases. J. Bacteriol. **155**:1001-1008

Simons RW, Kleckner N (1983) Translational control of IS10 transposition. Cell **34**:683-691

Smith GM, Jones P (1984) Effects of deletions in transposon Tn7 on its frequency of transposition. J. Bacteriol. **157**:962-964

Smith DW, Garland AM, Herman G, Enns RE, Baker TA, Zyskind JW (1985) Importance of state of methylation of oriC GATC sites in initiation of DNA replication in Escherichia coli.

EMBO J. 4:1319-1326

Sternberg N (1985) Evidence that adenine methylation influences DNA-protein interactions in E.coli. J. Bacteriol. 164:490-493

Strauss F, 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

Szyf M, Gruenbaum Y, Urieli-Shoval S, Razin A (1982) Studies on the biological role of DNA methylation: V. The pattern of E.coli DNA methylation. Nucl.Acid Res. 10:7247-7258

Temin HM (1980) Origin of retroviruses from cellular moveable genetic elements. Cell 21:599-600

Thomson JA, Henderson M, Megnes RM (1981) Mutagenesis by insertion of drug resistance transposon Tn7 into a Vibrio species. J. Bacteriol. 148:374-378

Tolias PP, Dubow MS (1985) The cloning and characterization of the bacteriophage D108 regulatory DNA-binding protein ner. EMBO J. 4:3031-3037

Tolias PP, Dubow MS (1986) The overproduction and characterization of the bacteriophage Mu regulatory DNA-binding protein ner. Virology 148:298-311

Toussaint A, Resibois A (1983) Phage Mu: Transposition as a life style. In Mobile Genetic Elements. Eds. JA Shapiro, pp105-158

Treismann R (1986) Identification of a protein-binding site that mediates transcriptional response of the c-fos gene to serum factors. Cell 46:567-574

Tu CD, Cohen SN (1980) Translocation specificity of Tn3:

Characteristics of sites of multiple insertion. Cell **19**:151-160

Tuggle CK, Fuchs JA (1986) Regulation of the operon encoding ribonucleotide reductase in E.coli: Evidence for both positive and negative control. EMBO J. **5**:1077-1085

Turner P, Barber C, Daniels M (1984) Behaviour of the transposons Tn5 and Tn7 in Xanthomonas campestris pv. campestris. Mol. Gen. Genet. **195**:101-107

Varshavsky A, Bakayev VV, Georgiev GP (1976) Heterogeneity of chromatin subunits and location of histone H1. Nucl. Acid Res. **3**: 477-492

Wallace LJ, Ward JM, Bennett PM, Robinson MK, Richmond MH (1980) Transposition immunity. Cold Spring Harbor Symp. Quant. Biol. **45**:183-188

Watson JD (1975) Molecular Biology of the Gene. WA Benjamin Inc.

Wu HM, Crothers DM (1984) The locus of sequence-directed and protein-induced DNA binding. Nature **308**:509-513

Zerbib B, Jakowec M, Prentki P, Galas D, Chandler M (1987) Expression of proteins essential for IS1 transposition: Specific binding of InsA to the ends of IS1. EMBO J. in press

GLASGOW UNIVERSITY LIBRARY