EXPRESSION OF TN1/3 TRANSPOSASE.

A thesis submitted for the degree of Doctor of

Philosophy at the University of Glasgow.

by

Simon John Herald Hettle.

Institute of Genetics,

Church Street,

Glasgow.

December, 1985.

For Diane

and our families.

Acknowledgements.

I thank my supervisor, Professor David Sherratt, for his help and enthusiasm and all members of the Genetics Department here in Glasgow, past and present, who have helped me in any way. Thanks in particular go to Miss Mary Burke for all her technical advice and a lot of her solutions, to Mark Rodgers, Dr.Chris Boyd and Dr.George Russell for their help with computing, to the 'media' ladies for their assistance, to Dr.David Summers and to Christine, especially for when I first came to Glasgow. Thank you too to Dr.David Onions of the Department of Veterinary Pathology, University of Glasgow. George gets a second mention in being thanked for proof-reading this thesis and I am really grateful too to Mrs. Gillian Hutchison for all her help with the typing. I was supported by a Glasgow University Research Scholarship for which I am grateful.

Further afield, thanks to all my other friends, especially to Dr.David Jackson of the Department of Anatomy and Experimental Pathology in the University of St.Andrews and to Dr. and Mrs. Alastair Mc.Blane for their constant friendship and support.

I could never express adequately my gratitude towards my mother, late father and my grandmother for all their love and support and to my wife's parents - Mr. and Mrs.D.B.R. Gray - for all they have done for me since I came to know them.

Finally, an especial expression of thanks to my wife Diane for all her love, encouragement and so much more.

Contents.

ABBREVIATIONS.

(i). Chemicals.

Ac - acetate.

APS - ammonium persulphate.

ATP - adenosine triphosphate.

BSA - bovine serum albumin.

CTP - cytosine triphosphate.

DNA - deoxyribonucleic acid.

d(NTP) - 2'-deoxy (nucleotide).

dd(NTP) - 2',3'-dideoxy (nucleotide).

DTT - dithiotreitol.

EDTA - ethylene diamine tetra-acetic acid (disodium salt, dihydrate).

EtBr - ethidium bromide.

EtOH - ethanol.

FSB - Final Sample Buffer.

GTP - guanosine triphosphate.

IPTG - isopropyl B-D thiogalactopyranoside.

mRNA - messenger ribonucleic acid.

PEG - polyethylene glycol 6000.

RNA - ribonucleic acid.

RNase - ribonuclease.

SDS - sodium dodecyl sulphate.

TEMED - N N N' N' tetramethylethylenediamine.

Tris - tris (hydroxymethyl) amino ethane.

TTP - thymidine triphosphate.

tRNA - transfer ribonucleic acid.

X-gal - 5-bromo, 4-chloro, 3-indolyl B-D galactoside.

(ii). **Antibiotics.**

Ap - ampicillin.

Cm - chloramphenicol.

Rif - rifampicin.

Spc - spectinomycin.

Str - streptomycin (chromosomal resistance).

Tc - tetracycline.

Tp - trimethoprim.

(iii). **Phenotype.**

X^r - resistance to X.

X^s - sensitivity to X.

ori - origin of replication.

res - resolution site.

Tra^+ - self-transmissible.

(v). **Measurements.**

mA. - milliamps. (10^{-3} amps).

bp. - base pair(s).

kb. - kilobase pair(s). (10^{3} bp.).

^{o}C. - degrees Celsius.

Ci. - Curie.

uCi. - microCurie. (10^{-6} Curie).

kd. - kilodaltons. (10^{3} daltons).

g. - centrifugal force equal to gravitational acceleration.

g. - gramme.

mg. - milligramme. (10^{-3} g.).

ug. - microgramme. (10^{-6} g.).

ng. - nanogramme. (10^{-9} g.).

l. - litre.

ml. - millilitre. (10^{-3} l.).

ul. - microlitre. (10^{-6} l.).

M. - Molar. (Moles per litre).

mM. - millimolar. (10^{-3} M.).

uM. - micromolar. (10^{-6} M.).

mmole. - millimole. (10^{-3} mole).

pmole. - picomole. (10^{-12} mole).

m. - metre.

cm. - centimetre. (10^{-2} m.).

mm. - millimetre. (10^{-3} m.).

nm. - nanometre. (10^{-9} m.).

mins. - minutes.

pH - acidity. ($-\log_{10}$[molar concentration H^+ ions]).

sec. - seconds.

V. - volts.

W. - Watts.

(v). Miscellaneous.

CAT - chloramphenicol acetyl transferase.

← - direction of transcription.

fig. - figure.

IS. - insertion sequence.

LH. - left hand.

log. - logarithm.

no. - number.

oc. - open circular plasmid.

OD_{650}. - optical density at a wavelength of 650nm.

RF. - replicative form.

RH. - right hand.

sc. - supercoiled plasmid.

Tn. - transposon.

w/v. - weight per volume.

(All percentage - % - solutions are weight per volume).

v/v. - volume per volume.

wt. - wild type.

Summary.

It has previously been shown that the only sequences near the termini of Tn1/3 essential for the first step of the transposition process of these elements are the two 38bp. perfect inverted repeats at the extreme ends of the elements. This finding has been confirmed and it has also been shown that both of these sequences are equally good substrates for transposition. Also, it was shown that Tn1/3 transposition normally acts to replicate the DNA sequences that correspond to the entire transposon sequences. Further indirect evidence for the binding of Tn1 transposase to the inverted repeat sequences of Tn1/3 was obtained, supporting the hypothesis that it is to and upon these sequences that transposase binds and acts.

The location of the HincII site present in Tn1 but absent in Tn3 has also been determined by DNA sequencing and found to be at a position in Tn1 equivalent to co-ordinates 3037-3042 in Tn3, with cleavage occurring between base pairs 3039 and 3040. Thus, this site lies just within the tnpA gene coding region, near its N-terminal end. The sequence of 114bp. of Tn1 lying to the left of this site (i.e. extending towards the C-terminus of the tnpA gene) was determined and was found to have 97/114 bp. (=85.1%) homology at the DNA level and 37/38 residue (=97.4%) homology at the amino acid level with Tn3. This confirmed the prediction of a high degree of homology between the two elements, which had been made on the basis of their very similar restriction profiles.

Using this information, a plamid (pSN015) was constructed that produced a fusion protein consisting of the whole Tn1 transposase protein with its three N-terminal amino acid residues removed and replaced by the first eleven N-terminal amino acid residues of B-galactosidase. As judged by the analysis of plasmid encoded proteins expressed in E.coli K-12 minicells, pSN015 did produce more protein than

the derepressed Tn1 derivative Tn103 carried on a plasmid (pMB9) with a copy number c.1/4 of that of pSNO15. Yet the transposition of Tn3651 mediated *in trans* by each of these plasmids occurred at very similar frequencies at both 30°C and 37°C. Thus the variation in transposition frequencies with temperature seen with both Tn1 and Tn3 was retained in pSNO15. It was also found that there was a variation with temperature in the amount of fusion protein expressed by pSNO15 : this level was at least five times greater at 30°C than at 37°C. Hence, it seemed likely that at least part of the reason for the different transposition frequencies observed with pSNO15 at 30°C and 37°C was a variation in the level of protein expressed at the two temperatures. These data, together with other observations made in this laboratory over a period of time, led to the formation of two hypotheses:-

(i) those DNA sequences responsible for the variation in transposition frequency seen with temperature lie between co-ordinates 1642 and 3040 in the Tn1/3 sequence and the reason for this variation is probably some feature(s) of the *tnpA* mRNA that varies with temperature.

(ii) an additional process seems to operate to regulate the transposition of Tn1/3 so that at high *tnpA* gene dosage (>20), transposition frequency does not significantly alter as *tnpA* gene dosage increases.

Various possible models for the mechanism of this additional process were considered, and one of these - a model involving antisense RNA regulation of *tnpA* gene expression, akin to that seen in the regulation of transposase gene expression in IS10-right of Tn10 - was investigated in detail. However, no evidence to support this model could be produced in an examination of that part of Tn3 extending from co-ordinates 2073 to 3094.

Chapter 1.

Introduction.

1.1. Historical perspective.

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.

The existence of such elements was first mooted by Mc.Clintock (1950) on the basis of her analysis of the patterns of pigmentation variation shown by maize kernels. However, further developments had to await the discovery in prokaryotes of a new class of highly polar mutations (Malamy, 1966; Saedler and Starlinger, 1967) and the characterisation of these by physical means as being due to the insertion of DNA (Jordan et al, 1968; Shapiro, 1969). Since then, these elements have been the subject of intensive investigation, and this has led to a much clearer understanding of their natures, effects on host genomes and mechanisms of transposition.

To date, transposable elements have been detected in bacteria, yeast, plants and invertebrates (e.g. Drosophila, Trypanosoma). See Calos and Miller, (1980); Kleckner, (1981); Shapiro, (1983a) and Doring, (1985) for reviews. Recently, suggestions have been made that such elements may also exist in the genomes of vertebrate species. e.g. Temin, (1980) suggested a close evolutionary relationship between known eukaryotic transposable elements and the proviral forms of RNA tumour viruses. This has been supported by the observation that transposition of the Ty1 element in yeast proceeds via reverse transcription (Boeke et al, (1985), is associated with the presence of virus-like particles in yeast cells (Garfinkel et al, 1985) and that Ty elements employ a similar protein coding strategy to some retroviruses (Mellor et al, 1985). Also intra-cisternal A particle (IAP) genes, which form a prominent middle repetitive component in the mouse genome, are structurally similar to mammalian and avian retroviruses and to the mobile, dispersed genetic

elements of Drosophila and yeast (Lueders and Kuff, 1977; Grigoryan et al, 1985). Finally, an element with the structural features of a transposable element has been found in human DNA (Paulson et al, (1985)).

1.2. Transposable elements.

The prokaryotic transposable elements have been classified into three groups by Kleckner (1981) according to structural and mechanistic criteria :

(i). class I elements : this class includes the IS-like elements and the compound transposons derived from them, in which a unique segment of DNA is flanked by two copies of an IS-sequence in either direct or inverted orientation (e.g. Tn9 - Mac.Hattie and Jackowski,1977; Tn903 - Grindley and Joyce, 1981). In these transposons, all the functions and sites necessary for transposition are carried in the IS-elements (Tn9 - Rosner and Guyer, 1980; Tn10 - Foster et al, 1981). Both these elements are structurally intact, but one is often functionally deficient (e.g. IS10-left in Tn10 - Foster et al, 1981). These transposons can transpose as a single unit, or one copy of the IS sequence can transpose alone (e.g. IS50 in Tn5 - Berg et al, 1982; IS903 in Tn903 - Grindley and Joyce,1980).

IS elements are quite small (between 0.77 and 1.75kb. in length; Kleckner, 1981) and, typically, the transposase gene spans almost the entire length of the element (e.g. in IS10-right of Tn10 - Halling et al, 1982).

Transposition of members of this class generates either a simple insertion or a transposition cointegrate, in which the whole of the donor replicon, together with two flanking copies of the complete element are inserted into the target replicon. See Fig.1.1 The proportions of these products varies from one element to another (Kleckner, 1981).

Fig.1.1.

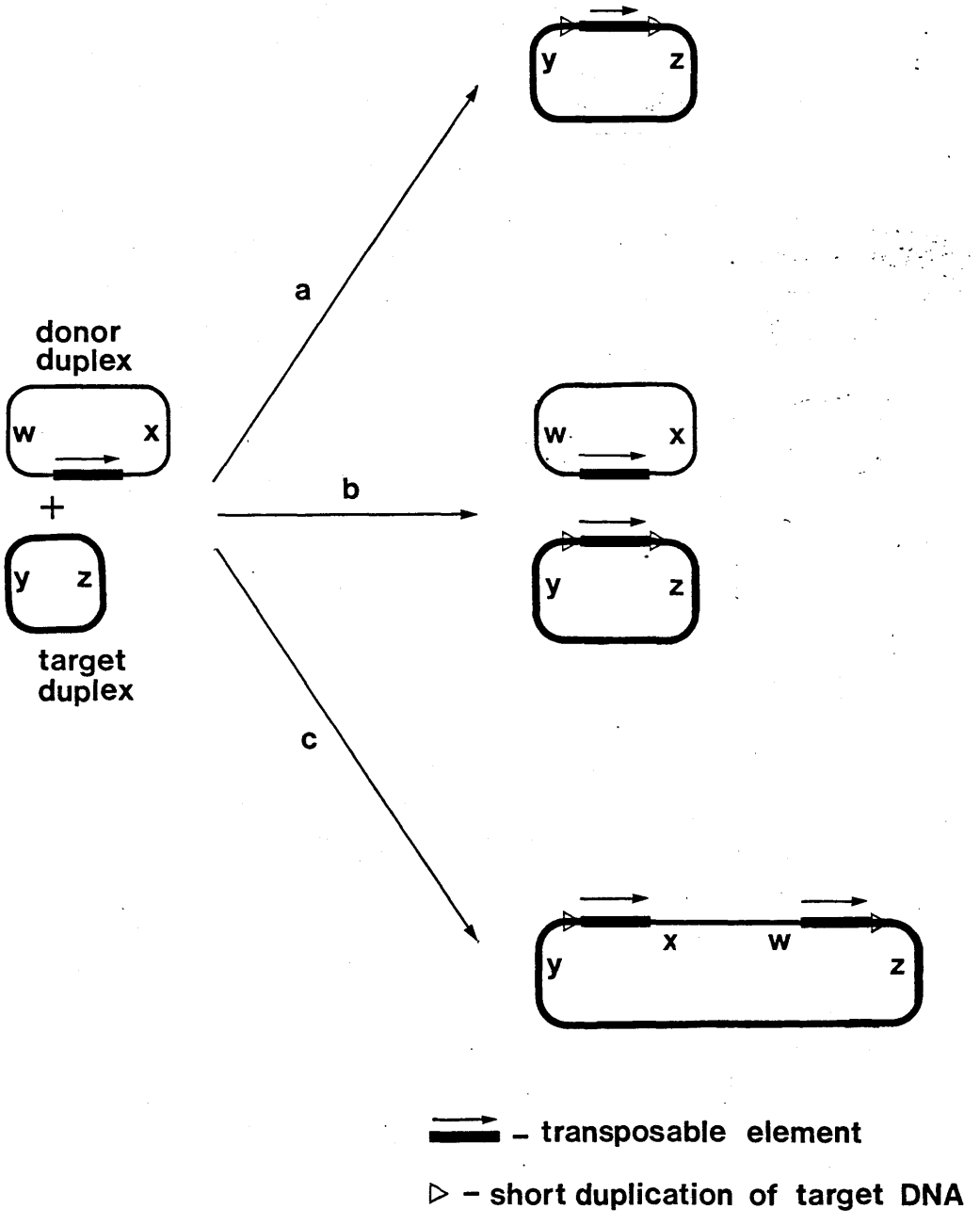

donor
duplex
w
x
+
y
z
target
duplex
a
b
c
y
z
w
x
y
z
y
x
w
z
– transposable element
▷ – short duplication of target DNA

Fig.1.1. The products of transposition.

a). non-replicative simple insertion with loss of donor duplex.

b). replicative simple insertion; one copy of element retained at original site.

c). cointegrate formation / replicon fusion: always replicative.

(ii). class II elements : the Tn3 family. This group consists of a large number (>20) of transposons that are closely related both structurally and functionally. Such close relationships are suggestive of a common evolutionary origin. This is a medically important group as most of its members can confer antibiotic resistance to many pathogenic bacteria. In particular, the B-lactamase enzyme encoded by Tn3 (Kopecko and Cohen, 1975) and its close relatives Tn1 and Tn2 (Hedges and Jacob, 1974; Heffron et al, 1975b) has been found in the pathogenic bacteria Haemophilus ducreyi (Brunton et al, 1981), H. influenzae (de Graffe et al, 1976) and Neisseria gonorrhoeae (Fayet et al, 1982). All the group's members have been isolated from Gram-negative bacteria, except Tn551, which was isolated from Staphylococcus aureus (Kahn and Novick, 1980). All can mediate their own transposition except two (IS101 and Tn951) which possess the sites but not functions necessary for transposition and therefore have to be complemented in trans for these (Heffron, 1983).

All these elements are long (>5kb.) and have 38-40bp. near perfect inverted repeats at their ends. These repeats are very similar among different elements (Kleckner, 1981; Heffron, 1983). All also contain genes necessary for the transposition process and most possess accessory determinants as well (Kleckner, 1981). The transcription of the genes may be unidirectional (e.g. Tn501) or divergent (e.g. Tn3. Heffron, 1983).

Much evidence has been accumulated that these elements transpose inter-molecularly in a two-step mechanism via an obligatory cointegrate intermediate that is resolved by an element encoded site-specific recombination system (Arthur and Sherratt, 1979; Kitts et al, 1982a.). However, Bennett et al, (1983), have reported transposition in the absence of element-encoded resolution of cointegrates at low frequency for Tn3 and Tn21, so it may be that some transposition can proceed by another mechanism too.

Class I and II elements share the feature that their ends are short

inverted repeats between 9 and 40bp. long (Kleckner,1981).

(iii). class III elements : the transposing bacteriophages Mu and D108. These are closely related (Hull et al, 1978) and both use replicative transposition as the normal method of genomic replication during lytic 'phage growth. Also, unlike other temperate 'phages, they both integrate into the host chromosome whether entering the lytic or lysogenic pathway of infection (Toussaint and Resibois, 1983). As well as their 'phage characteristics, these elements differ from class I and class II elements in that their ends are not perfect inverted repeats, though there is some inverted homology between the two (Kahmann and Kamp, 1979). The observation that lytic growth generates transposition cointegrates whilst the establishment of lysogeny generates simple insertions led to the suggestion that there may be two different pathways of Mu transposition (Chaconas et al, 1981). However, recent in vitro studies have shown that both simple inserts and cointegrates can arise from the same transposition intermediate (Craigie and Mizuuchi, 1985).

Among the eukaryotic transposable elements so far studied, terminal repeat sequences again seem to be a universal feature, but there may well be a variety of transposition mechanisms.

For example, the copia-like elements of Drosophila melanogaster possess long terminal direct repeats, reminiscent of some prokaryotic compound transposons. However, DNA-sequence analysis and the identification of virus-like particles containing a 5kb. RNA species that hybridises to copia DNA in cultured Drosophila cells suggests a closer relationship to vertebrate retroviruses (Shiba and Saigo, 1983; Saigo et al, 1984; Emori et al, 1985). The recent demonstrations that the Ty1 elements of yeast (that also have long terminal repeats) transpose via an RNA intermediate (Boeke et al, 1985), that their transposition is associated with the presence of virus-like particles in yeast cells (Garfinkel et al, 1985) and that they employ a similar protein coding strategy to some retroviruses (Mellor et al, 1985) are

similarly suggestive.

Another class of mobile elements in Drosophila, the P elements (responsible for one form of hybrid dysgenesis - see Bregliano and Kidwell (1983) for review), also possess terminal repeats (inverted in their case) but appear to transpose via a 'cut-and-paste' mechanism that may or may not be replicative and is dependent on an element-encoded function (O'Hare, 1985). A similar model has been put forward for the transposition of transposable elements in plants (Saedler and Nevers, 1985).

A third family of Drosophila transposable elements is the foldback elements (Finnegan, 1982). They, too, have long terminal inverted repeats, but differ from the above elements in that these repeats are composed primarily of tandem copies of simple sequence DNA (Rubin, 1983). Two fold-back elements have been observed to interact in the formation of a type of compound transposon - TE1 (Chia et al, 1985).

1.3. The consequences of transposition.

Transposable elements insert themselves into new target sites at frequencies between 10^{-4} and 10^{-7} per cell per generation (Kleckner, 1981). On doing so, all but one (Tn554 - Murphy and Lofdahl, 1984) generate a short duplication of host DNA (3-12bp. depending on the element - Kleckner, 1981) so that the inserted transposon is flanked by a direct repeat of host DNA. This duplication is not necessary for transposition to occur (Kleckner, 1979) and probably arises as a consequence of a staggered cleavage of the target DNA (Grindley and Sherratt, 1979).

A transposition event can generate either a simple insertion or a transposition cointegrate (see Fig 1.1) and, if stable, cointegrates can be the source of important events. For example, the F plasmid can become integrated into the host chromosome by transposition or by homologous recombination and can subsequently transfer chromosomal genes

to a new host (Cullum and Broda, 1979).

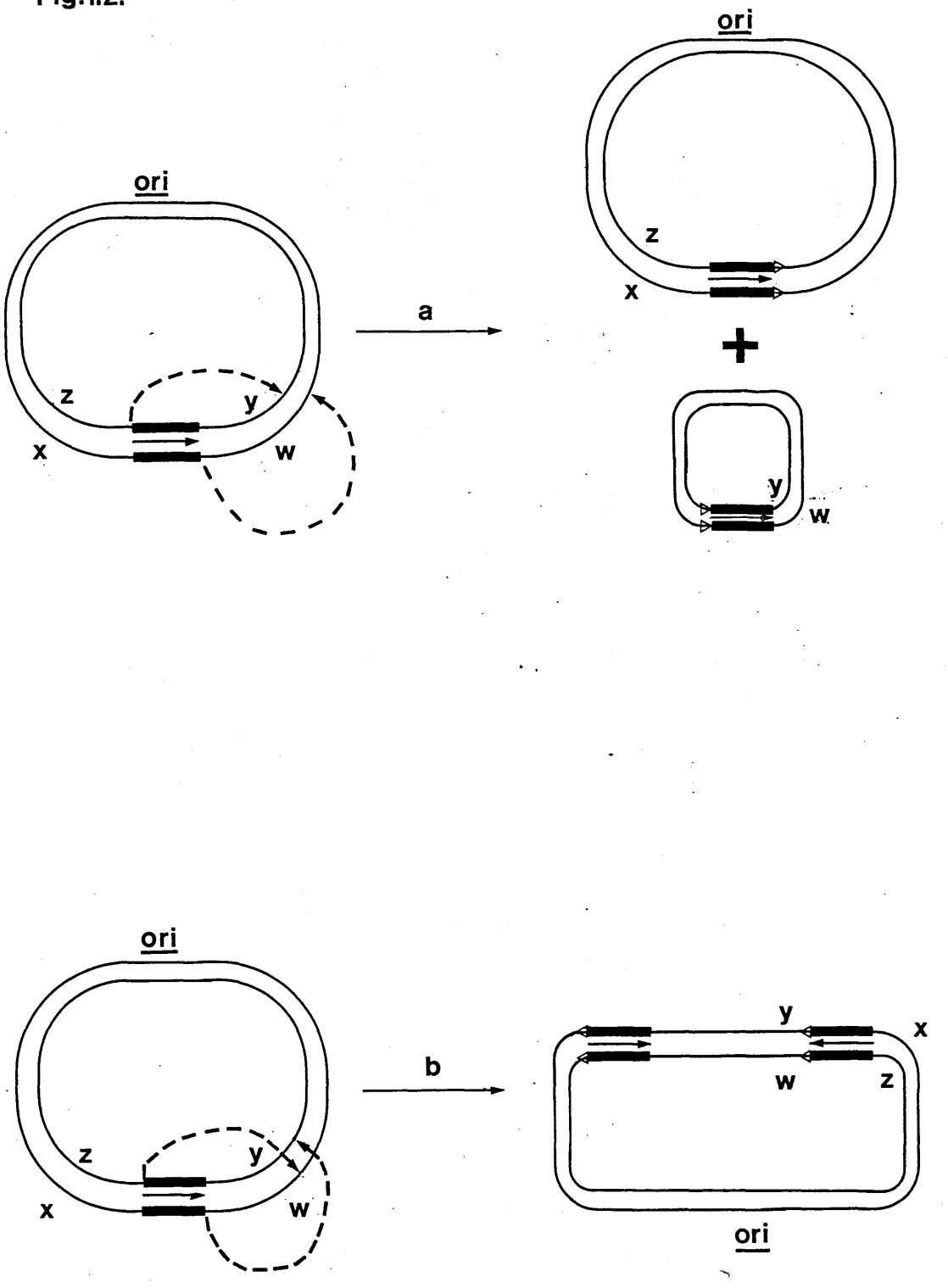
One of the most obvious consequences of transposon insertion is inactivation of the target gene due to the interruption of its coding sequence (e.g. Lieb, 1981). However, many transposable elements have also been observed to produce marked polar effects, abolishing or diminishing the expression of genes downstream of the site of insertion (e.g. Tn3 - Heffron et al, 1975a; Rubens et al, 1976; Tn5 - Berg et al, 1980; many IS elements - Iida et al, 1983). For IS elements, such effects are thought to be due to the presence of termination codons in all three reading frames of the elements (Adhya and Gottesman, 1978).

Additionally, adjacent genes may be transcriptionally activated by insertion of an element (e.g. Reynolds et al, 1981). Some such effects arise by transcriptional readthrough from a promoter carried on the transposon that is directed outwards and are therefore seen in only one orientation of element insertion (e.g. IS3 - Charlier et al, 1982). If the newly introduced promoter responds to effector molecules, then the adjacent gene(s) too becomes subject to control by these effectors e.g. Ty1 elements can insert into the 5' non-coding regions of genes and place them under the control of the mating-type control system (Errede et al, 1980; Roeder and Fink, 1982). Other effects are seen only at particular insertion sites and may be due to the juxtaposition of host and element sequences causing the formation of a new promoter (e.g. IS2 - Glansdorff et al, 1981; IS30 - Dalrymple and Arber, 1985) or to the rearrangement of sequences within the element (Besemer et al, 1979).

Once inserted in a genome, transposable elements can mediate further rearrangements of the adjacent DNA such as inversions, deletions and translocations (see Fig 1.2).

Deletion and duplicative inversion are the two possible results of an intramolecular transposition event predicted by the symmetric model of transposition (Arthur and Sherratt, 1979; Shapiro, 1979), and both have been reported : Chaing and Clowes, 1980; Kleckner et al, 1979; Ross et al, 1979; Weinstock et al, 1979; Bishop and Sherratt, 1984; Thorpe and

Fig.1.2.



– transposable element ori – origin of replication

– short duplication of target DNA

Fig.1.2. Intramolecular transposition events.

a). insertion of each strand of the element into the same strand generates a deletion. The segment without an origin of replication cannot replicate and is therefore lost.

b). insertion of each strand of the element into the opposite strand generates a duplicative inversion.

with respect to
Clowes, 1984. Tn1, both deletion and inversion require only the tnpA gene product, not that of tnpR (Bishop, 1983; Bishop and Sherratt, 1984). Deletions extend from one end of the element into adjacent, non-element, DNA, with a copy of the element remaining at the original point of insertion (Ross et al, 1979; Ohtsubo et al, 1979).

Additionally, transposable elements can mediate the translocation of segments of the host genome. In principle, any block of DNA flanked by two copies of an element, in direct or inverted repeat, can transpose. e.g. the whole E.coli K-12 chromosome has been shown to transpose when flanked by two copies of IS10 (Harayama et al, 1984a). However, it has been observed that the frequency of transposition of such compound elements decreases as their length increases (Chandler et al, 1982; Morisato et al, 1983). Also, it has been reported that a gene adjacent to a single insertion sequence can form a transposon if there is a sequence sufficiently homologous to the end of the IS element elsewhere on the host genome (Machida, Y. et al, 1982a).

It is not only element-mediated events that promote these DNA rearrangements. If there are two or more copies of an element present in a molecule or cell, these can act as substrates for the action of the host homologous recombination systems to cause significant levels of deletion, inversion, cointegration and translocation. e.g. variant Ty elements have been observed to replace each other by gene conversion at rates 10-100x higher than the rates of transposition (Roeder and Fink, 1982). Kleckner (1981) has stressed the importance of transposable elements acting in this way.

Finally, one recent report has suggested that the element Tn 1000 can repress gene expression by causing topological changes in the molecule of which it is a part (Stokes and Hall, 1984) and transposition has also been suggested to have a role in tumour progression (Grigoryan et al, 1985).

The net result of consideration of all these processes is that transposable elements are likely to have made, and to make, a

significant contribution to both mutational frequency and the range of genetic rearrangements, and thus to be the course of genome evolution itself. It has been suggested that such elements are fundamentally important in cellular differentiation in both development and evolution (Shapiro, 1983b). For reviews, see Baltimore (1985); Syvanen (1984); Cohen (1976); and see also Chao et al, 1983; Schwarz-Sommer et al, 1985.

1.4. The insertion specificity of transposable elements.

One of the main characteristics of transposable elements is their ability to insert at many sites in a DNA molecule. However, insertion is not entirely random, and different elements do show marked variations in specificity of insertion. This suggests that element encoded functions participate in the selection of target sites. At one end of the spectrum is IS4, which has only ever been found at one position in the galT gene of the E.coli K-12 chromosome (Habermann et al, 1979), whilst at the other are elements such as Mu, which show very little specificity, inserting at many different sites in a target (Bukhari and Zipser, 1972; Daniell et al, 1971).

Other elements show insertion specificities lying between these extremes, many showing what is called "regional specificity". That is, in some regions of a target, an element inserts rarely, whilst in others it inserts much more frequently, though not always at the same site. Tn3 shows this sort of specificity (Heffron et al, 1975a; Kretschmer and Cohen, 1977; Weinstock et al, 1979; Picken et al, 1985) and its preferred insertion sites seem to be A/T rich and have some homology with the ends of the element (Tu and Cohen, 1980). This observation has led to the suggestions that localised duplex denaturation must occur before insertion can take place and that the Tn3 transposase preferentially interacts with target sites that have some homology to its presumed normal substrate. For Tn9, similar observations have been made (Miller et al, 1980; Galas et al, 1980) and for Tn10, targets with

some homology to the element ends are favoured insertion sites (Halling and Kleckner, 1982). It has also been suggested for Tn3 that the physical structure of the target DNA may be an additional factor in determining insertion specificity (Heffron, 1983).

For Tn3, it has been found that interplasmidic transposition is a far more common event than transposition from a plasmid to the chromosome (Kretschmer and Cohen, 1977). This has been interpreted by Campbell(1981) to suggest that the elements of the Tn3 family are closely related due to the rapid dispersion across species boundaries of a particular successful transposon type rather than due to antiquity of the elements.

Much less is known about the insertion specificities of eukaryotic transposable elements. All Ty element insertions hitherto analysed occur in 5' non-coding rather than coding regions themselves, but the reason for this preference is obscure (Scherer et al, 1982; Eibel and Philippsen, 1984).

1.5. The mechanism of transposition.

The fundamental question of whether transposition is accompanied by specific element replication has not yet been answered for many eukaryotic transposable elements e.g. P elements in Drosophila (O'Hare, 1985).

Among prokaryotic elements, however, the situation is more clearly understood. It seems that all element-promoted genetic rearrangements can be understood as resulting from one of two types of element-target interaction (see Fig 1.1). In a simple insertion reaction, both ends of a single copy of the element become joined to target sequences. This process can either be replicative, when a copy of the element is retained at the donor site, or non-replicative, when the element is excised from the donor site without replication and there is concomitant loss of the donor replicon. In cointegrate formation, two

copies of the element are formed (hence the process is always replicative) and each retains one junction with the donor molecule and forms one with the target.

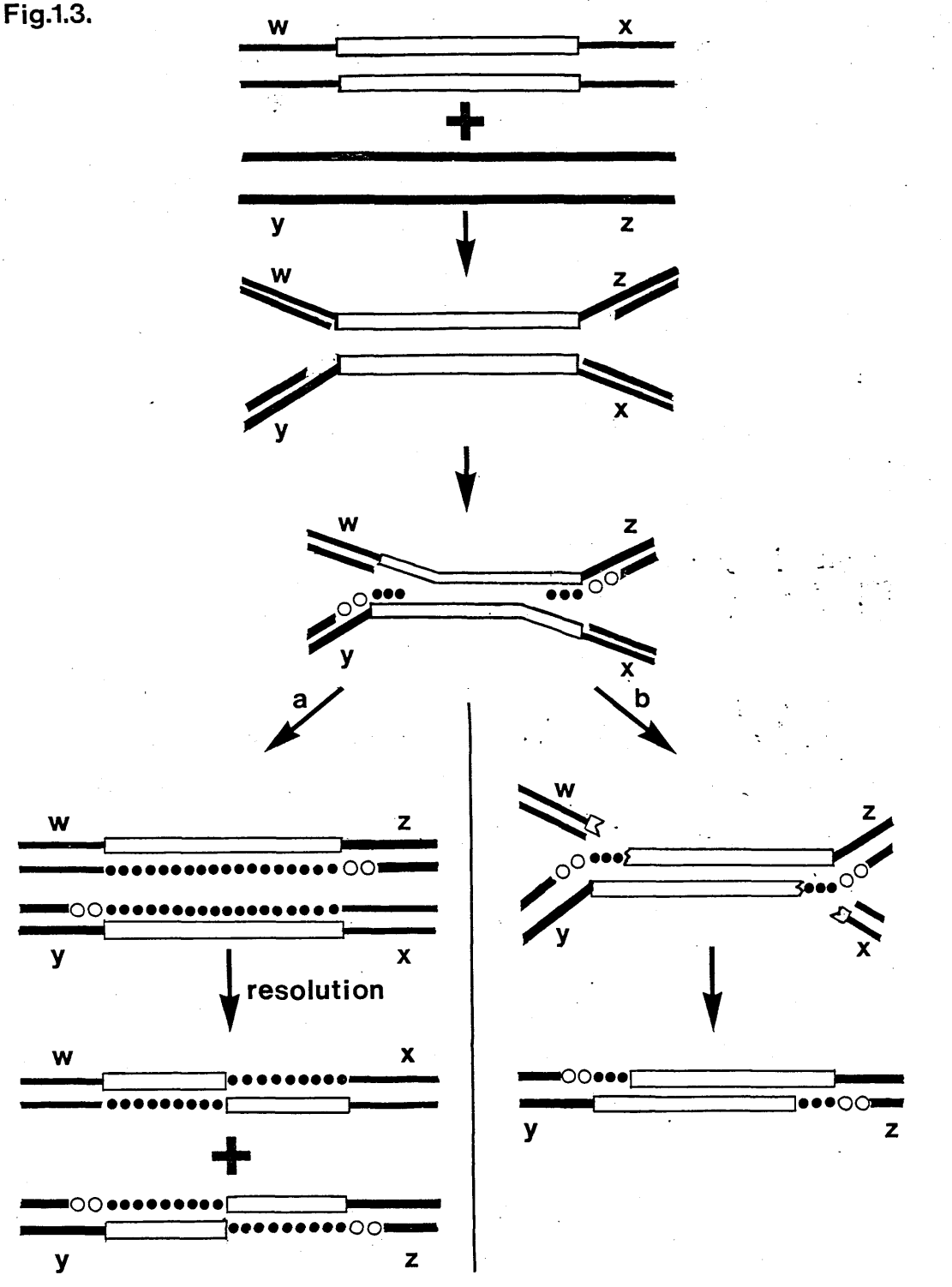
For Tn10, it seems that most transposition occurs by non-replicative simple insertion (Morisato and Kleckner, 1984; Kleckner et al,1984). For some IS elements (eg.IS1, IS903), however, transposition can proceed by non-replicative simple insertion or (at lower frequency) by cointegrate formation (Weinert et al, 1984; Biel and Berg, 1984).

For Tn3 a large body of evidence has been accumulated that inter-molecular transposition proceeds via an obligatory cointegrate intermediate and is thus replicative (Arthur and Sherratt, 1979; Muster and Shapiro, 1981; Mc.Cormick et al, 1981). However, Bennett et al, 1983, have reported low frequency Tn3 transposition in the absence of cointegrates. This may represent an alternative transposition pathway (e.g. simple insertion or tnpR-independent cointegrate resolution).

In order to explain the replicative nature of many transposition events, two basic models of transposition have been proposed. (See Figs 1.3 and 1.4). Both of these mechanistically couple the break/join and replication events involved in transposition and fundamentally involve the ends of the elements, in keeping with the experimental evidence on the need for those sequences (e.g. Tn3 - Mc.Cormick et al, 1981; Tn5 - Sasakawa et al, 1983, 1985; Tn10 - Way and Kleckner, 1984). In both also, strand transfer could involve either the 5' or 3' ends of the element.

The first of these models (Shapiro, 1979; Arthur and Sherratt, 1979) is a symmetric model, in that both ends of the element are treated in the same way, though not necessarily simultaneously. In this model, the same donor duplex is nicked on opposing strands, at opposite ends of the element, and the target duplex too is nicked on both strands in a staggered fashion. Attachment of the free ends of the element to the target DNA forms two replication forks and replication proceeds at both of these to replicate the whole element. This is folowed by ligation of

Fig.1.3.



▭ – parental element DNA

●●● – newly synthesised element DNA

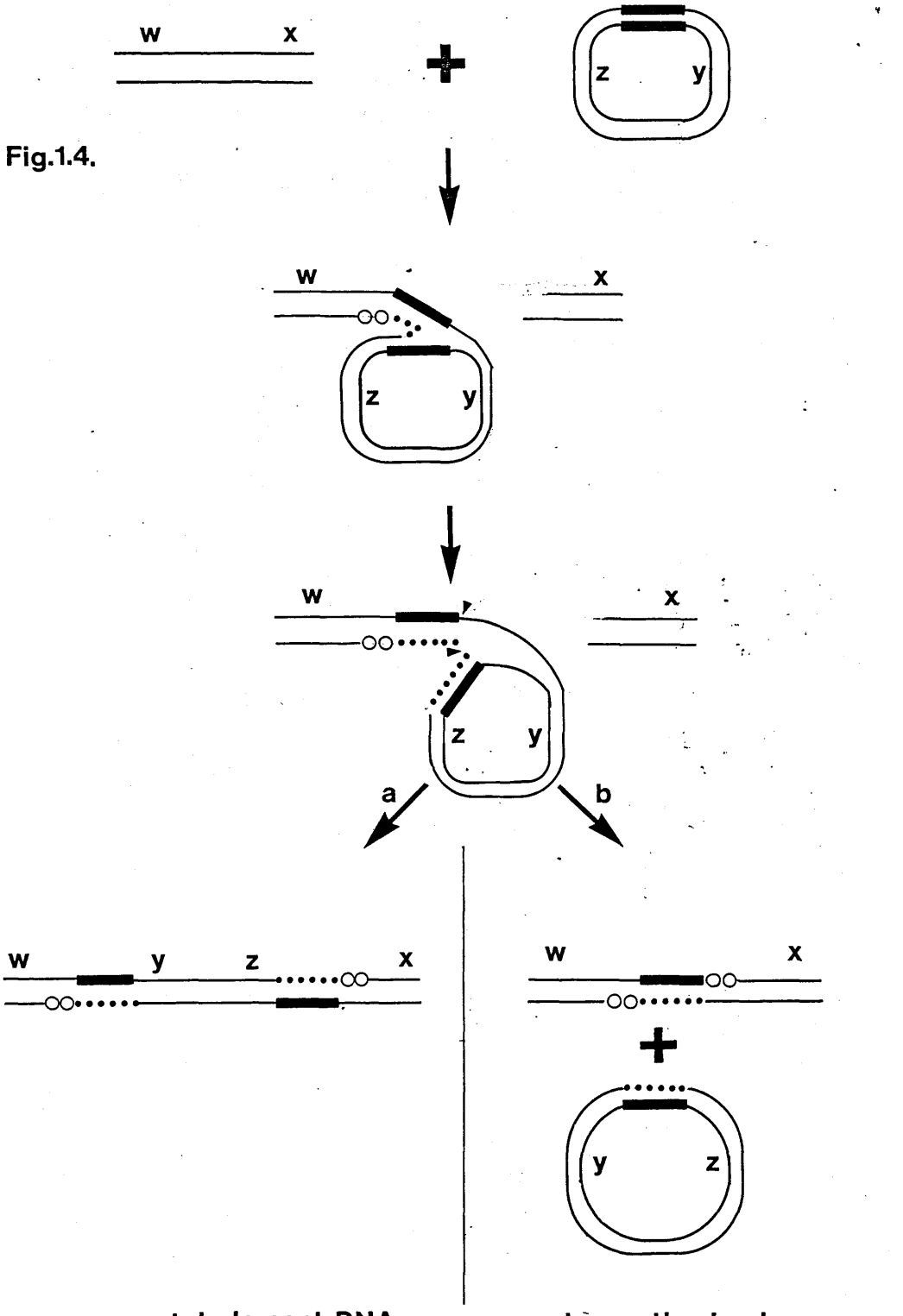
○○ – short duplication of target DNA

Fig.1.3. A symmetric model of transposition.

After Grindley and Reed, 1985.

This diagram depicts both the cointegrate forming pathway (a) and an alternative pathway by which simple inserts can be generated with concomitant loss of the donor duplex (b) - Ohtsubo *et al*, 1981.

Fig.1.4.



▬ – parental element DNA ••– newly synthesised element DNA

○○ – short duplication of target DNA

▲ – sites of cleavage leading to cointegration or simple insertion

Fig.1.4. An asymmetric model of transposition.

After Harshey and Bukhari, 1981; Galas and Chandler, 1981.

This diagram depicts both the cointegrate forming pathway (a) and the simple insert forming pathway (b).

replicated DNA to flanking sequences to generate a cointegrate which is resolved to yield the final transposition products. Thus, cointegrates are an obligatory part of this model, though a modification proposed by Ohtsubo et al, (1981) illustrated a means by which non-replicative simple insertions could be the outcome of this pathway, without cointegrate formation. This model was developed to explain the observations that Tn3 transposition seemed to proceed via an obligate cointegrate intermediate, but the above-mentioned modification of Ohtsubo et al could explain the data of Bennett et al, 1983.

The other model was derived from the observations that both insertion sequences and Mu transposition can produce either simple inserts or cointegrates. It is an asymmetric model, for the ends of the element are not treated in the same way (Grindley and Sherratt, 1979; Galas and Chandler, 1981; Harshey and Bukhari, 1981). In this model, transposition is initiated by nicking the donor duplex at one end of the transposon and ligating this end to one strand of a target duplex that has been cut in a staggered way. A replication fork is thus created and proceeds through the element until it reaches and replicates the opposite ends of the element. A second nick at the equivalent sequence to the original on the strand with opposite polarity will then result in cointegrate formation, whilst a second cleavage on the originally nicked parental strand will yield a replicative simple insertion.

However, evidence is accumulating against this model, for both Mu and the IS elements. For example, an intermediate in Mu transposition, that can be processed to yield either a non-replicative simple insert or a cointegrate, has been isolated in vitro and has the structure predicted by the symmetric model. (The strand transfer reaction involves interaction of the 3' ends of the element with the 5' ends of the target). Also, structures resulting from DNA strand transfer at only one end of the element are not seen (Mizuuchi, 1984; Craigie and Mizuuchi, 1985). Additionally, Tn10 seems to transpose principally by a non-replicative simple insertion process (Morisato and Kleckner, 1984;

Kleckner et al, 1984). Also, though initial data on the variation of transposition frequency with transposon length apparently supported the asymmetric model (Chandler et al, 1982; Morisato et al, 1983), later results with artificially constructed giant transposons did not (Rosner and Guyer, 1980; Harayama et al, 1984a). In passing, it is interesting to note that Tn3 transposition shows no such length dependence (Lee et al, 1985).

A novel type of transposition event is 'one-ended transposition', which has been observed for several transposons of the Tn3 family (e.g. Tn1 and Tn3 - Arthur et al,1984; Tn21 and Tn1721 - Avila et al, 1984; Motsch and Schmitt, 1984; Motsch et al, 1985). This sort of event appears to involve insertion of a donor replicon that contains only one copy of a transposon end into a target molecule, with a concomitant small and variable duplication of part of the inverted repeat sequence. Evidence that such events are replicative transposition events is presented in Section 3.2.1 below. It seems most likely that the mechanism of such events is one closely akin to the asymmetric model, since the transposase protein only has one inverted repeat sequence upon which to act. Thus, it may be that transposition of Tn3-like elements can proceed via either a symmetric or asymmetric mechanism. The asymmetric one would be the only mechanism of transposition available to one-ended elements, and may occur at low frequency in the transposition of two-ended elements. (This too would explain the data of Bennett et al, 1983). However, the symmetric mechanism seems to be highly favoured for the transposition of two-ended elements.

Finally, certain non-element factors may be important in the mechanism of transposition of some elements. e.g. DNA gyrase is a host factor required for the transposition of Tn5, apparently due to a requirement for the recipient DNA to be supercoiled (Isberg and Syvanen, 1982). Also, a decrease of c.40 fold is seen in the transposition frequency of this element in E.coli strains defective in DNA topoisomerase I. This could reflect either a direct or indirect role

for this protein in the transposition process. No such effect is seen for Tn3 (Sternglanz et al, 1981). For Ty elements, the intact product of the SPT3 gene is required for normal transposition (Winston et al, 1984).

1.6. The persistence of transposable elements.

Some elements can confer an obvious selective advantage on cells containing them in selective environments by reason of the accessory genes they carry. However, even in the absence of selection for the particular phenotype they confer, it was thought that transposable elements could confer a selective advantage on cells as a result of an increased mutation rate caused by transposition and the other genomic rearrangements mediated by them. Obviously, though, any such advantage must be balanced against the deleterious effects resulting from an increased mutation rate. Such a transposition-specific selective advantage has been observed for Tn10, and concluded to be due to just such a mechanism (Chao et al, 1983). Tn5 also confers a selective advantage on cells harbouring it in non-selective conditions, but this has been found to be independent of the occurence of transposition events, though dependent on Tn5 transposase activity. The mechanism of this advantage remains obscure (Hartl et al, 1983; Wurster-Biel and Hartl, 1983).

However, for replicative transposable elements (as most prokaryotic elements seem to be), the concept of 'selfish DNA' can explain their persistence. This hypothesis suggests that once a unique segment of DNA acquires the ability to duplicate itself independently of the replication of the rest of the genome, it will be maintained simply because of the difficulty in simultaneously eliminating all copies of it from the population (Doolittle and Sapienza, 1980; Orgel and Crick, 1980). Such an entity need not confer any advantage to host, and may even be disadvantageous, as long as it is able to propagate itself

rapidly enough to prevent its elimination from the genome. Thus, replicative transposable elements can be considered as intra-genomic parasites (Orgel and Crick, 1980).

The overall picture is probably made up of components from all these sources.

1.7. The regulation of transposition.

That regulation of transposition occurs is evidenced by the observation that uncontrolled proliferation of elements is rarely seen, at least in prokaryotes, despite the replicative nature of many transposition events.

Two obvious mechanisms for regulation are a limitation in the number of target sites for an element and/or excision of element copies. However, examination of the distribution of element insertions in closely related strains reveals more potential sites than are ever occupied (Nyman et al, 1981) and, for many elements, the rates of excision are less than the rates of transposition (Kleckner, 1981). Thus, additional regulatory mechanisms must exist.

Of these, one rather coarse one may be that the accumulation of many mutations in a cell by frequent transposition selectively disadvantages the cell, so that elements with low rates of transposition are favoured by selective processes.

Indeed, the transposition rates of most natural elements are very low: between 10^{-4} and 10^{-7} per cell per generation (Kleckner, 1981). Since artificially increasing the level of transposase protein results in an increased transposition frequency (Casadaban et al, 1982; Morisato et al, 1983) it seems probable that low level transposase expression is important in keeping this rate down. Indeed, data on the cellular levels of transposases seem to bear this prediction out (O'Hare, 1985; Casadaban et al, 1980,1982). Such low levels could arise by weak transcription and translation control signals (Simons et al, 1983;

Machida et al, 1984; Chou et al, 1979b), and other factors too are probably important. For example, instability of these proteins, their stoichiometric rather than catalytic use (both of which have been demonstrated for the Mu A protein : Pato and Reich, 1982, 1984) and the preferential cis action of many IS-element transposases (Machida et al, 1982b; Morisato et al, 1983; Isberg and Syvanen, 1981; Grindley and Joyce, 1981) all probably contribute to the low transposition rates that are observed. It is interesting to note Tn3 transposase appears to show no cis preference in its action (G. Russell, pers. comm.).

Additionally, the variation of transposition frequency with temperature seen for Tn3 and Tn951 (Kretschmer and Cohen, 1979; Cornelius, 1980) acts to reduce transposition frequency at normal body temperature (37°C). These data, together with the observation that the level of Tn3 transposase rises c.10 fold in stationary phase cells (Fennewald et al, 1981), may be interpreted to suggest that elements that transposase at higher frequencies in a stressed host may be selectively advantaged.

During the development of both Drosophila and maize, transposition seems to occur only at particular stages - i.e. occurs in bursts (Bregliano and Kidwell, 1983; Gerasimova et al, 1984; Fedoroff, 1983). This may reflect some transient change in DNA structure that favours transposition. For IS10, a mechanism of increased transposition frequency (by both increased transposase gene transcription and increased activity of the inside end of the element) associated with hemi-methylation of DNA has been described (Kleckner et al, 1984; N. Kleckner, pers. comm. to D. Sherratt). DNA exists in such a state during replication, so this may represent a means whereby IS10 transposition can also occur in bursts during DNA replication. Dam methylation sites have been found at similar positions to those in IS10 in other IS elements (IS50 - Sasakawa et al, 1983; IS903 - Grindley and Joyce, 1980) and these elements too exhibit increased transposition rates in the absence of Dam-methylation (Kleckner et al,1984; Sternberg,

1985). IS1 and Tn3,however, have no Dam sites in similar positions and thus it seems unlikely their transposition frequencies are regulated in this way (Gamas et al, 1985; Heffron et al, 1979).

An effect of transcription on the transposition frequency of some IS-elements has also been observed, and in IS1 the existence of two opposing transcriptional units within the element has been proposed as being a natural component of the transposition frequency regulatory process (Sasakawa et al, 1982; Machida et al, 1983, 1984).

Some elements encode their own molecules that help to regulate transposition frequency. Tn3 encodes the resolvase protein that acts to repress transcription of both the tnpA and tnpR genes (Gill et al, 1979); Tn10 transposition is regulated by an antisense RNA control mechanism (Simons et al, 1983; Simons and Kleckner, 1983); Mu transposition is regulated by a repressor that acts to repress transcription of the Mu A and B genes and also, at higher concentration, binds to the ends of the element (Craigie et al, 1984); and Tn5 transposition is regulated by a polypeptide that is a truncated form of the transposase protein (Lowe and Berg, 1983; Isberg et al, 1982; Johnson et al, 1982). Lastly, a new form of regulatory molecule has recently been described for Tn21 (Hyde and Tu, 1985). It is a modulator (M) protein that acts to both increase transcription of the Tn21 transposase gene and decrease expression of the resolvase gene (probably by acting on its mRNA). Because Tn21 resolvase also seems to repress its own transcription, however, this latter effect is not as great as might be at first expected. Thus, the M protein serves to increase transposition frequency and cointegrate formation. This is the first report of an element-encoded positive effector of transposase expression.

One regulatory mechanism unique to the Tn3 family of elements is transposition immunity. This is a transposon-encoded mechanism that acts only in cis to limit transposition of a second transposon copy into a replicon (by a factor of c.100) and is highly element specific. For

Tn1/3 itself, immunity can be conferred by the presence of only a single copy of the inverted repeat sequence, and other parts of element do not seem to confer immunity (Lee *et al*, 1983). It is inseparably linked to transposition proficiency and specificity and the strength of its expression seems to depend on the cellular level of transposase (Arthur *et al*, 1984; Heritage and Bennett, 1984). However, it seems to be a more complex phenomenon than this, depending also on which replicon carries the transposon sequences and the precise location of these sequences (Heffron, 1983).

Finally, in eukaryotic systems, it seems that external, rather than element-encoded, factors mainly control transposition frequency. For example, both hybrid dysgenesis in *Drosophila* and the movement of transposable elements in maize appear to occur only at specific stages of development, and the behaviour of some elements in maize is determined by their location within the organism as a whole or within a particular tissue (Bregliano and Kidwell, 1983; Federoff, 1983).

The aims of this work were to investigate further the roles of the ends of Tn1/3 in transposition and to attempt to develop a system for the over-expression of the Tn1 transposase protein with a view to developing an *in vitro* Tn1/3 transposition system. In pursuing this latter aim, evidence was obtained that suggested a novel regulatory process for the *tnpA* gene of Tn1/3. One possible mechanism for this, involving antisense RNA, was explored. Hence, this introduction has been extended to include both a general discussion and a specific example of antisense RNA regulation of gene expression.

1.8. The general features of antisense RNA regulation of gene expression.

There have recently been several reports of novel regulatory mechanisms of gene expression in prokaryotic systems, and one in an eukaryotic system, which act at the post-transcriptional level. All these mechanisms are similar in that they involve the hybridisation of a complementary RNA molecule to a target RNA (e.g. an mRNA, RNA acting as primer of DNA replication) and this interaction impairs the function of the target (e.g. an mRNA hybridised to a complementary polynucleotide cannot be translated in an *in vitro* system. Paterson *et al*, 1977). The complementary RNA is usually transcribed from the opposite strand to that giving rise to the target RNA, thus ensuring their complementarity, and may be known as antisense RNA. (There are now two examples - Coleman *et al*, 1984; Mizuno *et al*, 1984; Spena *et al*, 1985 - where the complementary RNA is actually produced from a region far removed from its target, but these are exceptional).

Of the systems so far studied, many are involved in the regulation of plasmid copy number. (See Scott, 1984; Cesareni and Banner, 1985 for reviews). And, of these, some act by regulating the expression of a protein essential for replication (e.g. plasmid pT181 in *Staphylococcus aureus* - Kumar and Novick, 1985; plasmids RI - Light and Molin, 1983, and NR-1 - Womble *et al*, 1984 in *E.coli*), whilst another acts by directly interacting with the RNA primer of DNA synthesis (plasmid ColE1 in *E.coli* - Tomizawa, 1984; Tamm and Polisky, 1985). In this latter case, the interaction of the two RNA species is catalyzed by a small protein - the Rop or Rom protein (Lacatena *et al*, 1984; Tomizawa and Som, 1984). Other systems too are regulated in this way - e.g. expression of the *ompF* gene (which encodes a major outer membrane protein) in *E.coli* (Mizuno *et al*, 1984) and expression of the transposase gene in IS10-right of Tn10 (Simons and Kleckner, 1983).

There has also recently been the first report of natural antisense RNA regulation (by a cis acting mechanism) in a eukaryotic system (the expression of zein genes in Zea mays - Spena et al, 1985), and such RNAs have also been implicated in RNA splicing processes (Lerner et al, 1980; Rogers and Wall, 1980).

Additionally, artificially constructed antisense RNA producing molecules have been used to modify the in vivo expression of a target gene in both prokaryotic (Pestka et al, 1984; Coleman et al, 1984; Coleman et al, 1985) and eukaryotic systems (Izant and Weintraub, 1984; Melton, 1985; Rosenberg et al, 1985; Kim and Wold, 1985).

The precise mechanism(s) of action of antisense RNAs is not yet clear, but it is thought that the postulated formation of RNA:RNA duplexes may cause inhibition of transcription initiation (Coleman et al, 1984), pausing of transcription (Tomizawa and Itoh, 1982), premature transcription termination (Carleton et al, 1984; Mizuno et al, 1984; Lactena et al, 1984) and/or decreased stability of mRNA, possibly due to a lack of ribosomes traversing the molecule (Coleman et al, 1984). These hypotheses would all explain the observations that the presence of an antisense RNA species in a cell causes a reduction in the amount of its target RNA (Mizuno et al, 1984; Coleman et al, 1984). Also, it is proposed that formation of these duplexes renders the ribosome binding site (Shine and Dalgarno, 1975) and translation initiation site inaccessible to the translation machinery, thus inhibiting translation. This could occur either by duplex formation involving this region of the target (Mizuno et al, 1983; Simons and Kleckner, 1983) or involving another region of the target such that a novel secondary structure is generated in which the ribosome binding and translation initiation sites are sequestered (Light and Molin, 1983; Womble et al, 1984; Kumar and Novick, 1985). Such a novel secondary structure, formed upstream of the translation initiation site, could also simply physically obstruct the movement of the translational machinery onto the coding region of an mRNA. In all natural cases so far studied, the antisense RNA is always

complementary to a region towards the 5' end of the target and outwith any coding sequence whose expression is being regulated (e.g. to the ribosome binding sites and translation initiation sites - Simons and Kleckner, 1983; Coleman et al, 1984 - or to a region further upstream than this, the action of the antisense RNA then being exerted from a distance (Kumar and Novick, 1985; Spena et al, 1985). And, it is interactions affecting the area where ribosomes attach to an mRNA that seem to be most effective (Coleman et al, 1984). In artificial constructs with some genes, this situation is reproduced (e.g. lpp gene of E.coli), but in others (e.g. ompA gene of E.coli) antisense RNAs complementary to the ribosome binding site and translation initiation site, to the ribosome binding site alone or to structural gene coding sequences were found to be equally inhibitory for gene expression. Also, the length of the antisense RNA was found to be important in determining its effectiveness (Coleman et al, 1984).

There clearly remains much to be elucidated about the precise mechanism(s) of action and in vivo roles of antisense RNAs. For reviews of the subject so far, including potential experimental applications, see Laporte, 1984; Travers, 1984 and Weintraub et al, 1985.

1.9. A specific example of antisense RNA regulation.

Of all the above systems, it was particularly the regulation of expression of the transposase gene of IS10-right of Tn10 that led to the suggestion that expression of the transposase gene of Tn1/3 might be similarly regulated.

The salient features of the Tn10 system are as follows:-
Tn10 is a typical composite transposon encoding tetracycline resistance. It is 9300bp. long and carries inverted repeats of an insertion sequence-like sequence at its ends (Kleckner et al, 1975). These two sequences co-operate to mediate the transposition of the intervening genetic material (Botstein & Kleckner, 1977), they contain all the sites and functions required for Tn10 transposition and have been

designated as IS10 sequences (Foster et al, 1981). The antibiotic resistance determinant is c.2500bp. long and located asymmetrically in the intervening 6500bp. of non-repetitive material (Jorgensen & Reznikoff, 1979). Both IS10 elements are structurally intact (Ross et al, 1979), c.1330bp. long, but are not functionally identical: IS10-right (IS10-R) specifies all the functions responsible for normal Tn10 transposition, whereas IS10-left (IS10-L) can only provide 1-10% of the transposition activity of IS10-R, and is therefore clearly functionally defective. (It seems most probable that IS10-R and IS10-L were originally identical, but have since evolved divergently with only one of them retaining essential transposition function(s). Foster et al, 1981 ; Halling et al, 1982). All the sites absolutely required for normal Tn10 transposition are located in the outermost 27bp. at each end of the element (Way & Kleckner,1984) and genetic data show that IS10-R encodes at least one diffusible function essential for Tn10 transposition that acts at the ends of the element (Foster et al, 1981). Transposition of Tn10 appears to proceed principally by a non-replicative mechanism (Kleckner et al, 1984 ; Morisato & Kleckner, 1984) rather than via a cointegrate intermediate (Foster et al, 1981; Harayama et al, 1984b ; Kleckner et al, 1984) as is the case for some other transposons (e.g. class II transposons such as Tn3 - Arthur & Sherratt, 1979).

The transposition frequency of Tn10 is low (c.$1x10^{-4}$ transposition events/element/generation - Simons et al, 1983) and at least three autogenous mechanisms act to keep it so. All act to keep the cellular level of transposase protein low, for the transposition frequency of Tn10 is directly related to the amount of transposase protein present in the cell (Morisato et al, 1983). Firstly, the promoter for the IS10 transposase gene (pIN - see below) is intrinsically weak (c.5% of induced placUV5 - Simons et al, 1983). Secondly the transposase protein itself acts preferentially in cis, thus ensuring that any transposition - defective mutants cannot be effectively complemented in trans and also

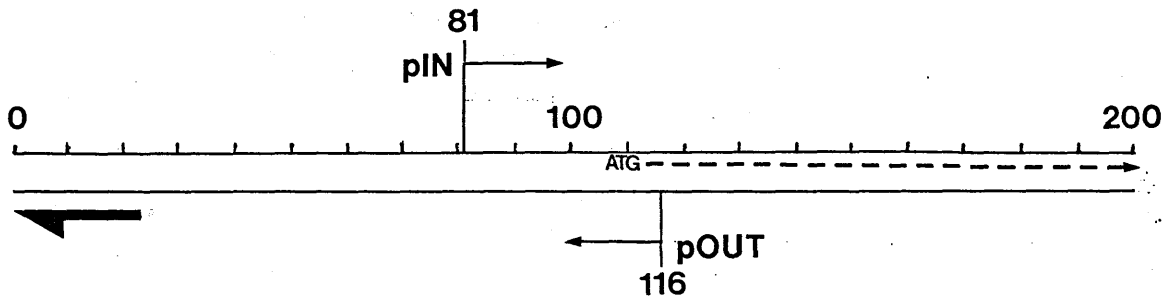
that , as the number of transposase genes increases, the effective cellular concentration of transposase does not increase (Morisato et al, 1983). Lastly, the antisense RNA regulatory mechanism already alluded to plays an important role.

This mechanism was elucidated in the study of 'multi-copy inhibition', a phenomenon by which an increase in the number of IS10 elements in a cell leads to a decrease in the transposition rate of a single copy chromosomal Tn10 element (Simons and Kleckner, 1983). Its workings can be summarised as follows:-

there are, near the termini of IS10, three promoters : pIN, pOUT and pIII. See Fig.1.5. pIN is the promoter for the transposase gene of IS10 located near the 'outside' end of IS10 and directs transcription 'inwards' towards the centre of Tn10. As already mentioned, this promoter is intrinsically weak. pIII is even weaker than pIN and also directs transcription 'inwards' towards the centre of Tn10, but it is situated near the 'inside' end of IS10 and its biological significance is unknown. pOUT is a much stronger promoter (c.33% of induced placUV5) and directs transcription 'outwards' from a point near the 'outside' end of IS10, very close, but internal, to pIN. Although a large proportion of the transcripts from pOUT terminate within IS10, a significant number pass on into adjacent host DNA sequences. The level of this readthrough transcription is c.10% of an induced placUV5 promoter and this is thought to be the mechanism of adjacent host gene 'turn-on' seen following some Tn10 insertions (Ciampi et al, 1982; Simons et al, 1983). pOUT does not appear to interfere with transcription initiation at pIN, but may interfere with elongation of the pIN transcript (Simons et al, 1983).

However, the most significant feature of this transcriptional arrangement is that the transcripts from pIN and pOUT overlap, and are therefore complementary, over a region of 36bp. that includes both the ribosome binding site and the translation initiation codon of the transposase mRNA. Thus it is proposed that the formation of an RNA:RNA

Fig.1.5.



– nearly perfect 23bp. terminal inverted repeat

ATG– –→ – transposase coding region

Fig.1.5. The outermost 200bp. of IS10-right of Tn10.

After Simons et al, 1983.

The diagram shows the exact start points for the transcripts generated from pIN and pOUT. It can thus be seen that these transcripts overlap by 36bp. over a region that includes both the ribosome binding site and the translation initiation codon of the transposase mRNA.

duplex between the pIN and pOUT transcripts will prevent translation of the transposase mRNA, thereby reducing the amount of transposase protein produced (Simons et al, 1983; Simons and Kleckner, 1983). Pairing between these transcripts has not yet been demonstrated in vivo ,though this has been done in vitro (N.Kleckner, pers. comm. to D. Sherratt), and that translation can be inhibited by sequestering of a ribosome binding site and/or translation initiation codon in a region of RNA:RNA duplex has also been demonstrated in many systems (Gold et al,1981; Hall et al, 1982; Bruckner & Matzura,1985). This therefore seems to be a highly plausible hypothesis.

Note that a consequence of this system is that the level of IS10 transposition is determined by the relative strengths of pIN and pOUT and the relative stabilities of their transcripts. Any change in either of these ratios will alter the amount of transposase protein synthesised and therefore the level of IS10 transposition.

Chapter 2.

Materials and methods.

2.1. Chemicals.

Chemical.	Source.
General chemicals and organic compounds.	B.D.H., Koch-light, Hopkins & Williams, B.C.L., May & Baker, Formachem Research Int., B.R.L.
Antibiotics.	Sigma.
Biochemicals.	Sigma, Koch-light.
Media.	Oxoid, Difco.
Agar.	Oxoid, Davis.
Agarose and low melting point agarose.	B.R.L.
S.D.S.	Serva.
Radiochemicals.	Amersham Int., New England Nuclear.
Lambda DNA.	Gift from M.Burke.
pKK223-3 plasmid DNA.	Gift from I.Anton.
pKO500, pKL500 and pKM1 plasmid DNAs.	Gifts from A.Lamond.

2.2. Enzymes.

All enzymes were obtained from B.R.L., with the exceptions of the following:- AhaIII - P&S Biochemicals Ltd.

calf-intestinal alkaline phosphatase - B.C.L.

lysozyme - Sigma.

M13 universal primer and yeast t-RNA were also obtained from B.R.L.

2.3. Bacterial strains.

The bacterial strains used were all derivatives of Escherichia coli K-12 and are listed in Table 2.1. Nomenclature is after Demerec et al, 1966.

2.4. Plasmids, bacteriophages and transposons.

Plasmids and bacteriophages are listed in Table 2.2. Symbols for genotype and phenotype are those recommended by Bachman et al, 1976 and Campbell et al,1976. Transposons are listed in Table 2.3, with the genotype of Tn1/3 derivatives being based on the nomenclature of Heffron et al,1979.

2.5. Culture media.

L broth: 10g.tryptone, 5g.yeast extract, 5g. NaCl, 1g.glucose, 20mg.thiamine made up to 1 litre with distilled water and adjusted to pH7.0 with NaOH.

L agar: above broth containing 1.5% agar.

Table 2.1: Bacterial strains.

Strain.	Relevant markers.	Source or reference.
△M15.	[*lac pro*]△ *thi* ϕd80*lacZ*⁻△M15.	Ruther *et al*, 1981.
JM101.	△ *lac pro supE thi* F'*traD*36 *proAB lacI*^{q} Z△M15.	Messing, 1983.
DS902.	*thr leu his pro arg recA*13 *rpsL* (=Str^{r}).	D.J. Sherratt.
DS903.	*thr leu thi lacY*1 *galK*2 *ara*14 *xyl*5 *proA*2 *his*4 *arg*C3 *tsx*33 *supE*44 Str^{r} *recF*.	D.J. Sherratt.
DS910.	*minA minB rpsL* (=Str^{r}).	D.J. Sherratt.
DS916.	*trp his recA*56 *rpoB* (=Rif^{r}).	D.J. Sherratt.
CJ100.	Spc^{r} Str^{s} P1 transductant derivative of DS903.	C.T. Jones, 1985.

Table 2.2: Plasmids and bacteriophages.

Plasmid/ 'phage.	Description.	Phenotype.	Size(kb.).	Source/reference.
pACYC184	Vector derived from P15A	Cm^r Tc^r	4.0	Chang & Cohen, 1978.
pUC8	Vector derived from pBR322	Ap^r	2.67	Vieira & Messing, 1982.
pUC9	Vector derived from pBR322	Ap^r	2.67	Vieira & Messing, 1982.
pUC18	Vector derived from pBR322	Ap^r	2.67	Yanisch-Perron *et al*, 1985.
pKK223-3	*tac* promoter expression vector	Ap^r	4.6.	I. Anton.
pKL500	Terminator probe vector	Ap^r	4.3	A. Lamond.
pKO-1	Promoter probe vector	Ap^r	4.0.	Mc.Kenney *et al*, 1981.
pKO500	Promoter probe vector	Ap^r.	4.1.	A. Lamond.
pKM-1	Promoter probe vector	Ap^r	4.4	A. Lamond.

Table 2.2 (cont.).

RSF1050	pMB8::Tn3	Ap^r	7.7	Heffron *et al*, 1977.
R388	Naturally occurring	Tp^r	33.0	Datta & Hedges, 1972.
pMB9:: Tn103	Insertion of Tn103 into pMB9	Tc^r	9.5	A. Arthur.
ColE1	Naturally occurring	*cea*$^+$ *iea*$^+$.	6.646.	Chan *et al*, 1985.
pDS4153	ColK::Tn1,Δ *Hae*II 53,*Hae*II scramble of pDS4101	Ap^r *tnpR*$^+$	8.1	Tacon *et al*, 1981.
pCB101*	Lambda dv vector	Cm^r	5.0	A.C. Boyd.
pKS400	pUC8 + 2.7kb. *Pst*I fragment from pDS1118	Ap^r	5.4	D.K.Summers; Arthur *et al*, 1984.
pKS401	pUC9+ 2.7kb. *Pst*I fragment from pDS1118	Ap^r	5.4	D.K. Summers; Arthur *et al*, 1984.
pPAK100	pAA231 *Bam*HI deletion	Cm^r	7.2	Kitts, 1982.

* - devoid of Tn1/3 sequences.

Table 2.2 (cont.).

pPAK200	pACYC184::Tn3651	Cm^r, Ap^r	7.9	Kitts, 1982.
pAA15	Resolvase-mediated resolution product of a pMB9::Tn103 derivative containing the Tn1000 *res* site	Tc^r	8.2	Arthur *et al*, 1984.
pAA32	pACYC184::Tn103	Cm^r, Tc^r	8.1	Arthur *et al*, 1984.
pAA35	4.7kb.*Bam*HI fragment of pAA30 + 4.5kb.*Bam*HI fragment of pAA32	Cm^r, Tc^r	9.2	Arthur *et al*, 1984.
pAA36	3.6kb.*Bam*HI fragment of pAA32 + 2.8kb.*Bam*HI fragment of pAA30	Cm^r.	6.4.	Arthur *et al*, 1984.
pAA361	As for pAA36, but with pAA30 fragment in opposite orientation	Cm^r	6.4	Arthur *et al*, 1984.
pJKA	282bp.*Eco*RI* fragment from Tn3 inserted into pKO-1	Ap^r	4.26.	Kelly, 1983.

Table 2.2 (cont.).

pJKR	As for pJKA, but fragment inserted in opposite orientation	Ap^r	4.26	Kelly, 1983.
pGR06	837bp.Sau3A fragment from Tn3 inserted into pUC8	Ap^r	3.51	G. Russell.
pSN015	4.0kb.HincII fragment from pPAK100 inserted into pUC8	Ap^r	6.7.	Chapter 4.
pSN016	As for pSN016, but fragment inserted in opposite orientation	Ap^r	6.7	Chapter 4.
pSN021	4.0kb.EcoRI-HindIII fragment from pSN015 inserted into pKK223-3	Ap^r	8.6	Chapter 4.
pSN070	4.0kb.BamHI-HindIII fragment from pSN015 inserted into pUC18	Ap^r	6.7.	Chapter 4.
pSN033	192bp.TaqI fragment from Tn3 + 3.66kb. AhaIII fragment of pACYC184	Tc^r	3.85	Chapter 5.

Table 2.2 (cont.).

pSNO34	As for pSNO33, but fragment inserted in opposite orientation	Tc^r	3.85	Chapter 5.
pSNO43	285bp.HaeIII-PvuII fragment from Tn3 inserted into pKO500	Ap^r	4.3	Chapter 5.
pSNO47	As for pSNO43, but fragment inserted in opposite orientation	Ap^r	4.3	Chapter 5.
pSNO48	340bp.HindIII-EcoRI fragment from pSNO47 inserted into pKM-1	Ap^r	4.42	Chapter 5.
pSNO52	859bp.EcoRI-HindIII fragment from pGRO6 inserted into pKO500	Ap^r	4.9	Chapter 5.
pSNO53	880bp.EcoRI-HindIII fragment from M13mpGRO41 inserted into pKM-1	Ap^r	4.97	Chapter 5.
M13mp18	'Phage sequencing vector	-	7.23	Yanisch-Perron et al, 1985.

Table 2.2 (cont.).

M13mpGRO41	837bp.Sau3A fragment from Tn3 inserted into M13mp18	-	8.07	G. Russell.
M13mpSNO7	4.0kb.HindIII-EcoRI fragment from pSNO16 inserted into M13mp18	-	11.3	Chapter 4.

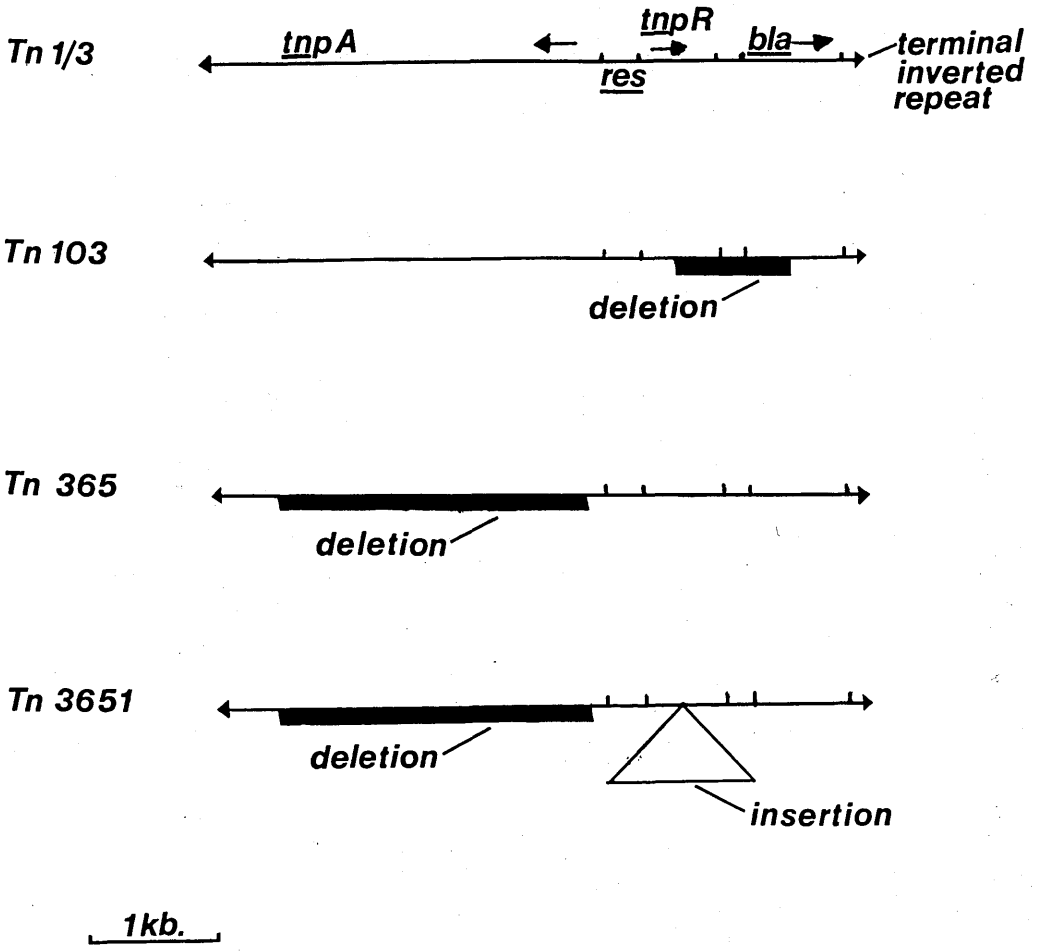


Table 2.3: Transposons.

Common name or location.	Systematic name.	Derivation.	Phenotype.	Reference.
Tn1.	Tn1.	Naturally occurring on RP4.	*tnpA*$^+$ *tnpR*$^+$ *res*$^+$ Ap^r.	Hedges & Jacob, 1974.
Tn3.	Tn3.	Naturally occurring on R1.	*tnpA*$^+$ *tnpR*$^+$ *res*$^+$ Ap^r.	Kopecko & Cohen, 1975.
Gamma-delta.	Tn1000.	Naturally occurring on F.	*tnpA*$^+$ *tnpR*$^+$ *res*$^+$.	Guyer, 1978.
Tn103.	Tn1ΔAp.	Spontaneous deletion of Tn1.	*tnpA*$^+$ *tnpR*$^-$ *res*$^+$ Ap^s.	Heffron *et al*, 1977.
Tn3651.	Tn3Δ3651.	Insertion into *Bam*HI site of Tn365.	*tnpA*$^-$ *tnpR*$^-$ *res*$^+$ Ap^r.	Kitts, 1982.

(See opposite).

L broth without glucose: above broth with glucose omitted.

2-YT: 16g.tryptone, 10g.yeast extract, 5g.NaCl made up to 1 litre with distilled water and adjusted to pH7.0 with NaOH.

Isosensitest broth: 23.4g.Isosensitest base made up to 1 litre with distilled water.

Isosensitest agar: above broth containing 1.25% agar.

Minimal broth: 7g.K_2HPO_4, 2g.KH_2PO_4, 4g.$(NH_4)_2SO_4$, 0.25g.tri-sodium citrate, 0.1g.$MgSO_4.7H_2O$ made up to 1 litre in distilled water.

Minimal agar: above broth containing 1.5% agar.

When required, supplements were added to minimal medium at the following concentrations:- glucose - 2mg./ml.
thiamine (vitamin B_1) - 20ug./ml.

Soft agar: 6g.agar in 1 litre of distilled water.

Mac.Conkey medium with galactose: 20g.peptone, 3g.bile salts, 5g.NaCl, 0.03g.neutral red, 0.001g.crystal violet, 10g.galactose, 13.5g.agar made up to 1 litre with distilled water and adjusted to pH7.1 with NaOH.
(Oxoid,1982).

Minicell medium: 3.2ml.glycerol, 6g.Na_2HPO_4, 3g.KH_2PO_4, 0.5g.NaCl, 1g.NH_4Cl made up to 1 litre with distilled water.

All growth media were sterilised by autoclaving at 121°C for 15mins.

2.6. Buffer and other solutions.

Buffer E: 40mM Tris HCl, 20mM NaAc, 1mM $Na_2EDTA.2H_2O$, pH adjusted to 8.2 with glacial acetic acid.

1xTE (general use): 10mM Tris HCl, 1mM $Na_2EDTA.2H_2O$, pH7.6.

1xTE (for use in sequencing reactions): 10mMTris HCl, 1mM $Na_2EDTA.2H_2O$, pH8.0.

1xTBE: 10.8g.Tris HCl, 5.5g.boric acid, 4ml.0.5M $Na_2EDTA.2H_2O$ made up to 1 litre with distilled water. pH8.3.

Gelatin in TE: 100ug. gelatin/ml. of TE.

Tris/glycine running buffer for protein gels: 14.41g.glycine, 3.03g.Tris HCl, 1g.SDS made up to 1 litre with distilled water.

DNA final sample buffer (FSB): 10% Ficoll, 0.5% SDS, 0.05% bromophenol blue, 0.06% orange G in buffer E.

Single colony lysis/gel buffer: 2% Ficoll, 1% SDS, 0.01% bromophenol blue, 0.01% orange G in buffer E.

Protein final sample buffer: 10% glycerol, 3% SDS, 5% B-mercaptoethanol, 0.01% bromophenol blue in 0.0625M Tris HCl, pH8.0.

Formamide dye (for sequencing gels): 100ml.formamide stirred with 5g. Amberlite MB1 resin for 30 mins. and then filtered to remove resin. 0.03g.xylene cyanol FF, 0.03g. bromophenol blue and 0.75g.$Na_2EDTA.2H_2O$ added and dissolved.

PEG/NaCl solution: 2.5M NaCl/20% polyethylene glycol 6000.

Lytic mix: 1% Triton X-100, 50mM Tris HCl, 60mM $Na_2EDTA.2H_2O$, pH adjusted to 8.0 with HCl.

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

Calcium chloride solution: 50mM $CaCl_2.2H_2O$ in distilled water.

Magnesium sulphate solution: 100mM $MgSO_4.7H_2O$ in distilled water.

Buffered saline gelatin (B.S.G.): 8.5g.NaCl, 0.3g.KH_2PO_4, 0.6g.Na_2HPO_4, 0.1g gelatin made up to 1 litre with distilled water.
: sucrose added to 5% or 20% if appropriate.

Upper buffer for SDS-PAGE gel (4x): 0.5M Tris HCl, pH6.8, 0.4% SDS.

Lower buffer for SDS-PAGE gel (4x): 1.5M Tris HCl, pH8.8, 0.4% SDS.

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.

10x ligation buffer: 660mM Tris HCl, pH 7.5, 100mM $MgCl_2$, 10mM DTT, 10mM $Na_2EDTA.2H_2O$.

10x nick translation buffer: 500mM Tris HCl, pH 7.8, 50mM $MgCl_2$, 10mM DTT.

10x Klenow buffer: 100mM Tris HCl, pH 8.0, 100mM $MgCl_2$.

Restriction buffers:

10x low salt: 100mM Tris HCl, pH 7.4, 100mM $MgSO_4.7H_2O$, 10mM DTT.

10x medium salt: 500mM NaCl, 100mM Tris HCl, pH 7.4, 10mM DTT, 100mM $MgSO_4.7.H_2O$.

10x high salt: 1M NaCl, 500mM Tris HCl, pH 7.4, 100mM $MgSO_4.7H_2O$, 10mM DTT.

10x AhaIII buffer: 250mM NaCl, 60mM Tris HCl, pH 7.6, 60mM $MgCl_2$, 60mM B-mercaptoethanol, 100ug./ml.BSA.

10x SmaI buffer: 200mM KCl, 100mM Tris HCl, pH 8.0, 10mM DTT, 100mM $MgSO_4.7H_2O$.

10x BamHI buffer: 1M NaCl, 200mM Tris HCl, pH 8.0, 70mM $MgCl_2$, 20mM B-mercaptoethanol.

All the above solutions were sterilised by filtration or by autoclaving at 108°C for 10 mins., except the $CaCl_2$ and $MgSO_4$ solutions which were autoclaved at 114°C for 10 mins.

2.7. Growth conditions.

Liquid cultures for transformations, matings, minicell experiments or plasmid DNA preparations were grown up from a single colony inoculum, taken from a fresh selective plate, with shaking at 30°C or 37°C according to the strain in use and the experiment being performed. The medium used was L broth, unless trimethoprim selection was being applied, in which case Isosensitest broth was used, or unless JM101 was the strain in use, in which case minimal broth + glucose and thiamine was used.

Stationary phase overnight cultures were grown up at 30°C or 37°C from a single colony inoculum without shaking.

Growth of cells on plates was on L, Isosensitest or minimal agar or Mac.Conkey's agar with galactose according to the strain and antibiotics in use. Plates contained c.25ml. agar solution and appropriate additives and were incubated for 24-36 hours at 30°C or 12-18 hours at 37°C unless otherwise stated. Dilution and washing of cells were carried out in 'phage buffer.

Bacterial strains were stored on L agar slopes at room temperature or frozen in the following way at -20°C:- 0.75ml. of a stationary phase overnight culture was mixed with 0.75ml. of a 40% glycerol, 2% peptone solution and the mixture stored at -20°C.

2.8. Antibiotic selections.

The standard concentrations listed in Table 2.4 were used throughout for both liquid and plate selections, except for chloramphenicol which was used at 25ug./ml. in plates and 50ug./ml. in liquid. Agar solutions were cooled to 50°C before adding the antibiotic solution(s). All antibiotic stock solutions were sterilised by filtration.

If trimethoprim was being used in selection, Isosensitest agar or broth was used as the growth medium. This was because trimethoprim acts by inhibiting thymine synthesis and is not effective in media containing exogenous thymine (Stacey and Simpson, 1965). All other antibiotics were used in L broth or agar or minimal broth or agar or Mac.Conkey's medium with galactose except when used in conjunction with trimethoprim.

2.9. Use of X-gal and IPTG.

X-gal (5-bromo, 4-chloro, 3-indolyl B-D galactoside) is a gratuitous substrate for the enzyme B-galactosidase. It is hydrolysed by this enzyme, giving colonies of cells possessing the active enzyme a blue colour when grown on a medium containing X-gal. Cells lacking the active enzyme produce white colonies. In the 'phage and plasmid vectors M13mp8 & 9, pUC18 & 19, etc., this colour difference is the basis of a screen for detecting plaques or colonies containing recombinant molecules.

IPTG (isopropyl B-D thiogalactopyranoside) is a gratuitous inducer of the *lac* promoter (and the related, hybrid *tac* promoter: de Boer *et al*, 1983) and its inclusion in a growth medium ensures that these promoters are fully induced. In the cases of these promoters, the absence of IPTG in the growth medium ensures that the promoters are not active in a *lacI*^{q} or *lacI*$^{super-q}$ strain. Thus, in such a strain, IPTG can be used as a 'switch' to regulate the expression of a gene cloned under the control of either of these promoters.

X-gal and IPTG were freshly made up at concentrations of 20mg./ml.

Table 2.4: Antibiotic selections.

Antibiotic.	Source of resistance.	Selective concentration.	Stock solution.
Ampicillin.	Transposon.	50ug./ml.	5mg./ml. in water.
Chloramphenicol.	Plasmid.	25ug./ml.	5mg./ml. in EtOH.
Rifampicin.	Chromosome.	50ug./ml.	5mg./ml. in methanol. (Made fresh).
Spectinomycin.	Chromosome.	50ug./ml.	5mg./ml. in water.
Streptomycin.	Chromosome.	50ug./ml.	5mg./ml. in water.
Tetracycline.	Plasmid.	10ug./ml.	1mg./ml. in 100mM.NaOH. (Made fresh).
Trimethoprim.	Plasmid.	50ug./ml.	5mg./ml. in 50% EtOH.

in dimethylformamide and distilled water respectively. In L agar plates, both were used at a concentration of 25ug./ml., and in soft agar overlays on minimal agar plates both were used at a concentration of 200ug./ml.

2.10. Plasmid and M13 RF DNA isolation.

Plasmid and M13 replicative form (RF) DNAs for all *in vitro* manipulations were prepared by the cleared lysate/CsCl method. Small amounts of such DNAs suitable for restriction analysis were prepared by the alkaline/SDS method. DNAs for use as gel markers or in transformations were prepared by either method, as convenient.

(i). cleared lysate/CsCl method.

a). plasmid DNAs: 100ml. of L broth plus appropriate antibiotic(s) was inoculated from a single colony of the plasmid-containing strain and grown up overnight (c.16 hours) with shaking. The cells were harvested by centrifugation (12000g, 2 mins., 4°C) and resuspended in 2.5ml. of an ice-cold 25% sucrose in 50mM Tris HCl, pH 8.0 solution. 0.5ml. of a 20mg./ml solution of lysozyme in 250mM Tris HCl, pH8.0, was thoroughly mixed in and the suspension allowed to stand on ice for 30-60 mins., being shaken every 10 mins. 1ml. of 250mM $Na_2EDTA.2H_2O$, pH 8.0, was added, mixed in and the suspension left on ice for a further 5 mins. 4ml. of lytic mix was gently added, mixed in and the mixture left at room temperature for 15-30 mins. to form a 'crude lysate'. The cell debris was pelleted (48000g, 30mins., 4°C) and the supernatant (the 'cleared lysate') carefully decanted. 50ul. of a 1 mg./ml. solution of ribonuclease in TE (previously boiled for two minutes to destroy any deoxyribonuclease activity) was mixed in and the solution allowed to stand at room temperature for 15 mins. 4.83ml. of this solution was added to 5.00g. of caesium chloride and the CsCl allowed to dissolve.

0.33ml. of a 3mg./ml. solution of ethidium bromide in TE was added, the density of the resultant solution checked and adjusted to 1.58-1.60g./ml. if necessary. The solution was poured into a 'Quick-Seal' ultracentrifuge tube, the tube filled with paraffin oil and sealed. After ultracentrifugation (200000g, 16 hours, 15°C in an angled rotor; or 270000g, 4 hours, 15°C in a vertical rotor), the DNA bands were visualised under long wavelength ultra-violet light (300-360nm.) and the lower band (containing the covalently closed, circular plasmid DNA) removed through the side of the tube. Three volumes of TE buffer were added to this material and the EtBr removed by repeated butanol extraction until the butanol layer was colourless. The DNA was then precipitated with two volumes of ethanol and resuspended in 250ul. of 2xTE buffer. Its concentration was estimated by running an aliquot on an agarose gel.

b). M13 RF DNAs: the treatment of the overnight culture was as described above for plasmid DNAs. However, the preparation of the overnight culture differed as follows:- a single plaque containing the desired 'phage was picked and placed in 1ml. of 'phage buffer for 3-4 hours at 4°C to form a 'phage suspension. A 5ml. liquid culture of JM101 cells was grown up to a density of c.$1x10^8$ cells/ml., 0.5ml. of the above 'phage suspension added to this and growth continued for a further 2 hours. This whole culture was then added to a 100ml. culture of JM101 cells (of density c.$1x10^8$ cells/ml.) and this was grown overnight at 37°C with vigorous shaking.

(ii). alkaline/SDS method.

This is a modification of the method of Birnboim and Doly, 1979.

a). plasmid DNAs: single colonies were used to inoculate 2.5ml. aliquots of L broth and these were grown to stationary phase overnight. The cells were harvested by centrifugation (12000g, 2 mins., 4°C) in a large (1.5ml.) microfuge tube, the supernatant removed by aspiration, the

cells resuspended in 100ul. of lysis buffer (50mM glucose, 25mM Tris HCl, pH 8.0, 10mM $Na_2EDTA.2H_2O$, 4mg./ml. lysozyme) and left at room temperature for 5 mins. 200ul. of a freshly made 200mM NaOH, 1% SDS solution was added and the mixture left on ice for a further 5 mins. 150ul. of pre-cooled 3M potassium acetate/2M acetic acid solution, pH 4.8, was added and left on ice for another 5 mins. Cell debris and chromosomal DNA were removed by centrifugation (12000g, 1 min., 4°C) and the supernatant carefully decanted. An equal volume of a phenol/chloroform (1:1 - v/v) solution was mixed in and the phases separated by centrifugation (12000g, 5 mins., 4°C). The aqueous phase was removed and the nucleic acids precipitated from it by adding two volumes of ethanol and leaving at room temperature for 2 mins. The precipitated material was spun down (12000g, 1 min., room temperature) and the supernatant discarded. The pellet was washed in 1ml. of 85% ethanol in TE, re-centrifuged (12000g, 5 mins., room temperature), dried and resuspended in 45ul. of 2xTE plus 5ul. of a 1mg./ml. solution of RNase in 1xTE. 10ul. aliquots were used for restriction analysis or transformation.

b). M13 RF DNAs: again, the isolation of DNA was just as described above but the growth of the cell culture was different as follows: a single plaque containing 'phage was placed in 1ml. of 'phage buffer and left at 4°C for 2-3 hours to form a 'phage suspension. A 2.5ml. culture of JM101 cells was grown to a density of c.$1x10^8$ cells/ml., 200ul of the above 'phage suspension added to it and growth continued for a further 5 hours.

2.11. Phenol extraction; precipitation of nucleic acids.

(i). phenol extraction: an equal volume of distilled phenol (saturated with 1M Tris HCl, pH 8.0) was added to the solution to be extracted, thoroughly mixed in and the two phases resolved by

centrifugation (12000g, 5mins., 4°C). The aqueous phase was carefully removed and extraction repeated once or twice more. The final aqueous phase was extracted two or three times with an equal volume of chloroform or diethyl ether to remove all traces of phenol.

(ii). precipitation of nucleic acids: this was performed with either isopropanol or ethanol. Unless there was a volume constraint, ethanol was used as there is less co-precipitation of non-nucleic acid materials (eg.salt, sugars) with ethanol and it is also more volatile and therefore more easily removed (Maniatis et al, 1982).

a). with ethanol: 0.1 volumes of 5M NaCl and two volumes of ethanol were added to the solution to be precipitated, mixed in and the mixture left at -20°C for 20-120 mins.

b). with isopropanol: 0.1 volumes of 5M NaCl and 0.54 volumes of isopropanol were added, mixed in and the mixture left at room temperature for 20-120 mins.

In either case, if only a small amount of DNA was being precipitated, 0.01 volumes of a 1mg./ml. solution of yeast tRNA in TE was also added to aid in the formation of a precipitate.

The precipitated material was spun down (27000g, 20 mins., 4°C), washed in 1ml. of 85% ethanol in TE, re-spun (27000g, 10mins., 4°C) and the washing and re-spinning repeated. The pellet was dried and resuspended in an appropriate volume of 1x or 2x TE.

2.12. Gel electrophoresis.

(i). standard agarose gels.

Vertical gel kits were used which held two 16.5 x 15.5cm. glass plates 0.3cm. apart. Agarose was dissolved in buffer E at 100°C and cooled to 55°C before being poured between the glass plates to fill

them. (If the gel was to be used for the analysis of single colony gel samples, RNase at a concentration of 1ug./ml. of agarose was mixed in with the agarose after cooling and before pouring). A 10 or 15-toothed 'Teflon' comb was quickly inserted into the top of the gel and, once the gel had set, this was removed and the top and bottom reservoirs of the gel apparatus filled with sufficient buffer E to cover the gel. Single colony gel supernatants were loaded directly onto the gel, whilst marker plasmid DNAs and restriction digests were made up to 20ul. with TE buffer (if necessary) and mixed with 5ul. of DNA final sample buffer before loading. Gels were run at constant voltage - 2 volts/cm. for 15 hours or 6 volts/cm. for 5 hours - at room temperature and were stained by soaking in an EtBr solution (0.5ug./ml. in buffer E) for 30-45 mins. They were viewed on a 260nm. ultra-violet transilluminator and photographed on Ilford HP-5 film using a 35mm. camera fitted with a red filter.

Interpretation of gels, based on Dugaiczyk et al, 1975, was as follows: the fastest moving band was the supercoiled (sc.) monomeric plasmid, which was usually the most abundant species. Behind this was another prominent band containing monomeric open circle (oc.) and supercoiled dimeric plasmid. Open circle dimers and other higher forms (not always clearly visible) ran more slowly still. Plasmid linears, which ran between the sc. and oc. forms of the plasmid, could sometimes be seen on single colony gels. Sheared fragments of chromosomal DNA ran as a single thick band towards the top of the gel.

0.8% agarose was used for all purposes: single colony gels, intact plasmid DNAs and analysis of restriction enzyme digests.

(ii). low melting point agarose gels.

These gels were used to isolate particular fragments from a restriction enzyme digest for use in a cloning procedure. A horizontal gel system was used for this type of agarose as it is considerably less

robust than general purpose agarose. The agarose concentration used was always 1%.

Agarose was dissolved in buffer E at 100°C and cooled to 37°C. A 1.5mm. thick comb was suspended over a plastic tray (19 x 11.5 x 0.4cm.) so that there was a gap of 0.5mm. between the bottom of the teeth and the floor of the tray. The agarose was poured in, allowed to set, the comb removed and the gel immersed in buffer E. The samples (15-25ul.) were loaded and the gel run at 2 volts/cm. for 15 hours at room temperature. Staining was as described for standard agarose gels above.

(iii). standard polyacrylamide gels (non-denaturing).

These gels were used to separate restriction fragments less than 1kb. in size. The acrylamide concentration could be varied, according to the size(s) of the fragment(s) being analysed. Throughout this work, 5% gels were used which gave good separation over a size range of 80-500 bp.

Vertical gels were again used, 16.5 x 15.5cm. with a 13-tooth plastic comb. The gel kit was first sealed using 0.8% agarose in TBE and 5% acrylamide prepared by mixing 25.3ml. distilled water, 4ml. 10x TBE buffer, 10ml. 20% acrylamide/1% N N' bis-methylene acrylamide, 240ul 10% TEMED (N N N'N' tetramethylethylenediamine) and 480ul.10% APS (ammonium persulphate - $(NH_4)_2S_2O_8$).

This solution was poured between the gel plates, the comb quickly inserted and the gel allowed to polymerise for at least one hour at room temperature. The comb was removed, the top and bottom reservoirs filled with TBE buffer, the wells washed out to remove any unpolymerised acrylamide, the samples (15-25ul.) loaded and the gel run at a constant current of 20mA. for c.3 hours at room temperature. Staining and photography were as described for standard agarose gels above, unless bands were being excised for further manipulations (see below).

(iv). SDS-polyacrylamide gels for analysis of protein samples.

These are described in the section on the use of E.coli minicells (2.22) below.

(v). denaturing polyacrylamide gels for analysis of DNA sequencing reaction products.

These are described in the section on DNA sequencing (2.23) below.

(vi). interpretation of restriction data.

The size of linear restriction fragments was estimated using the relationship:

$$\log M = c \times {}^{1}/_{D}$$ - Helling et al, 1974.

M = molecular size in base pairs.

D = distance migrated. c = a constant.

Molecular size standards were obtained by restriction of lambda C_I857 Sam7 DNA or pUC8 or pUC9 plasmid DNA (Philippsen et al, 1978; Haggerty and Scheif, 1976; Vieira and Messing, 1982).

2.13. Single colony / single plaque gel analysis.

(i). for plasmid DNAs: 'single colony' gel analysis: a single colony was streaked out onto a selective plate and grown overnight. Using a toothpick, a mass of cells (c.0.6cm^2 from an area of confluent growth) was removed from the plate and thoroughly resuspended in 150ul. of single colony gel buffer in a small (0.6ml.) microfuge tube. The cells were left to lyse for 15-30 mins. at room temperature, the cell debris pelleted by centrifugation (12000g, 15 mins., 4^{o}C) and 50ul. of the

supernatant loaded onto an agarose gel containing RNase (1ug./ml.).

(ii). for M13 RF DNAs; 'single plaque' gel analysis: cultures were grown up as described above for alkaline/SDS DNA preparations of M13 RF DNAs [2.10(ii)b] and 1ml. of culture removed to a large microfuge tube. The cells were pelleted (12000g, 2 mins., 4°C), the supernatant discarded and the cells resuspended in 150ul. of single colony gel buffer. Subsequent treatment was as described above for plasmid containing strains.

2.14. Extraction of DNA from gels.

(i). low melting point agarose gels.

After staining, the gel was illuminated using long wavelength ultraviolet light (300-360nm.) and the gel slice containing the fragment of interest excised. This slice was melted at 65°C and diluted by adding two volumes of buffer E, also at 65°C. After brief mixing, the solution was cooled to 37°C, phenol extracted three times, chloroform extracted twice, the DNA recovered by ethanol precipitation and resuspended in TE.

(ii). polyacrylamide gels.

After staining, the gel was illuminated using long wavelength ultraviolet light and the gel slice containing the fragment of interest excised. This slice was placed inside a piece of dialysis tubing (which had previously been boiled for c.20 mins. in a 1mM $Na_2EDTA.2H_2O$ solution) with 250ul. of TBE buffer, the ends sealed and the whole sac placed in a gel box and covered with buffer E. Electrophoresis was carried out at 50 volts for 90-180 mins. (depending on the size of the fragment involved) and at the end of this period the current was reversed for 30 seconds. The fluid from the sac was removed, the DNA

recovered by ethanol precipitation and resuspended in TE buffer.

2.15. Genetic transformation.

All plasmids, other than conjugative ones, and all M13 RF DNAs were introduced into different host strains by transformation.

(i). plasmid DNAs.

An overnight culture of the recipient strain was diluted 1 in 100 in liquid medium and grown with shaking to a density of c.$2x10^8$ cells/ml. The cells were harvested by centrifugation (12000g, 2 mins., 4°C), resuspended in 10ml. ice-cold 100mM $MgSO_4.7H_2O$ solution, re-harvested, resuspended in 10ml. ice-cold 50mM $CaCl_2.2H_2O$ solution and harvested again. They were then resuspended in 1ml. ice-cold 50mM $CaCl_2.2H_2O$ solution and 200ul. aliquots of this suspension dispensed to large microfuge tubes. The transforming DNA solution (1-50ul.in TE) was added and the tubes kept on ice for 15 mins. A heat shock was then carried out (30°C for 10 mins. or 37°C for 5 mins.) and the cells returned to ice for a further 15-60 mins. 1ml. of L broth was added to each tube and the cells incubated at 30°C or 37°C for 60-90 mins. to allow expression of the plasmid genes. Aliquots of the transformation mixture were spread onto appropriate selective plates and incubated as appropriate.

If the transforming DNA conferred the phenotype of ampicillin resistance on the recipient cells, the cells were spun down after the expression step (12000g, 2 mins., 4°C) and washed twice in 1ml. of 'phage buffer, being re-pelleted after each wash. They were finally resuspended in 500ul. of 'phage buffer and aliquots plated out on appropriate selective plates. This washing procedure removed exogenous B-lactamase from the liquid culture medium and thus prevented the growth of Ap^S cells on the selective plates.

The presence of plasmid DNA in transformed cells was checked by single colony gel analysis.

(ii). M13 RF DNAs.

The recipient strain used was always JM101, and cells were treated as described above for transformation with plasmid DNA up to and including the stage of addition of the transforming DNA. The cells were then left on ice for 1-2 hours, heat shocked at 37°C for 5 mins. and returned to ice for a further 1-2 hours. The whole transformation mixture was added to 2.5ml. molten (45°C) 0.6% agar, along with 200ul. of a log phase culture of JM101 cells (of density c.2x10^8 cells/ml.), 25ul. of a 20mg./ml. X-gal solution and 25ul. of a 20mg./ml. IPTG solution. All these components were briefly mixed and the solu tion poured over the surface of a minimal agar plate containing thiamine and glucose. After allowing the overlay to set, the plates were incubated at 37°C overnight.

2.16. Plate matings.

Plasmids of the IncW class like R388 produce fragile, rigid pili and hence conjugal transfer is c.20000x more efficient when conjugation occurs on a solid surface than when liquid culture is used (Bradley et al, 1980). Therefore, R388 matings were always carried out on the surface of an Isosensitest or L agar plate.

An overnight culture of the R388-containing (donor) strain was diluted 1 in 20 in fresh Isosensitest broth and grown up to a density of c.2x10^8 cells/ml. 5ml. of this culture was taken, mixed with 5ml. of an overnight culture of the recipient strain and the cells pelleted together (12000g, 2 mins., 20°C). They were resuspended in 500ul. Isosensitest broth and 200ul. of this suspension was spread over the

surface of a well dried agar plate. The plate was left to dry with the lid off for 20-30 mins. at 30°C or 37°C, the lid replaced and the plate incubated for a further 1-3 hours at the appropriate temperature. The cells were washed off the plate with two 10ml. aliquots of 'phage buffer, pelleted (12000g, 2 mins., 20°C), resuspended in 1ml. of 'phage buffer and dilutions plated out on appropriate selective plates and incubated as appropriate.

If the conjugative plasmid conferred the phenotype of ampicillin resistance on the recipient cells, the cells were washed twice with 10ml. of 'phage buffer, being re-pelleted after each wash, after being removed from the agar plate to remove exogenous B-lactamase from the system.

2.17. Restriction enzyme digestion.

1-10ul. of a DNA solution was mixed with 7ul. of gelatin in TE, 2ul. of 10x restriction buffer and the volume made up to 20ul. with TE. Restriction enzyme was added (at a ratio of 1 unit of enzyme per ug. of DNA; usually 0.1-1 ul. of enzyme solution), all the components mixed and the solution incubated at 37°C (65°C for TaqI digests) for 1-3 hours. If the sample was to be run on a gel, the reaction was terminated by adding 5ul. FSB. If the DNA was to be subject to further manipulation, the enzyme was inactivated by heating to 70°C for 10 mins. and the DNA fragments recovered by ethanol precipitation. (For TaqI digests, the reaction was terminated by phenol extraction twice, chloroform extraction twice and the fragments were then recovered by ethanol precipitation).

The restriction buffers used were:

a). low salt: HaeIII, HpaII, TaqI.
b). medium salt: AvaI, ClaI, HincII, HindIII, PstI, PvuII.
c). high salt: EcoRI, SalI.

AhaIII, BamHI and SmaI were used in their own particular buffers.

For multiple digests requiring different buffers, the lowest salt buffer was used first and restriction carried out for 1 hour. The salt concentration was then adjusted by adding an appropriate volume of 5M NaCl and the DNA incubated with the other enzyme(s).

For partial enzyme digests, the ratio of enzyme:DNA was reduced (to 0.1-0.5 units of enzyme per ug. of DNA) and restriction carried out for only a short time.

2.18. Calf-intestinal alkaline phosphatase (CIP) treatment.

If it was not possible to select or screen for the insertion of a DNA fragment into a plasmid, CIP was used to remove the 5' phosphate groups from the linearised vector to prevent its recircularisation. This was done by adding 2 units of CIP to the restriction digest mix after c.2 hours, mixing thoroughly and incubating for a further 10-15 mins. at 37°C. This procedure was effective in all restriction buffers used. The reaction was terminated by adding 5ul. of FSB (if the sample was to be loaded onto a gel) or by phenol extraction twice, chloroform extraction twice followed by ethanol precipitation to recover the DNA fragments (if the DNA was to be further manipulated).

2.19. 'Filling-in' of 'sticky ends' of DNA fragments.

After ethanol precipitation of restricted DNA, the fragments were resuspended in 10ul. of TE buffer and 2ul. 10x nick translation buffer added. 1mM aqueous solutions of each of the dNTPs needed for the 'filling-in' reaction were prepared, 1ul. of each added to the reaction and the total volume made up to 20ul. with TE. 2.5 units of T4 DNA polymerase were added, all the components mixed and incubated at 37°C for 30 mins. The reaction was terminated by two phenol extractions followed by two chloroform extractions and ethanol precipitation of the DNA fragments.

2.20. Ligation of DNA fragments

The fragments to be ligated were resuspended together in 16ul. of TE buffer, and 2ul. each of 10x ligation buffer and 4mM ATP (in 4mM Tris HCl, pH 7.5) added. T4 DNA ligase was added (0.01 units per ug. of DNA for 'sticky' ends; 1 unit per ug. for flush ends), all the components mixed and then incubated overnight at 18°C. The ligation mixture was used both undiluted and diluted 1 in 10 in the transformation of competent cells.

2.21. Galactokinase assays.

Assays were always performed in duplicate on any one day and assays were repeated at least twice so as to get a reliable mean figure for the galactokinase activity of each strain under test.

Overnight cultures of the strains to be assayed were grown up in L broth without glucose/containing ampicillin. These were diluted 1 in 20 in the same medium and grown up 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 incubated at 30°C for 10 mins. Assay reaction systems were made up in small microfuge tubes as follows:

5ul. of mix 1
+12.5ul. of mix 2
+2.5 ul. of mix 4.

(See Section 2.6 for the composition of these solutions).

The reactions were started by the addition of 5ul. of cell lysate. (Blank systems were set up by adding 5ul. of distilled water 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 2.3cm. diameter Whatman DE81 filter paper and the filters carrying the unwashed reaction systems dried. The other filters were washed 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 counter. Each filter was counted at least twice for two minutes each time.

Galactokinase activity was calculated using the following equation:

$$\text{Number of units of galactokinase activity} = \frac{(\text{Cpm from test systems} - \text{Cpm from blank systems})}{(\text{Average Cpm from 2 unwashed filters}) \times \text{Incubation time (mins.)} \times OD_{650}} \times 10400$$

Cpm = counts per minute. Ref.: Mc.Kenney *et al*, 1981.

2.22. Use of E.coli minicells.

The bacterial strain DS910 was used to examine specifically the expression of plasmid-encoded proteins, for the minicells budded off from whole cells in such minicell-producing strains contain only plasmid DNA and no chromosomal DNA (Adler et al, 1967; Frazer and Curtiss, 1975; Davie et al, 1984).

(i). preparation of minicell samples.

200ml. of L broth containing appropriate antibiotics was inoculated from a single colony of DS910 containing the plasmid under investigation and the culture grown up with shaking overnight at 37°C. (Some experiments were performed growing this culture at 30°C, but the yield of minicells was greatly decreased). The bulk of the whole cells were pelleted (1400g, 3 mins., 4°C), the supernatant decanted and re-centrifuged (15000g, 10 mins., 4°C) to pellet the minicells (and remaining whole cells). This pellet was resuspended in 1ml. of buffered saline gelatin (BSG), layered carefully onto a cold (4°C) 28ml. 5-20% sucrose/BSG gradient and centrifuged in a swing-out rotor (1660g, 20 mins., 4°C). The minicells formed a creamy band in the top $^1/_3$rd. of the gradient with the whole cells as a cloudy phase beneath. The minicell band was removed, the minicells pelleted (11500g, 2 mins., 4°C) and resuspended in 0.5ml. of BSG. The minicells were passed through a further two such sucrose/BSG gradients to ensure maximum purity. The final minicell pellet was resuspended in 0.5ml. of minicell medium and pre-incubated for 30 mins. at 30°C or 37°C. ^{35}S-methionine was added to each culture (5uCi per culture) and incubation continued for a further 30 mins. 5ml. L broth was added as a cold chase and incubation continued for a further 45mins. The minicells were pelleted (11500g, 2 mins., 4°C), excess labelled methionine removed by washing with 5ml. of 'phage buffer, the minicells re-pelleted and resuspended in 50-200ul. of

protein final sample buffer. This suspension was boiled for 5 mins. and either loaded straight onto a gel or kept frozen at -20°C until needed. After freezing, samples were always boiled for 5 mins. before loading onto a gel.

(ii). gel analysis of minicell samples.

The gel system used was a two-part SDS-polyacrylamide gel, a modification of that of Laemmli (1970). The upper part was a short stacking gel (acrylamide concentration = 4.5%) and the lower part was a longer separating gel (acrylamide concentration = 12.5%). The gel was poured in the same apparatus as used for standard polyacrylamide electrophoresis of DNA samples [see Section 2.12(iii)] and was poured in two parts:

a). separating gel: 10ml. of 4x lower gel buffer was mixed with 13.4ml. distilled water, 16.6 ml. 30% acrylamide / 0.8% N N' bis-methylene acrylamide, 20ul. TEMED and 120ul. 10% APS and the solution poured between the gel plates to a level 1cm. below the bottom edges of the teeth of the comb. The top of the gel was immediately sprayed with a 0.1% SDS solution to stop bubbles forming on it and the gel allowed to polymerise completely. The SDS solution was drained off.

b). stacking gel: 2.5ml. of 4x upper gel buffer was mixed with 6.0ml. distilled water, 1.5ml. 30% acrylamide / 0.8% N N' bis-methylene acrylamide, 20ul. TEMED and 30ul. 10% APS and the solution poured on top of the lower gel, ensuring there were no bubbles at the interface. The comb was quickly inserted, the gel allowed to polymerise and the comb removed. The gel was installed into the running apparatus, the reservoirs filled with protein gel Tris/glycine running buffer, the wells washed out with buffer to remove any unpolymerised acrylamide, the samples (2-30ul.) loaded and the gel run at a constant current of 30mA.

for 2.5-3 hours at room temperature. The gel plates were carefully separated, the stacking gel portion discarded and the separating gel portion fixed in a solution of methanol / glacial acetic acid / distilled water (50/7/50: v/v) for 45-60 mins. at room temperature. The gel was dried down under vacuum and the protein bands visualised by autoradiography of a sheet of Kodak 'X-Omat' S1 film at room temperature for 1-7 days. To increase sensitivity and ensure a linear relationship between the amount of radioactivity in a protein band and the intensity of the image produced by the band on the film, the film was pre-flashed at a distance of c.1 metre using a 'Magicube Eliminator' electronic flashgun covered with several sheets of tissue paper to decrease the light intensity and ensure diffuse, even illumination.

The apparent molecular weights of protein bands were calculated from their relative mobilities compared to ^{14}C-labelled size markers (concentration = 0.1ug./ml.) comprised of the following:

Myosin heavy chain - 200 kd.

Phosphorylase b - 92.5 kd.

Bovine serum albumin - 63 kd.

Ovalbumin - 46 kd.

Carbonic anhydrase - 30 kd.

Lysozyme - 14.5 kd.

If appropriate, the relative intensities of bands were determined by microdensitometry.

2.23. DNA sequencing.

The principal source of reference for this work was the 'M13 Cloning and Sequencing Handbook' published by Amersham International plc, Amersham, U.K. A more detailed account of the theory and practice of sequencing using M13 'phage vectors can be found there.

The aim of any DNA sequencing procedure is to generate a series of radio-labelled DNA fragments that share a common 5' terminus, differ in length at the 3' end by one defined nucleotide and which vary in length from 1-600 nucleotides. These can then be size-separated on a high resolution denaturing polyacrylamide gel and visualised by autoradiography.

Several methods for sequencing DNA have been developed (Sanger *et al*, 1977; Maxam and Gilbert, 1980) and, of these, it is the Sanger method of chain termination sequencing that is significantly more simple and rapid. One of its few drawbacks is that requires propagation of a single-stranded DNA template (the 'plus' strand of an M13 'phage) in an *E.coli* host cell (eg. strain JM101) and this is clearly a potentially mutagenic environment. However, assuming any mutation will be random and not site-specific, analysis of several independent isolates of each cloned DNA will allow correct determination of the DNA sequence. It is this method that was employed in this work.

This method makes use of the fact that the large ('Klenow') fragment of *E.coli* DNA polymerase I has an absolute requirement for a primer with a 3' hydroxyl group if it is going to perform nucleotide chain extension on a single-stranded template. Also, it will incorporate, albeit at lower efficiency, 2'3'-dideoxynucleotides into a growing polynucleotide chain. However, once such a nucleotide has been incorporated into a chain, no further extension can occur since this nucleotide lacks a 3' hydroxyl group. Thus, by varying the ratio of deoxy- : dideoxy-nucleotide for one type of nucleotide and having the other three deoxy-

nucleotides present in excess, it is possible to generate a series of fragments specifically terminated at each position in the nascent polynucleotide chain where this nucleotide occurs. To eliminate spurious bands resulting from failure to continue chain elongation from a 3' hydroxyl group, a cold chase is employed to produce very long oligonucleotides that do not substantially enter the gel. This can be done for each type of nucleotide and thus fragments of all possible lengths are produced to allow determination of the DNA sequence.

The primer used to initiate the chain elongation reaction was the commercially available 17bp. M13 universal primer.

Using the Klenow fragment of E.coli DNA polymerase I, rather than the holoenzyme, ensured that there was no degredation of fragments from the 5' end, which would have made the interpretation of results difficult.

The radio-labelled nucleotide used in this work was a-^{35}S-dATPaS. This was preferred to a-^{32}P-dATP since it is safer to work with, the longer half-life of ^{35}S allows samples to be kept for longer periods and, particularly, the lower energy B-emissions of ^{35}S produce tighter, less fuzzy bands on the autoradiograph, making it easier to read. dNTPaS's are incorporated less efficiently into a growing polynucleotide chain than dNTPs by E.coli DNA polymerase I (holoenzyme or Klenow fragment), but this problem is readily overcome by increasing the time allowed for the sequencing reaction (eg. from 15 mins. for dNTPs to 20 mins. for dNTPaS's).

All the procedures necessary for the generation, isolation and identification of recombinant M13 clones have been described above. The subsequent procedures needed for the determination of the sequence of the cloned fragment are as follows:

(i). preparation of single-stranded template DNA.

20ml. of 2-YT medium was inoculated with 200ul. of an overnight culture of JM101 cells and 1.5ml. aliquots of this mixture dispensed

into 10ml. culture tubes. Each tube was inoculated with a single M13 'phage [or 100ul. of a single 'phage suspension prepared as described in 2.10(i)b above] and grown at 37°C for 5 hours with vigorous shaking. The cells (containing all chromosomal, plasmid and M13 RF DNAs) were removed by centrifugaton (12000g, 5 mins., 4°C), the supernatant removed and re-centrifuged to remove all traces of host cells. 1.3ml. of this supernatant was added to 200ul. of 2.5M NaCl/20% PEG solution in a large microfuge tube and allowed to stand at room temperature for 15 mins. to precipitate viral particles. The solution was centrifuged (12000g, 5 mins., 4°C), the supernatant discarded, the remaining material re-centrifuged and all traces of supernatant scrupulously removed by aspiration, leaving only the white viral pellet. (If no viral pellet was visible at this stage, there was found to be no point in proceeding further with the prep.). The pellet was resuspended in 100ul. of TE buffer (pH 8.0) and this suspension phenol extracted three times, chloroform extracted twice to remove viral proteins and the single-stranded DNA recovered by overnight ethanol precipitation at -20°C. Finally, the DNA was spun down (12000g, 20 mins., 4°C), washed twice in 1ml. of 85% ethanol in TE buffer, being re-spun between each wash, dried and resuspended in 50ul. TE buffer. This method produces sufficiently pure single-stranded DNA for use in sequencing reactions (Schreier and Cortese, 1977).

(ii). annealing of primer to template.

To provide sufficient material for the four sequencing reactions for each clone, the following components were mixed together: 5ul. single-stranded template DNA, 2ul.(=4ng.) M13 universal primer, 1.5ul. 10x Klenow buffer and 1.5ul. distilled water. This mixture was incubated at 60°C for 60-90 minutes and then allowed to cool slowly (over a period of c.2 hours) to room temperature.

(iii). dideoxy sequencing reactions.

All solutions were made up in distilled water and kept frozen at -20°C.

Stock solutions at a concentration of 10mM of all four dNTPs were prepared and these used to make up working solutions (concentration = 0.5mM) and the cold chase mix needed for the sequencing reaction (a uniform mixture of all four dNTPs, each at a concentration of 0.5mM). dNTP (N^o) mixes were then prepared from these working solutions as shown in Table 2.5.

Stock solutions, also at a concentration of 10mM, of all four ddNTPs were prepared and used to make up working solutions. The optimum concentrations of these, when used with the above N^o solutions and the radio-labelled nucleotide described below, determined empirically, were as follows: 0.033mM ddATP

0.10mM ddCTP

0.15mM ddGTP

0.25mM ddTTP.

Equal volumes of each N^o mix and the corresponding ddNTP working solution were mixed to produce the N^o/ddNTP reaction mixes.

The radio-labelled nucleotide used was a-^{35}S-dATPaS of specific activity 1000-1300 Ci/mmole. The concentration of the solution supplied varied from 7.9-9.22uM. and this was adjusted by the addition of an appropriate volume of a 100uM. cold dATPaS solution to give a final overall concentration of 16uM dATPaS. Such a volume of this solution was taken (1.5ul.) so that each reaction system contained 7.60-8.87 pmoles (= 8.07-11.54 uCi.) of a-^{35}S-dATPaS.

To carry out the sequencing reactions, 1.5ul. of the above labelled nucleotide mix was added to the annealed primer/template mix and 1ul. of Klenow enzyme solution (1 unit/ul.) mixed in carefully by pipetting in

Table 2.5: N° mixes for sequencing reactions.

	A°.	C°.	G°.	T°.
0.5mM. dCTP.	30ul.	1.5ul.	30ul.	30ul.
0.5mM. dGTP.	30ul.	30ul.	1.5ul.	30ul.
0.5mM. dTTP.	30ul.	30ul.	30ul.	1.5ul.
1x TE buffer.	30ul.	30ul.	30ul.	30ul.

and out. Four small microfuge tubes were labelled A,G,C, or T and 2.5ul. of the annealed primer/template/labelled nucleotide/enzyme mix put into each. 2ul. of appropriate N°/ddNTP reaction mix (A°/ddATP to A tube, etc.) was placed just inside the rim of each tube, the contents mixed and the reaction started by a brief spin. The reaction was allowed to proceed for 20 mins. at room temperature, 2ul. of the chase mix placed just inside the rim of each tube, mixed in by a brief spin and the reaction allowed to proceed for a further 15 mins. at room temperature. The reaction was stopped by the addition of 4ul. of formamide dye mix to each tube. Samples were then either immediately prepared for loading onto a gel or stored at -20°C until convenient.

(iv). high resolution polyacrylamide gel electrophoresis.

The gel apparatus used consisted of two 45 x 23cm. siliconised glass plates held apart by 0.4mm. thick 'Plasticard' spacers along each long side and sealed with vinyl tape. Combs were also cut from 0.4mm. thick 'Plasticard' and had 29 teeth.

For a 50ml. 6% gel, the following components were mixed together: 21g. 'Ultrapure' urea, 5ml. 10x TBE buffer, 7.5ml. 38% acrylamide / 2% N N' bis-methylene acrylamide (deionised by the addition of 5g. of 'Amberlite' MB1 resin and stirring at room temperature for 30 mins., followed by filtration to remove the resin) and distilled water up to a total volume of 50ml. These were shaken together till the urea was dissolved, 300ul. of 10% APS and 50ul. of TEMED mixed in, the gel poured, the comb inserted, the plates clamped together firmly and polymerisation allowed to proceed to completion (c.2 hours at room temperature).

The gel was installed in the running apparatus, the comb removed, the reservoirs filled with TBE buffer and the wells washed out with buffer to remove any unpolymerised acrylamide. Just before loading, the wells were again washed out with buffer to remove any urea that may have

leached into them, as this would have prevented even loading of the samples.

The samples were heated to 90-95°C for at least 3 mins. and 2ul. of each loaded immediately onto the gel. The gel was run at a constant power (40 Watts) which maintained it at a sufficiently high temperature (60-65°C) to keep the DNA samples denatured. After running, the gel was fixed in a 10% acetic acid (v/v) solution for 20-30 mins. at room temperature, carefully transferred to Whatman 3mm. filter paper and dried down under vacuum. Bands were visualised by autoradiography of a sheet of Kodak 'X-Omat' S1 film for 1-4 days at room temperature.

Chapter 3.

The role of transposon ends in the transposition of Tn1/3 derivatives.

3.1 Introduction.

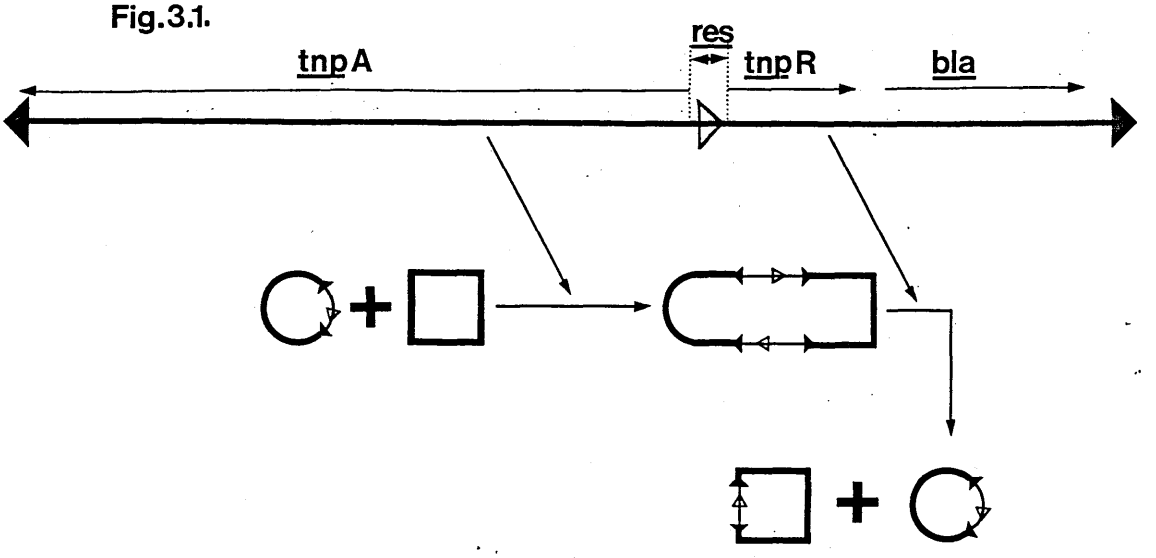
The work of Heffron and others has shown that there are three sites and three genes in Tn3 which are involved in its transposition (Heffron *et al* 1977, 1979; Gill *et al* 1978). Of the sites, two are the ends of the transposon, which include perfect 38 bp. terminal inverted repeats and at least one of which is absolutely required *in cis* for transposition, whilst the third is the *res* site : the site of the site-specific recombination event that occurs during resolvase-mediated cointegrate resolution (Reed 1981). The three genes are the *bla* gene, which encodes the enzyme B-lactamase and provides the transposon with its phenotypic marker, and the *tnpA* and *tnpR* genes, whose products are involved in the transposition process itself. (Fig : 3.1).

Inter-replicon transposition of Tn3 occurs in two sequential steps that are genetically distinguishable (Kitts *et al* 1982a). The *tnpA* gene product (transposase), *in trans* or *in cis*, and the ends of the element, *in cis*, are essential for the first step, in which transposon replication occurs and donor and recipient replicons become fused between directly repeated copies of the transposon to form a cointegrate (Kitts *et al* 1982a). This cointegrate is then resolved in the second step to yield the final transposition products, a process requiring two copies of the *res* site in direct repeat *in cis* and the product of the *tnpR* gene (resolvase) *in cis* or *in trans* (Arthur & Sherratt, 1979; Kitts *et al*, 1982b). The resolvase protein also acts to regulate the cellular levels of both itself and the transposase protein, acting at the level of transcription (Chou *et al*, 1979a).

Intra-replicon transposition of Tn3 is a one-stage process and requires only the ends of the element and the transposase protein (*in cis* or *in trans*) (Bishop, 1983; Bishop & Sherratt, 1984).

Heffron *et al*, (1979), have demonstrated an absolute requirement *in cis* for the ends of Tn3 if transposition is to occur and it was

Fig.3.1.



◀ - perfect 38bp. terminal inverted repeats

▷ - res region

Fig.3.1. Genetic organisation of Tn1/3 and the pathway of intermolecular transposition of these elements.

After Arthur et al, 1984.

The roles of the tnpA and tnpR gene products are indicated.
See text for further details.

therefore of interest to find out more detailed information about the roles of, and requirements for, these sequences. Especially, were the sequences internal to the inverted repeats important in transposition; were the ends interchangeable; was their relative orientation significant; can the "inside" ends of the inverted repeats function as a substrate in transposition?

For some compound transposons, it has been shown that the inside ends of the terminal inverted repeat sequences can act as sites for the initiation of transposition, sometimes with a comparable efficiency to the outside ends (e.g. IS10 in Tn10 : Foster et al, 1981), sometimes with a lower efficiency (e.g. IS1 in Tn9 : Ishizaki and Ohtsubo, 1984; Gamas et al, 1985; IS50 in Tn5 : Sasakawa & Berg, 1982). However, it should be borne in mind that such terminal inverted repeat sequences are structurally intact insertion sequences, and that such sequences are often capable of independent transposition if provided with the necessary functions either in cis or in trans (Kleckner, 1981).

3.2 Construction and assaying of substrate plasmids, and analysis of transposition products.

3.2.1 Construction and assaying of substrate plasmids.

To address these questions, various plasmids were constructed containing different combinations of transposon ends in different orientations. The construction of these is described by Arthur et al, 1984 and their salient features, and those of their parental plasmids, are listed in Table 3.1.

In determining transposition frequencies throughout this work, a system based on the conjugal transfer of small, non-mobilisable plasmids carrying transposons or their derivatives by the Tp^r IncW conjugative plasmid R388 was used. R388 is particularly useful for this purpose since it contains no known transposable elements and its relatively

Table 3.1: Salient features of plasmids possessing different combinations of transposon ends.

Plasmid.	Size(kb.).	Resistance.	Ends.	Transposon: Genotype (all tnpR$^-$).
pACYC184.	4.0.	Cm. Tc.	-	-
pAA32.	8.1.	Cm. Tc.	LH + RH.	Tn1. tnpA$^+$ res$^+$
pAA35.	9.2.	Cm. Tc.	LH + LH.	Tn1. tnpA$^+$ res$^+$
pAA36.	6.4.	Cm.	RH + RH.	Tn1. tnpA$^-$ res$^+$
pAA361.	6.4.	Cm.	RH + RH.*	Tn1. tnpA$^-$ res$^+$
pUC8.	2.67.	Ap.	RH.	Tn3. tnpA$^-$ res$^-$
pUC9.	2.67.	Ap.	RH.	Tn3. tnpA$^-$ res$^-$
pKS400.	5.4.	Ap.	RH + RH.	Tn1/3. tnpA$^-$ res$^-$
pKS401.	5.4.	Ap.	RH + RH.*	Tn1/3. tnpA$^-$ res$^-$

Table 3.1 legend:

(i). * - transposon ends in direct repeat.

(ii). it should be noted that preliminary restriction analysis data now indicate that the plasmid pAA33 (used in the construction of pAA32, pAA35, pAA36 and pAA361) appears to be pACYC184::Tn1 rather than pACYC184::Tn3 as previously reported (Sherratt et al, 1981; Arthur et al, 1984). It therefore seems likely that all the pACYC184-derived replicons listed here contain only Tn1 ends and that only the transposon ends present in pUC8 and pUC9 (and the corresponding ones in pKS400 and pKS401) are Tn3 ends.

small size permits ready analysis of transposition products. Such conjugal transfer can only occur if the small plasmid has formed a transposition cointegrate with R388, in a non-essential part of the R388 molecule. Hence, a measure of the frequency of transposition into non-essential R388 sequences can be obtained by mating a donor strain containing both these plasmids (and a *tnpA*$^+$ complementing plasmid if necessary) with a suitable recipient strain and calculating the fraction of the recipients that receive R388(Tp^r) that also receive the small plasmid marker, and transposon marker if present. This system is clearly only useful in a *tnpR*$^-$ background and, even then, has certain limitations - e.g. transposition events into essential R388 sequences (eg. transfer gene region, Tp^r gene) are not detected; cointegrate molecules will use the origin of replication of their small plasmid component and therefore have a much higher copy number than R388 and tend to out-replicate R388 and displace it by incompatibility. However, it still allows valid comparisons between different molecules to be made. Donor strains were always freshly constructed, care being taken to ensure that each different type of molecule (e.g. transposition target; *tnpA*$^+$ molecule) was introduced at the same stage in the constructions of different strains to allow valid comparisons between them. c.40 colonies were used to grow up a liquid culture of each donor strain to get a representative mean figure for transposition frequency. The presence of cointegrates in exconjugant cells was checked by single colony gel analysis (cointegrates being readily recognisable by their increased size and high copy number) and by their resolution in the presence of a *tnpR*$^+$ plasmid. The transposition assay data presented throughout this work are the mean values obtained from at least two separate experiments. The transposition frequencies, except where otherwise stated, are those obtained from cells that have undergone 20-30 generations of growth. That these figures were maximal was shown by allowing some transposition proficient donor cells to grow for many more generations before mating with the recipient strain. No significant

differences in transposition frequencies in such cells were seen.

All the molecules listed in Table 3.1 were thus assayed for their proficiencies to act as transposition substrates, and the results of these assays are presented in Table 3.2. A photograph of typical cointegrates isolated from these systems is also shown in Fig 3.2. That these molecules shown were true cointegrates was demonstrated by their resolution in the presence of a tnpR$^+$ plasmid (pDS4153) to the original donor replicon plus an R388 derivative containing a copy of the transposing sequence. (Data not shown).

It can be seen from these data that the presence of both at least one inverted repeat sequence *in cis* and of transposase protein *in cis* or *in trans* is absolutely required for any transposition to occur. Also the combination of ends used made no significant difference to the transposition frequency of these Tn1/3 derivatives, as long as they were inverted with respect to one another. Also, the right and left ends appeared to be equally good substrates for transposition. Therefore, as there is no significant homology between the right and left ends of Tn3 other than the 38bp. perfect inverted repeats, it appears that only these inverted repeats are essential for the first step of Tn1/3 transposition.

It is also interesting to note the high degree of specificity conferred by these inverted repeats : they are recognised only by Tn3 (or Tn1) transposase. Not even the transposase protein from the closely related element Tn1000, whose inverted repeats differ at only seven non-contiguous positions from those of Tn3, will recognise and act at them (Heffron, 1983; Kitts *et al*, 1982b).

It can also be seen from the above results that some apparent transposition was detected at 30°C in those systems where only one transposon end (the Tn3 right end) or two right ends in direct repeat (one from Tn1, one from Tn3) were present. These events occurred at very low frequency, even at 30°C, and the inability to detect them at 37°C was probably a consequence of the generally lower transposition

Table 3.2: Transposition assay data obtained using plasmids possessing different combinations of transposon ends.

Plasmid.	Transposition frequency at a temperature of: 30°C. +	30°C. −	37°C. +	37°C. −	Ends.
pUC8.	1.4×10^{-7}.	$<10^{-8}$.	$<10^{-8}$.	$<10^{-8}$.	RH.
pUC9.	3.6×10^{-7}.	$<10^{-8}$.	$<10^{-8}$.	$<10^{-8}$.	RH.
pKS400.	8.32×10^{-1}.	$<10^{-8}$.	4.4×10^{-3}.	$<10^{-8}$.	RH + RH.
pKS401.	1.83×10^{-6}.	$<10^{-8}$.	$<10^{-8}$.	$<10^{-8}$.	RH + RH*.
pACYC184.	NT.	NT.	$<10^{-8}$.	$<10^{-8}$.	—
pAA32.	NT.	-	1.03×10^{-3}.	-	LH + RH.
pAA35.	NT.	-	4.0×10^{-3}.	-	LH + LH.
pAA36.	NT.	NT.	4.7×10^{-3}.	$<10^{-8}$.	RH + RH.
pAA361.	NT.	NT.	$<10^{-8}$.	$<10^{-8}$.	RH + RH*.

Table 3.2 legend:

(i). NT - not tested.

(ii). the '+' and '-' columns indicate assays carried out in the presence or absence of transposase respectively.

(iii). the only molecules used above that were also tnpA$^+$ were pAA32 and pAA35. In all other cases, tnpA$^+$ molecules were used as follows to provide transposase *in trans*:

in the pUC8, pUC9, pKS400 and pKS401 systems - pPAK100.
in the pACYC184, pAA36 and pAA361 systems - pAA15.

(iv). in the pACYC184, pAA32, pAA35, pAA36 and pAA361 systems, transposition frequencies were determined by calculating the ratio of $Str^rTp^rCm^r$: Str^rTp^r exconjugants after mating with DS902. In the pUC8, pUC9, pKS400 and pKS401 systems, transposition frequencies were determined by calculating the ratio of $Str^rTp^rAp^r$: Str^rTp^r exconjugants after mating with DS902.

(v). frequencies given as $<10^{-n}$ indicate that no transposition was detected in the examination of at least $5x10^n$ exconjugants. At least 150 colonies, all containing cointegrates, were examined to calculate positive transposition frequencies.

(vi). The transposition frequencies given throughout this work were calculated as described in the text (p.60) and are thus an indication of the number of transposition events of the transposing sequence under investigation into non-essential parts of the R388 plasmid molecule over a period of c.25 - 30 generations.

(vii). * - transposon ends in direct repeat.

Fig.3.2.

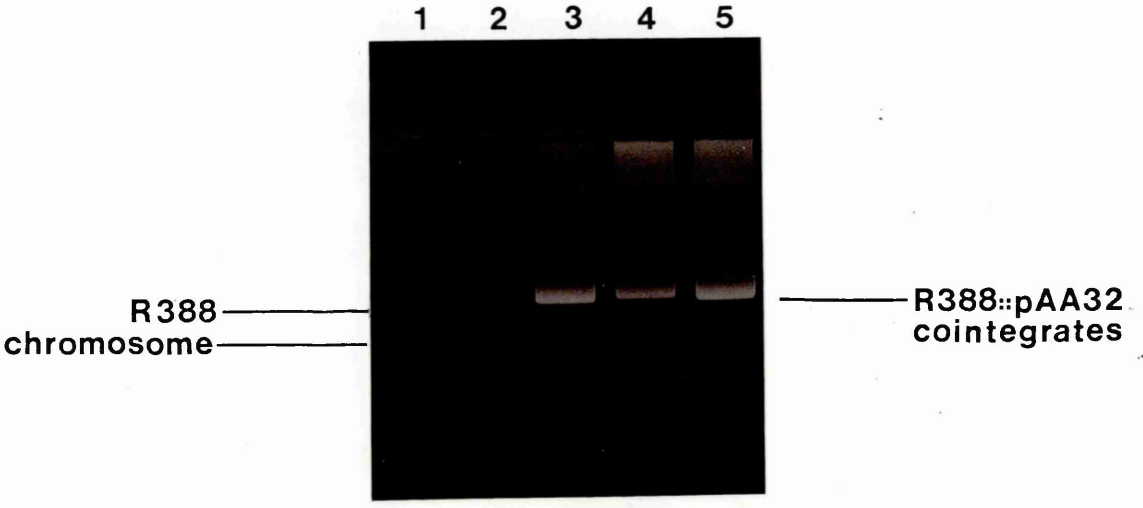


Fig.3.2. Cointegrates formed between pAA32 and R388.

0.8% agarose gel showing the larger size and higher copy number of cointegrates.

1,2: DS916, R388.

3-5: DS902, R388/pAA32 cointegrate.

frequency of Tn1/3 at this temperature (Kretschmer & Cohen, 1979). Similar findings have also been observed in the apparent transposition of pPAK100 (which contains only the left end of Tn1) and pBR322 (which contains only the right end of Tn3: Bolivar et al, 1977) and single colony gel analysis of exconjugants for such systems showed that they contained high copy number plasmids slightly larger than R388. However, subsequent restriction analysis and DNA sequence determination showed that these plasmids were not true transposition cointegrates. Rather, they appeared to be an insertion of the donor replicon (ie. that carrying the transposon end) into the recipient with some small variable duplication of the inverted repeat sequence. The region of inserted donor DNA began exactly at the outside end of the inverted repeat, extended completely through the entire donor replicon and ended with this small and variable duplication (14 - 32 bp.) of inverted repeat sequence. There was also a duplication (3 or 5bp.) of host DNA immediately flanking the insertion site and insertions occured at many sites in the target replicon, just as in the transposition of wild-type Tn1/3 (Arthur et al, 1984; A.Arthur & E.Nimmo, pers. comms.). Thus, this type of event shares three structural features exactly with the transposition of wild-type Tn1/3 : the small duplication of host DNA in the target replicon immediately flanking the transposon insertion, the insertion of transposon DNA beginning precisely at the outside end of the inverted repeat sequence and the distribution of different insertions over many sites in the target replicon. Also, such events are dependent on both transposase protein (in cis or in trans) and the presence of of an inverted repeat sequence in cis (Arthur et al,1984). All these features strongly suggest that these events are a form of transposition event rather than some other form of recombination, and the small duplication of inverted repeat sequences seen suggests too that this is a replicative process, as is the transposition of wild-type Tn1/3 (Heffron, 1983).

This type of event is known as one-ended transposition but, as it is

not one of the areas of investigation of this work, it will not be discussed in great detail here.

In summary, it seems probable that this form of transposition proceeds by a mechanism similar to the postulated asymmetric mechanism of transposition (Grindley & Sherratt, 1979; Galas & Chandler, 1981; Harshey & Bukhari, 1981), as the transposase protein has only one inverted repeat sequence upon which to act. If this is so, then it seems that the inverted repeat sequence, on being replicated and therefore encountered a second time by the replication fork, acts as a terminator of replication, as is proposed in the asymmetric transposition model. This is suggested by the observation that all known duplications of donor DNA in these events involving Tn1/3 derivatives duplicate only a small, and variable portion of the inverted repeat. This variation in the precise extent of the duplication could possibly be due to the replication fork, perhaps with transposase protein complexed to it, encountering the newly-replicated inverted repeat in the inverse orientation compared to the wild-type situation. This could conceivably result in an impairment of specificity of recognition of the inverted repeat by transposase protein.

If this is the mechanism of one-ended transposition, it suggests that transposition of wild-type Tn1/3 (or any derivative consisting of a DNA segment flanked by two Tn1/3 ends in inverted repeat) could also proceed by this mechanism. However, it does not preclude such events occurring by the postulated symmetric mechanism of transposition, either in all, or only a proportion, of cases (Arthur & Sherratt, 1979; Shapiro, 1979). And, indeed, there is good evidence from studies on both inter- and intra-replicon Tn1/3 transposition that wild-type Tn1/3 transposition does proceed by a mechanism closely akin to that of the symmetric model (Kitts, 1982; Kitts et al,1982a; Bishop, 1983; Bishop & Sherratt,1984).

It has been recently suggested (Grindley & Reed,1985) that one-ended transposition may proceed via a symmetric mechanism. However, this requires the invocation of a secondary site on the donor molecule which

* - the group of transposons collectively known as TnA does not include Tn21, Tn1721 and other elements of their sub-group..

the transposase can bind to and use in transpositional recombination, albeit at a reduced frequency. While such an event has been described for IS102 (Machida Y. et al, 1982a), no evidence from either DNA sequence analysis or experimental observations has ever been produced in this laboratory to support the existence of such a sequence in the cases of one-ended transposition studied here.

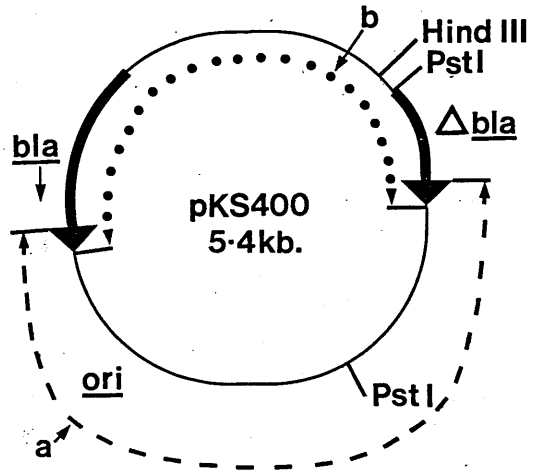
Lastly, extremely similar events and structures have been observed in the transposition of one-ended derivatives of other class II transposable elements - eg. TnA (Heritage & Bennett, 1985. TnA was the general designation given to several independently isolated and closely related transposable elements encoding a TEM-type B-lactamase. These elements included Tn1 and Tn3 (Gill et al, 1978)); Tn21 and Tn1721 (Avila et al, 1984; Motsch & Schmitt, 1984; Motsch et al, 1985) - so it seems that this is not a phenomenon peculiar to Tn1/3.*

3.2.2 Analysis of transposition products from pKS400 system.

Since transposing sequences are duplicated in the transposition of Tn1/3 (Heffron, 1983), the cointegrates generated between pKS400 and R388 allowed an investigation to be carried out to determine just which sequences were being transposed. (See Fig 3.3.). Was it the segment of DNA bounded by the inside ends of the two inverted repeats (ie. that including the origin of replication -(a)in Fig 3.3) or another segment - that bounded by the outside ends of the inverted repeats ((b)in Fig 3.3.)?

pKS400 was constructed by inserting a 2.7kb. PstI fragment from pDS1118 (ColE1::Tn1) into the PstI site of the multiple cloning site of pUC8 in such an orientation that the two transposon ends were inverted with respect to one another. This fragment included the right-most 466bp. of Tn1 (Arthur et al,1984). Therefore, the unique HindIII site in pKS400 is located outside the segment bounded externally by the inside ends of the inverted repeats, but inside that bounded externally

Fig.3.3.



◀ – right inverted repeat of Tn1 or Tn3

ori – origin of replication

Fig.3.3. Plasmid pKS400.

The diagram depicts the plasmid pKS400, indicating the PstI fragment cloned from pDS1118 and the inverted orientation of the two transposon ends.

The regions a and b were two sequences that it was thought might be the transposing sequences in the transposition events seen with this plasmid.

See text for further details.

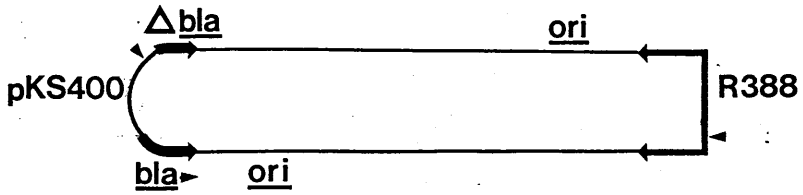
by their outside ends. (See Fig 3.3.), R388 contains one HindIII site and therefore, if the former segment had transposed, HindIII restriction of the cointegrates produced would yield simply two novel fragments. If the latter segment had transposed, however, with a concomitant duplication of the HindIII site of pKS400, HindIII restriction of the cointegrates would yield unit length linear pKS400 fragments as well as two novel fragments. (See Fig 3.4.). DNA from ten cointegrates was restricted with HindIII and, in all cases, unit length linear pKS400 fragments were generated, as well as two novel fragments. (See Fig 3.5.), This result suggests that tnpA mediated transposition normally acts to replicate those DNA sequences corresponding to those in the position of the transposon, and that the inside ends of the inverted repeat sequences are highly disfavoured as sites of action for transposase.

3.3 Effect of the presence of large numbers of transposon ends in trans on transposition frequency.

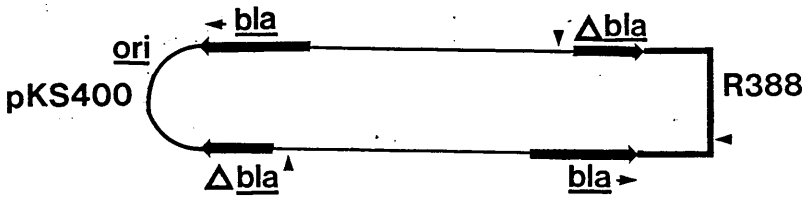
Much evidence has been accumulated that the 38bp. terminal inverted repeats of Tn1/3 are the sites at which the transposase protein binds to and acts upon the DNA. The work described above points to this conclusion, as do the findings of Heffron et al, 1977; M^{c}Cormick et al, 1981 and Wishart et al,1985. Heffron and co-workers showed that deletion of the whole inverted repeat sequence abolishes transposition under all circumstances, and M^{c}Cormick et al, demonstrated that a mutation which removes all but the outermost 14bp. of the right inverted repeat decreases transposition to a level which is 0.1% of the wild-type level. Wishart and co-workers showed that Tn3 transposase binds specifically to linear double stranded DNA fragments containing Tn3 inverted repeat sequences in an ATP-dependent manner. Data obtained using other transposable elements also suggest the ends of such elements are the sites at which transposase proteins bind to and act upon DNA.

Fig.3.4.

i).



ii).



▲ – Hind III site
ori – origin of replication
◀ – right inverted repeat of Tn1 or Tn3

Fig.3.4. The cointegrate molecules that would be generated by the (i).& transposition of the segments a or b respectively of (ii). pKS400 shown in Fig.3.3.

Thus, it can be seen that HindIII restriction of the cointegrates produced if segment a transposed would yield two novel fragments, whilst HindIII restriction of the cointegrates produced if segment b transposed would yield two novel fragments and also unit length linear pKS400 fragments.

Fig.3.5.

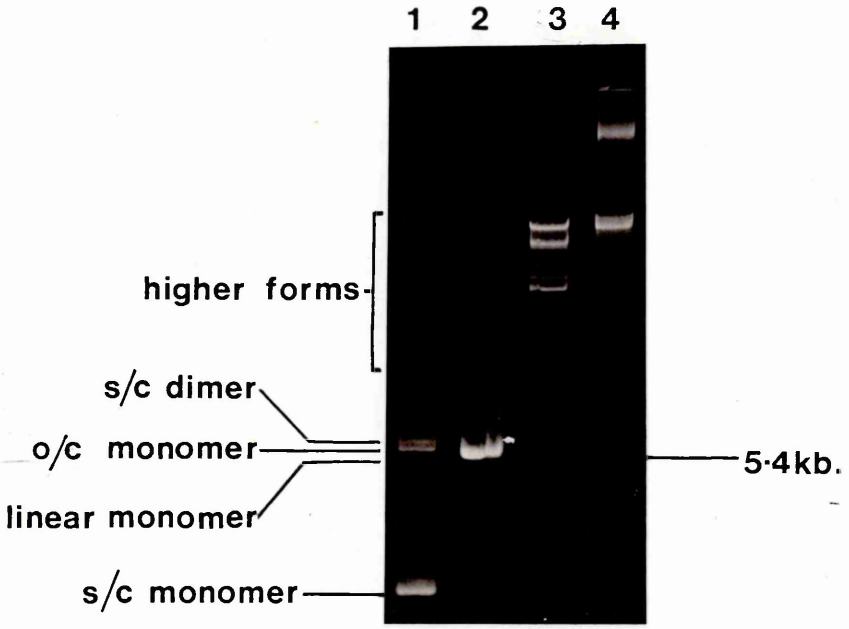


Fig.3.5. HindIII restriction analysis of cointegrates isolated from DS916, R388, pKS400 , pPAK100 cells.

0.8% agarose gel showing that HindIII restriction of these cointegrates yields unit length linear pKS400 fragments (and two other, novel fragments). This indicates that it is that part of pKS400 bounded externally by the outside ends of the inverted repeats (segment b in Fig.3.3) that is replicated in the tnpA-mediated transposition of this plasmid.

1: pKS400.
2: pKS400, HindIII.

3: cointegrate DNA, HindIII. (partial digest).

4: cointegrate DNA.

eg. Craigie et al, 1984 have shown that the Mu A protein (which is thought to be the Mu transposase) recognises and binds specifically to the ends of the Mu genome. However, direct evidence for the binding of Tn1/3 transposase to Tn1/3 inverted repeats has yet to be obtained.

Further indirect evidence for this binding was obtained during this work by assaying the effect of the presence of high copy number replicons containing inverted repeat sequences (but no tnpA or tnpR sequences) on the transposition of a transposition-proficient transposon (Tn103) on a low copy number plasmid in the same cell. The assay system used was the transposition of Tn103 from R388::Tn103 into pACYC184 in DS916. Parallel assays were performed at 30°C and 37°C.

pACYC184 plasmid DNA alone or together with pKS400 or pKS401 plasmid DNA was transformed into DS916, selection for transformants being on L agar containing rifampicin and chloramphenicol or L agar containing rifampicin, chloramphenicol and ampicillin respectively. Each of these strains was then used as a recipient in a mating with DS903, R388::Tn103 and exconjugants were selected as follows:- DS916, pACYC184, R388::Tn103 cells on Isosensitest agar containing rifampicin, chloramphenicol, and trimethoprim; DS916, pACYC184, pKS400 or pKS401, R388::Tn103 cells on Isosensitest agar containing rifampicin, chloramphenicol, ampicillin and trimethroprim. By putting in the transposition proficient molecule last, any effect of pKS400 and pKS401 on transposition frequency was exerted right from the start of the period during which transposition could occur. Transposition of Tn103 into pACYC184 yields an R388::Tn103::pACYC184::Tn103 cointegrate ($Tp^rCm^rTc^r$) and the frequency of this event in each of the above systems was therefore determined by calculating the ratio of $Str^rTp^rCm^r$:Str^rTp^r exconjugants after mating with DS902. The results of these assays are shown in Table 3.3.

It can be seen from these results that the presence of large numbers of inverted repeat sequences on pKS400 or pKS401 in trans causes a decrease of approximately 100 fold in the transposition frequency of

Table 3.3: The effect of the presence of pKS400 or pKS401 on the transposition frequency of Tn103.

Cross.	Transposition frequency of Tn103 into pACYC184 at: 30°C.	37°C.
DS916, pACYC184, R388::Tn103 x DS902.	4.39×10^{-2}.	3.62×10^{-4}.
DS916, pACYC184, R388::Tn103, pKS400 x DS902.	1.14×10^{-4}.	1.77×10^{-6}.
DS916, pACYC184, R388::Tn103, pKS401 x DS902.	1.13×10^{-4}.	3.05×10^{-6}.

Table 3.3 legend:

(i). at least 150 colonies, all containing cointegrates were examined in each case to establish these transposition frequencies.

Tn103 from R388::Tn103 into pACYC184 at both 30°C and 37°C. This effect is seen regardless of the orientation of these sequences with respect to one another and regardless of the fact that pKS401, unlike pKS400, does not contain a transposable segment that could transpose into pACYC184 and block the transposition of Tn103 into the resulting cointegrate by transposition immunity. This is interpreted to mean that these inverted repeat sequences act to titrate out some of the available transposase protein. It is, of course, possible that other sequences in the pKS400 and pKS401 plasmids somehow affected the transposition of Tn103, but an experiment to assay the effect of a plasmid containing no Tn3 sequences (pACYC184) on the transposition of Tn103 from R388::Tn103 to pCB101 showed no detectable effect. (Data presented in Table 5.1). It is not useful to perform the obvious control experiment of assaying the effect of pUC8 and pUC9 (the parental plasmids of pKS400 and pKS401 respectively) on the transposition of Tn103, since they each contain the right inverted repeat of Tn3.

It therefore seems reasonable to conclude that the decreased transposition frequency of Tn103 in the presence of pKS400 or pKS401 is a result of activity of the inverted repeat sequences, most probably by their titrating out transposase protein. Similar observations have been made with TnA by Heritage and Bennett (1984) and they have put forward a similar explanation.

3.4 Discussion.

By assaying the ability of sequences flanked by two copies of either end of Tn1/3 to transpose, and comparing these to the wild-type situation, it has been shown that the transposition frequencies of Tn1/3 derivatives do not appear to differ significantly, no matter what combination of transposon ends they possess, as long as these are inverted with respect to one another. Thus, the right and left ends of Tn1/3 appear to be equally good substrates for transposition. A similar

situation is observed with some insertion sequences (eg. IS10 in Tn10 - Foster et al,1981), while others show a marked variation in the ability of their two ends to act as substrates in transposition (eg. IS50 in Tn5 - Sasakawa and Berg,1982; IS1 in Tn9 - Ishizaki and Ohtsubo,1984; Gamas et al,1985). In the transposition of bacteriophage Mu also, only DNA sequences flanked by one left and one right Mu end have been observed to transpose. Though electron-microscopic analysis has revealed the existence of DNA segments flanked by two left or two right Mu ends, such segments have never been observed to transpose (Toussaint and Faelen,1973; Resibois et al,1981; Toussaint and Resibois,1983). It is also interesting to note that only sequences flanked by one left and one right Mu end in one particular orientation have been observed to transpose. Transposition of sequences flanked by the same two ends in the opposite orientation has not been seen. This is the case both in vivo and in vitro (Toussaint and Resibois, 1983; Mizuuchi,1983).

As there is no significant homology between the two ends of Tn1/3 other than the 38bp. perfect inverted repeats, these results suggest that only these inverted repeat sequences are essential for the first step of Tn1/3 transposition. Also, it would seem that there is no specialised role for the sequences immediately internal to the inverted repeats in transposition.

Additionally, it has been shown that tnpA mediated transposition normally acts to replicate those DNA sequences bounded externally by the outside ends of the inverted repeats and that the inside ends of the inverted repeats are highly disfavoured as sites for the initiation of transposition.

These data also confirm the findings of other workers (Heffron et al, 1977, 1979; Gill et al, 1978; Kitts et al, 1982a.) that the first step of Tn1/3 transposition absolutely requires the presence in cis of the 38bp. terminal inverted repeat sequence, as well as the transposase protein, in cis or in trans, of Tn1/3. Even one copy alone of this sequence is sufficient to mediate a novel type of recombination event

between a replicon carrying the sequence and a suitable target. This appears to be a form of transposition ('one-ended' transposition) in which the whole donor replicon is inserted into the target with a small and variable duplication of the inverted repeat sequence. This type of event occurs at a far lower frequency - c.1000 times less - than the transposition of sequences flanked by two transposon ends in inverted repeat.

The findings on the apparent equivalence of the two ends of Tn1/3 are at variance with the data of Heritage et al (1984 - hitherto only reported as reference 65 in Grindley & Reed, 1985) obtained using derivatives of the transposon Tn802. Tn802 appears to be equivalent to Tn1 (Wallace et al, 1980) and in plasmids containing different combinations of three inverted repeat sequences of Tn802, it was found that there was a hierarchy of usage of these sequences which was almost independent of their order on the DNA. In summary, sequences flanked by one left end and one right end in inverted repeat transposed c.7-fold less frequently than those flanked by two left ends in inverted repeat but more than 100-fold more frequently than those flanked by two right ends in inverted repeat. These differential frequences were seen even in cases where the third end on the plasmid lay between the two ends that were active in transposition and defined a shorter transposable segment. These results suggest that the Tn802 transposase interacts preferentially with the left end of the transposon, implying a specific role in transposition for those sequences immediately within the left inverted repeat.

The experiments performed in this laboratory closely resemble the wild-type situation - i.e. each of the plasmids used contained a maximum of two transposon ends - and it could be that the data of Heritage et al are a consequence of some peculiarity of plasmids containing three transposon ends. Also, they could be a specific feature of Tn802 transposition.

Experiments analogous to theirs with plasmids constructed in this

laboratory have yet to be performed, but the results of these will clearly throw more light on the preferential or non- preferential use of the two ends of Tn1/3 in transposition.

Chapter 4.

Attempted over-expression, and regulation of the tnpA gene of Tn1.

4.1. Introduction.

A valuable aid in the elucidation of the precise mechanism of Tn1/3 transposition would be the development of an *in vitro* Tn1/3 transposition system. Such systems have recently been developed for the transposing bacteriophage Mu and have led to a much clearer understanding of the mechanism of Mu transposition. The available data support the theory that Mu transposition proceeds by a mechanism closely akin to the symmetric model of transposition proposed by Shapiro (Shapiro, 1979; Mizuuchi, 1983,1984; Craigie and Mizuuchi, 1985; Higgins *et al*, 1983; Higgins and Olivera, 1984).

For such a system, it is clearly necessary to have a source of purified Tn1/3 transposase protein. However, initial studies indicated that the cellular levels of this protein were rather low. For example, several workers studying expression of Tn1/3 encoded polypeptides in *E.coli* K-12 minicells did report visualising, either by autoradiography of ^{35}S-methionine labelled material or Coomassie brilliant blue staining of SDS-polyacrylamide gels, a band corresponding to a polypeptide of a size appropriate for transposase as determined by analysis of the nucleotide sequence of Tn3 (c.111 Kd). However, this band was only seen in a derepressed background (*tnp*R$^-$) and, even then it was far less intense than the bands of the other proteins from the transposon : resolvase and B-lactamase. (Chou *et al*, 1979a and b; Gill *et al*, 1979; Wishart *et al*, 1985). Dougan *et al* (1979) reported being unable to detect any expression from the transposase gene region of Tn1/3 in *E.coli* K-12 minicells, but re-examination of their gel data revealed a faint band of a polypeptide of the anticipated size for transposase, but again only in a *tnp*R$^-$ background. Experiments undertaken during this work to study the expression of polypeptides from the plasmid pMB9::Tn103 in *E.coli* K-12 minicells failed to show any band of a size appropriate for transposase.

The final cellular level of a protein is determined by several

factors, including the level of transcription of the gene encoding the protein, the stability of the mRNA, the efficiency of the translation of the mRNA and the stability of the protein itself. Experiments to study the strength of the transposase gene promoter and translation initiation signals in Tn3 have shown that, while the promoter is of a reasonable strength:- c.55% the strength of the derepressed wild-type lac promoter of E.coli, the translation initiation signals are rather weak:-only 3% the strength of the corresponding lac signals (see below and Chou et al, 1979b.) Thus, while the gene may be transcribed at quite a high level, the efficiency of translation of the mRNA seems likely to be low, and this could obviously contribute to the low cellular level seen for this protein.

In the light of these observations on the cellular level of the Tn1/3 transposase protein and the data on the gene's promoter and translation initiation signals, it seemed wise that any method for isolating usable quantities of this protein should take as its starting point a system for its over-production e.g. an appropriately modified Tn1/3 derivative.

Shortly before the start of this work, Fennewald et al (1981) reported a purification method for Tn3 transposase using a mutant Tn3 that was both a tnpR$^-$ (due to the generation of a UAG - amber - termination codon at the position of the 130th. amino acid residue of resolvase) and had a mutation in the Shine/Dalgarno sequence (Shine & Dalgarno, 1975) of the tnpA gene that resulted in an increased efficiency of translation of tnpA mRNA (Chou et al, 1979a; Casadaban et al, 1980). Very recently, and near the end of this work, Wishart et al (1985) reported a simpler purification method for this protein using a Tn3 derivative that was mutant only in being tnpR$^-$. This mutation involved deletion of the 12bp. between 3243 and 3254 in the Tn3 sequence. (Heffron et al, 1979).

It therefore seemed quite feasible to be able to develop a similar system for the over-expression of the Tn1 transposase protein, as a first step in developing a protocol for its purification.

It was decided to use Tn1 in this work because restriction analysis data had shown the presence of an additional HincII site in Tn1 compared to Tn3. This extra site was found to be in the vicinity of the initiation codon of the transposase gene (P. Kitts and D. Sherratt, pers. comm.). It was thought therefore that cutting the DNA at this site would separate all or most of the tnpA gene coding region from either just the promoter or both the promoter and translation initiation signals of the gene. This would allow the coding region to be placed under the control of other known strong transcriptional and translational control signals.

Neither the exact site nor the exact nature of the mutation that gave rise to this extra HincII site in Tn1 was known. Examination of the nucleotide sequence of Tn3 (Heffron et al, 1979) revealed four 6bp. sequences around the tnpA gene initiation codon that could each, by a single base pair substitution mutation, give rise to a HincII site. See Fig 4.1. However, it was quite possible that the extra site could arise by some other mutational event (e.g. single base pair insertion; mutations affecting more than one base pair). To determine the nature of the mutation that generated this additional HincII site and, especially, to precisely localise this site, it was decided to carry out DNA sequence determination on this region of Tn1.

4.2. DNA sequence determination of part of Tn1.

Since there is a HincII site in Tn1 (as in Tn3) at position 1385, isolation of the intact tnpA gene on a HincII fragment from Tn1 necessitates a partial restriction digest. It was therefore decided to firstly clone such a HincII fragment into the multiple cloning site of a suitable high copy number vector, as this would provide a plentiful source of the fragment for other manipulations. Additionally, the tnpA gene sequence would then be flanked by sites for various other

Fig.4.1.

Hinc II site: -GTPy↓PuAC-

i). CTCAAC ⟶ GTCAAC – transversion
3013 3018

ii). GTCAAA ⟶ GTCAAC – transversion
3029 3034

iii). ATCAAC ⟶ GTCAAC – transition
3037 3042

iv). TTTAAC ⟶ GTTAAC – transversion
3074 3079

Fig.4.1. Four 6bp. sequences near the translation initiation codon of Tn3 that, with a one bp. substitution mutation, could give rise to a HincII site.

The figures indicate the co-ordinates of these sequences in Tn3.

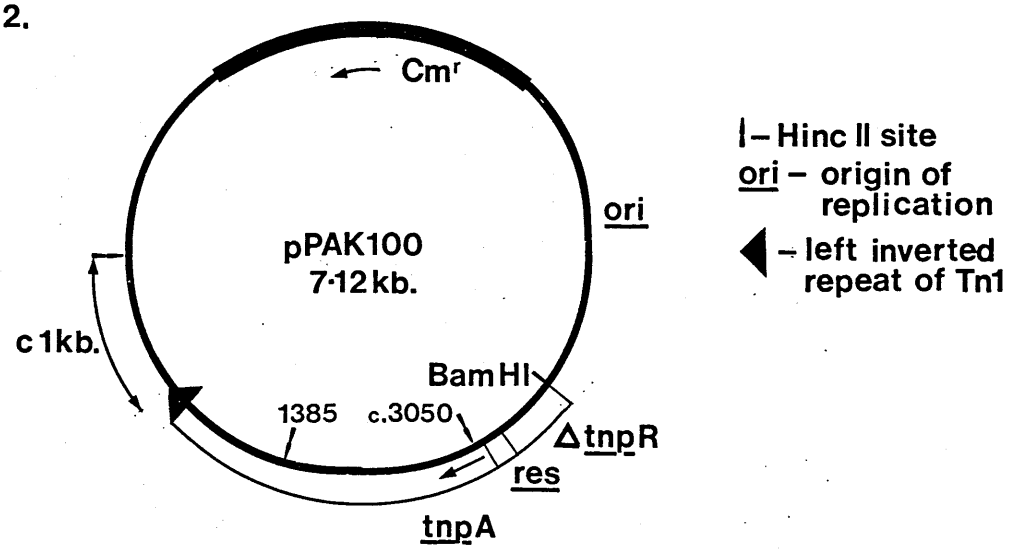
restriction enzymes that did not have sites within the tnpA gene itself, thus facilitating these manipulations.

It was appreciated that the sequence determination of a construct obtained in this way would not identify the mutation responsible for the presence of the additional HincII site in Tn1 if this lay 3' to the cleavage site. However the exact nature of this mutation was not of prime importance - most significant was the location of the site. Also, it was considered that it may be difficult to exactly localise the region of Tn1 being analysed. However, as it was known that the extra site lay in the region of the tnpA gene initiation codon and that Tn1 and Tn3 differ by only a few base changes (Heffron, 1983), it was presumed that the position of the region of Tn1 under examination could be determined by comparing it with the Tn3 nucleotide sequence and finding the area that gave greatest homology between the two. Therefore, this approach was pursued, the chosen vector being pUC8 and the source of the Tn1 tnpA gene-containing fragment being pPAK100.

pUC8 has only one HincII site (in its multiple cloning site) but pPAK100 has three : two in the Tn1 part of the molecule and one in the pACYC184 part. See Fig 4.2.

Three 2ug aliquots of pPAK100 plasmid DNA were partially digested with HincII as described in Section 2.17 and the products of digestion separated on a 1% low melting point agarose gel. The gel slices containing the fragment of interest were cut out and pooled. The DNA was purified from the agarose and used in ligation with pUC8 plasmid DNA that had been linearised with HincII and phosphatased. This ligation mix was used to transform $\triangle$M15 cells to Ap^r, and the selective plates also contained X-gal. Of 456 transformant colonies obtained, 34 had a blue phenotype and were found to contain monomeric pUC8 plasmid. The remaining 422 colonies all had a white phenotype, suggesting they contained recombinant plasmids. Single colony gel analysis, however, revealed that all but three of them contained plasmids of approximately the same size as pUC8. Most probably these plasmids were small deletion

Fig.4.2.



Hinc II restriction fragment sizes (kb.):

i). partial:	ii). complete:
4·050	1·665
4·736	2·385
5·456	3·071

Fig.4.2. Plasmid pPAK100.

The diagram shows the HincII restriction sites in the plasmid. The small figures indicate the co-ordinates of the HincII sites in the part of Tn1 that is retained in pPAK100.

Also shown is a list of all the possible restriction fragments that can be generated by HincII restriction of the plasmid, complete or partial. It is the c.4.05kb. partial fragment that contains the tnpA gene. Notice that this fragment also includes c.1kb. of pACYC184 sequences beyond the left inverted repeat of Tn1.

derivatives of pUC8 in which the reading frame of the lacZ gene fragment downstream of the HincII site had been changed, preventing intragenic-complementation between the plasmid and chromosomal lacZ gene fragments occurring. The remaining three white colonies, however, each contained a plasmid of the expected size of the desired recombinant. See Fig 4.3. Restriction analysis of plasmid DNA prepared by the alkaline/SDS method from each of these clones using HincII complete digestion, ClaI digestion and ClaI/BamHI co-digestion showed that all these plasmids contained the fragment from pPAK100, one of them having it inserted in one orientation and the remaining two having it inserted in the other. These plasmids were designated pSNO15 and pSNO16, depending on the orientation of the inserted fragment. See Figs 4.4 and 4.5.

It can be seen that pSNO15 contains the pPAK100 fragment orientated in such a way that transcription initiated by the lac promoter and translation of the resultant mRNA would be expected to generate a functional transposase protein. This would be the case as long as the protein product were wild-type transposase (though the lack of a translational stop codon in the B-galactosidase sequence makes this unlikely) or a B-galactosidase/transposase in-frame fusion protein that had not lost some crucial part of the transposase protein and that was not, for some reason, inactive. pSNO16, on the other hand, contains the pPAK100 fragment inserted in the opposite orientation and hence would not have been expected to produce functional transposase protein by the same mechanism. Transposition assays were carried out to assess the ability of each of these plasmids to supply transposase protein *in trans* and the results of these are given later in this chapter.

The immediate use of these plasmids, however, was to determine the exact position of the extra HincII site in Tn1. The method used for DNA sequencing in this work was the chain-terminator method of Sanger *et al* (1977). It was therefore necessary to clone a suitable fragment into an M13 bacteriophage vector before its sequence could be determined. The vector chosen was M13mp18.

Fig.4.3.

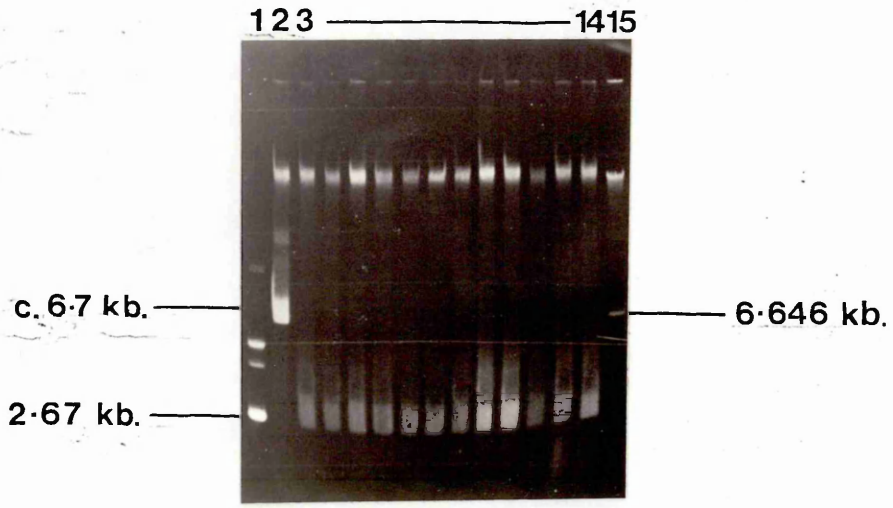


Fig.4.3. Single colony gel analysis of possible recombinant clones obtained in the construction of pSNO15 and pSNO16.

0.8% agarose gel showing plasmid DNAs present in thirteen white transformant colonies. Of these, twelve (3-14 incl.) are very probably small deletion derivatives of pUC8, but one (2) is of the anticipated size of the desired recombinant.

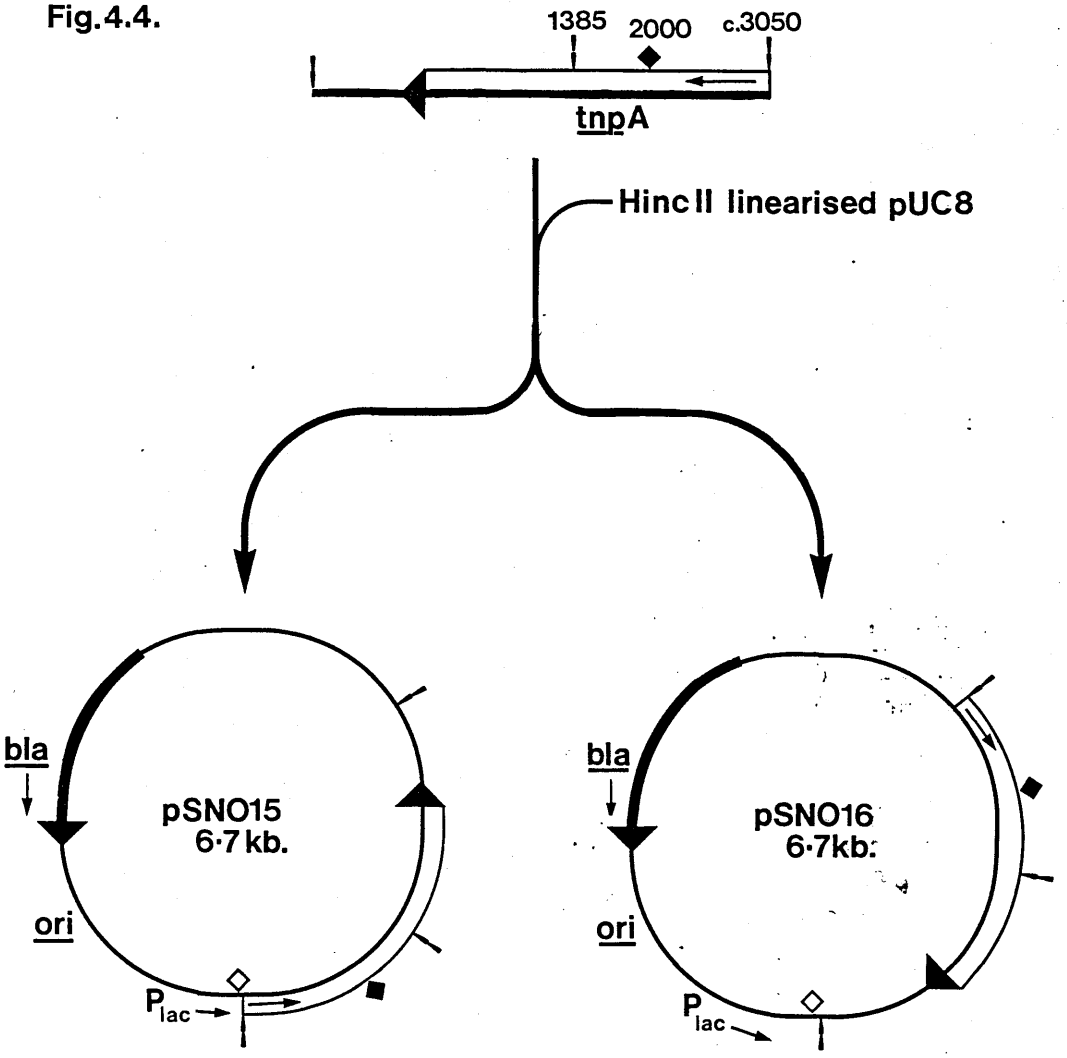
Note the very high copy number of the plasmid in (2).

1: pUC8 plasmid DNA.

2-14: plasmids in white transformant colonies.

15: ColE1 plasmid DNA.

Fig.4.4.



◀ – inverted repeat of Tn1 or Tn3
ori – origin of replication
P_{lac} – lac promoter
| – HincII site
◆ – ClaI site
◇ – BamHI site

Fig.4.4. The construction of pSNO15 and pSNO16.

Ligation of the c.4.05kb. HincII fragment from pPAK100 to HincII linearised pUC8 generated two products, differing in the orientation of the inserted fragment. These were designated pSNO15 and pSNO16.

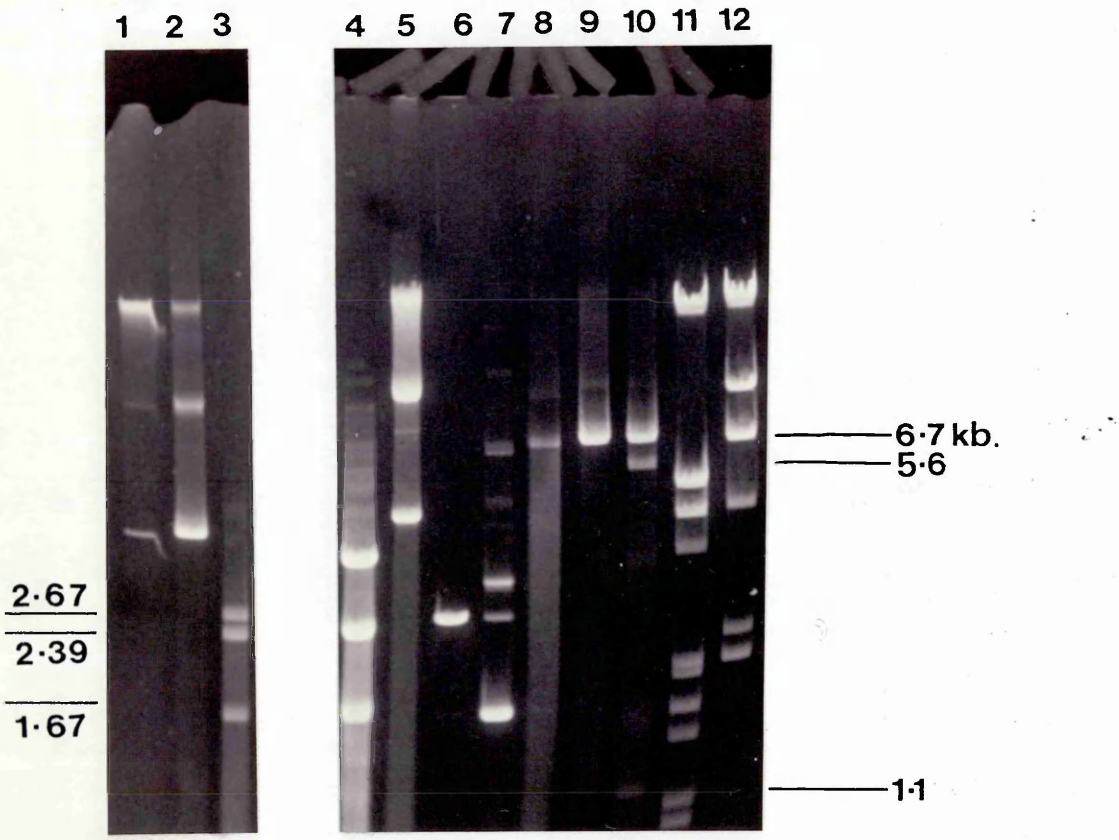


Fig.4.5.a). Restriction analysis of pSNO15 plasmid DNA.

0.8% agarose gel showing that HincII digestion of pSNO15 yielded the 2.39kb. and 1.67kb. HincII fragments derived from pPAK100 (and a unit length linear pUC8 fragment). Also, ClaI/BamHI co-digestion yielded c.1.04 and 5.6kb. fragments, showing that the fragment from pPAK100 was inserted in such a way that transcription from the lac promoter proceeded through it in the correct direction to generate a functional transposase protein.

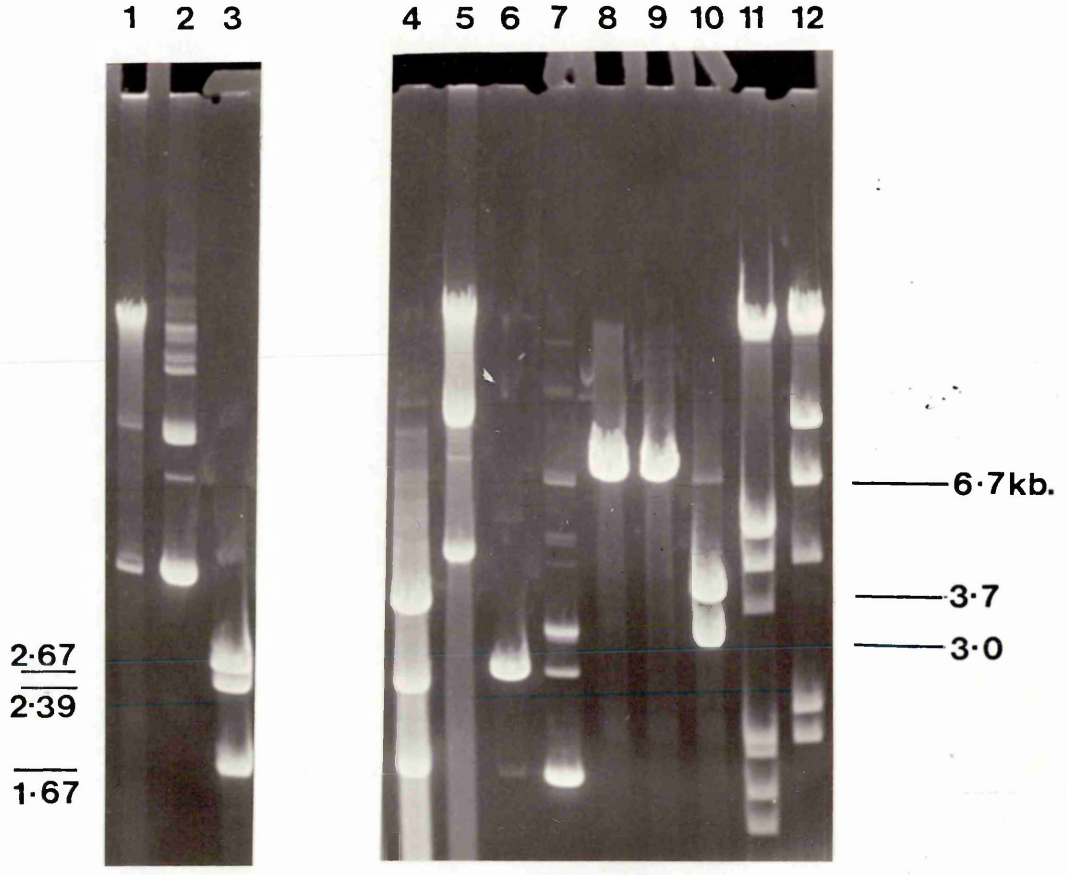


Fig.4.5.b). Restriction analysis of pSNO16 plasmid DNA.

0.8% agarose gel showing that HincII digestion of pSNO16 yielded the 2.39kb. and 1.67kb. HincII fragments derived from pPAK100 (and a unit length linear pUC8 fragment). Also, ClaI/BamHI co-digestion yielded c.3.0 and 3.7kb. fragments, showing that the fragment from pPAK100 was inserted in such a way that transcription from the lac promoter proceeded through it in the incorrect direction to generate a functional transposase protein.

The lane allocations for Figs.4.5.a). & b). are as follows:-

1: ColE1 DNA.

2: pSNO15 or pSNO16 DNA.

3: pSNO15 or pSNO16, HincII.

4: pPAK100, HincII.

5: pPAK100 DNA.

6: pUC8, HincII. (partial digest).

7: pUC8 DNA, with many higher forms.

8: pSNO15 or pSNO16, BamHI.

9: pSNO15 or pSNO16, ClaI.

10: pSNO15 or pSNO16, ClaI/BamHI.

11: Lambda DNA, HindIII/EcoRI.

12: Lambda DNA, HindIII.

The sizes (kb.) of the lambda DNA fragments are as follows:-

i). lambda DNA, HindIII,EcoRI:

21.2, 5.15, 4.97, 4.28, 3.53, 2.03, 1.91, 1.58, 1.33, 0.98, 0.83.

ii). lambda DNA, HindIII:

23.1, 9.42, 6.68, 4.36, 2.32, 2.03.

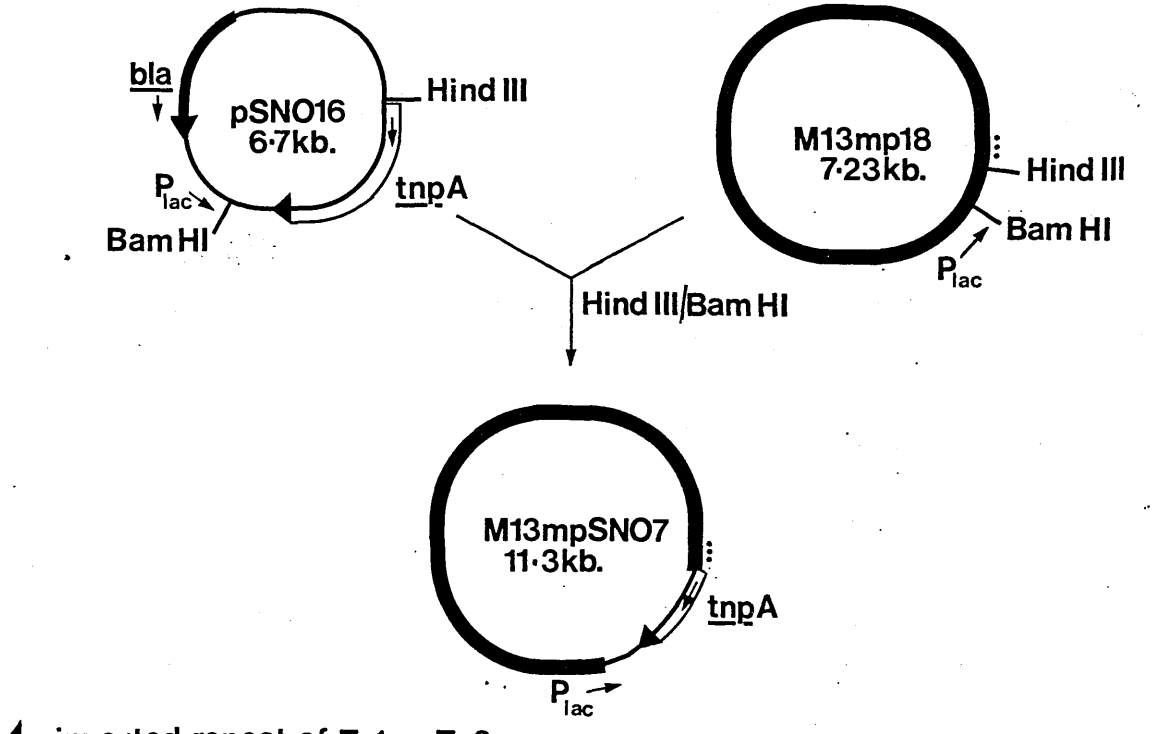
pSNO16 plasmid DNA was chosen as the source of the fragment to be sequenced. This was because ligation of the HindIII-BamHI fragment containing the Tn1 tnpA sequences from this plasmid to the largest HindIII-BamHI fragment of M13mp18 replicative form (RF) DNA would mean that the region immediately downstream of the primer annealing site in the resultant recombinant would be that part of the Tn1 sequence adjacent to the additional HincII site that is retained in pSNO16 (and pSNO15). See Fig 4.6.

pSNO16 plasmid DNA and M13mp18 RF DNA were each co-digested with HindIII and BamHI and the resultant fragments mixed and ligated together. Five recombinant 'phages were isolated by M13 transformation of JM101 cells, screening the transformants using X-gal and IPTG. Single plaque gel and restriction analysis of RF DNA from these five recombinants (data not shown) showed that they all were of the desired structure. This recombinant was designated M13mpSNO7.

Template DNA from each of these five independently isolated clones was prepared and used for sequencing reactions as described in Section 2.23. Sequencing was carried out three times. On each occasion, the gels could be read clearly and unequivocally and the sequence determined from each of the five clones was identical.

Comparing the sequence data obtained from Tn1 with the known sequence of Tn3 (Heffron et al, 1979) using the computer programme ZSEQFIT (C. Boyd, pers. comm.) revealed an area in the Tn3 sequence where there was 97/114 bp.(=85.1%) homology between the two sequences. This region extended from co-ordinates 2926-3040 in the Tn3 sequence and the Tn1 sequence included, at a position equivalent to 3037 in the Tn3 sequence, a T to C transition mutation. This is one mutation that would generate a HincII site in Tn1 that extended over those base pairs equivalent to base pairs 3037-3042 in Tn3 as shown in Fig 4.1.(iii). At such a site, cleavage by the enzyme would occur between base pairs 3039 and 3040, and the sequence data showed this to be the site of cleavage. It therefore was concluded that it was this mutation that generated the extra HincII

Fig.4.6.



◀ – inverted repeat of Tn1 or Tn3
P_{lac} – lac promoter
⋮ – primer binding site

Fig.4.6. The construction of M13mpSNO7.

Ligation of the c.4.1kb. HindIII/BamHI fragment from pSNO16 to the largest HindIII/BamHI fragment of M13mp18 generated M13mpSNO7. In this M13 derivative, the tnpA gene sequences of Tn1 were immediately downstream of the binding site for the M13 universal primer.

* See also the photograph of the sequencing gels from which the DNA sequence was determined inside the back cover.

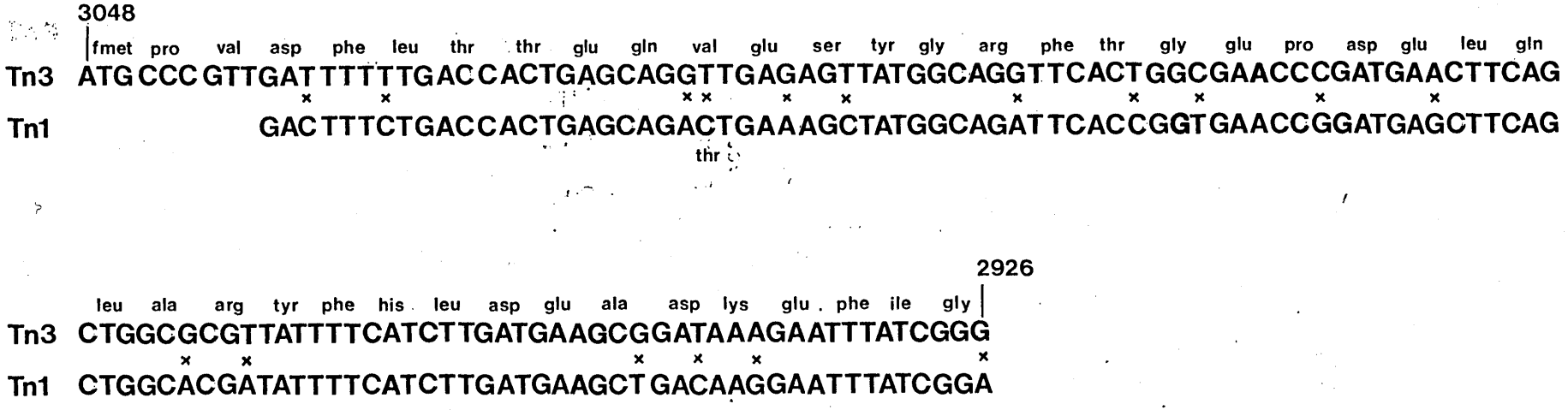
site in Tn1 and that this site extended over the above region in Tn1. The region of Tn1 sequence that was determined is shown and compared with the corresponding region of Tn3 in Fig 4.7. Also shown are the predicted translation products of the two regions.*

It can be seen that those nucleotide differences that exist between Tn1 and Tn3 in this area include eleven transition and three transversion mutations. These are arranged to include eleven third base pair changes, all of which are synonymous ('same-sense') mutations, one separate first base pair change, which is also a synonymous mutation and two adjacent mutations that change the first and second base pairs of one codon and cause a mis-sense mutation at this point. (A valine residue in Tn3 is replaced by a threonine residue in Tn1). Thus, over the region sequenced, only one amino acid residue in thirty-eight is altered between Tn1 and Tn3 (=97.4% homology), and hence the predictions of close homology between Tn1 and Tn3 at the DNA, and therefore the amino acid, level are borne out in this area.

4.3. Galactokinase assays to determine the relative strengths of the transposase, resolvase and lac promoters.

A previous worker in this laboratory had constructed derivatives of the promoter probe vector pKO-1 (McKenney *et al*,1981) where the galactokinase gene of the plasmid was under the control of the *tnpA* or *tnpR* gene promoter. These plasmids were designated pJKA and pJKR respectively (Kelly,1983). Galactokinase assays of these plasmids were undertaken in comparison with pKL500 where the galactokinase gene is under the control of the *lac* promoter (A. Lamond, pers. comm.). The results of these assays are shown in Table 4.1.

Fig.4.7.



× – nucleotide differences

Fig.4.7. Comparison of the nucleotide sequences of Tn1 and Tn3 between the co-ordinates 3040 & 2926 of Tn3.

Also shown is the amino acid translation product of the Tn3 region. The corresponding Tn1 amino acid sequence is identical, except for one change as indicated.

Nucleotide sequence homology between Tn1 and Tn3 over this region = $^{97}/_{114}$. (85.1%).

Amino acid sequence homology between Tn1 and Tn3 over this region = $^{37}/_{38}$. (97.4%).

Table 4.1: Galactokinase assay data on plasmids pKL500, pJKA and pJKR.

Plasmid.	Description.	Units of galactokinase activity.
pKL500.	lac promoter driving galK gene.	247.2.
pJKA.	tnpA gene promoter driving galK gene.	140.3.
pJKR.	tnpR gene promoter driving galK gene.	54.4.

It can be seen that both the tnpA and tnpR gene promoters are weaker than the lac promoters as follows:

tnpA promoter : lac promoter= 0.567 : 1

tnpR promoter : lac promoter= 0.22 : 1

Also, it can be seen that the tnpA gene promoter is c.2.58x stronger than the tnpR gene promoter.

4.4. Transposition assays using pSNO15 and pSNO16.

With this sequence information the precise structure of the predicted protein product from pSNO15 can be deduced. This is shown in Fig 4.8.

The region encoding the protein begins at the translation initiation codon of B-galactosidase, includes the first eleven amino acid residues of this protein and then is fused, in frame, to that part of the tnpA coding sequence downstream of the HincII cleavage site at 3039/3040. Therefore, the protein produced from pSNO15 would be expected to consist of the entire Tn1 transposase protein with its three N-terminal amino acid residues removed and the first eleven N-terminal amino acid residues of B-galactosidase inserted in their place. In fact, the eleventh amino acid residue of B-galactosidase is the same as the third residue of transposase (both valine residues) and hence it is legitimate to say that this fusion protein effectively consists of the entire transposase protein with its two N-terminal amino-acid residues removed and replaced by the first ten N-terminal residues of B-galactosidase. This proposed chimaeric protein would have a molecular weight of c.111kd. and both its transcriptional and translational initiation would be expected to be directed by lac control signals. This would be advantageous since it has been shown that both the transposase promoter and translation initiation signals are weaker than the corresponding lac signals. (4.3 above and Chou et al, 1979b.). It would therefore be

Fig.4.8.

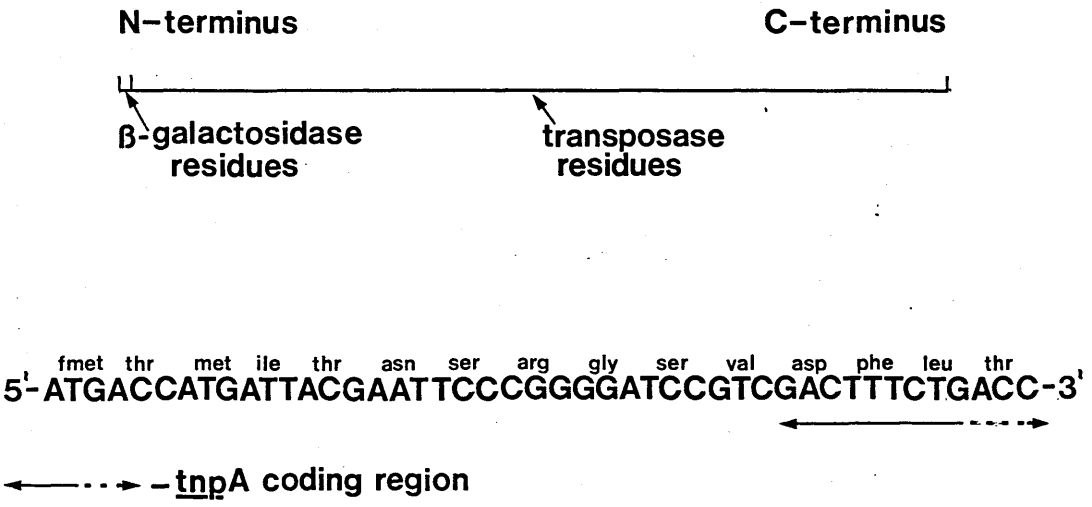


Fig.4.8. The beta-galactosidase/transposase fusion protein encoded by pSNO15.

This consists of the entire Tn1 transposase protein with its three N-terminal amino acid residues removed and replaced by the first eleven N-terminal amino acid residues of beta-galactosidase. Its predicted molecular weight is c.111kd.

The nucleotide sequence of the N-terminus is also shown, and it can be seen from this that translation initiated at the beta-galactosidase translation initiation codon enters the tnpA coding region in the correct reading frame for transposase.

anticipated that this protein would be expressed at a higher level than native transposase protein produced from its own transcriptional and translational control signals in a derepressed background. And hence, if it retains full transposase activity, it would be anticipated that one copy of a pSNO15 plasmid molecule would promote transposition at a higher level than one copy of a molecule containing a derepressed wild-type tnpA gene. Additionally, as the copy number of pSNO15 is high (c.87 per genome equivalent: C. Jones, pers. comm. See also Fig 4.3 above), and higher than the copy numbers of previously assayed plasmids carrying Tn1/3 or a derivative (e.g. pMB9::Tn103 - c.20-25 per genome equivalent, Twigg and Sherratt, 1980; C. Jones, pers. comm.) the level of transposition in a cell containing pSNO15 would be expected to be very high indeed. It was interesting to notice, however, that cells containing pSNO15 grew quite normally, since Heffron (1983) has stated that overproduction of transposase had a deleterious effect on the cell.

To test this hypothesis, transposition assays were undertaken to determine the ability of both pSNO15 and pSNO16 to promote the transposition of Tn3651 (on pPAK200) into R388. Parallel assays were performed at 30°C and 37°C, and a system assaying the same transposition event mediated by pMB9::Tn103 was used as a positive control. To ensure the events detected were transposon sequence specific, negative control systems using a molecule devoid of transposon sequences (pACYC184) instead of pPAK200 were also set up.

DS916 cells containing R388 were transformed to Ap^r with either pSNO15 or pSNO16, transformants being selected on Isosensitest agar containing rifampicin, trimethoprim and ampicillin. It was important to transform these DNAs in before pPAK200, since this plasmid carries determinants for both ampicillin and chloramphenicol resistance. Hence, if it were transformed in first, there would be no simple selection available for subsequent introduction of pSNO15 or pSNO16.

Cells of each of these strains were then transformed to Cm^r with pPAK200 or pACYC184, transformants being selected on Isosensitest agar

containing rifampicin, trimethoprim, ampicillin and chloramphenicol. Whilst DS916 cells containing only R388 and pPAK200 could grow on this medium, it was presumed that the high copy numbers of both pSNO15 and pSNO16 would ensure their stable maintenance, and single colony gel analysis did indeed confirm their presence in the transformant strains (data not shown).

To construct the pMB9::Tn103 containing strains, DS916 cells containing R388 were first transformed to Tc^r with pMB9::Tn103 and then this strain transformed to Cm^r with pPAK200 or pACYC184.

Transposition of Tn3651 into R388 yields a pACYC184::Tn3651:: R388::Tn3651 cointegrate ($Tp^rCm^rAp^r$) and therefore the frequency of this transposition event in all the above systems could be determined by calculating the ratio of $Str^rTp^rCm^r$:Str^rTp^r exconjugants after mating with DS902.

The results of these assays are shown in Table 4.2.

It can be seen from these data that pSNO15 did indeed promote the transposition of Tn3651 at high frequency : 1.02×10^{-1} at 30°C and 4.43×10^{-4} at 37°C. In the pSNO16-containing strains, however, no transposition of Tn3651 was detectable at frequencies greater than 10^{-7} (30°C) and 10^{-8} (37°C). Neither pSNO15 nor pSNO16 promoted transposition of pACYC184 at frequencies greater than 10^{-7} (30°C) and 10^{-8} (37°C).

In the pMB9::Tn103 strains, some apparent transposition of pACYC184 was detected, though at much lower frequencies: 3.92×10^{-4} (30°C) and 2.6×10^{-7} (37°C). Such data have been obtained before and have been shown to be the result of double transposition events, causing the formation of a pMB9::Tn103::pACYC184::Tn103:: R388::Tn103 cointegrate (A. Arthur, pers. comm.) These were not investigated further. Analogous events were not detected in the pSNO15 strains, presumably because pSNO15 contains no sequences flanked by two transposon ends in inverted repeat, nor in the pSNO16 strains, presumably because pSNO16 does not produce a functional transposase protein.

Table 4.2: Transposition assay data on plasmids pMB9::Tn103, pSNO15 and pSNO16.

Complementing plasmid.	Substrate plasmid.	Transposition frequency at: 30°C.	37°C.
pMB9::Tn103	pACYC184	3.92×10^{-4}	2.60×10^{-7}
pMB9::Tn103	pPAK200	1.92×10^{-1}	8.73×10^{-4}
pSNO15	pACYC184	$<10^{-7}$	$<10^{-8}$
pSNO15	pPAK200	1.02×10^{-1}	4.43×10^{-4}
pSNO16	pACYC184	$<10^{-7}$	$<10^{-8}$
pSNO16	pPAK200	$<10^{-7}$	$<10^{-8}$

Table 4.2 legend:

(i). frequencies given as $<10^{-n}$ indicate that no transposition was detected in the examination of at least $5x10^{n}$ exconjugants. At least 100 colonies, all containing cointegrates, were examined to calculate positive transposition frequencies.

(ii). notice that pMB9::Tn103, pSNO15 and pSNO16 each contain two transposon ends. Hence, any effect due to the presence of transposon ends on these molecules will be the same for each.

(iii). assays carried out in the presence or absence of IPTG showed no significant differences in transposition frequencies. This is presumably because the copy number of pSNO15 is so high that all the lac repressor produced from the wt. chromosomal lacI gene of DS916 is titrated out by a small proportion of the plasmid molecules, leaving the majority of the plasmids de-repressed.

(iv). Isosensitest agar contains 0.2% glucose, but assays were not also carried out on a glucose-free medium, since the decrease in gene expression caused by catabolite repression is only c.50% (Magasnik, 1970), and such a difference is less than the experimental variation observed in assaying Tn1/3 transposition in the same system from day to day.

In the transposition of Tn3651 mediated by pMB9::Tn103, the frequencies of transposition observed were also high : 1.92×10^{-1} ($30^{o}C$) and 8.73×10^{-4} ($37^{o}C$). These figures are very similar to those obtained using pSNO15.

Photographs of cointegrates obtained from the DS916, R388, pSNO15, pPAK200 and DS916, R388, pMB9::Tn103, pPAK200 strains and their resolution to pPAK200 and R388::Tn3651 in the presence of the tnpR^{+} plasmid pDS4153 are shown in Fig. 4.9 and 4.10.

Thus, despite the higher copy number of pSNO15, and the stronger transcriptional and translational control signals of the hybrid gene producing the B-galactosidase/transposase fusion protein encoded on it, there is no significant difference in the transposition frequency of Tn3651 in a cell containing pSNO15 and in a cell containing a derepressed transposon producing native transposase expressed from its own natural transcriptional and translational control signals. It is interesting to notice that the temperature variability of transposition is maintained in pSNO15. This variability was not simply due to the rate of transposition being slower at $37^{o}C$: transposition assays using donor strains grown for 1 - 4 days at $37^{o}C$ showed no significant variations in transposition frequency.

4.5. Is the protein produced by pSNO15 a fusion protein?

Whilst the DNA sequence of pSNO15 suggested that the protein produced as a result of the insertion of the pPAK100 fragment into pUC8 was a fusion protein between B-galactosidase and transposase, it was possible that transposase activity expressed by this plasmid arose by some other mechanism. e.g. by translation initiation at an in-frame translation start lying within the tnpA sequences in pSNO15 and therefore downstream of the normal translation start. (Even though the Tn1 sequence data obtained in this work showed no codon within it at which this would

Fig.4.9.

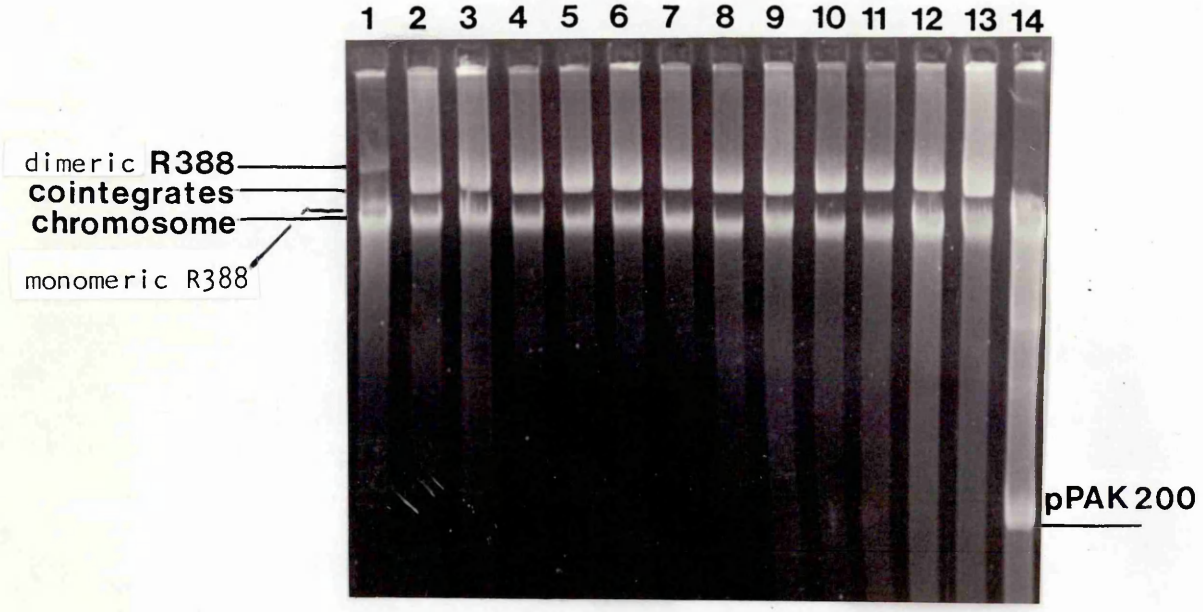


Fig.4.9. Single colony gel analysis of R388::Tn3651::pACYC184:: Tn3651 cointegrates isolated from pSNO15 and pMB9:: Tn103 transposition assay systems.

0.8% agarose gel showing higher copy number and larger size of cointegrates compared to R388.

1: DS916, R388.

2-4: DS902, cointegrates from pSNO15 - 30°C assay systems.

5-7: DS902, cointegrates from pSNO15 - 37°C assay systems.

8-10: DS902, cointegrates from pMB9::Tn103 - 30°C assay systems.

11-13: DS902, cointegrates from pMB9::Tn103 - 37°C assay systems.

14: DS903, pPAK200.

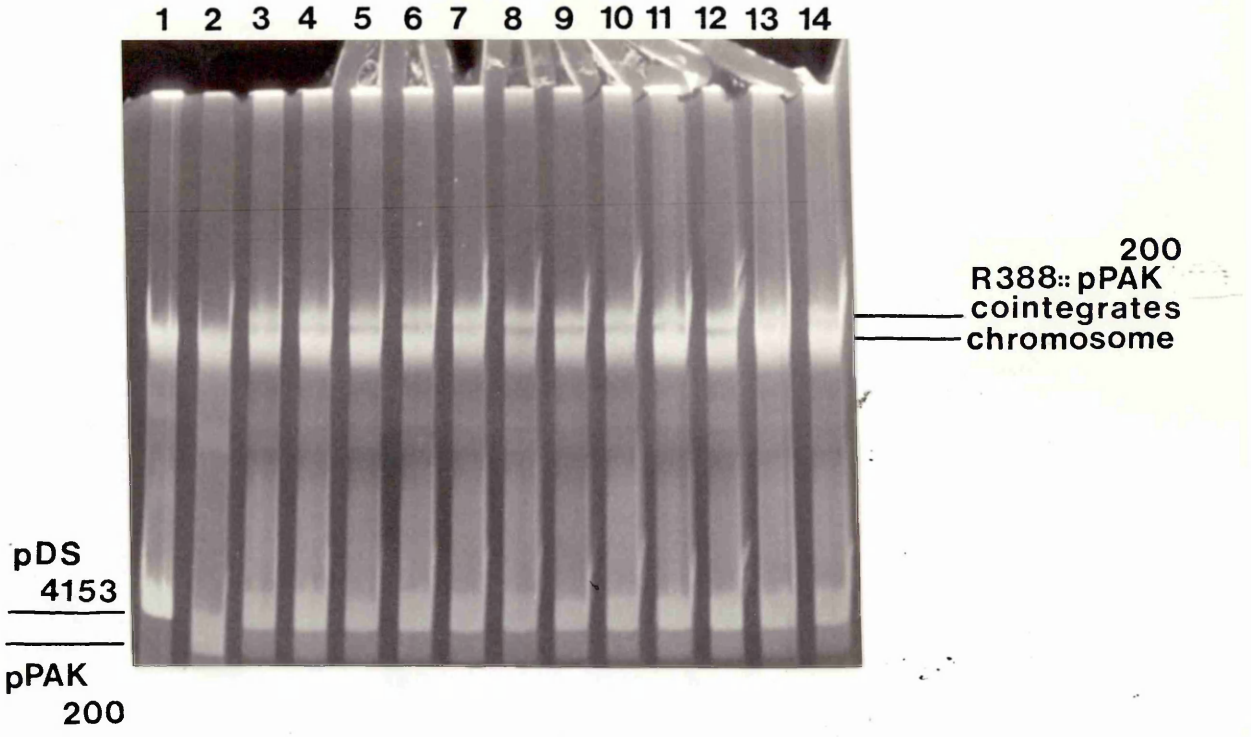


Fig.4.10. Single colony gel to show resolution of cointegrates shown in Fig.4.9. to pPAK200 and R388::Tn3651 in the presence of pDS4153.

0.8% agarose gel showing generation of pPAK200 (and R388::Tn3651) from the cointegrates seen in Fig.4.9.

Note that a significant proportion of cointegrates remained unresolved. (This was after c.20-25 generations of growth).

1: CJ100, pDS4153.

2: DS903, pPAK200.

3-5: CJ100, pDS4153, cointegrates from pSN015 - 30°C assay systems.

6-8: CJ100, pDS4153, cointegrates from pSN015 - 37°C assay systems.

9-11: CJ100, pDS4153, cointegrates from pMB9::Tn103 - 30°C assay systems.

12-14: CJ100, pDS4153, cointegrates from pMB9::Tn103 - 37°C assay systems.

currently be thought possible to occur, it was still conceivable that such a codon could lie downstream of the sequenced region and that translation initiation at this position could generate a protein that retained all or some of its biological activity).

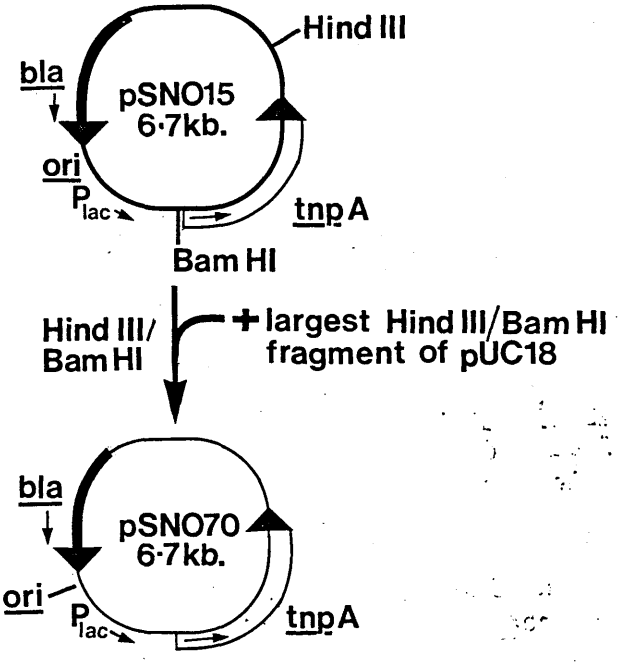
Other plasmids were therefore constructed using the same tnpA gene sequences as in pSNO15, to investigate this.

These were :

(i) pSNO70 : this was constructed by cloning the tnpA gene sequence-containing, BamHI-HindIII fragment from pSNO15 into pUC18. pUC18 differs from pUC8 only in the number and arrangement of restriction sites in its multiple cloning site, but these differences mean that translation initiated at the B-galactosidase initiation codon in pSNO70 enters the tnpA coding region out of frame and is terminated 7bp. into the tnpA sequence by a UGA (opal)codon. UGA is one of the two major *in vivo* stop codons, and therefore termination at this site would be expected to be efficient in a non-suppressor strain (Glass, 1982). See Fig 4.11. Thus, there is no possibilty that a fusion protein could be produced from this construct without frameshift suppression occurring. Hence any protein active in promoting transposition that is produced must be the result of in frame translation initiation within the tnpA coding sequence or of the natural frameshift suppression causing the tnpA sequence to be translated in the correct frame after translation initiation had occurred at the B-galactosidase initiation codon. This would lead to the production of another B-galactosidase/transposase fusion protein, different from the one generated by pSNO15.

(ii) pSNO21 : this was constructed by cloning the tnpA gene sequence-containing EcoRI-HindIII fragment from pSNO15 into the vector pKK223-3 where it was under the control of the very strong hybrid trp-lac (tacI) promoter. This vector contains no translation initiation codons between the promoter and the multiple cloning site (de Boer *et al*, 1983; Pharmacia PL Biochemicals, 1984). See Fig.4.12. Any expression of a protein active in promoting transposition from this construct would

Fig. 4.11.



◀– inverted repeat of Tn1 or Tn3

P_{lac} – lac promoter ori – origin of replication

fmet thr met ile thr asn ser ser ser val pro gly asp pro ser

5'-ATGACCATGATTACGAATTCGAGCTCGGTACCCGGGGATCCGTCG-

thr phe stop

-ACTTTCTGA-3'

⟷ – tnpA coding region

Fig.4.11. The construction of pSNO70.

Ligation of the c.4.1kb. HindIII/BamHI fragment from pSNO15 to the largest HindIII/BamHI fragment from pUC18 generated pSNO70.

Also shown is the N-terminal nucleotide sequence of the polypeptide whose translation initiation occurs at the translation initiation codon of beta-galactosidase. It can be seen from this that the reading frame of this polypeptide enters the tnpA coding region out of the correct reading frame for transposase and is terminated seven bases into the tnpA sequence by a TGA (opal) codon.

Fig.4.12.

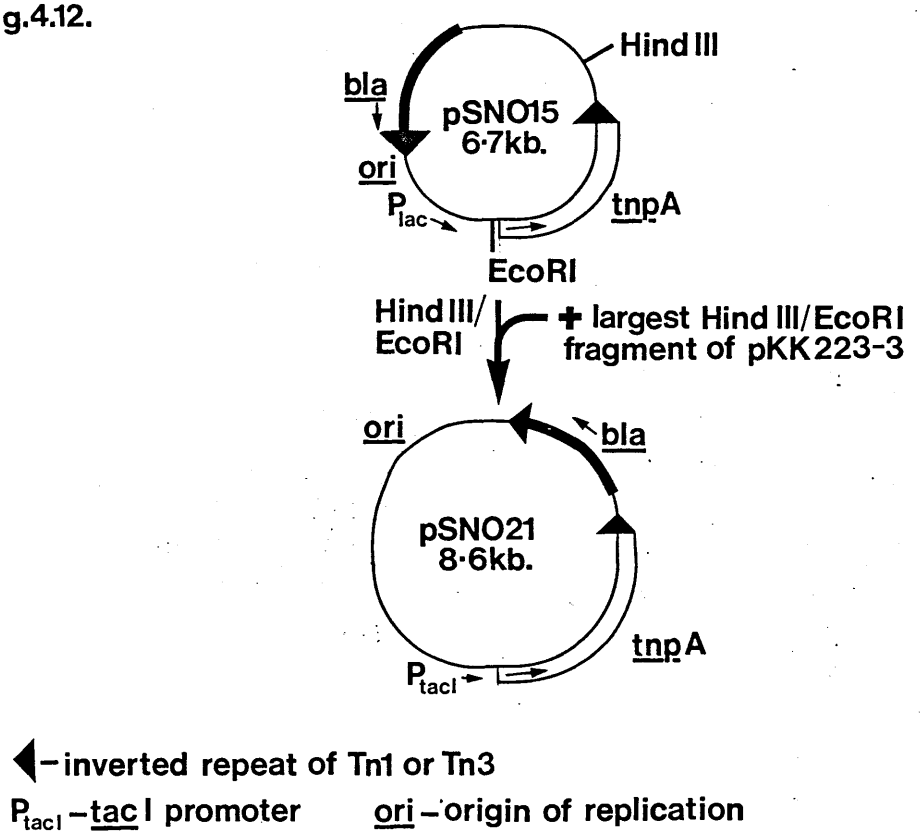


Fig.4.12. The construction of pSNO21.

Ligation of the c.4.1kb. HindIII/EcoRI fragment from pSNO15 to the largest HindIII/EcoRI fragment of pKK223-3 generated pSNO21.

There are no recognised prokaryotic translation initiation signals between the tacI promoter and the multiple cloning site in pKK223-3. Hence, any expression of polypeptide(s) from a pKK223-3 derivative will only occur if the inserted fragment contains a functional translational initiation region.

therefore have to arise from in frame translation initiation within the tnpA sequences.

Both these plasmids were constructed as described and their structures confirmed by restriction analysis. (Data not shown). They were then assayed for their ability to provide transposase *in trans* by determining the transposition frequency of Tn3651 into R388 in the presence of each plasmid. Again, a similar system using pMB9::Tn103 as a $tnpA^+$ plasmid was used as a positive control system. The donor strain for these assays was again DS916 in which it would be expected that both pSNO21 and pSNO70 would be derepressed, even in the absence of IPTG. However, parallel assays in the presence or absence of IPTG were carried out at both 30°C and 37°C. Donor strain constructions in these assays were exactly as described in Section 4.4 for the pSNO15 and pSNO16 assays, except that pSNO70 and pSNO21 were used instead of pSNO15 and pSNO16. After their constructions, donor strains were mated with DS902 and the transposition frequency of Tn3651 determined by calculating the ratio of $Str^rTp^rCm^r$: Str^rTp^r exconjugants. The results of these assays are shown in table 4.3.

It can be seen that both pSNO21 and pSNO70 do promote the transposition of Tn3651, but at a much lower frequency than native transposase produced from Tn103. Because the transposition frequencies mediated by pSNO21 and pSNO70 are very similar, it seems most likely that the mechanism of expression of the transposition-promoting protein they produce is the same. Therefore, it seems probable that this mechanism is in frame translation initiation at a site within the tnpA coding region. It is not possible to say whether or not the low level of transposition promoted by both pSNO21 and pSNO70 is due to very low level expression of a fully active protein or to higher level expression of a less active protein. Neither plasmid produced a visible protein band of a size appropriate for transposase (or a slightly smaller protein) in *E.coli* minicells. (Data not shown.)

These data clearly show that the levels of transposition mediated by

Table 4.3: Transposition assay data on plasmids pMB9::Tn103, pSNO21 and pSNO70.

Complementing plasmid.	Substrate plasmid.	+/- IPTG.	Transposition frequency at: 30°C.	37°C.
pMB9::Tn103	pACYC184	+	1.44×10^{-4}	4.38×10^{-6}
pMB9::Tn103	pACYC184	-	2.49×10^{-4}	7.86×10^{-6}
pMB9::Tn103	pPAK200	+	9.61×10^{-2}	4.15×10^{-5}
pMB9::Tn103	pPAK200	-	4.07×10^{-3}	9.43×10^{-5}
pSNO21	pACYC184	+	$<10^{-7}$	$<10^{-8}$
pSNO21	pACYC184	-	$<10^{-7}$	$<10^{-8}$
pSNO21	pPAK200	+	9.49×10^{-6}	1.59×10^{-6}
pSNO21	pPAK200	-	5.48×10^{-5}	2.17×10^{-6}
pSNO70	pACYC184	+	$<10^{-7}$	$<10^{-8}$
pSNO70	pACYC184	-	$<10^{-7}$	$<10^{-8}$
pSNO70	pPAK200	+	5.98×10^{-5}	1.57×10^{-6}
pSNO70	pPAK200	-	2.86×10^{-5}	2.13×10^{-6}

Table 4.3 legend:

(i). frequencies given as $<10^{-n}$ indicate that no transposition was detected despite examining at least $5x10^{n}$ exconjugants. At least 100 colonies, all containing cointegrates, were examined to calculate positive transposition frequencies.

(ii). the presence of cointegrates in $Str^{r}Tp^{r}Cm^{r}$ exconjugants was confirmed by single colony gel analysis and by their resolution in the presence of a *tnpR*$^{+}$ plasmid (pDS4153) to pPAK200 and R388::Tn103. (Data not shown).

pSNO21 and pSNO70 are much lower than those promoted by pSNO15. This strongly suggests that the majority of transposase activity produced from pSNO15 is the above-described B-galactosidase/transposase fusion protein and that this is active in mediating transposition. It is also possible that a small amount of a protein whose translation is initiated in frame at an undefined site within the tnpA coding region of Tn1 downstream of the position in Tn1 equivalent to co-ordinate 2926 in Tn3 is produced from this plasmid.

4.6. Minicell analysis of proteins produced from pSNO15.

The proteins produced from pSNO15 were analysed using the minicell producing strain DS910 as described in section 2.22. They were compared with the proteins produced from the parental plasmid pUC8. The results of this analysis are shown in Fig.4.13.

It can be seen from these data that there is a protein of c.111Kd. in size produced specifically by pSNO15 at both 30°C and 37°C. This was taken to be the B-galactosidase/transposase fusion protein. It can also be seen that its level is considerably lower (more than 5x lower by comparing the two different loadings on the gel and c.7x lower by microdensitometry) at 37°C than at 30°C. See Fig.4.14. Thus, it seems likely that at least part of the reason for the variation in the level of transposition mediated by pSNO15 at 30°C and 37°C is a variation in the amount of the fusion protein expressed at these temperatures. The possibility still remains open, however, that the activity of the protein varies with temperature as well.

Fig.4.13.

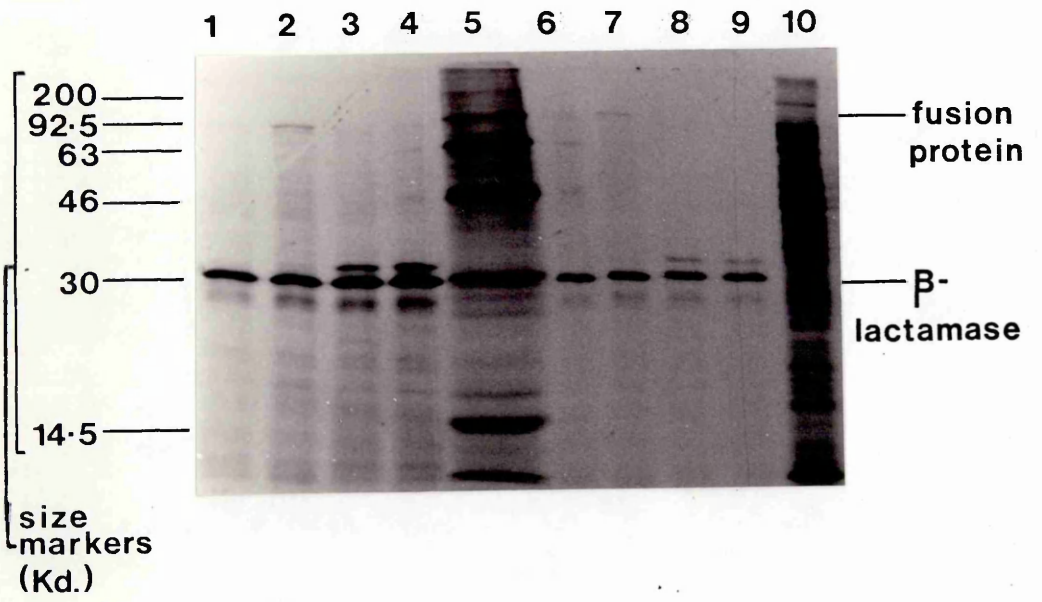


Fig.4.13. Autoradiograph of ^{35}S-methionine labelled proteins encoded by pUC8 and pSNO15 in E.coli minicells at 30°C and 37°C.

4.5% / 12.5% stacking polyacrylamide gel showing a band of a polypeptide of the anticipated size (c.111kd.) of the beta-galactosidase/ transposase fusion protein produced specifically by pSNO15. This polypeptide is at least 5x more plentiful at 30°C than at 37°C.

The three bands around c.30kd. in size in both the pUC8 and pSNO15 lanes are probably three different forms of the beta-lactamase protein: precursor, active form and breakdown product from largest to smallest (Dougan et al, 1979). The putative precursor band can be seen to be much fainter in the pSNO15 than in the pUC8 lanes. The reason for this is not known.

1: DS910 minicells, pSNO15 - 37°C: 20ul. of sample.

2: DS910 minicells, pSNO15 - 30°C: 20ul. of sample.

3: DS910 minicells, pUC8 - 37°C: 20ul. of sample.

4: DS910 minicells, pUC8 - 30°C: 20ul. of sample.

5: ^{14}C-labelled size standards. (See below).

6: DS910 minicells, pSNO15 - 37°C: 4ul. of sample.

7: DS910 minicells, pSNO15 - 30°C: 4ul. of sample.

8: DS910 minicells, pUC8 - 37°C: 4ul. of sample.

9: DS910 minicells, pUC8 - 30°C: 4ul. of sample.

10: DS910 whole cellls, plasmid-free.

The sizes of the ^{14}C-labelled size standards are as follows:-

200kd. - myosin heavy chain.	46kd. - ovalbumin.
92.5kd. - phosphorylase b.	30kd. - carbonic anhydrase.
63kd. - bovine serum albumin.	14.5kd. - lysozyme.

Fig.4.14.

i). 30°C.

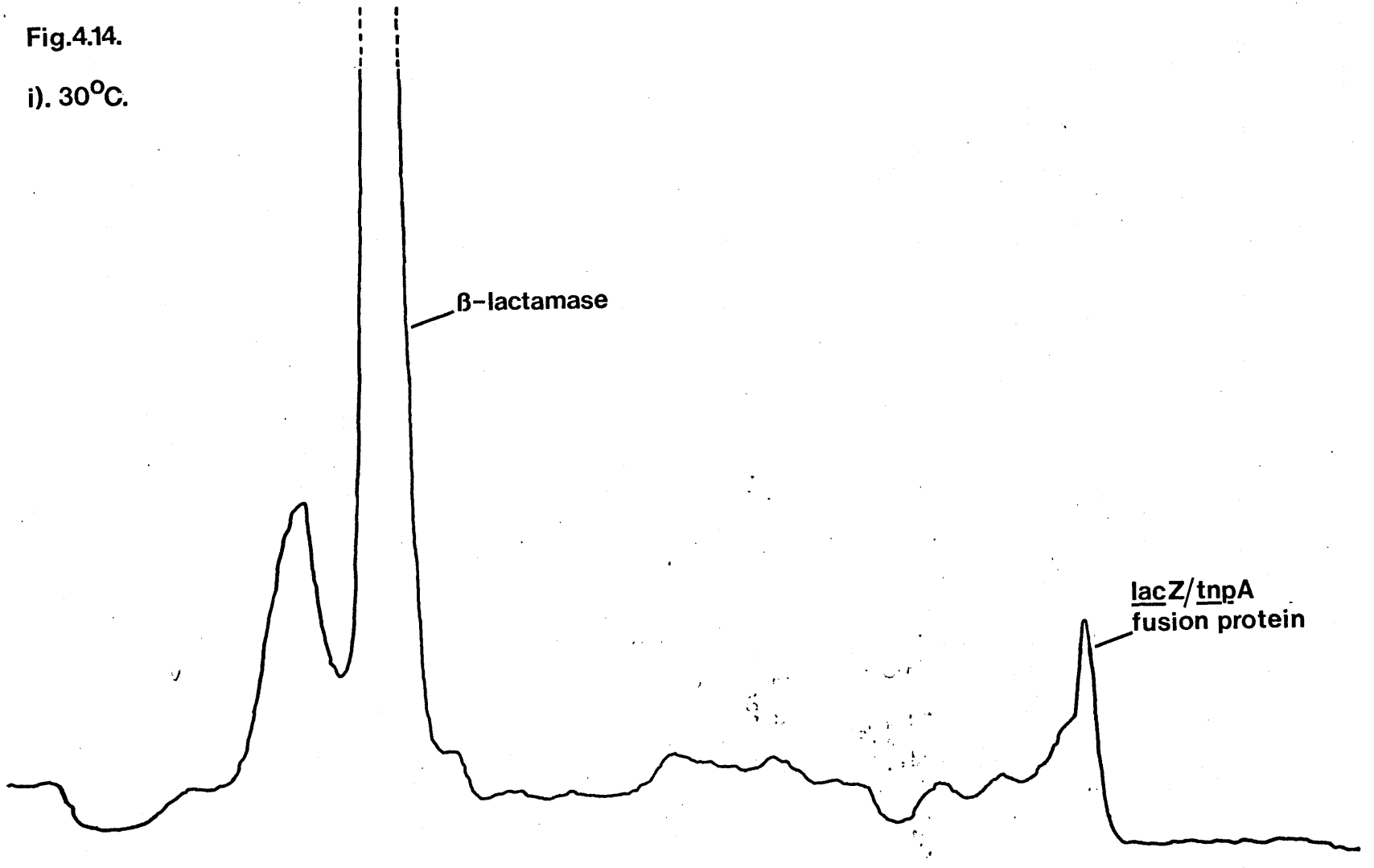


Fig.4.14.

ii). 37^{o}C.

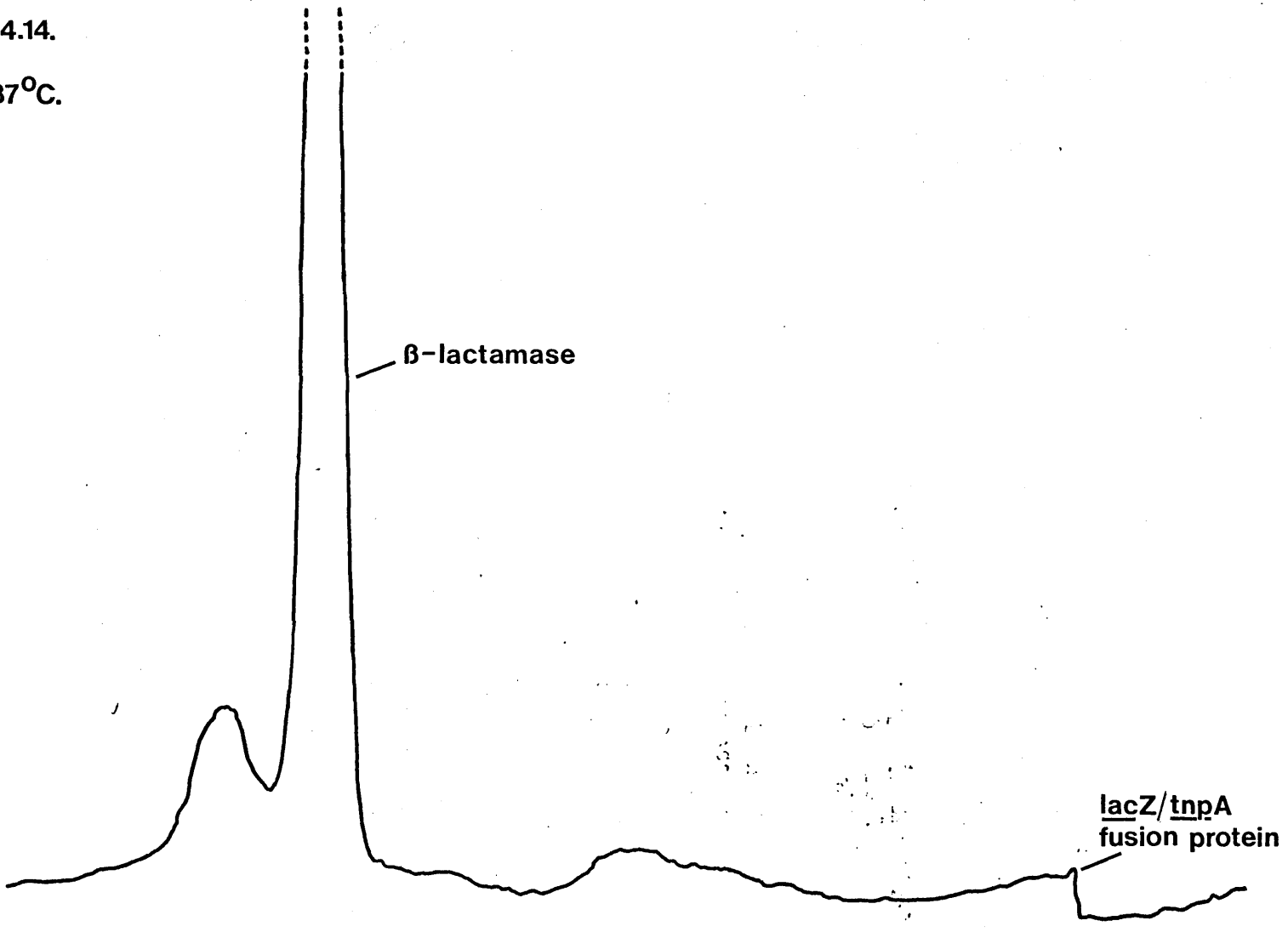


Fig.4.14. Microdensitometer tracings of the pSN015 tracks in the autoradiograph shown in Fig.4.13.

The tracks used for this were the DS910 minicells, pSN015 - 20ul. of samples at 30°C and 37°C.

From these tracings, it was calculated that the beta-galactosidase/ transposase fusion protein was c.7x more abundant at 30°C than at 37°C.

4.7. Discussion.

The exact location of the Hinc II restriction site present in Tn1 but absent in Tn3 was determined by DNA sequence analysis. It was found to extend over a region in Tn1 equivalent to base-pairs 3037-3042 in Tn3, with the cleavage occurring between base-pairs 3039 and 3040. This additional site has been generated by an A to G (transition) mutation at base-pair 3037 in Tn3. Examination of the Tn1 sequence data obtained revealed 97/114 (= 85.1%) homology at the nucleotide level and 37/38 (= 97.4%) homology at the amino acid level with Tn3.

A HincII fragment extending leftwards from this additional site in Tn1 to a HincII site outwith the tnpA gene was cloned into pUC8 to produce pSNO15. This plasmid encoded a fusion protein consisting of the Tn1 transposase protein with its three N-terminal amino acid residues removed and replaced by the first eleven N-terminal amino acid residues of B-galactosidase. The initiation of both the transcription and translation of the hybrid gene encoding this protein would be expected to be under the control of lac regulatory elements. This protein was found to be active in promoting the transposition of Tn3651 *in trans*, and the level of transposition in a cell containing pSNO15 was very similar to that in a cell containing pMB9::Tn103 at both 30°C and 37°C. (In Tn103, wild-type transposase protein was produced under the control of the tnpA gene transcriptional and translational regulatory signals). Analysis of the proteins produced by pSNO15 in *E.coli* K-12 minicells showed a band of size appropriate for this fusion protein, and this was more than 5x more abundant at 30°C than at 37°C.

Thus, the initial aim of generating a bacterial strain capable of over-producing Tn1 transposase sufficiently for it to be a source from which to isolate quantities of the protein necessary to set up an *in vitro* transposition system was not achieved. However, in attempting to do this, data were produced that shed some light on the regulation of tnpA gene expression and on the possible source of the variation of

Tn1/3 transposition frequency seen with temperature (this work: Tn1; Kretschmer and Cohen (1979): Tn3).

It was an unexpected observation that the transposition frequency of Tn3651 was much the same in the presence of pSNO15 or pMB9::Tn103. For, in pSNO15 , the tnpA coding sequences were expected to be under the control of lac transcriptional and translational initiation signals, both of which were stronger than the corresponding tnpA signals (Table 4.1 above and Chou et al, 1979b). The copy number of pSNO15 was also considerably higher - probably c.4x so - than that of pMB9::Tn103.

The most obvious explanation for these findings is that the fusion protein from pSNO15 is less active than the wild-type transposase protein. This is quite feasible. However, it is also coincidental that the degree of reduction in activity of the fusion protein is such that it almost exactly balances out the increased expression anticipated due to the lac regulatory elements present on, and the higher copy number of, pSNO15. An alternative explanation could be that there is a hitherto unknown mechanism(s) of regulation of tnpA gene expression that acts at a stage other than the initiation of transcription or translation.

It is also interesting to note that the variation in transposition frequency with temperature persists in cells containing pSNO15, and this observation provides some clues as to the source of this effect. Possible explanations for both these findings are given below, but it is firstly valuable to consider other features of Tn1/3 transposition and expression of the tnpA gene not directly addressed by this work:

(i). the presence of an active tnpR gene in a cell decreases the transposition frequency of a Tn1/3 derivative (in which tnpA gene transcription is directed by the tnpA promoter) by c.5-100fold (Heffron et al, 1979; Casadaban et al, 1982; Heffron, 1983). This observation is explained by the repressive effect of resolvase protein on the activity of the tnpA gene promoter: in the absence of resolvase, the activity of this promoter is c.3.5x more than in its presence, as measured by galactokinase assays (Kelly, 1983; Dyson, 1984).

(ii). very little is known about the stability or secondary structure of the tnpA gene mRNA. However, computer analysis of the nucleotide sequence of the 5' end of this message reveal no obvious areas of significant and stable secondary structure, suggesting that modifications of any such structure by other factors are not very important in regulating tnpA gene expression.

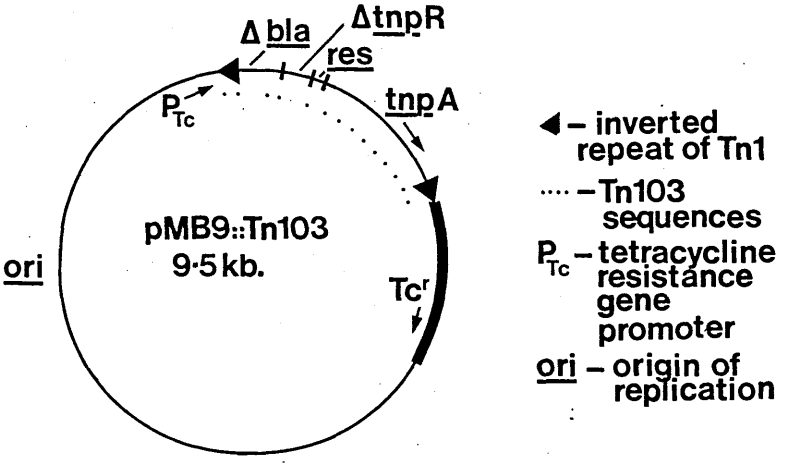
(iii). the transposase proteins of Tn1 and Tn3 show no cis preference (G. Russell, pers. comm.) and are thus different from the transposase proteins of 'phage Mu, and many IS elements, which do show a cis preference in their actions e.g. Mu - Pato and Reich, 1982; IS1 - Machida, Y. et al, 1982b.; IS10 - Morisato et al, 1983; IS50 - Isberg and Syvanen, 1981; IS903 - Grindley and Joyce, 1981.

With regard to the variation of transposition frequency with temperature, the observation that this persists in cells containing pSNO15 indicates that the cause of this variation cannot reside in the tnpA gene elements responsible for regulating transcriptional and translational initiation, unless the corresponding lac elements show an identical pattern of variation. No direct data is currently available on any variation in the activity of these latter elements with temperature. However, theoretical analyses suggest that promoter activity will increase with increasing temperature (due to displacement of the equilibrium between open and closed transcription initiation complexes towards the open form) and studies on the interaction of E.coli RNA polymerase holoenzyme with 'phage T7 promoters showed there is little change in the stability of the open initiation complex between 20°C and 37°C (von Hippel et al, 1984; Chamberlin, 1974). Hence, it seems unlikely that the lac promoter will decrease in activity between 30°C and 37°C and it seems most probable that the sequences responsible for the variation in transposition frequency seen with temperature lie in that part of the Tn1 tnpA gene present in pSNO15. Similarly, it seems most probable that the variation in protein levels seen with

temperature in cells containing pSNO15 is due to some feature of those tnpA sequences retained in this plasmid. The mechanism of this variation remains unclear. However, it seems likely to be due to some features of the tnpA mRNA or protein - e.g. different stabilities of either, or both, at different temperatures; an attenuation system of transcription showing different activity at different temperatures.

Other data have been obtained in this laboratory that have a bearing on the effect of temperature on tnpA gene expression. A plasmid encoding a transposase/B-galactosidase fusion protein was constructed by inserting an active lacZ gene fragment in frame into the tnpA coding region at the PstI site at ~~1642~~ 984 in Tn3. Hence, the transcription and translation of this protein are under the control of tnpA gene regulatory elements. B-galactosidase assays on this plasmid showed that the level of expression of this protein varies with temperature, being c.3-4x greater at 30°C than at 37°C. (M. Burke, pers. comm.). This suggests that the sequences responsible for these temperature differences lie to the right of this PstI site - i.e. towards the N-terminus of the tnpA gene. Also, in the pMB9::Tn103 molecule used in these experiments, Tn103 is inserted between the promoter and translation initiation codon of the Tc^r gene. Thus, this gene's transcription is directed by the tnpA gene promoter by transcriptional readthrough causing the production of a dicistronic mRNA (See Fig. 4.15). It is found that the resistance to tetracycline conferred on a cell by this plasmid varies with temperature in a similar way to the frequency of transposition of Tn3651 mediated in trans by pMB9::Tn103 e.g. level of tetracycline resistance at 30°C = c.30 ug/ml; level of resistance at 37°C = c.12 ug/ml (A. Arthur, pers.comm.). Since, in this plasmid, translation of the mRNA for the tetracycline resistance gene product is under the control of tetracycline resistance gene control elements, it seems most likely that the source of this temperature variation is either the tnpA gene promoter or some feature of the mRNA. (The tetracycline resistance level conferred by pMB9 itself does not

Fig.4.15.



mRNA produced from tnpA promoter:

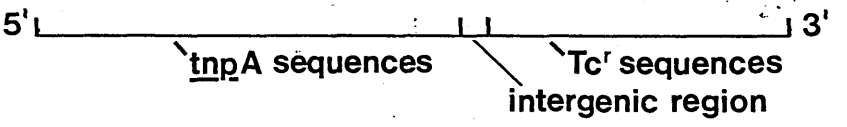


Fig.4.15. Map of the plasmid pMB9::Tn103 used throughout this work.

Tn103 is inserted betwen the tetracycline resistance gene coding region and its promoter. Therefore, the transcription of this gene is under the control of the tnpA gene promoter by transcriptional readthrough, involving the production of the dicistronic mRNA that is also shown.

The translation of the tetracycline resistance gene part of this mRNA, however, remains under the control of that gene's own translational control signals.

vary with temperature and unpublished data from this laboratory shows the copy number of pMB9::Tn103 does not differ significantly at 30°C and 37°C.) As mentioned above, it seems unlikely that the tnpA gene promoter's activity varies with temperature. Hence, it seems most probable that the reason for this variation is some feature of, or action upon, the mRNA.

As there is a variation with temperature of both transposition frequency mediated by, and amount of fusion protein produced by pSNO15, it seems highly probable that at least part of the reason for this variation in transposition frequency is a difference in the amount of protein present in the the cell at different temperatures.

Other workers too have made the observation with Tn3 that the presence of more transposase in the cell is associated with an increase in transposition frequency. e.g. in the comparison of a tnpR$^+$ cells with tnpR$^-$ ones - Heffron et al, 1977, 1978; Casadaban et al, 1982. For Tn10 also, it has been shown that an increase in the amount of transposase is associated with an increase in the level of transposition (Morisato et al, 1983).

The relationship between the amount of protein and transposition frequency for Tn1 is, however, clearly non-linear, and it could be, therefore, that another effect(s) contributes to the variation of transposition frequency with temperature - e.g. altered activity of the protein.

Taking all these data into account, it seems most likely that the difference in transposition frequency seen with temperature is caused, at least partly, by a difference in the amount of transposase protein present in the cell and that the sequences responsible for these variations lie between the co-ordinates ~~1642~~ 984 and 3040 in Tn1/3. The possible mechanism(s) of action of these sequences is not yet clear - it could be that a polypeptide is encoded within this region that has different effects at 30°C and 37°C. Or, it could be that the transposase protein is less stable at 37°C than at 30°C, due to some

feature of this region. Or it could be that there is some feature of the tnpA mRNA secondary structure present in the region that renders the mRNA less stable at 37°C than at 30°C e.g. by rendering it more accessible to enzyme degredation. Or, perhaps some secondary structure is necessary for higher level expression of transposase (eg. between two parts of one RNA molecule, between two separate RNA molecules or between an RNA and a DNA molecule) and this structure is more stable at 30°C than at 37°C. Or, perhaps this is the region where a mechanism of attenuation of transcription occurs, active to different degrees at different temperatures. Overall, however, the available data suggest that it is some feature(s) of the tnpA mRNA between these co-ordinates that is responsible for these variations.

Several additional pieces of data support the hypothesis that the lack of significant variation in the transposition frequency of Tn3651 mediated by pSNO15 or pMB9::Tn103 is not solely due to an appropriate decrease in activity of the fusion protein encoded by pSNO15 in comparison to wild-type transposase. For instance, minicell analysis of proteins encoded by Tn1/3 showed that ^{35}S-methionine labelled resolvase protein is much more abundant than transposase protein labelled in the same way (Chou et al, 1979a and b; Gill et al, 1979). Yet when the relevant factors are taken into account, it would be anticipated that the two bands would be of approximately equal intensity. See Table 4.4.

Also, minicell analysis of proteins encoded by pMB9::Tn103 and pSNO15 consistently failed to show a band corresponding to a polypeptide of a size appropriate for wild-type transposase from pMB9::Tn103, whilst it did show a band corresponding to a polypeptide of the size predicted for the fusion protein encoded by pSNO15. And, single colony gel analysis of donor strains used in transposition assays showed that, in both pMB9::Tn103 and pSNO15 containing cells, transposition did not occur up to a level where all the potential donor replicons had been involved in

Table 4.4: Factors affecting the intensities of transposase and resolvase ^{35}-S methionine labelled protein bands seen in E.coli K-12 minicells.

Factor.	Ratio.	Source.
tnpA gene promoter: tnpR gene promoter.	2.58: 1.	This work.
tnpA gene translation initiation signals: tnpR gene translation initiation signals.	0.136: 1.	M.Burke, pers. comm.
No. of internal methionine residues in transposase: no. of internal methionine residues in resolvase.	4: 1.	Heffron et al, 1979.
Overall ratio.	1.4:1.	

Table 4.4 legend:

(i). no data are currently available concerning the relative stabilities of the tnpA and tnpR mRNAs and proteins, both of which are also relevant to this matter.

a transposition event. See, for example, Fig.4.16. Also, preliminary data has indicated that transposition of Tn3651 mediated by a chromosomally located tnpR$^-$ Tn3 derivative reproducibly occurs at a c.10-fold lower frequency than when mediated by the same transposon carried on the plasmid pRSB31 (copy number = 30-40 per genome equivalent) or by pSNO15 (G.Russell and R.Slatter, pers.comms.). Lastly, transposition assay data obtained using R388::Tn103 (copy number = 3-4 per genome equivalent) as the donor replicon showed that the transposition frequency of Tn103 in such a system was consistently lower in systems using pMB9::Tn103 as the tnpA$^+$ replicon. On average, the transposition frequency of Tn103 from R388::Tn103 was 0.33x and 0.62x that of Tn3651 mediated *in trans* by pMB9::Tn103 at 30°C and 37°C respectively. (See Tables 3.3, 4.2, 4.3 and 5.1). Clearly, comparisons of these two systems can only be made with caution as there are certain differences between them. For instance, the immediate sequence environment of the transposon in the two donor replicons is not the same, the sequences of the two target replicons are different and pACYC184 probably has a greater proportion of essential DNA sequences per plasmid than R388. (But, it also has a copy number that is c.10x higher than R388 and hence the amount of non-essential DNA in a cell containing pACYC184 or R388 is probably much the same). However, despite these differences, the observed variation may be significant.

Hence, there is some evidence to suggest that, at low gene dosage (between one and twenty genes per cell), there is an increase in transposition frequency with increasing gene dosage. However, at higher gene dosage (greater than twenty genes per cell), there is no further significant increase in transposition frequency with increasing gene dosage. See Fig. 4.17.

In turn, this suggests that there is a ceiling imposed on the level of transposition in a cell so that there is no increase in this level above a certain gene dosage (>3/4; <20). This could be due to some transposon encoded product(s), or to a host factor(s).

Fig.4.16.

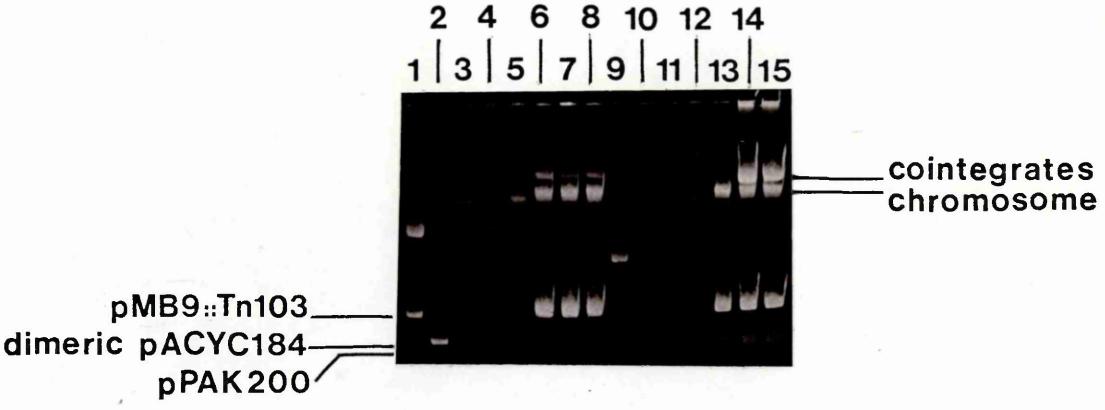


Fig.4.16. Single colony gel analysis of donor strains in pMB9:: Tn103 transposition assay systems.

0.8% agarose gel showing that cointegrates were specifically generated in cells containing R388, pMB9::Tn103 and pPAK200, but that not all potential donor replicons were involved in a transposition event.

Very similar findings were observed in donor strains in pSNO15 transposition assay systems.

1: pMB9::Tn103 DNA.

2: dimeric pACYC184 DNA.

3-5: DS916, R388, pMB9::Tn103, pACYC184 - 30°C.

6-8: DS916, R388, pMB9::Tn103, pACYC184 - 37°C.

9: pPAK200 DNA.

10-12: DS916, R388, pMB9::Tn103, pPAK200 - 30°C.

13-15: DS916, R388, pMB9::Tn103, pPAK200 - 37°C.

Fig.4.17.

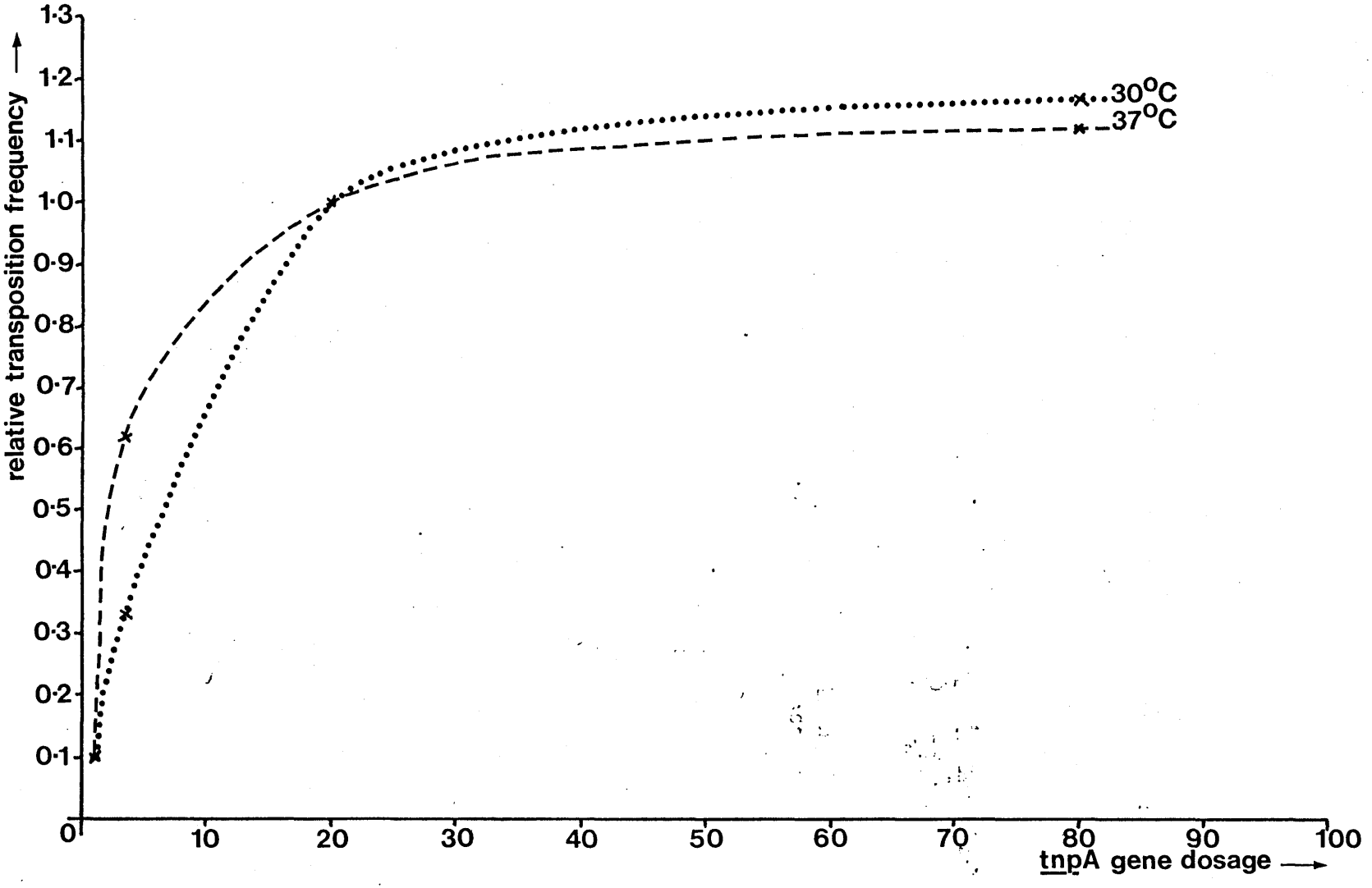

relative transposition frequency
1·3
1·2
1·1
1·0
0·9
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
0
10
20
30
40
50
60
70
80
90
100
tnpA gene dosage
30°C
37°C

Fig.4.17. Graph showing the variation in transposition frequency of Tn1 derivatives with tnpA gene dosage at 30°C and 37°C.

With a tnpA gene dosage of 1-c.20 copies, transposition frequency increases as gene dosage increases. Above a tnpA gene dosage of c.20 copies, however, no further significant increase in transposition frequency is seen as gene dosage rises.

For example, each transposon could produce a product inhibitory to transposition at a higher rate than transposase. Thus, an increase in the number of transposons in a cell would increase the difference in total amounts of these two molecules. Hence, as the copy number of the transposon rises, the transposition rate per element would fall, until it was negligible, when any further increase in transposon copy number would cause no further increase in transposition frequency. Such a system, where the inhibitory molecule is an antisense RNA species that inhibits translation of the transposase mRNA, has been shown to exist in Tn10 - Simons et al, 1983; Simons and Kleckner, 1983. Also, in Tn5, a regulatory system involving a second polypeptide has been shown to exist. This protein acts by some mechanism other than repression of the synthesis of Tn5 proteins (Isberg et al, 1982; Johnson et al, 1982). Any such molecules would clearly have to be encoded by that part of Tn1/3 lying between co-ordinates 1 and 3040. The observation that pSNO15 appears to produce more transposition - promoting protein than pMB9::Tn103, yet the transposition frequencies mediated by the two are very similar suggests that such a molecule may act either on the protein itself, on the protein's site of action or on a host factor (or factors) also needed for the transposition process.

Alternatively, some host factor(s) needed for transposition could become limiting at high transposon copy number. It is known that an unknown number of host proteins are needed for Mu transposition (Mizuuchi, 1983; Craigie and Mizuuchi, 1985) and that both DNA polymerase I and DNA gyrase are needed for Tn5 transposition (Sasakawa et al, 1981; Isberg and Syvanen, 1982), that one or more host factors are involved in Tn9 transposition (Ilyina et al, 1981) and integration host factor has been suggested as being involved in IS1 transposition (Gamas et al, 1985), so this is quite conceivable. It could, of course, simply be that some factor necessary for transcription or translation becomes limiting at moderate/high copy number. However, this would seem unlikely since very high expression of proteins is possible in certain

specially constructed vectors (e.g. pKK223-3).

This proposed ceiling effect seems to be separate from the effect of temperature on transposition frequency : the maximum transposition frequency seen at 30°C is higher than at 37°C. This could be interpreted to mean that the expression of the factor responsible for this ceiling varies with temperature in a similar way to that of transposase itself.

Overall then, it seems likely that several factors contribute to keep down the transposition frequency of Tn1/3, many by regulating the cellular level of the transposase protein. The effects of resolvase protein, transposition immunity, low frequency transposition to the chromosome and temperature are well documented, and evidence is presented here for an additional form of regulation that appears to cause the imposition of a maximum transposition frequency in any given system.

Chapter 5.

Does antisense RNA contribute to the regulation of *tnpA* gene expression?

* See also the computer print-out of the sequence of Tn3 from co-ordinates 1 - 3200. (Inside the back cover.)

* Note that no search for features characteristic of ribosome binding sites was carried out on the Tn3 sequence.

5.1. Introduction.

As described in Chapter 1, there have now been several well-documented examples of post-transcriptional regulation of gene expression by antisense RNA. Most interesting in relation to this work was the system employed in IS10-right of Tn10, by which the level of transposase protein (and hence the transposition rate per element) is carefully controlled. This mechanism also means that an increase in the number of IS10 elements in the cell leads to a decrease in the rate of transposition per element, a phenomenon known as 'multicopy inhibition' (Simons & Kleckner, 1983).

In view of data obtained from Tn10 (Simons *et al*, 1983, Simons and Kleckner, 1983), a hypothesis was formed that a similar mechanism might operate in the regulation of the transposase gene of Tn1/3. From the data obtained in experiments with a B-galactosidase/transposase fusion protein construct (pSNO15) described in Chapter 4, it seemed plausible that some form of regulation of transposase gene expression, apart from the effect of temperature and at a level other than the initiation of transcription or translation, was occurring in Tn1.

Analysis of the nucleotide sequence of Tn3 (Heffron *et al*, 1979) using both computer programmes and visual examination revealed a number of sequences, in addition to those controlling the three known genes of Tn3, with considerable homology to the optimal prokaryotic consensus promoter sequence (Dahl, 1983; de Boer *et al*, 1983; Hawley and Mc.Clure, 1983). Also revealed were a number of potentially significant open reading frames, some of which were in a correct structural relationship to promoter-like sequences. (The computer programmes used were:- ZPROM : to search for prokaryotic promoter-like sequences ; C. Boyd pers. comm. ; and DSPLAY : to search for open reading frames beginning with an AUG or GUG codon ; P. Taylor pers. comm.).*

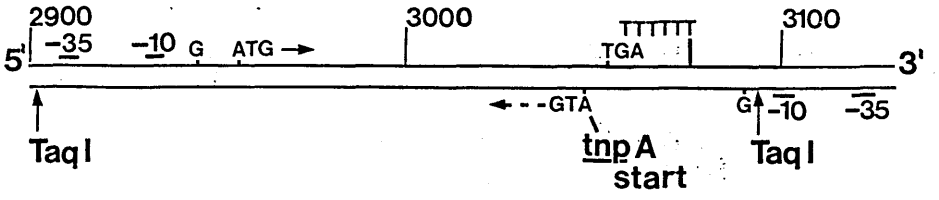
In particular, an area of potential transcriptional and translational activity was located on the antisense strand of the *tnpA* structural

gene around the amino-terminus of the gene. See Fig. 5.1.

Transcription of this region would generate a 131 base RNA species and translation of this would yield a 36 amino acid polypeptide. Such an RNA transcript would be of a similar size to other, known, antisense RNA molecules (e.g. RNA1 of plasmid ColE1 : 108 bases - Tomizawa and Itoh, 1981 ; pOUT transcript of IS10-right of Tn10 : 101 bases - Simons and Kleckner, 1983) and, especially, it would be complementary to the transposase mRNA over an area including both the ribosome binding and translation initiation sites. It was also speculated that the possible translation product from this region could act in a similar way to the Rom protein of ColE1 (Tomizawa and Som, 1984). Comparison of this possible translation product with the Rom protein, however, showed no significant amino acid homology other than a $^{4}/_{5}$ residue area of near perfect homology. See Fig. 5.1. The existence of a potential open reading frame on the 'antisense' strand of this region of Tn3 was not, however, held to be of great significance since the presence of an open reading frame on one strand of a piece of DNA always constrains the arrangement of nucleotides on the other strand in such a way that the generation of sequences with the structural characteristics of open reading frames is favoured. This is because the preferred codon usage in E.coli is for codons of the general structure RNY (where R = a purine nucleotide, N = any nucleotide and Y = a pyrimidine nucleotide). (M. Rogers. pers. comm.).

It is worth noticing that any antisense RNA regulation of tnpA gene expression would persist, at least qualitatively, in the B-galactosidase/transposase fusion protein produced by pSNO15. For, even though the sequence of the mRNA for this fusion protein differs from that of wild-type Tn1 at its 5' end, any antisense RNA transcript will also differ in a complementary way, and all the elements within the tnpA coding region postulated to be important in generating this transcript are retained in pSNO15. The only real difference between pSNO15 and wild-type Tn1 would be that the postulated antisense translation product would be of altered composition beyond amino acid residue 28 and

Fig.5.1.



N-terminus

Rom val thr lys gln glu lys thr ala leu asn met ala arg phe ile arg ser gln thr leu thr leu leu glu lys leu-

Possible Tn3 peptide met lys ile thr arg gln leu lys phe ile gly phe ala ser glu pro ala ile thr leu asn leu leu ser gly gln-

◇ ◇ ◇ ◇

C-terminus

Rom -asn glu leu asp ala asp glu gln ala asp ile ⟶ leu STOP

Possible Tn3 peptide -lys ile asn gly his ile gly leu pro ala STOP

◇ - homologous residues

Fig.5.1. The region of Tn3 lying between co-ordinates 2900 and 3130.

The areas of known and putative transcriptional and translational significance are indicated.

Also shown is a comparison between the Rom protein of plasmid ColE1 and the possible 36 amino acid residue polypeptide encoded on the antisense strand of Tn3 in this region. There is no significant homology between these two other than a $^{4}/_{5}$ residue region of near perfect homology as shown.

extended at its C terminus by 15 amino acid residues in pSN015 due to the above sequence differences.

However, it is commonly found that identification of possibly significant features during analysis of sequence data is by no means a reliable indication of *in vivo* activity for such features. A number of experiments were therefore designed to see if any *in vivo* evidence could be produced to support this antisense RNA model. These were of two kinds:-

(i) cloning of a region containing the intact putative antisense RNA producing region into a new, transposon-sequence free, high copy number vector and assaying this new construct for its effect *in trans* on the frquency of transposition of a derepressed Tn1 derivative.

(ii) cloning of fragments that would be expected to contain the promoter sequences needed for the production of an antisense RNA into promoter probe vectors (Mc.Kenney *et al*, 1981) and assaying these constructs for promoter activity.

5.2. Cloning and assaying of putative RNA producing region.

5.2.1. Cloning of putative antisense RNA producing region.

In choosing a cloning vector for this construction it was important to select a molecule devoid of any known Tn3 sequences, particularly terminal sequences, for it has been shown that such additional sequences, in high copy number, can cause a drop in the transposition frequency of the transposon being analysed, presumably by titrating out active transposase protein (this work - Chapter Three; Heritage and Bennett, 1984). Also, it was important to use a vector that did not have any areas of homology to the putative antisense RNA transcript, so that its only possible target would be the transposase mRNA 5' region. The vector chosen was pACYC184, which fulfils both the above criteria with the proviso that it has been shown, by Southern hydridisation

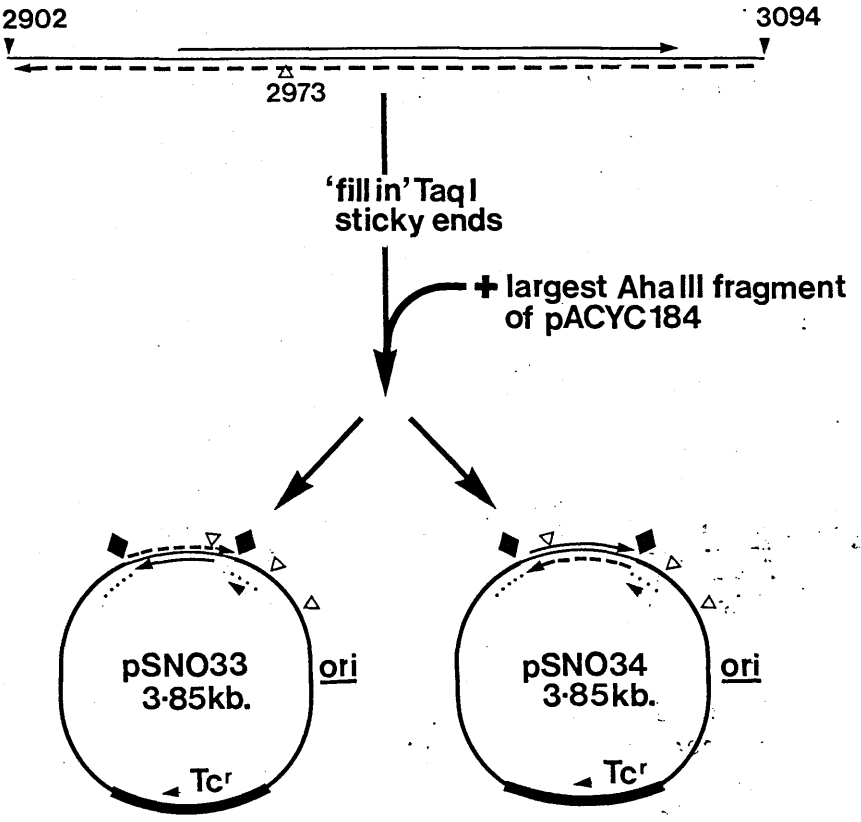
techniques, to contain some IS1 sequences that show homology with the right inverted repeat of Tn3 (E. Nimmo, pers. comm.). However, the possibility that these sequences would interfere with the experiment as planned was examined and allowed for by the design of an appropriate control experiment (see below). pACYC184 contains two AhaIII sites, located 339bp. apart in the chloramphenicol acetyl transferase (CAT) gene (Alton & Vapnek, 1979). AhaIII digests DNA to yield flush-ended fragments.

An examination of the sequence of Tn3 shows that there is a 192 bp. TaqI fragment, extending from 2902 - 3094, that would neatly enclose the whole region thought to be responsible for the generation of the antisense RNA. See Fig 5.1.

RSF1050 was chosen as the source of this 192bp. TaqI fragment: plasmid DNA was digested to completion with TaqI and the resultant fragments separated on a 5% polyacrylamide gel. The 192bp. band was cut from the gel, the DNA electroeluted, ethanol precipitated and resuspended in TE buffer. The 'sticky' ends generated by TaqI digestion were filled in using T4 DNA polymerase, the reaction mixture phenol extracted once, chloroform extracted twice and the DNA again precipitated with ethanol. It was then used in ligation with the largest (3661bp.) AhaIII fragment of pACYC184, previously isolated from a low melting point agarose gel. $\triangle$M15 cells were transformed to Tc^r with this ligation mix and 37 transformants were obtained. Single colony gel analysis revealed that all these clones contained a plasmid of the required size - 3.85kb. See Fig 5.2.

It can be seen that ligation of an AhaIII fragment to a filled in TaqI fragment regenerates only the TaqI site. Therefore, digestion of the desired construct with TaqI will yield the 192bp. fragment. The orientation of the inserted fragment can be determined by PvuII digestion. See Fig 5.2. Plasmid DNA from 5 clones was prepared by the alkaline/SDS method and restricted with TaqI or PvuII. All the five clones analysed yielded the 192bp. TaqI fragment and two of them

Fig. 5.2.



▲ – TaqI site △ – PvuII site ◆ – 'filled in' TaqI/AhaIII junction
ori – origin of replication ····· – remnants of Cm^r gene
---→ – tnpA transcription → – antisense RNA transcription (putative)

5'–TTTCGA — Tn3 sequences — TCGAAA–3'

'filled in' TaqI/AhaIII junctions

TaqI site: TCGA
AhaIII site: TTTAAA

Fig.5.2. The construction of pSNO33 and pSNO34.

The 192bp. TaqI fragment from Tn3 (2902-3094) was isolated, its 'sticky' ends filled in and the resulting piece of DNA ligated to the largest AhaIII fragment of pACYC184. Insertion of the fragment in either orientation was possible and both types of recombinant molecule were recovered. These were designated pSNO33 and pSNO34.

Also shown are the filled-in TaqI/AhaIII junction sequences, showing that these sequences are targets for TaqI but not for AhaIII.

Note that the 192bp. TaqI fragment does not carry the tnpA gene promoter.

contained the fragment inserted in one orientation, whilst the other three contained it in the opposite orientation See Fig 5.3.

These plasmids were designated pSNO33 and pSNO34 according to the orientation of the inserted fragment.

It would be postulated that the promoter of an antisense RNA species would be quite strong, as is the case in known antisense RNA control systems -eg. pOUT in ISIO -right of Tn10 : 1/3rd. as strong as induced placUV5 (Simons et al , 1983). And it is worth considering that, since the TaqI fragment is inserted into the middle of the CAT gene of pACYC184, there may be an effect of CAT gene transcription on the expression of sequences in the fragment, See Fig 5.4.

As the CAT promoter is quite strong (Stueber & Bujard, 1981), and assuming no secondary structures exist in the CAT-promoter generated RNA transcripts from pSNO33 and pSNO34, as seems likely from examination of their sequences, then in pSNO33, transcription from the CAT promoter would generate a large amount of a transcript that would include the antisense region from Tn3, presumably terminating at the transcription terminator of this region. In pSNO34, a corresponding amount of a transcript containing the complementary sequence to the antisense RNA region would be produced and it would be postulated that this would tend to titrate out any antisense RNA transcript generated from its own promoter.

5.2.2. Assay of the effect of pSNO33 and pSNO34 on the transposition frequency of Tn103.

In the assay, constructs putatively producing an antisense RNA species from Tn3 were assayed for their effect on the transposition of Tn103, a Tn1 derivative, into pCB101. Though Tn1 and Tn3 do differ in their nucleotide sequences, the Tn1 sequence data obtained in this work shows that there is $^{82}/_{95}$ (=86.3%) homology between Tn1 and Tn3 at the

Fig. 5.3.

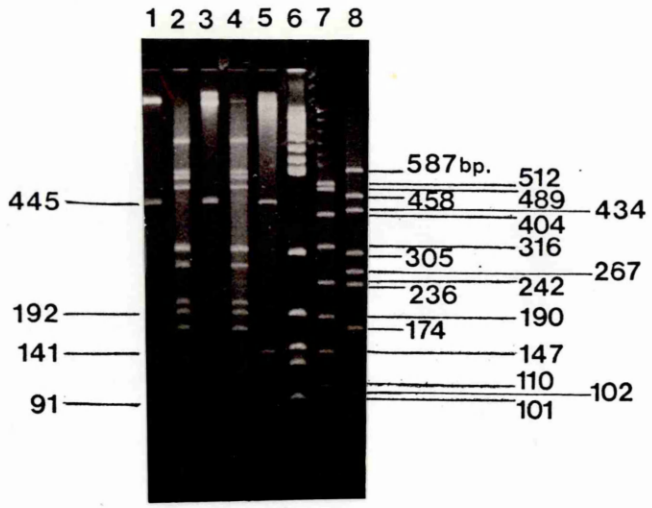


Fig.5.3. Restriction analysis of plasmids pSNO33 and pSNO34.

5% polyacrylamide gel showing the TaqI and PvuII restriction profiles of pSNO33 and pSNO34, the TaqI restriction profile of RSF1050 and the PvuII restriction profile of pACYC184.

It can be seen that TaqI restriction of pSNO33 and pSNO34 yielded the 192bp. fragment derived from RSF1050 (and many other, pACYC184-specific, fragments). Also, PvuII restriction yielded a 91bp. fragment from pSNO33 and a 141bp. fragment from pSNO34, indicating these molecules were of the structures shown in Fig.5.2. (PvuII restriction of each of these molecules also generated c.445 and 3216bp. pACYC184-specific fragments).

1: pACYC184, PvuII.

2: pSNO33, TaqI.

3: pSNO33, PvuII.

4: pSNO34, TaqI.

5: pSNO34, PvuII.

6: RSF1050, TaqI.

7: pUC9, HpaII.

8: pUC9, HaeIII.

Fig.5.4.

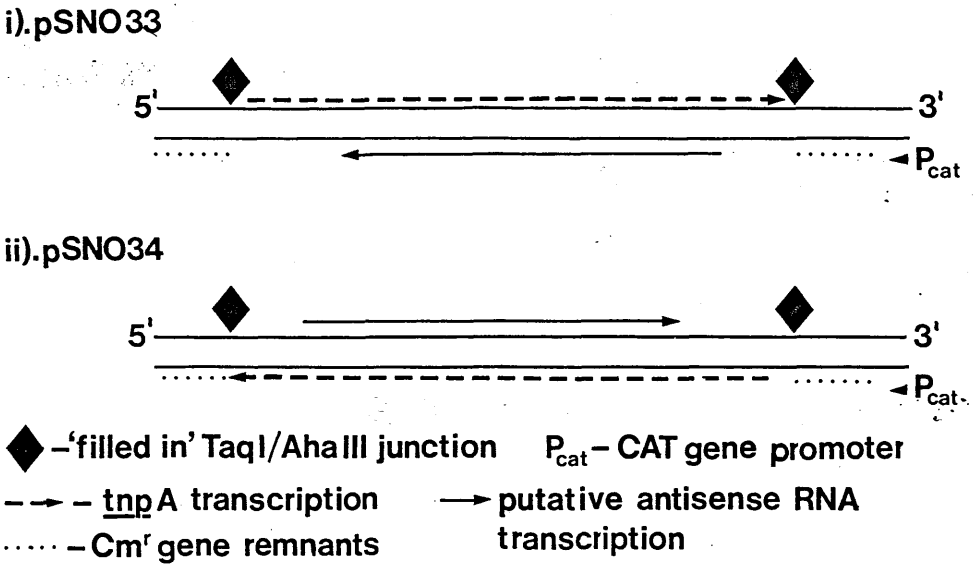


Fig.5.4. The effects of transcription from the CAT gene promoter on the inserted DNA fragment in pSNO33 and pSNO34.

In pSNO33, transcription initiated at the CAT gene promoter will generate a large amount of a transcript that will include the whole putative antisense region of Tn3.

In pSNO34, transcription initiated at this promoter will generate a large amount of a transcript that will contain the complementary sequence to the putative antisense region of Tn3.

nucleotide level in the region of the putative RNA transcript from 2945-3039. Thus, an antisense RNA transcript from Tn1 would be expected to stably hybridise to DNA, or an mRNA from the corresponding region, of Tn3, and any effect on transposition frequency due to the action of antisense RNA would be detected in this assay system.

Parallel assays were performed at 30°C and 37°C to assess the effect of temperature on any effect due to pSNO33 or pSNO34.

CJ100 was transformed to Cm^r with pCB101 and this strain then used to construct four further strains as follows:-

(i) control systems : a). R388::Tn103 was introduced from DS903, R388::Tn103 by conjugal transfer, exconjugants being selected on Isosensitest agar containing chloramphenicol and trimethoprim.

b). pACYC184 was used to transform CJ100, pCB101 to Tc^r and then R388::Tn103 was introduced as above, exconjugants being selected on Isosensitest agar containing chloramphenicol, tetracycline and trimethoprim.

(ii) test systems : pSNO33 or pSNO34 was used to transform CJ100, pCB101 to Tc^r and then R388::Tn103 was introduced as above, exconjugants again being selected on Isosensitest agar containing chloramphenicol, tetracycline and trimethoprim. (pSNO33 or pSNO34 were introduced before R388::Tn103 so that any effect they caused would be operative throughout the whole period the transposition proficient molecule was in the cell).

Transposition of Tn103 into pCB101 would yield an R388::Tn103::pCB101::Tn103 cointegrate (Tp^rCm^r). The frequency of this transposition event in each of the above systems was therefore determined by calculating the ratio of $Str^rTp^rCm^r$: Str^rTp^r exconjugants after mating of each of the strains with DS902.

The results of these assays are shown in Table 5.1 and the conclusions that can be drawn from them are as follows:-

(i). the remnant of IS1 in pACYC184 appears to have no effect on the transposition frequency of Tn103 into pCB101 -ie. the IS1 sequence does not seem to titrate out transposase to any significant extent. This

Table 5.1: The effect of the presence of pSNO33 or pSNO34 on the transposition frequency of Tn103.

Cross.	Transposition frequency of Tn103 into pCB101 at: 30°C.	37°C.
CJ100, pCB101, R388::Tn103 x DS902.	2.58×10^{-2}.	1.21×10^{-4}.
CJ100, pCB101, R388::Tn103, pACYC184 x DS902.	2.55×10^{-2}.	2.43×10^{-4}.
CJ100, pCB101, R388::Tn103, pSNO33 x DS902.	2.45×10^{-2}.	2.23×10^{-4}.
CJ100, pCB101, R388::Tn103, pSNO34 x DS902.	2.61×10^{-2}.	2.67×10^{-4}.

Table 5.1 legend:

(i). in each case, at least 100 colonies, all containing cointegrates, were examined to establish the transposition frequencies shown.

(ii). the presence of cointegrates in $Str^rTp^rCm^r$ exconjugants was confirmed by single colony gel analysis and by their resolution in the presence of a *tnpR*$^+$ plasmid (pDS4153) to pCB101::Tn103 and R388::Tn103. (Data not shown).

(iii). in the CJ100, pCB101, R388::Tn103, pACYC184 system, selection for $Str^rTp^rCm^r$ exconjugants detected both R388::Tn103::pCB101::Tn103 and R388::Tn103::pACYC184::Tn103 cointegrates. (Also detected were any 'double' cointegrates: eg.R388::Tn103::pACYC184::Tn103::pCB101::Tn103). To determine specifically the number of R388::Tn103::pCB101::Tn103 cointegrates, aliquots of the mating mixes were also plated out on Isosensitest agar containing streptomycin, trimethoprim, chloramphenicol and tetracycline. This supported only the growth of cells containing R388::Tn103::pACYC184::Tn103 cointegrates (and any cells containing 'double' cointegrates, but double events are so rare as to be statistically insignificant:- c. $^1/_{100}$-$^1/_{1000}$x the frequency of 'single' events. See Tables 4.2 and 4.3). Therefore, by subtracting the number of colonies on the Str, Tp, Cm, Tc-containing plates from the number on the Str, Tp, Cm-containing plates, the transposition frequency of Tn103 into pCB101 in this system was specifically determined.

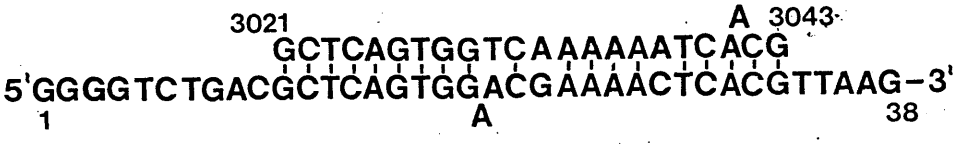
finding is in keeping with other data obtained during this work where this IS1 sequence was never found to act as a substrate for the action of transposase in a one-ended transposition event (See transposition assay data in Tables 4.2 & 4.3).

(ii). the cloned TaqI fragment of Tn3 appears not to produce any molecule(s) that decreases the transposition frequency of Tn103 into pCB101 when acting in trans. As the CAT promoter is strong, the results in the pSNO34 system are inconclusive, since it would be argued that any antisense transcript produced from this plasmid might be titrated out by the transcript generated from the CAT promoter. The result from the pSNO33 system, however, provides evidence against the activity of the above antisense RNA species for, even when the region held to be responsible for its production is under the control of a strong promoter, no decrease in the transposition frequency of Tn103 is observed. This suggests that there is no stage in the production of the transposase protein at which this antisense RNA molecule could act. This could be because the tnpA mRNA is very rapidly occupied by ribosomes once its transcription has begun or, if this does not occur, it is rapidly degraded from its 5' end.

(iii). the area of homology to the inverted repeat sequence of Tn3 that lies just within the amino terminus of the transposase gene and is carried on the cloned fragment (suggested to be possibly the internal resolution site by Heffron et al, 1979; see Fig. 5.5) does not act to bind and therefore titrate out transposase. In turn, it seems logical to conclude that this sequence will not act as a substrate for the action of transposase, and that this protein needs more than this amount of sequence information to recognise and act on its target sequence.

These results, however, still leave open several possibilities about any possible antisense RNA regulation of transposase gene expression. For example, it could be that an antisense RNA molecule is transcribed from a start site 5' to the cloned region; or, the cloned region does not include all the regulatory elements necessary for production of the

Fig.5.5.



| – homologous nucleotides

Fig.5.5. The area of partial homology near the N-terminus of the tnpA gene to the 38bp. perfect terminal inverted repeat of Tn3.

After Heffron et al, 1979.

There are $^{19}/_{23}$ perfect matches between these two regions over an area extending from base pairs 11-33 within the inverted repeat.

antisense RNA (there have recently been several reports of sequences upstream of the -35 promoter region being needed for optimal gene expression -eg. in E.coli tyrT promoter - Lamond and Travers, 1983; Lamond and Travers,1985); or, any molecule produced by the antisense region is only active in cis; or such a molecule actively acts to increase the transposition frequency of Tn103 but this effect is not seen because of some other, overall, ceiling that is imposed on transposition frequency.

To investigate further these possibilities and to see if any in vivo evidence of activity of a promoter that could generate an antisense RNA transcript could be produced, experiments were undertaken to clone appropriate fragments of Tn3 into promoter probe vectors.

5.3. Promoter probe vector experiments.

5.3.1. Background information.

These experiments were undertaken using two vectors that rely on the same principle for the detection of in vivo prokaryotic promoter activity. Both contain the galactokinase (galK) gene cloned separately from its own promoter, lying downstream of a number of unique restriction sites. Cells of a galactokinase-deficient strain containing either vector alone produce white colonies on Mac.Conkey's agar containing galactose (Mac.gal agar) for the galactokinase gene cannot be expressed in these vectors and therefore no metabolism of galactose can occur. In contrast, if a fragment of DNA possessing promoter activity is inserted into one of the unique restriction sites in the correct orientation, then transcription and translation of the galactokinase gene can occur, even in a chromosomally galactokinase-deficient strain, and the metabolism of galactose proceeds (assuming the strain also possesses galactose-epimerase (galE) and galacto-transferase (galT) activities) to yield acidic products, with the result that cells

containing such plasmids produce bright pink colonies on the above medium (Mc.Kenney et al, 1981).

Plasmid pKO500 is a derivative of pKO-1 (Mc.Kenney et al, 1981) which contains the multiple cloning site from the 'phage vector MI3mp11 (Messing and Vieira, 1982) inserted upstream of the promoter-less galactokinase gene (A. Lamond, pers. comm). See Fig.5.6.

Plasmid pKM-1 (Lamond and Travers, 1983) is similar to pKO-1, but as well as having unique AvaI, EcoRI and HindIII restriction sites upstream of the promoterless galactokinase gene, it contains a fragment of DNA from phage lambda carrying the lambda tR1 terminator inserted between the HindIII site and the galactokinase gene. This vector was constructed as it had been found that the cloning of some very strong promoters into vectors such as pKO-1 produced either non-viable plasmids, or plasmids with greatly reduced copy numbers, presumably because the high level of transcription disrupted plasmid replication processes. Such constructs, if viable, often produce white colonies in cells of a galactokinase-deficient strain on Mac.gal medium. (eg. E.coli tyrT promoter : Lamond and Travers, 1983; E.coli leuT promoter : Duester et al, (1982); Adams and Hatfield, 1979; Stueber and Bujard, 1982). The lambda tR1 terminator is rho-dependent and N-suppressible but is only c.50% efficient, even in the absence of N protein (Friedman and Gottesman, 1983). Therefore, c.50% of any transcripts emanating from a cloned fragment will pass through this terminator and on through the galactokinase gene. This arrangement has been empirically found to allow the cloning and detection of strong promoters in this sort of promoter-probe system, for it reduces transcription through the galactokinase gene to a level that still allows identification of cells containing plasmids with strong promoters by the above chromogenic screen whilst also allowing plasmid replication to proceed satisfactorily. See Fig. 5.6.

For each fragment being tested for promoter activity, therefore, the

Fig.5.6.

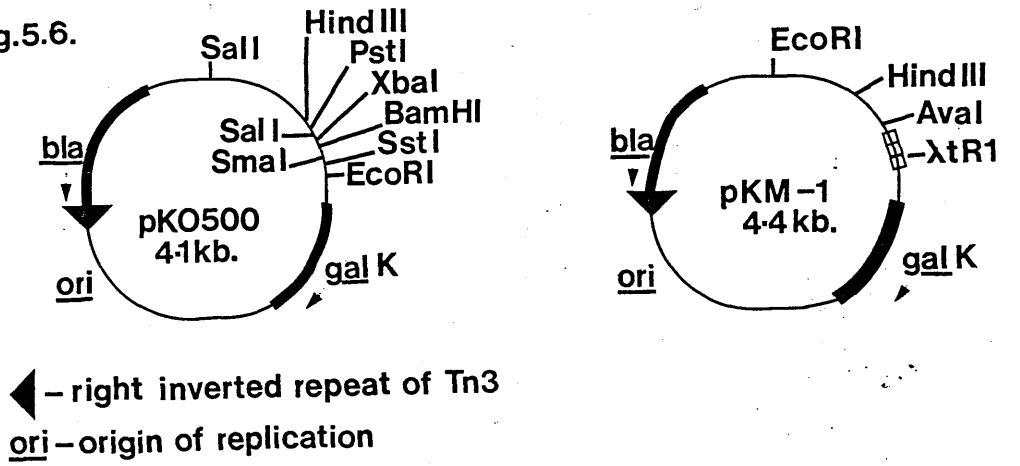


Fig.5.6. The plasmids pKO500 and pKM-1.

Refs.: A. Lamond, pers. comm. and Lamond and Travers, 1983, respectively.

Both plasmids contain a promoterless galactokinase gene to allow them to be used to detect prokaryotic promoter sequences.

In pKO500, the multiple cloning site from the 'phage vector M13mp11 (Messing and Vieira, 1982) lies upstream of the galactokinase gene.

In pKM-1, there are three unique cloning sites (HindIII, EcoRI, AvaI) upstream of the galactokinase gene, but separated from it by the t_{R1} terminator of 'phage lambda.

fragment was cloned into both pKO500 and pKM-1 to check out thoroughly whether or not it possessed any promoter activity. Following the construction of each different test plasmid, galactokinase assays were carried out to quantitate the level of promoter activity.

5.3.2. Construction of plasmids.

It was decided to clone two different fragments from Tn3 into pKO500 and pKM-1. These were;

(i) a 285bp. HaeIII-PvuII fragment from 2688-2973 in the Tn3 sequence. This would contain the intact promoter region (including any important sequences upstream of the -35 region) for the possible antisense RNA transcript described above.

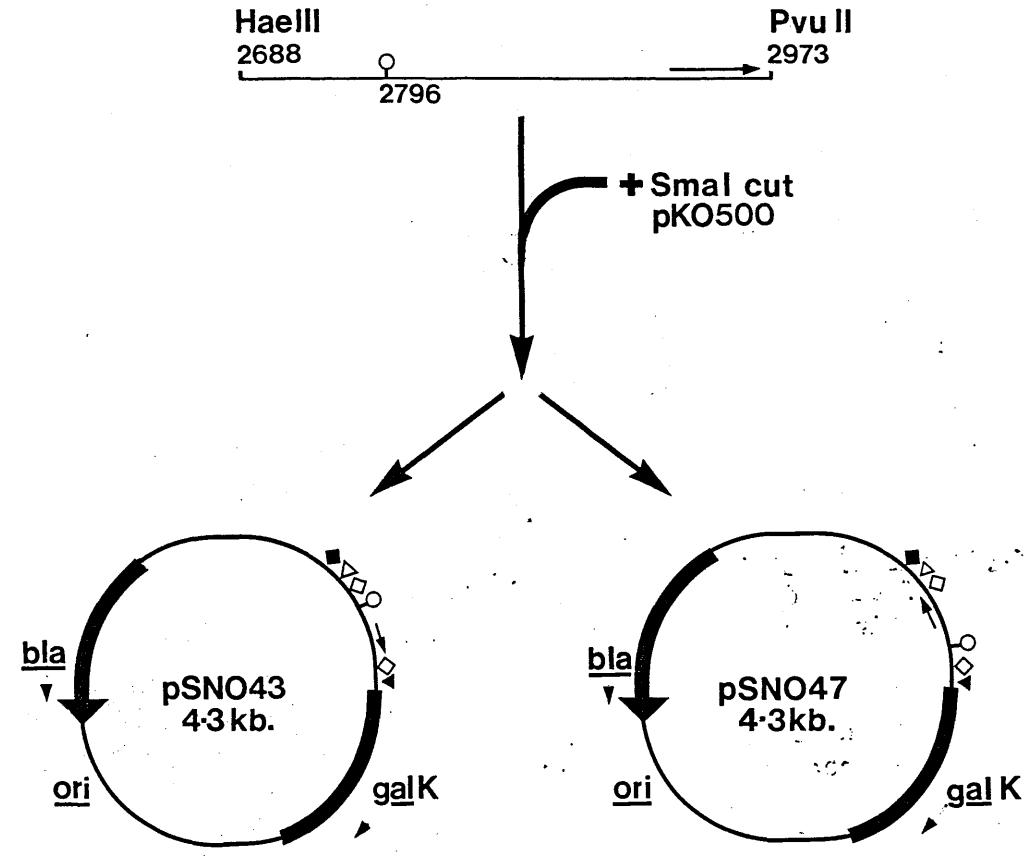
(ii) an 837bp. Sau3A fragment from 2073-2910 in the Tn3 sequence. This would not contain the whole promoter for the possible antisense RNA species described above but would include any other promoters responsible for generating transcripts from this region.

5.3.2(i). construction of plasmids using 285bp. HaeIII-PvuII fragment of Tn3.

See Fig 5.7. RSF1050 plasmid DNA was digested to completion with both HaeIII and PvuII, the resultant fragments separated on a 5% polyacrylamide gel, the 285bp. fragment-containing slice cut out from the gel, the DNA electro-eluted, ethanol-precipitated and resuspended in TE buffer. pKO500 plasmid DNA was linearised with SmaI, phosphatased, ethanol precipitated and resuspended in TE buffer. The two fragments were ligated together and the ligation mixture used to transform DS903 cells to Ap^r. Cells were plated on Mac.gal agar containing ampicillin and streptomycin to screen for promoter activity in the transformants.

160 transformant colonies were obtained, and all of them were white in colour. Single colony gel analysis revealed that 16 of these

Fig.5.7.



◇ – HaeIII/SmaI, PvuII/SmaI junctions
△ – BamHI site ▲ – EcoRI site ♀ – AvaI site ◆ – HindIII site
ori – origin of replication
◀ – right inverted repeat of Tn3
→ – direction of any antisense transcription

Fig.5.7. The construction of plasmids pSNO43 and pSNO47.

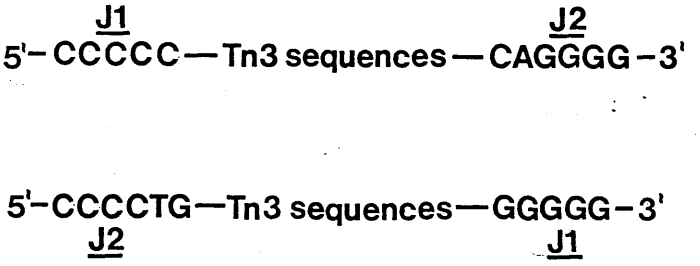
Ligation of the 285bp. HaeIII/PvuII fragment from Tn3 (2688- 2973) to SmaI-linearised pKO500 generated two classes of recombinant molecules. In the first (pSNO43), any transcription from the antisense strand of this region of Tn3 would be directed towards the galactokinase gene. In the second (pSNO47), any such transcription would be directed away from this gene.

contained a plasmid of the size of the desired construct (data not shown) and DNA was prepared from each of these by the alkaline/SDS method for restriction analysis.

It can be seen (Fig. 5.8) that hybrid HaeIII/SmaI and SmaI/PvuII sites are targets for none of these enzymes. Hence, the size of any inserted fragment had to be checked by digesting the DNA with restriction enzymes having sites immediately adjacent to the SmaI site of pKO500: BamHI and EcoRI were the enzymes chosen. The orientation of the inserted fragment was determined by AvaI and BamHI co-digestion. 14 of the 16 plasmid DNAs tested yielded a c.307bp. fragment on BamHI/EcoRI co-digestion, as anticipated for the desired construct, yet only one of these 14 contained the fragment orientated in such a way so that any promoter on its antisense strand would drive transcription of the galactokinase gene. This plasmid was designated pSNO43, whilst one of the other thirteen - with the insert in the opposite orientation - was chosen and designated pSNO47. See Fig. 5.9. This led to speculation that there may be some physiological bias against insertion in the former orientation, possibly as a result of promoter activity within the fragment. All 14 of the plasmids produced a white colony phenotype in DS903 cells. This was as expected for pSNO47, whilst for pSNO43, this finding could be interpreted to mean that there was either no promoter activity in this direction on the cloned fragment or that the promoter there was a strong one.

To clone this same region into pKM-1, pSNO47 plasmid DNA was digested to completion with HindIII and EcoRI, the resultant fragments separated on a 5% polyacrylamide gel, the desired 340bp. fragment containing slice cut out from the gel, the DNA electro-eluted, ethanol precipitated and resuspended in TE buffer. It was then used in ligation with the 4088bp. HindIII-EcoRI fragment of pKM-1, previously isolated from a low melting point agarose gel. DS903 cells were transformed to Ap^r with this ligation mixture and the cells plated out on a Mac.gal agar containing ampicillin and streptomycin. 197 transformant colonies were obtained

Fig.5·8.



J1 - HaeIII/SmaI junctions

J2 - PvuII/SmaI junctions

HaeIII site: GGCC
PvuII site: CAGCTG
SmaI site: CCCGGG

Fig.5.8. HaeIII/SmaI and PvuII/SmaI junction sequences found in pSNO43 and pSNO47 (top and bottom respectively).

It can be seen that none of these sequences are targets for any of the three enzymes.

Fig.5.9a.

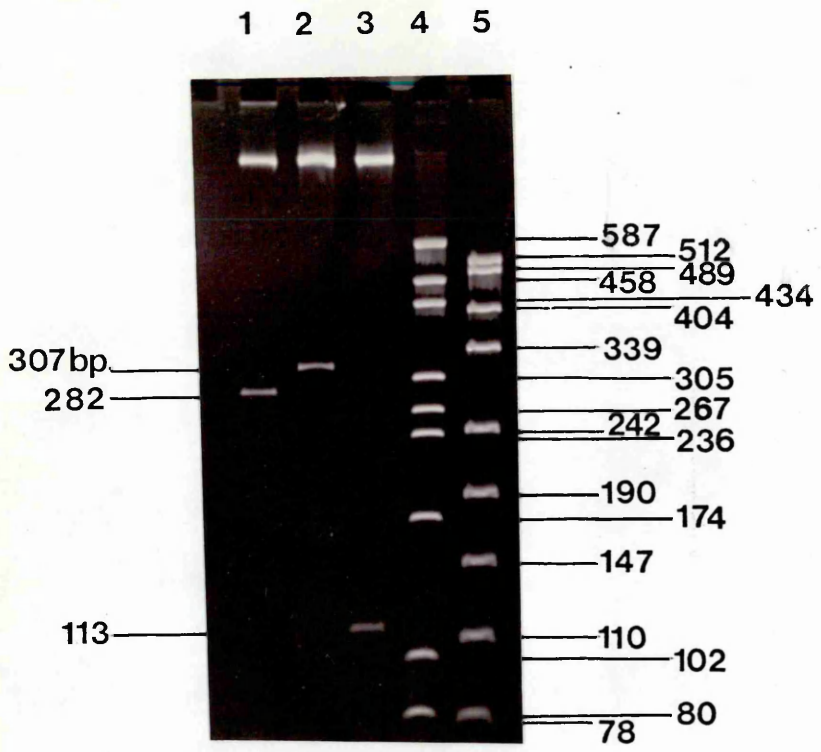


Fig.5.9. Restriction analysis of plasmids pSNO43 and pSNO47.

a). pSNO43: 5% polyacrylamide gel showing BamHI/EcoRI and BamHI/AvaI restriction profiles of pSNO43.

It can be seen that BamHI/EcoRI co-digestion of this plasmid generated a 307bp. fragment (and a c.4.03kb. one) and that BamHI/AvaI co-digestion generated a 113bp. fragment (and a c.4.23kb. one). This showed that pSNO43 contained the 285bp. HaeIII/PvuII fragment from Tn3 inserted in such a way that any transcription emanating from the antisense strand of this region of Tn3 would be directed towards the galactokinase gene.

1: pJKR, EcoRI.

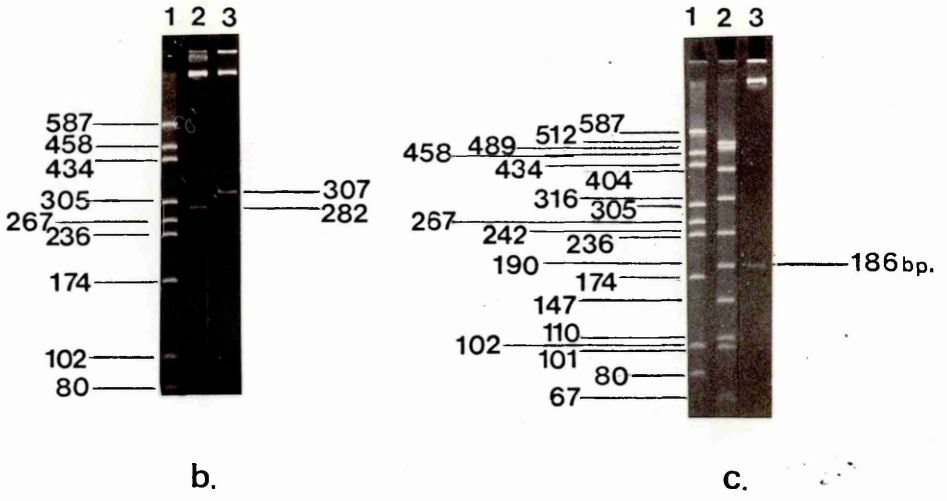
2: pSNO43, BamHI/EcoRI.

3: pSNO43, BamHI/AvaI.

4: pUC8, HaeIII.

5: pUC8, HpaII.

Fig.5.9b&c.



b). & c). pSNO47: 5% polyacrylamide gels showing the BamHI/EcoRI and BamHI/AvaI restriction profiles of pSNO47 respectively.

It can be seen that BamHI/EcoRI co-digestion of pSNO47 generated a 307bp. fragment (and a c.4.03kb. one) and BamHI/AvaI co-digestion generated a 186bp. fragment (and a c.4.16kb. one). This indicated that pSNO47 contained the 285bp. HaeIII/PvuII fragment from Tn3 inserted in the opposite orientation to pSNO43, so that any transcription emanating from the antisense strand of this region of Tn3 would be directed away from the galactokinase gene.

b).

1: pUC9, HaeIII.

2: pJKR, EcoRI.

3: pSNO47, BamHI/EcoRI.

c).

1: pUC9, HaeIII.

2: pUC9, HpaII.

3: pSNO47, BamHI/AvaI.

and all were white in colour. 42 of these were examined by single colony gel analysis and all but one contained a plasmid of the desired size (data not shown). HindIII/EcoRI co-digestion and AvaI digestion on DNA prepared by the alkaline/SDS method from five of these clones showed that they all contained the desired fragment inserted in the desired orientation. This plasmid was designated pSNO48. See Figs. 5.10, 11. It therefore seemed most probable that this fragment of Tn3 contained no promoter activity in the antisense direction.

5.3.2(ii). construction of plasmids using 837bp. Sau3A fragment of Tn3.

This fragment had previously been cloned from Tn3 into both BamH1-digested pUC8 and M13mp18 RF DNA, in both possible orientations in each case, by another worker in this laboratory and it was two of these constructs that were used as the source of this fragment for these constructions (G. Russell, pers. comm.). This was done to allow the Tn3 insert to be isolated on a restriction fragment with two different "sticky ends" which, in turn, would allow the orientation of insertion into pKO500 and pKM-1 to be pre-determined by an appropriate cloning strategy. See Fig. 5.12.

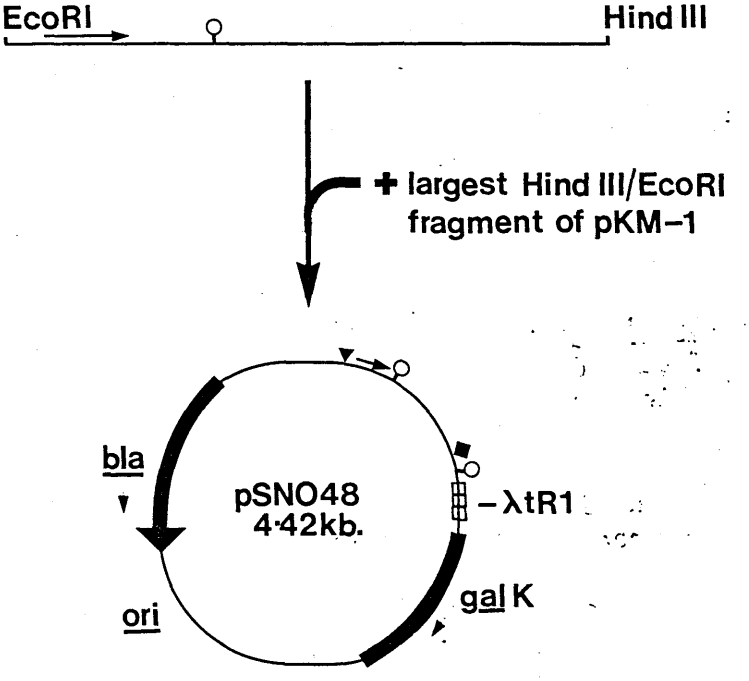
M13mpGRO41 RF DNA, pGRO6, pKO500 and pKM-1 plasmid DNAs were each restricted to completion with HindIII and EcoRI and the resultant fragments separated on a 1% low melting point agarose gel. The 880bp fragment from M13mpGRO41, the 859bp fragment from pGRO6, the 4088bp fragment from pKM-1 and the largest fragment from pKO500 were all purified from low melting point agarose, ethanol precipitated, resuspended in TE and used in the following ligations:

pGRO6 and pKO500 fragments.

M13mpGRO41 and pKM-1 fragments.

These ligation mixtures were used to transform DS903 cells to Ap^r, the cells again being plated out on Mac.gal. agar containing ampicillin and streptomycin.

Fig.5.10.



▲ – EcoRI site ◆ – Hind III site ♀ – AvaI site
ori – origin of replication
◀ – right inverted repeat of Tn3
→ – direction of any antisense transcription

Fig.5.10. The construction of pSNO48.

The 340bp. HindIII/EcoRI fragment from pSNO47 (which included all the Tn3 sequences cloned into pSNO47 in its construction) was ligated to the largest (4.1kb.) HindIII/EcoRI fragment of pKM-1 to generate pSNO48.

In this plasmid, any transcription emanating from the antisense strand of this region of Tn3 would be directed towards the galactokinase gene, but would have to pass through the t_{R1} terminator of 'phage lambda before reaching this gene.

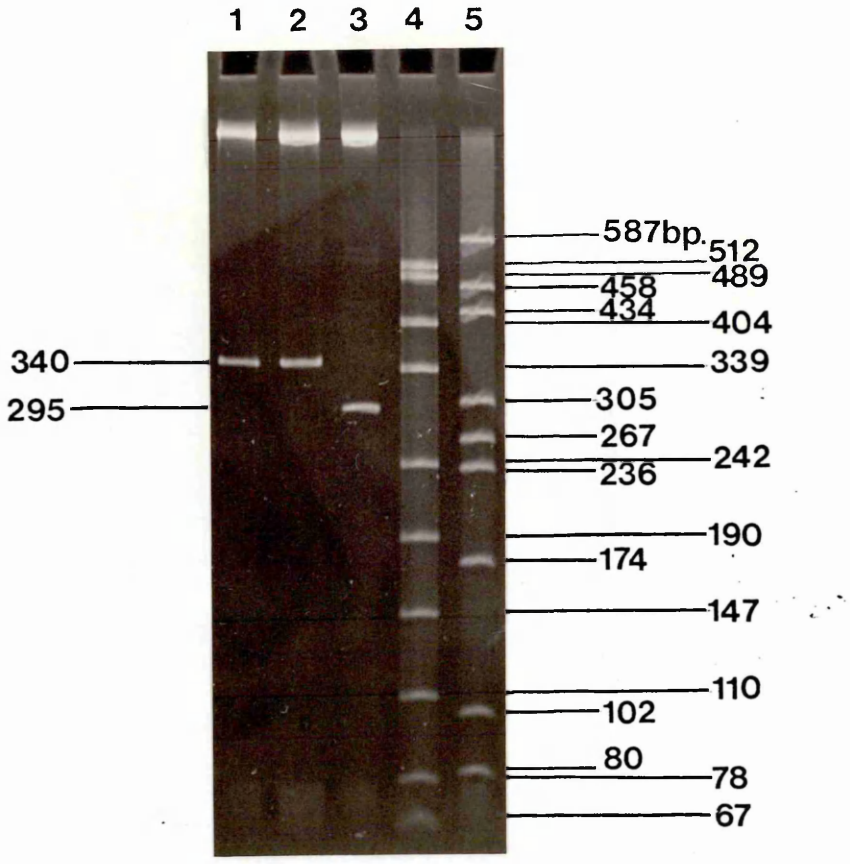


Fig.5.11. Restriction analysis of pSNO48.

5% polyacrylamide gel showing the HindIII/EcoRI restriction profiles of pSNO47 and pSNO48 and the AvaI restriction profile of pSNO48.

It can be seen that HindIII/EcoRI co-digestion of pSNO48 generated the 340bp. fragment derived from pSNO47 and that AvaI restriction generated a 295bp. fragment (and a c.4.13kb. one). This indicated that the fragment from pSNO47 was orientated in such a way that any transcription emanating from the antisense strand of the region of Tn3 included in it would be directed towards the galactokinase gene.

1: pSNO47, HindIII/EcoRI.

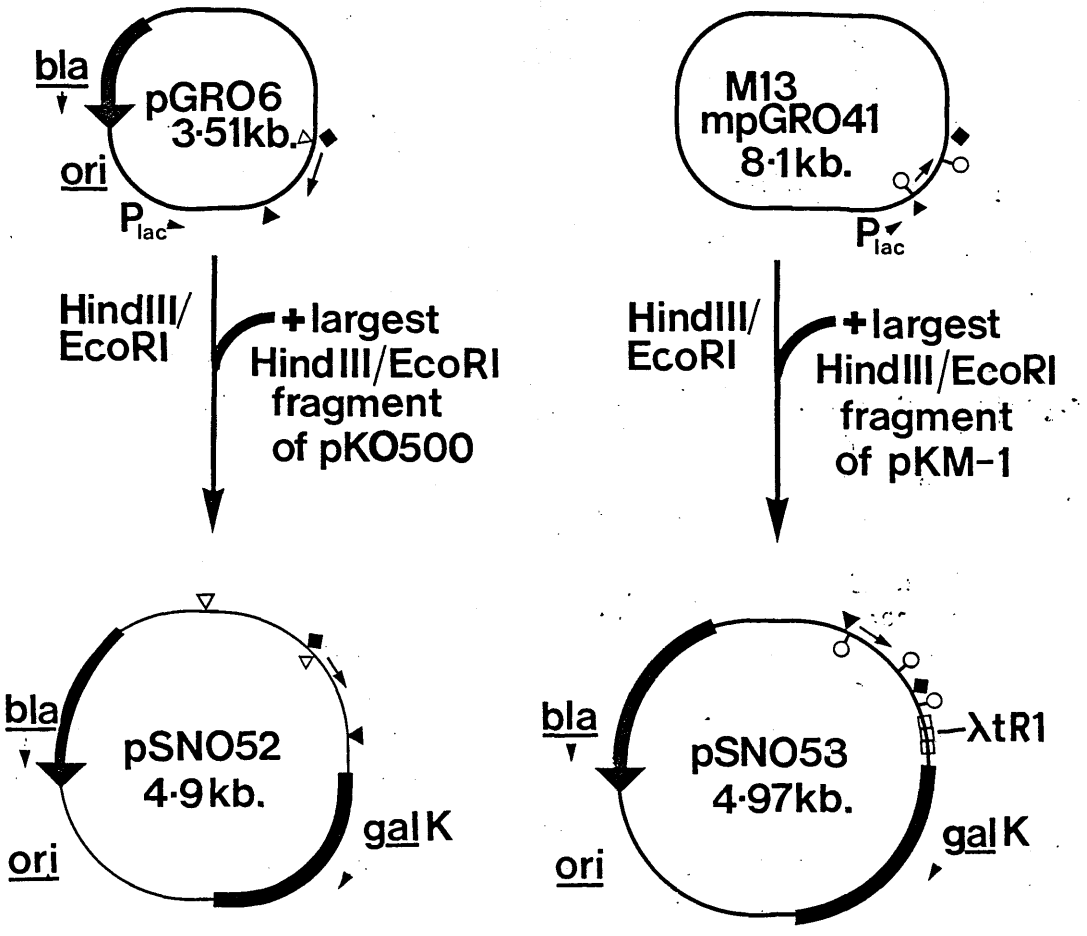
2: pSNO48, HindIII/EcoRI.

3: pSNO48, AvaI.

4: pUC8, HpaII.

5: pUC8, HaeIII.

Fig.5.12.



♀-AvaI site ◆-HindIII site ▲-EcoRI site △-SalI site
ori-origin of replication P_{lac}-lac promoter
◀-right inverted repeat of Tn3
→-direction of any antisense transcription

Fig.5.12. The construction of pSNO52 and pSNO53.

a). pSNO52: the 859bp. HindIII/EcoRI fragment from pGRO6 (which includes those Tn3 sequences between co-ordinates 2073 and 2910) was ligated to the largest HindIII/EcoRI fragment of pKO500 to generate pSNO52.

In this plasmid, any transcription emanating from the antisense strand of this region of Tn3 would be directed towards the galactokinase gene.

b). pSNO53: the 880bp. HindIII/EcoRI fragment from M13mpGRO41 (which includes the same Tn3 sequences as are present in pGRO6, but inserted in the opposite orientation) was ligated to the largest HindIII/EcoRI fragment of pKM-1 to generate pSNO53.

In this plasmid, any transcription emanating from the antisense strand of this region of Tn3 would be directed towards the galactokinase gene, but would have to pass through the t_{R1} terminator of 'phage lambda before reaching this gene.

361 transformant colonies were obtained with the pGRO6/pKO500 ligation mix; all were white in colour. 19 out of 22 of these checked by single colony gel analysis contained a plasmid of the expected size of the desired construct and restriction analysis of plasmid DNA, made by the alkaline/SDS method, from six of these clones by EcoRI/HindIII co-digestion and SalI digestion showed that all of them contained the required insert in the correct orientation. This plasmid was designated pSNO52. See Fig. 5.13.

Only 12 transformant colonies were obtained with the M13 mpGRO41/pKM-1 ligation mix, and again all were white in colour. All were found, by single colony gel analysis, to contain a plasmid of the size expected of the desired construct, and restriction analysis by EcoRI/HindIII co-digestion and AvaI digestion of plasmid DNAs prepared by the alkaline/SDS method from six of these clones showed that all contained the desired insert in the correct orientation. This plasmid was designated pSNO53. See Fig. 5.14.

It therefore seemed most probable that this fragment of Tn3 also possessed no promoter activity in the antisense direction.

5.3.3. Galactokinase assays of constructed plasmids.

Strains containing each of the above plasmids, and appropriate control strains, were constructed by transformation and galactokinase assays carried out on each of them, as described in Section 2.21. The results of these assays are presented in Table 5.2.

The main conclusion to be drawn from these data is that there is no significant *in vivo* promoter activity in the 'antisense' direction of the *tnpA* coding region of Tn3 between the co-ordinates of 2073 and 2973. Nor is there any such activity in the 'sense' direction between the co-ordinates of 2688 and 2973. Thus, no 'antisense' transcription or translation products are produced from this region and, in particular,

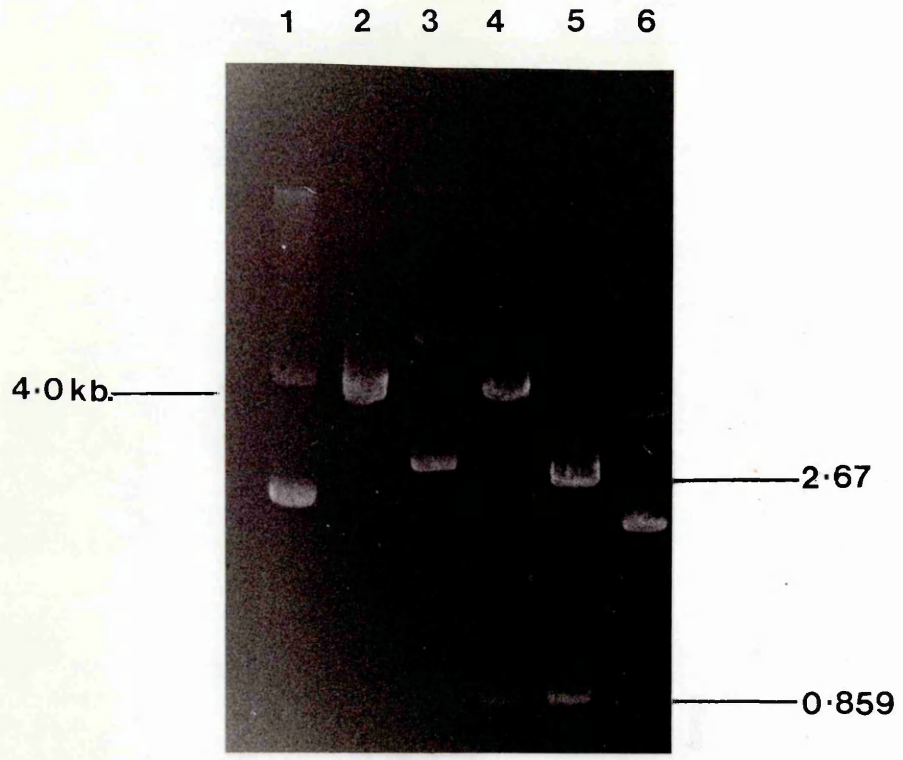


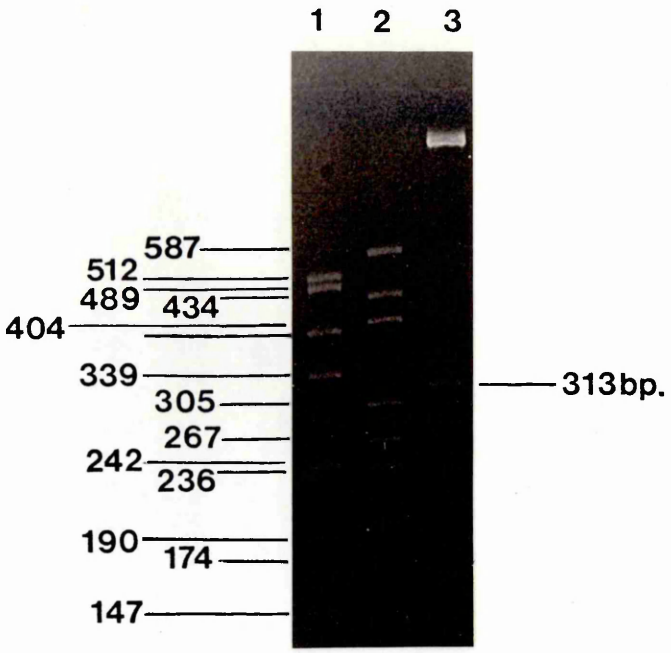
Fig.5.13. Restriction analysis of plasmid pSNO52.

a). 0.8% agarose gel showing the HindIII/EcoRI restriction profile of pSNO52.

b). 5% polyacrylamide gel showing the SalI restriction profile of pSNO52.

It can be seen that HindIII/EcoRI co-digestion of pSNO52 generated the 859bp. fragment derived from pGRO6 and the c.4.0kb. fragment derived from pKO500. Also, SalI restriction of pSNO52 generated a 313bp. fragment (and a c.4.55kb. one). This indicated that the fragment from pGRO6 was orientated in such a way that any transcription emanating from the antisense strand of the region of Tn3 included on this fragment would be directed towards the galactokinase gene.

It was appreciated that the SalI site on the inserted fragment lay in the pUC8 multiple cloning site and not in the Tn3 sequences. However, the orientation of the insert in pGRO6 had previously been clearly established (data not shown) and hence SalI restriction was a satisfactory means by which to establish the orientation of the inserted fragment in pSNO52.



a).

1: pKO500.

2: pKO500, HindIII/EcoRI.

3: pSNO52.

4: pSNO52, HindIII/EcoRI.

5: pGRO6, HindIII/EcoRI.

6: pGRO6.

b).

1: pUC8, HpaII.

2: pUC8, HaeIII.

3: pSNO52, SalI.

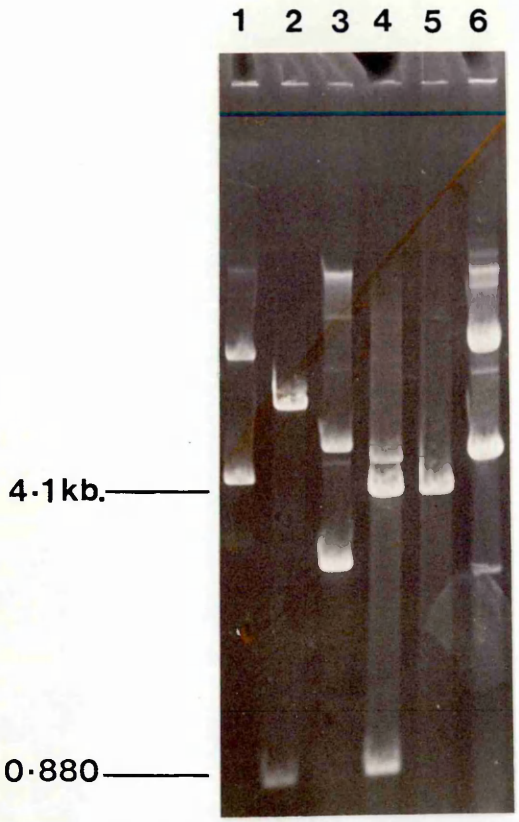


Fig.5.14. Restriction analysis of plasmid pSNO53.

a). 0.8% agarose gel showing the HindIII/EcoRI restriction profile of pSNO53.

b). 5% polyacrylamide gel showing the AvaI restriction profile of pSNO53.

It can be seen that HindIII/EcoRI co-digestion of pSNO53 generated the 880bp. fragment derived from M13mpGRO41 and the c.4.1kb. fragment derived from pKM-1. Also, AvaI restriction of pSNO53 generated 232 and 724bp. fragments (and a c.4.0kb. one).

This indicated that pSNO53 contained the fragment from M13mpGRO41 orientated in such a way that any transcription emanating from the antisense strand of the region of Tn3 included in this fragment would be directed towards the galactokinase gene, but would have to pass through the t_{R1} terminator of 'phage lambda before reaching this gene.

a).

1: M13mpGRO41.

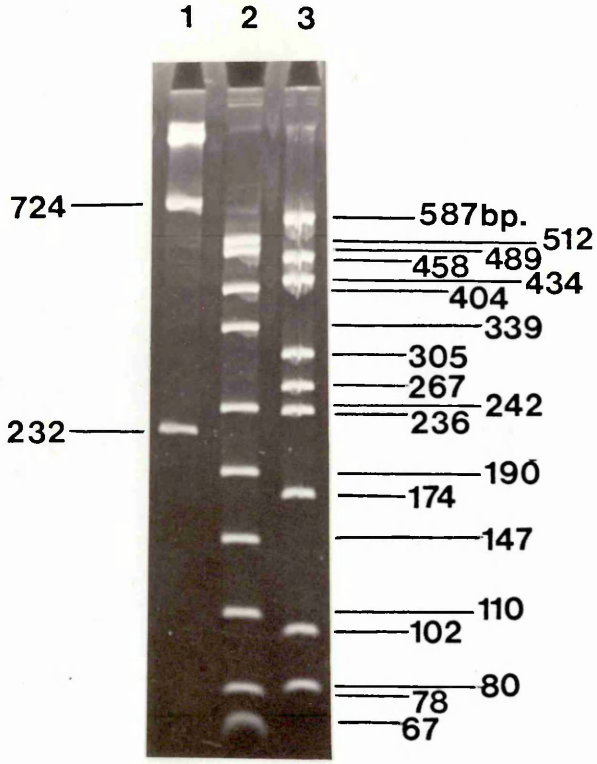
2: M13mpGRO41, HindIII/EcoRI.

3: pSNO53.

4: pSNO53, HindIII/EcoRI.

5: pKM-1, HindIII/EcoRI.

6: pKM-1.



b).

1: pSNO53, AvaI.

2: pUC8, HpaII.

3: pUC8, HaeIII.

Table 5.2: Galactokinase assay data on plasmids pSNO43, pSNO47, pSNO48, pSNO52 and pSNO53.

Strain under test.	Units of galactokinase activity.
DS903.	0.
DS903 + pKO-1.	5.82.
DS903 + pKO500.	7.95.
DS903 + pKL500.	237.82.
DS903 + pKM-1.	1.19.
DS903 + pSNO43.	3.48.
DS903 + pSNO47.	5.74.
DS903 + pSNO48.	1.72.
DS903 + pSNO52.	1.67.
DS903 + pSNO53.	0.725.

Table 5.2 legend:

(i). pKL500 is a plasmid identical to pKO500 except that it has the lac promoter upstream of the multiple cloning site and galactokinase gene. It can therefore be used as a terminator probe vector, but its use in these assays was as a positive control system.

neither the putative antisense RNA transcript, nor its possible translation product, described above are produced *in vivo*. These data are entirely consistent with the results obtained previously in assaying the effects of pSN033 and pSN034 *in trans* on the transposition frequency of Tn103.

5.4. Discussion.

Overall, then, the experiments described here reveal that the 192bp *Taq*1 fragment, extending from 2902 - 3094 in Tn3, has no effect *in trans* on the transposition frequency of Tn103; that there is no *in vivo* promoter activity in the antisense direction of the *tnp*A coding region between co-ordinates 2073 and 2973 and that therefore the above model of antisense regulation of *tnp*A gene expression is incorrect. Further confirmatory data from this conclusion has come from preliminary analysis of total RNA samples prepared from cells containing Tn3 using Northern blotting techniques and strand specific probes. No antisense RNA transcripts from the region of the amino-terminus of the *tnp*A gene were detectable in such experiments (G. Russell, pers. comm.)

It therefore seems most probable that some other mechanism must operate to regulate the expression of the *tnp*A gene. Plausible possibilities that remain to be explored include premature termination of the *tnp* mRNA transcript, with or without an attenuation system (eg. as seen in many metabolic operons : Lee and Yanofsky, 1977; Keller and Calvo, 1979), and/or specific or non-specific processing of this message. Any mechanism must clearly include an explanation of the variation in this effect seen with temperature. In an attenuation system, this could be explained by invoking temperature variability in a mechanism for overcoming the attenuating element(s), whilst a mechanism involving RNA processing could include a processing enzyme with different levels of activity at 30°C and 37°C.

Examination of the sequence of the tnpA mRNA produced data of great complexity with no really obvious areas of potential secondary structure. A region of dyad symmetry that could possibly give rise to a hairpin loop structure centered on base pair 2875 was located, and it was followed by a short 'T-rich' nucleotide region. This led to speculation that this region could act as a terminator of transcription. However, experiments using transposase-B galactosidase fusion protein constructs have provided no evidence at all for in vivo terminator activity in this region (M.Burke, pers.comm.).

Thus, the additional mechanism(s) of regulation of tnpA gene expression remains obscure and open to further investigation, and it could be that a host factor (or factors) that may be involved in transposition are involved in this regulation.

Chapter 6.

Concluding remarks.

In studies to investigate the regulation of the tnpA gene of Tn1/3 and the action of its product - the transposase protein - it has been shown that the only sequences essential for the first step of Tn1/3 transposition at the ends of the element are the 38bp. perfect inverted repeats and that both of these are equally good substrates for transposition. As mentioned in the discussion of Chapter 3, these results are at variance with those of Heritage et al, (1984), obtained using plasmids containing three ends of Tn802. Analogous experiments to theirs with Tn1/3 derivatives are soon to be carried out in this laboratory and should produce more information on the relative importance of the two ends of Tn1/3 in transposition. Additionally it was shown that Tn1/3 transposition normally acts to replicate the DNA sequences corresponding to those in the position of the transposon. Further indirect evidence for the binding of Tn1 transposase to Tn1/3 terminal inverted repeat sequences was also obtained.

Data obtained using a new plasmid that produced a B-galactosidase/ transposase fusion protein, together with data obtained using pMB9::Tn103 and R388::Tn103 and other observations made in this laboratory over a period of time led to the formulation of two hypotheses concerning the regulation of Tn1/3 transposition. One of these concerned the mechanism of the variation in the transposition frequency of these elements with temperature and the other a new postulated process that seemed to keep transposition frequency at a maximum level, despite further increases in the number of copies of the tnpA gene. Some possible mechanisms for this latter process were considered and one of these - involving antisense RNA regulation of tnpA gene expression - was investigated in detail. However, no evidence for this model could be produced.

The average transposition frequency of the derepressed transposon Tn3651 mediated in trans by transposase protein from pMB9::Tn103 obtained in this work was 8.8×10^{-2} at 30°C and 3.95×10^{-4} at 37°C.

The average transposition frequency of Tn103 (also

a derepressed transposon) from R388::Tn103 was 4.0×10^{-2} at 30°C and 2.5×10^{-4} at 37°C per cell per generation. The presence of a functional tnpR gene in a cell is known to cause a decrease in transposition frequency by a factor of c.5-100 (see, for example, Heffron, 1983; Heffron et al, 1977, 1978). Bearing in mind that the wild-type situation of Tn1/3 is in a tnpR$^+$ cell most often at 37°C its transposition frequency in the wild would be expected to be c.$10^{-6}/10^{-7}$. Thus it would seem likely that the transposition of these elements makes a significant contribution to the level of spontaneous mutation seen in bacterial cells. Also it seems plausible that the various mechanisms that act to regulate the transposition frequency of Tn1/3 (e.g. the repression of tnpA gene transcription by resolvase; transposition immunity; the temperature-dependent variation in transposition frequency; low frequency of transposition to the chromosome) have evolved to keep this frequency at such a level that the host cell is not selectively disadvantaged by the accumulation of many mutations caused by transposon insertion or by recombination events between transposon copies following multiple insertions.

Following on from this work, it seems most appropriate to investigate further the new postulated mechanism of regulation of transposition frequency. Initially in this respect, it would be valuable to determine transposition frequencies obtained with a tnpA$^+$ tnpR$^-$ transposon (preferably the same one in the same sequence environment in all cases) on a variety of other replicons of different copy numbers to provide more data to substantiate the hypothesis. Also, the possible in vivo significance of any additional potential open reading frames or RNA transcripts from the tnpA gene area of Tn1/3 could be investigated to see if they were involved in this process. It should also be possible to determine in more detail the mechanism of variation of transposition frequency of Tn1/3 with temperature. In this respect, studies could be particularly directed towards the structure and properties of the 5' end of the tnpA mRNA and any interactions it may have with other molecules

in the cell (eg. nucleic acid and/or protein).

Also, it is obviously important to continue to try and develop a system to overproduce Tn1 or Tn3 transposase to facilitate the setting-up of a Tn1/3 *in vitro* transposition system. An experiment that may well be valuable in this area would be to produce a transcriptional fusion of the *tnpA* gene under the control of a known strong promoter (e.g. lambda P_L; P_{tac}). However, as the *tnpA* gene promoter seems to be quite strong, other steps too should be taken to this end. For example, modification of the Shine/Dalgarno sequence (eg. using synthetic oligonucleotides) could be carried out to try and improve the efficiency of the translation initiation process. This has, in fact, been reported for Tn3, with dramatic effects on *tnpA* gene expression (Casadaban *et al*, 1980, 1982; Fennewald *et al*, 1981). Clearly too, a means of circumventing the proposed additional regulatory mechanism of the *tnpA* gene will have to be found if its expression is to be maximised, and other factors also may prove to be important (eg. host factors). Once such a system has been established, more detailed investigations into the mechanism and requirements of the Tn1/3 transposition process could be carried out and the availability of usable quantities of transposase protein would also allow analyses of its possible abilities to bind to and nick DNA at the Tn1/3 inverted repeat sequences to be undertaken.

Ultimately, it is conceivable that an inhibitor of the first step of Tn1/3 transposition could be developed that could be of value clinically in reducing the spread of drug resistance by transposition.

Bibliography.

Bibliography.

Adams, C.W. and Hatfield, G.W. 1984.
Effects of promoter strengths and growth conditions on copy number of transcription fusion vectors.
J. Biol. Chem. Vol.259. pp.7399-7403.

Adhya, S. and Gottesman, M. 1978.
Control of transcription termination.
Ann. Rev. Biochem. Vol.47. pp.967-996.

Adler, H.I., Fisher, W.D., Cohen, A. and Hardigree, A. 1967.
Miniature Escherichia coli cells deficient in DNA.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.57. pp.321-326.

Alton, N.K. and Vapnek, D. 1979.
Nucleotide sequence analysis of the chloramphenicol resistance transposon Tn9.
Nature. Vol.282. pp.864-869.

Amersham International plc.
M13 Cloning and Sequencing Handbook.
Amersham International plc, Amersham, U.K.

Arthur, A., Nimmo, E., Hettle, S. and Sherratt, D. 1984.
Transposition and transposition immunity of transposon Tn3 derivatives having different ends.
E.M.B.O. Journal. Vol.3. pp.1723-1729.

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

Avila, P., de la Cruz, F., Ward, E. and Grinsted, J. 1984.
Plasmids containing one inverted repeat of Tn21 can fuse with other plasmids in the presence of Tn21 transposase.
Mol. Gen. Genet. Vol.195. pp.288-293.

Bachman, B.J., Low, K.B. and Taylor, A.L. 1976.
Re-calibrated linkage map of Escherichia coli K-12.
Bacteriol. Rev. Vol.40. pp.116-167.

Baltimore, D. 1985.
Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome.
Cell. Vol.40. pp.481-482.

Bennett, P.M., de la Cruz, F. and Grinsted, J. 1983.
Cointegrates are not obligatory intermediates in the transposition of Tn3 and Tn21.
Nature. Vol.305. pp.743-744.

Berg, D.E., Johnsrud, L., Mc.Divitt, L., Ramabhadran, R. and Hirschel, B.J. 1982.
The inverted repeats of Tn5 are transposable elements.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.79. pp.2632-2635.

Berg, D.E., Weiss, A. and Crossland, L. 1980.
The polarity of Tn5 insertion mutations in Escherichia coli.
J. Bacteriol. Vol.142. pp.439-446.

Besemer, J., Charlier, D., Gortz, G. and Delius, H.J. 1979.
DNA sequence rearrangements at a deletion end-point within the insertion sequence IS2 element.
Hoppe-Seyler's Z. Physiol. Chem. Vol.360. pp.1019 ff.

Biel, S.W. and Berg, D.E. 1984.
Mechanism of IS1 transposition in Escherichia coli: choice between simple insertion and cointegration.
Genetics. Vol.108. pp.319-330.

Birnboim, H.C. and Doly, J. 1979.
A rapid alkaline extraction procedure for screening recombinant DNA.
Nucl. Acids Res. Vol.7. pp.1513-1523.

Bishop, R.P. 1983.
Intra-molecular transposition of Tn1.
Ph.D. thesis. University of Glasgow.

Bishop, R. and Sherratt, D. 1984.
Transposon Tn1 intra-molecular transposition.
Mol. Gen. Genet. Vol.196. pp.117-122.

Boeke, J.D., Garfinkel, D.J., Styles, C.A. and Fink, G.R. 1985.
Ty elements transpose through an RNA intermediate.
Cell. Vol.40. pp.491-500.

Bolivar, F., Rodriguez, R., Greene, P.J., Betlach, M., Heynecker, H.L., Boyer, H.W., Crosa, J. and Falkow, S. 1977.
Construction and characterisation of new cloning vehicles: II: a multipurpose cloning system.
Gene. Vol.2. pp.95-113.

Botstein, D. and Kleckner, N. 1977.
Translocation and illegitimate recombination by the tetracycline-resistance element Tn10.
In:'DNA Insertion Elements, Plasmids and Episomes.'
Eds.: Bukhari, A.I., Shapiro, J.A. and Adhya, S.
Cold Spring Harbor Press, Cold Spring Harbor, New York.
pp.185-204.

Bradley, D.E., Taylor, D.E. and Cohen, D.R. 1980.
Specification of surface mating systems among conjugative drug resistance plasmids in Escherichia coli K-12.
J. Bacteriol. Vol.143. pp.1466-1470.

Bregliano, J-C. and Kidwell, M.G. 1983.
Hybrid dysgenesis determinants.
In:'Mobile Genetic Elements.'
Ed.: Shapiro, J.
Academic Press, New York.
pp.363-410.

Bruckner, R. and Matzura, H. 1985.
Regulation of the inducible chloramphenicol acetyl transferase gene of the Staphyloccus aureus plasmid pUB112.
E.M.B.O. Journal. Vol.9. pp.2295-2300.

Brunton, J., Bennett, P. and Grinsted, J. 1981.
Molecular nature of a plasmid specifying beta-lactamase production in Haemophilus ducreyi.
J. Bacteriol. Vol.148. pp.788-795.

Bukhari, A.I. and Zipser, D. 1972.
Random insertion of Mu-1 DNA within a single gene.
Nature New Biology. Vol.236. pp.240-243.

Calos, M. and Miller, J. 1980.
Transposable elements.
Cell. Vol.20. pp.579-595.

Campbell, A. 1981.
Evolutionary significance of accessory DNA elements in bacteria.
Ann. Rev. Microbiol. Vol.35. pp.55-83.

Carleton, S., Projan, S.J., Highlander, S.K., Moghazeh, S.M. and Novick, R.P. 1984.
Control of pT181 replication II: mutational analysis.
E.M.B.O. Journal. Vol.3. pp.2407-2414.

Casadaban, M.J., Chou, J. and Cohen, S.N. 1982.
Overproduction of the Tn3 transposition protein and its role in DNA transposition.
Cell. Vol.28. pp.345-354.

Casadaban, M.J., Chou, J., Lemaux, P., Tu, C.D. and Cohen, S.N. 1980.
Tn3 transposition and control.
Cold Spring Harbor Symp. Quant. Biol. Vol.45. pp.269-273.

Cesareni, G. and Banner, D.W. 1985.
Regulation of plasmid copy number by complementary RNAs.
Trends Biochem. Sci. Vol.10. pp.303-306.

Chaconas, G., Harshey, R.M., Sarvetnick, M. and Bukhari, A.I. 1981.
Mechanism of bacteriophage Mu transposition.
Cold Spring Harbor Symp. Quant. Biol. Vol.45. pp.311-318.

(Chiang)
Chaing, S.J. and Clowes, R.C. 1980.
Intramolecular transposition and inversion in plasmid R6K.
J. Bacteriol. Vol.142. pp.668-682.

Chamberlin, M.J. 1974.
The selectivity of transcription.
Ann. Rev. Biochem. Vol. 43. pp721-775.

Chan, P.T., Ohmori, H., Tomizawa, J. and Lebowitz, J. 1985.
Nucleotide sequence and gene organisation of ColE1 DNA.
J. Biol. Chem. Vol.260. pp.8925-8935.

Chandler, M., Clerget, M. and Galas, D.J. 1982.
The transposition frequency of IS1-flanked transposons is a function of their size.
J. Mol. Biol. Vol.154. pp.229-243.

Chang, A.C.Y., and Cohen, S.N. 1978.
Construction and characterisation of amplifiable, multi-copy DNA cloning vehicles derived from the P15A cryptic miniplasmid.
J. Bacteriol. Vol.134. pp.1141-1156.

Chao, L., Vargas, V., Spear, B.B. and Cox, E.C. 1983.
Transposable elements as mutator genes in evolution.
Nature. Vol. 303. pp.633-635.

Charlier, D., Piette, J. and Glansdorff, N. 1982.
IS3 can function as a mobile promoter in Escherichia coli.
Nucl. Acids Res. Vol.10. pp.5935-5948.

Chia, W., Mc.Gill, S., Karp, R., Gubb, D. and Ashburner, M. 1985.
Spontaneous excision of a large composite transposable element of Drosophila melanogaster.
Nature. Vol.316. pp.81-83.

Chou, J., Casadaban, M.J., Lemaux, P.G. and Cohen, S.N. 1979a.
Identification and characterisation of a self-regulated repressor of translocation of the Tn3 element.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.76. pp.4020-4024.

Chou, J., Lemaux, P.G., Casadaban, M.J. and Cohen, S.N. 1979b.
Transposition protein of Tn3: identification and characterisation of an essential repressor-controlled gene product.
Nature. Vol.282. pp.801-806.

Ciampi, M.S., Schmid, M.B. and Roth, J.R. 1982.
Transposon Tn10 provides a promoter for transcription of adjacent sequences.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.79. pp.5016-5020.

Cohen, S.N. 1976.
Transposable genetic elements and plasmid evolution.
Nature. Vol.263. pp.731-738.

Coleman, J., Green, P.J. and Inouye, M. 1984.
The use of RNAs complementary to specific mRNAs to regulate the expression of individual bacterial genes.
Cell. Vol.37. pp.429-436.

Coleman, J., Hirashima, A., Inokuchi, Y., Green, P.J. and Inouye, M. 1985.
A novel immune system against bacteriophage infection using complementary RNA (micRNA).
Nature. Vol.315. pp.601-603.

Cornelius, J. 1980.
Transposition of Tn951 (Tn*lac*) and cointegrate formation are thermosensitive processes.
J. Gen. Microbiol. Vol.117. pp.243-247.

Craigie, R. and Mizuuchi, K. 1985.
Mechanism of transposition of bacteriophage Mu: structure of a transposition intermediate.
Cell. Vol.41. pp.867-876.

Craigie, R., Mizuuchi, M. and Mizuuchi, K. 1984.
Site specific recognition of the bacteriophage Mu ends by the Mu A protein.
Cell. Vol.39. pp.387-394.

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

Dahl, H-H.M. 1983.
Constructing the optimal promoter.
Trends Biochem. Sci. Vol.8. p.230.

Dalrymple, B. and Arber, W. 1985.
Promotion of RNA transcription on the element IS30 of Escherichia coli K-12.
E.M.B.O. Journal. Vol.4. pp.2687-2693.

Daniell, E., Roberts, R. and Abelson, J. 1971.
Mutations in the lactose operon caused by bacteriophage Mu.
J. Mol. Biol. Vol.69. pp.1-8.

Datta, N. and Hedges, R.W. 1972.
Trimethoprim resistance conferred by W plasmids in Enterobacteriacae.
J. Gen. Microbiol. Vol.72. pp.349-355.

Davie, E., Sydnor, K. and Rothfield, L.I. 1984.
Genetic basis of minicell formation in Escherichia coli K-12.
J. Bacteriol. Vol.158. pp. 1202-1203.

de Boer, H.A., Comstock, L.J.and Vasser, M. 1983.
The tac promoter: a functional hybrid derived from the trp and lac promoters.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.80. pp. 21-25.

de Graaf, J., Elwell, L.P. and Falkow, S. 1976.
Molecular nature of two beta-lactamase specifying plasmids isolated from Haemophilus influenzae type b.
J. Bacteriol. Vol.126. pp.439-446.

Demerec, M., Adelberg, E.A., Clark, A,J. and Hartman, P.E. 1966.
A proposal for a uniform nomenclature in bacterial genetics.
Genetics. Vol.54. pp.61-76.

Doolittle, W.F. and Sapienza, C. 1980.
Selfish genes, the phenotype paradigm and genome evolution.
Nature. Vol.284. pp.601-603.

Doring, H-P. 1985.
Plant transposable elements.
Bioessays. Vol.3. pp.164-166.

Dougan, G., Saul, M., Twigg, A., Gill, R. and Sherratt, D. 1979.
Polypeptides expressed in Escherichia coli K-12 minicells by the transposition elements Tn1 and Tn3.
J. Bacteriol. Vol.138. pp.48-54.

Duester, G., Elford, R.M. and Holmes, W.M. 1982.
Fusion of the Escherichia coli $tRNA^{leu}_1$ promoter to the gal K gene: analysis of sequences necessary for growth rate-dependent regulation.
Cell. Vol.30. pp.855-864.

Dugaiczyk, A., Boyer, H.W. and Goodman, H.M. 1975.
Ligation of EcoR1 endonuclease generated DNA fragments into linear and circular structures.
J. Mol. Biol. Vol.96. pp.171-184.

Dyson, P.J. 1984.
The mechanism of site-specific recombination encoded by Tn3.
Ph.D. thesis, University of Glasgow.

Eibel, H. and Philippsen, P. 1984.
Preferential integration of yeast transposable element Ty into a promoter region.
Nature. Vol.307. pp.386-388.

Emori, Y., Shiba, T., Kanaya, S., Inouye, S., Yuki, S. and Saigo, K. 1985.
The nucleotide sequences of copia and copia-related RNA in Drosophila virus-like particles.
Nature. Vol.315. pp.773-776.

Errede, B., Cardillo, T.S., Sherman, F., Dubois, E., Deschamps, J. and Wiame, J.M. 1980.

Mating signals control expression of mutations resulting from insertion of a transposable repetitive element adjacent to diverse yeast genes.

Cell. Vol.22. pp.427-436.

Fayet, O., Froment, Y. and Piffaretti, J.C. 1982.

Beta-lactamase specifying plasmids isolated from Neisseria gonorrhoeae have retained an intact right part of a Tn3-like transposon.

J. Bacteriol. Vol.149. pp.136-144.

Fedoroff, N. 1983.

Controlling elements in maize.

In: Mobile Genetic Elements.

Ed.: Shapiro, J.A.

Academic Press, New York.

pp. 1-63.

Fennewald, M.A., Gerrard, S.P., Chou, J., Casadaban, M.J. and Cozzarelli, N. 1981.

Purification of the Tn3 transposase and analysis of its binding to DNA.

J. Biol. Chem. Vol.256. pp.4687-4690.

Finnegan, D.J. 1982.

The 'fold-back' elements of Drosophila.

Nature. Vol.297. pp.181-182.

Foster, T.J., Davis, M., Roberts, D., Takeshita, K. and Kleckner, N. 1981.
Genetic organisation of the transposon Tn10.
Cell. Vol.23. pp.201-213.

Frazer, A.C. and Curtiss, R. 1975.
Production, properties and utility of bacterial minicells.
Curr. Top. Immunol. Microbiol. Vol.69. pp.1-84.

Friedman, D.I. and Gottesman, M. 1983.
Lytic mode of lambda development.
In: Lambda II.
Cold Spring Harbor Laboratory,
Cold Spring Harbor, New York.
pp. 21-51.

Galas, D., Calos, M.P. and Miller, J.H. 1980.
Sequence analysis of Tn9 insertions in the lacZ gene.
J. Mol. Biol. Vol.144. pp.19-41.

Galas, D. and Chandler, M. 1981.
On the molecular mechanisms of transposition.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.78. pp.4858-4862.

Gamas, P., Galas, D. and Chandler, M. 1985.
DNA sequence at the end of IS1 required for transposition.
Nature. Vol.317. pp.458-460.

Garfinkel, D.J., Boeke, J.D. and Fink, G.R. 1985.
Ty element transposition: reverse transcriptase and virus-like particles.
Cell. Vol.42. pp.507-517.

Gerasimova, T.I., Mizrokhi, L.J. and Georgiev, G.P. 1984.
Transposition bursts in genetically unstable Drosophila melanogaster.
Nature. Vol.309. pp.714-716.

Gill, R., Heffron, F., Dougan, G. and Falkow, S. 1978.
Analysis of sequences transposed by complementation of two classes of transposition-deficient mutants of Tn3.
J. Bacteriol. Vol.136. pp.742-756.

Gill, R.E., Heffron, F. and Falkow, S. 1979.
Identification of the protein encoded by the transposable element Tn3 which is required for its transposition.
Nature. Vol.282. pp.797-801.

Glansdorff, N., Charlier, D. and Zafarullah, M. 1981.
Activation of gene expression by IS2 and IS3.
Cold Spring Harbor Symp. Quant. Biol. Vol.45. pp.153-156.

Glass, R.E. 1982.
Gene function: E.coli and its heritable elements.
Croom Helm, London.

Gold,L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B., and Stormo, G. 1981.
Translation initiation in prokaryotes.
Ann. Rev. Microbiol. Vol.35. pp.365-403.

Grigoryan, M.S., Kramerov, D.A., Tulchinsky, E.M., Revasova, E.S. and Lukanidin, E.M. 1985.
Activation of putative transposition intermediate formation in tumour cells.
E.M.B.O. Journal. Vol.4 pp.2209-2215.

Grindley, N.D.F. and Joyce, C.M. 1980.
Genetic and DNA sequence analysis of the kanamycin resistance transposon Tn903.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.77. pp.7176-7180.

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

Grindley, N.D.F. and Reed, R.R. 1985.
Transpositional recombination in prokaryotes.
Ann. Rev. Biochem. Vol.54. pp.863-896.

Grindley, N.D.F. and Sherratt. D.J. 1979.
Sequence analysis at IS1 insertion sites: models for transposition.
Cold Spring Harbor Symp. Quant. Biol. Vol.43. pp.1257-1261.

Guyer, M.S. 1978.
The gamma-delta sequence of F is an insertion sequence.
J. Mol. Biol. Vol.126. pp.347-365.

Habermann, P., Klaer, R., Kuhn, S. and Starlinger, P. 1979.
IS4 is found between 11 or 12 base pair duplications.
Mol. Gen. Genet. Vol.175. pp.369-373.

Haggerty, D.M.and Scheif, R.F. 1976.
Location in bacteriophage lambda DNA of cleavage sites of the site-specific endonuclease from Bacillus amyloliquefaciens H.
J. Virol. Vol.18. pp.659-663.

Hall, M.N., Gabay, J., Debarbouille, M. and Schwartz, M. 1982.
A role for mRNA secondary structure in the control of translation initiation.
Nature. Vol.295. pp.616-618.

Halling, M.S. and Kleckner, N. 1982.
A symmetrical 6bp. target site sequence determines Tn10 insertion specificity.
Cell. Vol.28. pp.155-163.

Halling, S., Simons, R., Way, J., Walsh, R. and Kleckner, N. 1982.
DNA sequence organisation of Tn10's IS10-right and comparison with IS10-left.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.79. pp.2608-2612.

Harayama, S., Oguchi, T. and Iino, T. 1984a.
The Escherichia coli K-12 chromosome flanked by two IS10 sequences transposes.
Mol. Gen. Genet. Vol.197. pp.62-66.

Harayama, S., Oguchi, T. and Iino, T. 1984b.
Does Tn10 transpose via the cointegrate molecule?
Mol. Gen. Genet. Vol.194. pp.444-450.

Harshey, R.M. and Bukhari, A. 1981.
A mechanism of DNA transposition.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.78. pp.1090-1094.

Hartl, D.L., Dykhuizen, D.E., Miller, R.D., Green, L. and de Framond, J. 1983.
Transposable element IS50 improves growth rate of Escherichia coli cells without transposition.
Cell. Vol.35. pp.503-510.

Hawley, D.K. and Mc.Clure, W.R. 1983.
Compilation and analysis of Escherichia coli promoter DNA sequences.
Nucl. Acids Res. Vol.11. pp.2237-2255.

Hedges, R.W. and Jacob, A.E. 1974.
Transposition of ampicillin resistance from RP4 to other replicons.
Mol. Gen. Genet. Vol.132. pp.31-40.

Heffron, F. 1983.
Tn3 and its relatives.
In: Mobile Genetic Elements.
Ed.: Shapiro, J.A.
Academic Press, New York.
pp.223-260.

Heffron, F., Bedinger, P., Champoux, J.J. and Falkow, S. 1977.
Deletions affecting the transposition of an antibiotic resistance gene.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.74. pp.702-706.

Heffron, F., Mc.Carthy, B.J., Ohtsubo, H. and Ohtsubo, E. 1979.
DNA sequence analysis of the transposon Tn3: three genes and three sites involved in transposition of Tn3.
Cell. Vol.18. pp.1153-1163.

Heffron, F., Rubens, C. and Falkow, S. 1975a.
The translocation of a DNA sequence which mediates ampicillin resistance: molecular nature and specificity of insertion.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.72. pp.3623-3627.

Heffron, F., So, M. and Mc.Carthy, B.J. 1978.
In vitro mutagenesis of a circular DNA molecule by using synthetic restriction sites.
Proc. Nat. Acad. Sci. (U.S.A.). Vol. 75. pp.6012-6016.

Heffron, F., Sublett, R., Hedges, R.W., Jacob, A. and Falkow, S. 1975b.
Origin of the TEM beta-lactamase gene found on plasmids.
J. Bacteriol. Vol.122. pp.250-256.

Helling, R.B., Goodman, H.M. and Boyer, H.W. 1974.
Analysis of EcoRI fragments of DNA from lambdoid bacteriophage and other viruses by agarose gel electrophoresis.
J. Virol. Vol.14. pp.1235-1244.

Heritage, J. and Bennett, P. 1984.
The role of TnA transposase in transposition immunity.
Plasmid. Vol.12. pp.218-221.

Heritage, J. and Bennett, P.M. 1985.
Plasmid fusions mediated by one end of TnA.
J. Gen. Microbiol. Vol.131. pp.1131-1140.

Heritage, J., Perry, A.C.F. and Bennett, P.M. 1984.
Hitherto only reported by Grindley and Reed, 1985.

Higgins, N.P., Moncecchi, D., Manlapaz-Ramos, P. and Olivera, B.M. 1983.
Mu DNA replication *in vitro*.
J. Biol. Chem. Vol.258. pp.4293-4297.

Higgins, N.P. and Olivera, B.M. 1984.
Mu DNA replication *in vitro*: criteria for initiation.
Mol. Gen. Genet. Vol. 194. pp. 60-64.

Hull, R.A., Gill, G.S. and Curtis, R. 1978.
Genetic characterisation of Mu-like bacteriophage D108.
J. Virol. Vol.27. pp.513-518.

Hyde, D.R. and Tu, C.P.D. 1985.
*tnp*M: a novel regulatory gene that enhances Tn21 transposition and suppresses cointegrate resolution.
Cell. Vol.42. pp.629-638.

Iida, S., Meyer, J. and Arber, W. 1983.
Prokaryotic IS elements.
In: Mobile Genetic Elements.
Ed.: Shapiro, J.A.
Academic Press, New York.
pp.159-221.

Ilyina, T.S., Nechaeva, E.V., Romanov, Y.A. and Smirnov, G.B. 1981.
Isolation and mapping of Escherichia coli K-12 mutants defective in Tn9 transposition.
Mol. Gen. Genet. Vol.181. pp.384-389.

Isberg, R.R., Lazaar, A.L. and Syvanen, M. 1982.
Regulation of Tn5 by the right-repeat proteins: control at the level of the transposition reaction?
Cell. Vol.30. pp.883-892.

Isberg, R.R. and Syvanen, M. 1981.
Replicon fusions promoted by the inverted repeats of Tn5.
J. Mol. Biol. Vol.150. pp.15-32.

Isberg, R.R. and Syvanen, M. 1982.
DNA gyrase is a host factor required for transposition of Tn5.
Cell. Vol.30. pp.9-18.

Ishizaki, K. and Ohtsubo,E. 1984.
Functional difference in the two ends of insertion sequence IS1 in transposition and cointegration.
Cold Spring Harbor Symp. Quant. Biol. Vol.49. pp.207-213.

Izant, J.G. and Weintraub, H. 1984.
Inhibition of thymidine kinase gene expression by antisense RNA: a molecular approach to genetic analysis.
Cell. Vol.36. pp.1007-1015.

Jones, C.T. 1985.

Factors affecting the stability of E.coli plasmid vectors.

Ph. D. thesis. University of Glasgow.

Johnson, R.C., Yin, J.C.P. and Reznikoff, W.S. 1982.
Control of Tn5 transposition in Escherichia coli is mediated by protein from the right repeat.
Cell. Vol.30. pp.873-882.

Jordan, E., Saedler, H. and Starlinger, P. 1968.
O^c and strong polar mutations in the gal operon are insertions.
Mol. Gen. Genet. Vol.102. pp.353-363.

Jorgensen, R.A. and Reznikoff, W.S. 1979.
Organisation of structural and regulatory genes that mediate tetracycline resistance in transposon Tn10.
J. Bacteriol. Vol.138. pp.705-714.

Kahmann, R. and Kamp, D. 1979.
The nucleotide sequences of the attachment sites of bacteriophage Mu.
Nature. Vol.280. pp.247-250.

Kahn, S.A. and Novick, R.P. 1980.
Terminal nucleotide sequences of Tn551, a transposon specifying erythromycin resistance in Staphyloccus aureus: homology with Tn3.
Plasmid. Vol.4. pp.148-154.

Keller, E.B. and Calvo, J.M. 1979.
Alternative secondary structures of leader RNAs and the regulation of the trp, phe, his, thr and leu operons.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.76. pp.6186-6190.

Kelly, P.J. 1983.
The fine control of Tn3.
Honours Project Report.
Institute of Genetics, University of Glasgow.

Kim, S.K. and Wold, B.J. 1985.
Stable reduction of thymidine kinase activity in cells expressing high levels of antisense RNA.
Cell. Vol.42. pp.129-138.

Kitts, P.A. 1982.
The transposition mechanism of Tn3.
Ph.D. thesis. University of Glasgow.

Kitts, P.A., Lamond, A. and Sherratt, D.J. 1982a.
Inter-replicon transposition of Tn1/3 occurs in two sequential, genetically separable steps.
Nature. Vol.295. pp.626-628.

Kitts, P.A., Symington, L., Burke, M., Reed, R. and Sherratt, D.J. 1982b.
Transposon-encoded site-specific recombination.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.79. pp.46-50.

Kleckner, N. 1979.
DNA sequence analysis of Tn10 insertions: origin and role of 9bp. flanking repetitions during Tn10 translocation.
Cell. Vol.16. pp.711-720.

Kleckner, N. 1981.
Transposable elements in prokaryotes.
Ann. Rev. Genet. Vol.15. pp.341-404.

Kleckner, N., Chan R.K., Tye, B.K. and Botstein, D. 1975.
Mutagenesis by insertion of a drug resistance element carrying an inverted repetition.
J. Mol. Biol. Vol.97. pp.561-575.

Kleckner, N., Morisato, D., Roberts, D. and Bender, B. 1984.
Mechanism and regulation of Tn10 transposition.
Cold Spring Harbor Symp. Quant. Biol. Vol.49. pp.235-244.

Kleckner, N., Reichardt, K. and Botstein, D. 1979.
Inversions and deletions of the Salmonella chromosome generated by the translocatable tetracycline resistance element Tn10.
J. Mol. Biol. Vol.127. pp.89-115.

Kopecko, D.J. and Cohen, S.N. 1975.
Site specific recA-independent recombination between bacterial plasmids: involvement of palindromes at the recombinational loci.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.72. pp.1373-1377.

Kretschmer, P.J. and Cohen, S.N. 1977.
Selected translocation of plasmid genes: frequency and regional specifity of translocation of the Tn3 element.
J. Bacteriol. Vol.130. pp.888-899.

Kretschmer P.J. and Cohen, S.N. 1979.
Effect of temperature on translocation frequency of the Tn3 element.
J. Bacteriol. Vol.139. pp.515-519.

Kumar, C.C. and Novick, R.P. 1985.
Plasmid pT181 replication is regulated by two counter transcripts.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.82. pp.638-642.

Lacatena, R.M., Banner, D.W., Castagnoli, L. and Cesareni, G. 1984.
Control of initiation of pMB1 replication: purified Rop protein and RNAI affect primer formation in vitro.
Cell. Vol.37. pp.1009-1014.

Lamond. A.I. and Travers, A.A. 1983.
Requirement for an upstream element for optimal transcription of a bacterial tRNA gene.
Nature. Vol.305. pp.248-250.

Lamond A.I. and Travers, A.A. 1985.
Stringent control of bacterial transcription.
Cell. Vol.41. pp.6-8.

Laporte, D.C. 1984.
Antisense RNA: a new mechanism for the control of gene expression.
Trends Biochem. Sci. Vol.9. p463.

Lee, C-H., Bhagwat, A. and Heffron, F. 1983.
Identification of a transposon Tn3 sequence required for transposition immunity.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.80. pp.6765-6769.

Lee, C-H., Twu, J.S. and Schloemer, R.H. 1985.
Transposition frequency of Tn3 is not affected by the size of the transposon.
Abstract H137 in 'Abstracts of the Annual Meeting, 1985, of the American Society for Microbiology.'

Lee, F. and Yanofsky, C. 1977.
Transcription termination at the trp operon attenuators of Escherichia coli and Salmonella typhimurium: RNA secondary structure and regulation of termination.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.74. pp.4365-4369.

Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L. and Steitz, J.A. 1980.
Are snRNPs involved in splicing?
Nature. Vol.283. pp.220-224.

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. Vol.184. pp.364-371.

Light, J. and Molin S. 1983.
Post-transcriptional control of expression of the repA gene of plasmid R1 mediated by a small RNA molecule.
E.M.B.O. Journal. Vol.2. pp. 93-98.

Lowe, J.B. and Berg, D.E. 1983.
A product of the Tn5 transposase gene inhibits transposition.
Genetics. Vol.103. pp.605-615.

Lueders, K.K. and Kuff, E.L. 1977.
Sequences associated with intra-cisternal A particles are reiterated in the mouse genome.
Cell. Vol.12. pp.963-972.

Mac.Hattie, L.A. and Jackowski, J.B. 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, J.A. and Adhya, S.
Cold Spring Harbor Laboratory,
Cold Spring Harbor, New York.
pp.219-228.

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

Machida, C., Machida, Y., Wang, H.C., Ishizaki, K. and Ohtsubo, E. 1983.
Repression of cointegrate ability of insertion element IS1 by transcriptional readthrough from flanking regions.
Cell. Vol.34. pp.135-142.

Machida, Y., Machida, C. and Ohtsubo, E. 1982a.
A novel type of transposon generated by insertion element IS102 present in a pSC101 derivative.
Cell. Vol.30. pp.29-36.

Machida, Y., Machida, C. and Ohtsubo, E. 1984.
Insertion element IS1 encodes two structural genes required for its transposition.
J. Mol. Biol. Vol.177. pp.229-245.

Machida, Y., Machida, C., Ohtsubo, H. and Ohtsubo, E. 1982b.
Factors determining frequency of plasmid cointegration mediated by the insertion sequence IS1.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.79. pp.277-281.

Magasnik, B. 1970.
Glucose effects: inducer exclusion and repression.
In: The Lactose Operon.
Eds.: Beckwith, J.R. and Zipser, D.
Cold Spring Harbor Laboratory,
Cold Spring Harbor, New York.
pp.189-219.

Malamy, M.H. 1966.
Frameshift mutations in the lactose operon of Escherichia coli.
Cold Spring Harbor Symp. Quant. Biol. Vol.31. pp.189-203.

Maniatis, T., Fritsch, E.F. and Sambrook, J. 1982.
Molecular Cloning - A Laboratory Manual.
Cold Spring Harbor Laboratory,
Cold Spring Harbor, New York.

Maxam, A. and Gilbert, W. 1980.
Sequencing end-labelled DNA with base specific chemical cleavages.
Methods Enzymol. Vol.65. pp.499-560.

Mc.Clintock, B. 1950.
The origin and behaviour of mutable loci in maize.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.36. pp.344-355.

Mc.Cormick, M., Wishart, W., Ohtsubo, H., Heffron, F. and Ohtsubo, E. 1981.
Plasmid cointegrates and their resolution mediated by transposon Tn3 mutants.
Gene. Vol.15. pp. 103-118.

Mc.Kenney, K., Shimatake, H., Court, D., Schmeissner, U., Brady, C. and Rosenberg, M. 1981.
A system to study promoter and terminator signals recognised by Escherichia coli RNA polymerase.
In: Gene Amplification and Analysis, Vol.II.
Eds.: Chirikjian, J.G. and Papas, T.S.
Elsevier / North Holland, Amsterdam.
pp.383-415.

Mellor, J., Fulton, S.M., Dobson, M.J., Wilson, W., Kingsman, S.M. and Kingsman, A.J. 1985.
A retrovirus-like strategy for expression of a fusion protein encoded by yeast transposon Ty1.
Nature. Vol.313. pp.243-246.

Melton D.A. 1985.
Injected antisense RNAs specifically block messenger RNA translation in vivo.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.82. pp.144-148.

Messing, J. 1983.
New M13 vectors for cloning.
Methods Enzymol. Vol.101. pp.20-78.

Messing, J. and Vieira, J. 1982.
A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments.
Gene. Vol.19. pp.269-276.

Miller, J.H., Calos, M.P., Galas, D., Hofer, M., Buchel, D.E. and Muller-Hill, B. 1980.
Genetic analysis of transposition in the lac region of Escherichia coli.
J. Mol. Biol. Vol.144. pp.1-18.

Mizuno, T., Chou, M-Y. and Inouye, M. 1983.
Regulation of gene expression by a small RNA transcript (mic RNA) in Escherichia coli K-12.
Proc. Japan. Acad. Vol.59. pp.335-338.

Mizuno, T., Chou, M-Y. and Inouye, M. 1984.
A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (mic RNA).
Proc. Nat. Acad. Sci. (U.S.A.). Vol.81. pp.1966-1970.

Mizuuchi, K. 1983.
In vitro transposition of bacteriophage Mu: a biochemical approach to a novel replication reaction.
Cell. Vol.35. pp.785-794.

Mizuuchi, K. 1984.
Mechanism of transposition of bacteriophage Mu: polarity of the strand transfer reaction at the initiation of transposition.
Cell. Vol.39. pp.395-404.

Morisato, D. and Kleckner, N. 1984.
Transposase promotes double strand breaks and single strand joints at transposon Tn10 termini *in vivo*.
Cell. Vol.39. pp.181-190.

Morisato, D., Way, J.C., Kim, H.J. and Kleckner, N. 1983.
Tn10 transposase acts preferentially on nearby transposon ends *in vivo*.
Cell. Vol.32. pp.799-807.

Motsch, S. and Schmitt, R. 1984.
Replicon fusion mediated by a single-ended derivative of transposon Tn1721.
Mol. Gen. Genet. Vol.195. pp.281-287.

Motsch, S., Schmitt, R., Avila, P., de la Cruz, F., Ward, E. and Grinsted, J. 1985.
Junction sequences generated by one-ended transposition.
Nucl. Acids Res. Vol.9. pp.3335-3342.

Murphy, E. and Lofdahl, S. 1984.
Transposition of Tn554 does not generate a target duplication.
Nature. Vol.307. pp.292-294.

Muster, C.J. and Shapiro. J.A. 1981.
Recombination involving transposable elements: on replicon fusion.
Cold Spring Harbor Symp. Quant. Biol. Vol.45. pp.239-242.

Novick, R.P., Clowes, R.C., Cohen, S.N., Curtiss, R., Datta, N. and Falkow, S. 1976.
Uniform nomenclature for bacterial plasmids: a proposal.
Bacteriol. Rev. Vol.40. pp.168-189.

Nyman, K., Nakamura, K., Ohtsubo, H. and Ohtsubo, E. 1981.
Distribution of the insertion sequence IS1 in Gram-negative bacteria.
Nature. Vol.289. pp.609-612.

O'Hare, K. 1985.
The mechanism and control of P-element transposition in Drosophila melanogaster.
Trends Genet. Vol.1. pp.250-254.

Ohtsubo, E., Zenilman, M., Ohtsubo, H., Mc.Cormick, M., Machida, C. and Machida, Y. 1981.
Mechanism of insertion and cointegration mediated by IS1 and Tn3.
Cold Spring Harbor Symp. Quant. Biol. Vol.45. pp.283-295.

Ohtsubo, H., Ohmori, H. and Ohtsubo, E. 1979.
Nucleotide sequence analysis of Tn3 (Ap): implications for insertion and deletion.
Cold Spring Harbor Symp. Quant. Biol. Vol.43. pp.1269-1277.

Orgel, L.E. and Crick, F.H.C. 1980.
Selfish DNA: the ultimate parasite.
Nature. Vol.284. pp.604-610.

Oxoid Ltd.
The Oxoid Manual (5th. edition).
Oxoid Ltd., Basingstoke, Hampshire, U.K.

Paterson, B.M., Roberts, B.E. and Kuff, E.L. 1977.
Structural gene identification and mapping by DNA:mRNA hybrid-arrested cell-free translation.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.74. pp.4370-4374.

Pato, M.L. and Reich, C. 1982.
Instability of transposase activity: evidence from bacteriophage Mu replication.
Cell. Vol.29. pp.219-225.

Pato, M.L. and Reich, C. 1984.
Stoichiometric use of the transposase of bacteriophage Mu.
Cell. Vol.36. pp.197-202.

Paulson, K.E., Deka, N., Schmid, C.W., Misra, R., Schindler, C.W., Rush, M.G., Kadyk, L. and Leinwand, L. 1985.
A transposon-like element in human DNA.
Nature. Vol.316. pp.359-361.

Pestka, S., Daugherty, B.L., Jung, V., Hotta, K. and Pestka, R.K. 1984.
Anti-mRNA: specific inhibition of translation of single RNA molecules.
Proc. Nat. Acad. Sci. (U.S.A.). Vol. 81. pp.7525-7528.

Pharmacia P-L Biochemicals. 1984.
Expression vector pKK223-3.
Analects. Vol.12. No.3. pp.1-2.

Philippsen, P., Kramer, R.A. and Davis, R.W. 1978.
Cloning of the yeast ribosomal DNA repeat unit in SstI and HindIII lambda vectors using genetic and physical size selections.
J. Mol. Biol. Vol.123. pp.371-386.

Picken, R.N., Mazaitis, A.J. and Maas, W.K. 1984.
High incidence of transposon Tn3 insertions into a replication control gene of the chimaeric R/Ent plasmid pCG86 of Escherichia coli.
J. Bacteriol. Vol.160. pp.430-433.

Reed, R.R. 1981.
Resolution of cointegrates between transposons gamma-delta and Tn3 defines the resolution site.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.78. pp.3428-3432.

Resibois, A., Toussaint, A., van Gijsegem, F. and Faelen, M. 1981.
Physical characterisation of mini-Mu and mini-D108.
Gene. Vol.14. pp.103-113.

Reynolds, A.E., Felton, J. and Wright, A. 1981.
Insertion of DNA activates the cryptic bgl operon in Escherichia coli K-12.
Nature. Vol.293. pp.625-629.

Roeder, G.S. and Fink, G.R. 1982.
Movement of yeast transposable elements by gene conversion.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.79. pp.5621-5625.

Rogers, J. and Wall, R. 1980.
A mechanism for RNA splicing.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.77. pp.1877-1879.

Rosenberg, U.B., Preiss, A., Seifert, E., Jackle, H. and Knipple, D.C. 1985.
Production of phenocopies by Kruppel antisense RNA injection into Drosophila embryos.
Nature. Vol.313. pp.703-706.

Rosner, J.L. and Guyer, M.S. 1980.
Transposition of IS1-lambda bio-IS1 from a bacteriophage lambda derivative carrying the IS1-cat-IS1 transposon Tn9.
Mol. Gen. Genet. Vol.178. pp.111-120.

Ross, D.G., Swan, J. and Kleckner, N. 1979.
Physical structures of Tn10-promoted deletions and inversions: role of 1400bp. inverted repetitions.
Cell. Vol.16. pp.721-731.

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

Rubin, G.M. 1983.
Dispersed repetitive DNAs in Drosophila.
In: Mobile Genetic Elements.
Ed.: Shapiro, J.A.
Academic Press, New York.
pp.329-361.

Ruther, W., Koenen, M., Otto, K. and Muller-Hill, B. 1981.
pUR222, a vector for cloning and rapid chemical sequencing of DNA.
Nucl. Acids Res. Vol.9. pp.4087-4098.

Saedler,. and Nevers, P. 1985.
Transposition in plants: a molecular model.
E.M.B.O. Journal. Vol.4. pp.585-590.

Saedler., H. and Starlinger, R.P. 1967.
Mutations in the galactose operon in Escherichia coli: genetic characterisation.
Mol. Gen. Genet. Vol.100. pp.178-191.

Saigo, K., Kugimiya, W., Matsuo, Y., Inouye, S., Yoshioka, K. and Yuki, S. 1984.
Identification of the coding sequence for a reverse transcriptase-like enzyme in a transposable genetic element in Drosophila melanogaster.
Nature. Vol.312. pp.659-661.

Sanger, F., Nicklen, S. and Coulson, A.R. 1977.
DNA sequencing with chain-terminating inhibitors.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.74. pp.5463-5467.

Sasakawa, C. and Berg, D.E. 1982.
IS50-mediated inverse transposition: discrimination between the two ends of an IS element.
J. Mol. Biol. Vol.159. pp.257-271.

Sasakawa, C., Carle, G.F. and Berg, D.E. 1983.
Sequences essential for transposition at the termini of IS50.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.80. pp.7293-7297.

Sasakawa, C., Lowe, J.B., Mc.Divitt, L. and Berg, D.E. 1982.
Control of transposon Tn5 transposition in Escherichia coli.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.79.pp.7450-7454.

Sasakawa, C., Phadnis, S.H., Carle, G.F. and Berg, D.E. 1985.
Sequences essential for IS50 transposition.
J. Mol. Biol. Vol.182. pp.487-493.

Sasakawa, C., Uno, Y. and Yoshikawa, M. 1981.
The requirement for both DNA polymerase and 5' to 3' exonuclease activities of DNA polymerase I during Tn5 transposition.
Mol. Gen. Genet. Vol.182. pp.19-24.

Scherer, S., Mann, C. and Davis, R.W. 1982.
Reversion of a promoter deletion in yeast.
Nature. Vol.298. pp.815-819.

Schreier, P.H. and Cortese, R. 1979.
A fast and simple method for sequencing DNA cloned in the single-stranded bacteriophage M13.
J. Mol. Biol. Vol.129. pp.169-172.

Schwarz-Sommer, Z., Gierl, A., Cuypers, H., Peterson, P.A. and Saedler, H. 1985.
Plant transposable elements generate the DNA sequence diversity needed in evolution.
E.M.B.O. Journal. Vol.4. pp.591-597.

Scott, J.R. 1984.
Regulation of plasmid replication.
Microbiol. Rev. Vol.48. pp.1-23.

Shapiro, J.A. 1969.
Mutations caused by the insertion of genetic material into the galactose operon of Escherichia coli.
J. Mol. Biol. Vol.40. pp.93-105.

Shapiro, J.A. 1979.
Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.76. pp. 1933-1937.

Shapiro, J.A. 1983a.
Mobile Genetic Elements.
Academic Press, New York.

Shapiro, J.A. 1983b.
Genomic reorganisation in cell lineages.
In: Mobile Genetic Elements.
Ed.: Shapiro, J.A.
Academic Press, New York.
pp.xi-xvi.

Sherratt, D., Arthur, A. and Burke, M. 1981.
Transposon-specified site-specific recombination systems.
Cold Spring Harbor Symp. Quant. Biol. Vol.45. pp.275-281.

Shiba, T. and Saigo, K. 1983.
Retrovirus-like particles containing RNA homologous to the transposable element copia in Drosophila melanogaster.
Nature. Vol.302. pp.119-124.

Shine, J. and Dalgarno, L. 1975.
Determinant of cistron specificity in bacterial ribosomes.
Nature. Vol.254. pp.34-38.

Simons, R.W., Hoopes, B.C., Mc.Clure, W.R. and Kleckner, N. 1983.
Three promoters near the termini of IS10: pIN, pOUT and pIII.
Cell. Vol.34. pp.673-682.

Simons, R.W. and Kleckner, N. 1983.
Translational control of IS10 transposition.
Cell. Vol.34. pp.683-691.

Spena, A., Krause, E. and Dobberstein, B. 1985.
Translation efficiency of zein mRNA is reduced by hybrid formation between the 5' and 3' untranslated regions.
E.M.B.O. Journal. Vol.4. pp.2153-2158.

Stacey, K.A. and Simpson, E. 1965.
Improved method for the isolation of thymine requiring mutants of Escherichia coli.
J. Bacteriol. Vol.90. pp.554-555.

Sternberg, N. 1985.
Evidence that adenine methylation influences DNA-protein interactions in Escherichia coli.
J. Bacteriol. Vol.164. pp.490-493.

Sternglanz, R., Dinardo, S., Voelkel, K.A., Nishimura, Y., Hirota, Y., Becherer, K., Zumstein, L. and Wang, J.C. 1981.
Mutations in the gene coding for Escherichia coli DNA topoisomerase I affect transcription and transposition.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.78. pp.2747-2751.

Stokes, H.W. and Hall, B.G. 1984.
Topological repression of gene activity by a transposable element.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.81. pp.6115-6119.

Stueber, D. and Bujard, H. 1981.
Organisation of transcriptional signals in plasmids pBR322 and pACYC184.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.78. pp. 167-171.

Stueber, D. and Bujard, H. 1982.
Transcription from efficient promoters can interfere with plasmid replication and diminish expression of plasmid specified genes.
E.M.B.O. Journal. Vol.1. pp.1399-1404.

Syvanen, M. 1984.
The evolutionary implications of mobile genetic elements.
Ann. Rev. Genet. Vol.18. pp.271-293.

Tacon, W., Bhamra, B., Sunar, B. and Sherratt, D.J. 1981.
Structure and function of plasmid ColK.
Plasmid. Vol.6. pp.358-359.

Tamm, J. and Polisky, B. 1985.
Characterisation of the ColE1 primer-RNAI complex: analysis of a domain of ColE1 RNAI necessary for its interaction with primer RNA.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.82. pp.2257-2261.

Temin, H.M. 1980.
Origin of retroviruses from cellular moveable genetic elements.
Cell. Vol.21. pp.599-600.

Thorpe, P.A. and Clowes, R.C. 1984.
Absence of direct repeats flanking transposons resulting from intra-molecular transposition.
Gene. Vol.28. pp.103-112.

Tomizawa, J. 1984.
Control of ColE1 plasmid replication: the process of binding RNAI to the primer transcript.
Cell. Vol.38. pp.861-870.

Tomizawa, J. and Itoh, T. 1981.
Plasmid ColE1 incompatibility determined by interaction of RNAI with primer transcript.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.78. pp.6096-6100.

Tomizawa, J. and Itoh, T. 1982.
The importance of RNA secondary structure in ColE1 primer formation.
Cell. Vol.31. pp.575-583.

Tomizawa, J. and Som, T. 1984.
Control of plasmid ColE1 replication: enhancement of binding of RNAI to the primer transcript by the Rom protein.
Cell. Vol.38. pp.871-878.

Toussaint, A. and Faelen, M. 1973.
Connecting two unrelated DNA sequences with a Mu dimer.
Nature New Biology. Vol.242. pp.1-4.

Toussaint, A. and Resibois, A. 1983.
'Phage Mu: transposition as a life-style.
In: Mobile Genetic Elements.
Ed.: Shapiro, J.A.
Academic Press, New York.
pp.105-158.

Travers, A. 1984.
Regulation by antisense RNA.
Nature. Vol.311. p410.

Tu, C.D. and Cohen, S.N. 1980.
Translocation specificity of Tn3: characterisation of sites of multiple insertions.
Cell. Vol.19. pp.151-160.

Twigg, A.J. and Sherratt, D. 1980.
Trans-complementable copy number mutants of plasmid ColE1.
Nature. Vol.283. pp.216-218.

Vieira, J. and Messing, J. 1982.
The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers.
Gene. Vol.19. pp.259-268.

von Hippel, P.H., Bear, D.G., Morgan, W.D. and McSwiggen, J.A. 1984.
Protein-nucleic acid interactions in transcription: a molecular analysis.
Ann. Rev. Biochem. Vol.53. pp.389-446.

Wallace, L.J., Ward, J.M., Bennett, P.M., Robinson, M.K. and Richmond, M.H. 1980.
Transpositon immunity.
Cold Spring Harbor Symp. Quant. Biol. Vol.45. pp.183-188.

Way, J. and Kleckner, N. 1984.
Essential sites at Tn10 termini.
Proc. Nat. Acad. Sci. (U.S.A.). Vol.81. pp.3452-3456.

Weinert, T.A., Derbyshire, K., Hughson, F.M. and Grindley, N.D.F. 1984.
Replicative and conservative transpositional recombination of insertion sequences.
Cold Spring Harbor Symp. Quant. Biol. Vol.49. pp.251-260.

Weinert, T.A., Schaus, N.A. and Grindley, N.D.F. 1983.
Insertion sequence duplication in transpositional recombination.
Science. Vol.222. pp.755-765.

Weinstock, G.M., Susskind, M.M. and Botstein, D. 1979.
Regional specifity of illegitimate recombination by the translocable ampicillin-resistance element Tn1 in the genome of 'phage P22.
Genetics. Vol.92. pp.685-710.

Weintraub, H., Izant, J.G. and Harland, R.M. 1985.
Antisense RNA as a molecular tool for genetic analysis.
Trends Genet. Vol.1. pp.22-25.

Winston, F., Durbin, K.J. and Fink, G.R. 1984.
The SPT3 gene is required for normal transposition of Ty elements in Saccharomyces cerevisiae.
Cell. Vol.39. pp.675-682.

Wishart, W.L., Broach, J.R. and Ohtsubo, E. 1985.
ATP-dependent specific binding of Tn3 transposase to Tn3 inverted repeats.
Nature. Vol.314. pp.556-558.

Womble, D.D., Dong, X., Wu, R.P., Luckow, V.A., Martinez, A.F. and Rownd, R.H. 1984.
IncFII plasmid incompatibility product and its target are both RNA transcripts.
J. Bacteriol. Vol. 160. pp.28-35.

Wurster-Biel, S. and Hartl, D.L. 1983.
Evolution of transposons - natural selection for Tn5 in Escherichia coli K-12.
Genetics. Vol.103. pp.581-592.

Yanisch-Perron, C., Vieira, J. and Messing, J. 1985.
Improved M13 'phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors.
Gene. Vol.33. pp.103-119.

Tn3 sequence between co-ordinates 1 & 3200.

```
        10          20          30          40          50          60
GGGGTCTGAC GCTCAGTGGA ACGAAAACTC ACGTTAAGCA ACGTTTTCTG CCTCTGACGC
CCCCAGACTG CGAGTCACCT TGCTTTTGAG TGCAATTCGT TGCAAAAGAC GGAGACTGCG

        70          80          90         100         110         120
CTCTTTTAAT GGTCTCAGAT GACCTTTGGT CACCAGTTCT GCCAGCGTGA AGGAATAATG
GAGAAAATTA CCAGAGTCTA CTGGAAACCA GTGGTCAAGA CGGTCGCACT TCCTTATTAC

       130         140         150         160         170         180
GCCGAGCATA TTGATATGTC CGTGGCAAAG CGGGGAGAGG CGTGCGATAT CTTCATCATT
CGGCTCGTAT AACTATACAG GCACCGTTTC GCCCCTCTCC GCACGCTATA GAAGTAGTAA

       190         200         210         220         230         240
CAGTGTTTCA CCCTGGGCCC GGAGATGATC CAGAGCTGCC TGCATATAAA TAGTGTTCCA
GTCACAAAGT GGGACCCGGG CCTCTACTAG GTCTCGACGG ACGTATATTT ATCACAAGGT

       250         260         270         280         290         300
TAACACGACG GCGTTAGTGA CCAGCCCCAG AGCTCCCAAC TGATCTTCCT GACCGTCGGT
ATTGTGCTGC CGCAATCACT GGTCGGGGTC TCGAGGGTTG ACTAGAAGGA CTGGCAGCCA

       310         320         330         340         350         360
ATATCGTTTT CTTATCTCAC CTTTTTGACC GTGACAGATG GCTCTGGCAA CTGCATGACG
TATAGCAAAA GAATAGAGTG GAAAAACTGG CACTGTCTAC CGAGACCGTT GACGTACTGC

       370         380         390         400         410         420
GCTTTCTCCC CGATTAAGCT GGGTCAGAAT GCGCCGGCGG TAATCTTCAT CATCAATATA
CGAAAGAGGG GCTAATTCGA CCCAGTCTTA CGCGGCCGCC ATTAGAAGTA GTAGTTATAT

       430         440         450         460         470         480
ATTAAGCAGA TACAGCGTTT TGTTGATGCG CCCCACTTCA ATGATTGCCT GAGTCAGTCC
TAATTCGTCT ATGTCGCAAA ACAACTACGC GGGGTGAAGT TACTAACGGA CTCAGTCAGG
```

```
       490         500         510         520         530         540
GGAGGGACGT TCACTTTTCA GCAATGAACG GACCAGCACT GAGGCCTGTA CTTTGCCCAG
CCTCCCTGCA AGTGAAAAGT CGTTACTTGC CTGGTCGTGA CTCCGGACAT GAAACGGGTC

       550         560         570         580         590         600
CTTCAGGGAG CCTGCGGTCC GGATCATTTC GTCCCACTGA AGGACTATTT TTCGGGGATC
GAAGTCCCTC GGACGCCAGG CCTAGTAAAG CAGGGTGACT TCCTGATAAA AAGCCCCTAG

       610         620         630         640         650         660
TGATTGCCCT CTGGCAATAT CATTCAGCAC GCCATAGTCG GCATCATGGT CCATTCGCCA
ACTAACGGGA GACCGTTATA GTAAGTCGTG CGGTATCAGC CGTAGTACCA GGTAAGCGGT

       670         680         690         700         710         720
GAAAACCGAA GCACCGGCAT CAGCCAGGCG TGGAGAAAAC TGGTATCCCA GCAGCCAGAA
CTTTTGGCTT CGTGGCCGTA GTCGGTCCGC ACCTCTTTTG ACCATAGGGT CGTCGGTCTT

       730         740         750         760         770         780
AAGGCCAAAG ACAAGATCGC TGGCACCTGC CGTATCGGTC ATAATTTCGG TTGGATTCAG
TTCCGGTTTC TGTTCTAGCG ACCGTGGACG GCATAGCCAG TATTAAAGCC AACCTAAGTC

       790         800         810         820         830         840
CCCGGTCTCC TGTTCCAGAA GGCCTTCCAG CACAAAGATA GAGTCCCTCA GCGTCCCCGG
GGGCCAGAGG ACAAGGTCTT CCGGAAGGTC GTGTTTCTAT CTCAGGGAGT CGCAGGGGCC

       850         860         870         880         890         900
TATAACGATG CCATGAAAGC CATGAAAGCC GGAATACTGA TCGGACACAA AGTTGTACCA
ATATTGCTAC GGTACTTTCG GTACTTTCGG CCTTATGACT AGCCTGTGTT TCAACATGGT

       910         920         930         940         950         960
GGTGATCCCT CTGTTATTAC CAAAGTATTT GCGGTTCGGT CCGGCATTGA TTGTTCTGAC
CCACTAGGGA GACAATAATG GTTTCATAAA CGCCAAGCCA GGCCGTAACT AACAAGACTG
```

```
      970         980         990        1000        1010        1020
TGGCGTAACA AAGCGCATTC CATCTGCAGA TGCCACTTCT CCTCCACCCC ATATCTGTGC
ACCGCATTGT TTCGCGTAAG GTAGACGTCT ACGGTGAAGA GGAGGTGGGG TATAGACACG

                  0.76
-35   1030        1040 -10    1050        1060        1070        1080
CAGTGGCAGC GTTGCCTGAA AATCAACCAG TCTGGCATTA GCGCTGGTGA TAGTTTCAGC
GTCACCGTCG CAACGGACTT TTAGTTGGTC AGACCGTAAT CGCGACCACT ATCAAAGTCG

      1090        1100        1110        1120        1130        1140
CCGCAGATAG TTCGCTTTTG TCCAGTTCAG CCGGTGTCGG GTCAGTGCAG GAACATTTGA
GGCGTCTATC AAGCGAAAAC AGGTCAAGTC GGCCACAGCC CAGTCACGTC CTTGTAAACT

      1150        1160        1170        1180        1190        1200
TCTGATCAGT GGTTCCAGAC CGATATTGCA GGCTTCAGCC ATCAGCACGG CGCTGATGCT
AGACTAGTCA CCAAGGTCTG GCTATAACGT CCGAAGTCGG TAGTCGTGCC GCGACTACGA

      1210        1220        1230        1240        1250        1260
GACGGGCAGA TCATCAACTC TGGCACTGGC TTCACTGGCA TGGAAAAACT CATCAGCAAA
CTGCCCGTCT AGTAGTTGAG ACCGTGACCG AAGTGACCGT ACCTTTTTGA GTAGTCGTTT

      1270        1280        1290        1300        1310        1320
TCCGGTATGG GCGTTAATTT CGAGCAGCAA CTCCGTTAAA TCCACCGGAG GGAGCAGATC
AGGCCATACC CGCAATTAAA GCTCGTCGTT GAGGCAATTT AGGTGGCCTC CCTCGTCTAG

      1330        1340        1350        1360        1370        1380
ACTGATCATT TTGCTCAGTC GTTTCAGACT GTCCGGCTCA TCAAGACTGG CGAGGGGAGA
TGACTAGTAA AACGAGTCAG CAAAGTCTGA CAGGCCGAGT AGTTCTGACC GCTCCCCTCT

      1390        1400        1410        1420        1430        1440
AATTGTCAAC CGGGGCTTCG GGCCAGAAAC ATCGAGTTCG ACAGCCTCAT TTTCGCCAAG
TTAACAGTTG GCCCCGAAGC CCGGTCTTTG TAGCTCAAGC TGTCGGAGTA AAAGCGGTTC
```

```
      1450        1460        1470        1480        1490        1500
ACGTGCAGCA ACCTGTCTGT AACGACTATC AAGCTGATGG CCCAGAGATT TTATTGCTTC
TGCACGTCGT TGGACAGACA TTGCTGATAG TTCGACTACC GGGTCTCTAA AATAACGAAG

      1510        1520        1530        1540        1550 -35    1560
CTGCGGGTCT GTCGGGTGCC CCAAAGAACG ATAAACCTTA ATCCGATTTG CCTGCCAGTC
GACGCCCAGA CAGCCCACGG GGTTTCTTGC TATTTGGAAT TAGGCTAAAC GGACGGTCAG

0.77
      1570    -10 1580        1590        1600        1610        1620 -3!
AGCACCCTGT AGTAATCTTG CACGAGGATC TCCCCACCGG TTACTGCCGG TAACGTAGAC
TCGTGGGACA TCATTAGAAC GTGCTCCTAG AGGGGTGGCC AATGACGGCC ATTGCATCTG

          0.81
      1630        1640 -10    1650        1660        1670        1680
ATCCCTCCGT CTCAGACTAT CCTGCAGTTT ACTGAGAAAG CAGAGCGTGT ATCCCCTTCG
TAGGGAGGCA GAGTCTGATA GGACGTCAAA TGACTCTTTC GTCTCGCACA TAGGGGAAGC

      1690        1700        1710        1720        1730        1740
GGTGATATGT TTTTCCTTGT TAATCACCAG CCGTTTCCAT GACTGACTGA TGATTTCCGT
CCACTATACA AAAAGGAACA ATTAGTGGTC GGCAAAGGTA CTGACTGACT ACTAAAGGCA

      1750        1760        1770        1780        1790        1800
TGGTGCGTCG TCAAAAAACT GCCGCCGTGA GCTGAACTCC CGGCTGAGGT AGTCACAGGC
ACCACGCAGC AGTTTTTTGA CGGCGGCACT CGACTTGAGG GCCGACTCCA TCAGTGTCCG

                  0.71
      1810 -35    1820        1830    -10 1840        1850        1860
ATTCAGAGTG GTAACCCCGG CAGGTGCGGA TGAAAATTTA ACGGTATTCA GCAGATGGGG
TAAGTCTCAC CATTGGGGCC GTCCACGCCT ACTTTTAAAT TGCCATAAGT CGTCTACCCC

      1870        1880        1890        1900        1910        1920
CAGGAAACGA CGAACGCGCC CGTACTGCTC CACCATTTCG TCATGAAAAT TATCGTCTGA
GTCCTTTGCT GCTTGCGCGG GCATGACGAG GTGGTAAAGC AGTACTTTTA ATAGCAGACT
```

```
      1930        1940        1950        1960        1970        1980
GGGCCGGGCA ATTTCACGGA CAAGCGTGAT GATTTCAGCC AGCTTTTGCC TAGGGATGTA
CCCGGCCCGT TAAAGTGCCT GTTCGCACTA CTAAAGTCGG TCGAAAACGG ATCCCTACAT

      1990        2000        2010        2020        2030        2040
GCTGAACACC TCAGCACGAA TCGATTCGTC CGGTGTTTCT TCTTTCAGCA AGTACGAACA
CGACTTGTGG AGTCGTGCTT AGCTAAGCAG GCCACAAAGA AGAAAGTCGT TCATGCTTGT

      2050        2060        2070        2080        2090        2100
TGCGCTGGCG AGCGCCAATG CAGATTTATC CAGATCCTTC AGCGAGCGGA GCCGTTTTTT
ACGCGACCGC TCGCGGTTAC GTCTAAATAG GTCTAGGAAG TCGCTCGCCT CGGCAAAAAA

      2110        2120        2130        2140        2150        2160
CTGCCCAATC TTTCTGGCGT CACGGATGAT AACGGCCAGC ATGGCGTCCA GAACGTCCAG
GACGGGTTAG AAAGACCGCA GTGCCTACTA TTGCCGGTCG TACCGCAGGT CTTGCAGGTC

      2170        2180        2190        2200        2210        2220
TGCATCATCC AGCGCCAGCG TTTCCCATGC AAGGACAAAG GCAACCAGAA CCGCCATCCT
ACGTAGTAGG TCGCGGTCGC AAAGGGTACG TTCCTGTTTC CGTTGGTCTT GGCGGTAGGA

                                                0.72
      2230        2240        2250  -35   2260        2270  -10   2280
TTTCTGCGGT GACATCCTGG CAATATTGAA CACCGAAGTC ATACCAGCAT AACGTGCGAG
AAAGACGCCA CTGTAGGACC GTTATAACTT GTGGCTTCAG TATGGTCGTA TTGCACGCTC

      2290        2300        2310        2320        2330        2340
ATTTTTCAGG CGCACAGCCG GGAGTGTACT CAGGTTTTCA GCATGCAGGC CAAAATCGTT
TAAAAAGTCC GCGTGTCGGC CCTCACATGA GTCCAAAAGT CGTACGTCCG GTTTTAGCAA

      2350        2360        2370        2380        2390        2400
CAGAGTTTTC CAGCGTTCAA TTGCTTCATT AAACGCCGGA CCACTGATGG TCACAGGGCC
GTCTCAAAAG GTCGCAAGTT AACGAAGTAA TTTGCGGCCT GGTGACTACC AGTGTCCCGG
```

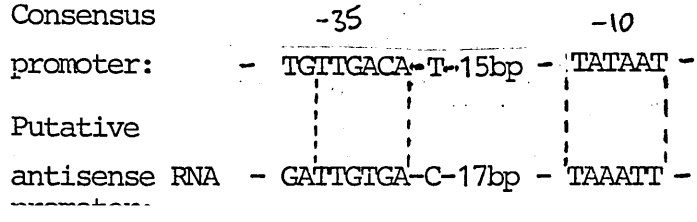
```
      2410        2420        2430        2440        2450        2460
TTTTTTCAGT GATTCCAGTA AAGACAGGCG GCTGCAATCA GTTGGCCCCA GCAGCATCTC
AAAAAAGTCA CTAAGGTCAT TTCTGTCCGC CGACGTTAGT CAACCGGGGT CGTCGTAGAG

      2470        2480        2490        2500        2510        2520
CAGCTGTGAA CGCTGTTCGG CTGACGGTAT CAGTGCCAGT TTGTTCCACA GGCGCAACGT
GTCGACACTT GCGACAAGCC GACTGCCATA GTCACGGTCA AACAAGGTGT CCGCGTTGCA

      2530        2540        2550        2560        2570        2580
CGCCTTTTCC CTTACCTCTG AAATCAACCG GGTCAGCGTG GTGGCTCCGG GGAGAATAAT
GCGGAAAAGG GAATGGAGAC TTTAGTTGGC CCAGTCGCAC CACCGAGGCC CCTCTTATTA

      2590        2600        2610        2620        2630        2640
ACGATGTTGC ATAAGCCACC CTGTCGCCAG GTCGAAAAGC AGCCCAGGAC GTTCGTTGCT
TGCTACAACG TATTCGGTGG GACAGCGGTC CAGCTTTTCG TCGGGTCCTG CAAGCAACGA

      2650        2660        2670        2680        2690        2700
TATCCAGCTC CGGGTATATA AAAGACGGGT AAGGCGAAAT GTCCAGGGCC AGGCAAATTC
ATAGGTCGAG GCCCATATAT TTTCTGCCCA TTCCGCTTTA CAGGTCCCGG TCCGTTTAAG

      2710        2720        2730        2740        2750        2760
ACGATACTGA TAGTGCTGAC GTATCAGCGC TGCATGCTCA CGGCGGGTAT TTTCCCTCTG
TGCTATGACT ATCACGACTG CATAGTCGCG ACGTACGAGT GCCGCCCATA AAAGGGAGAC

                                                0.74.
      2770        2780        2790  -35   2800        2810  -10   2820
ACCGTATTCT GCAAGAACGG TGATATCACG AATCCCGAGC TGTCTGGCGG TAAAATGCCG
TGGCATAAGA CGTTCTTGCC ACTATAGTGC TTAGGGCTCG ACAGACCGCC ATTTTACGGC

      2830        2840        2850        2860        2870        2880
GACGCCGGAA GGAATATGAT TCATGTCAGT AAGAAAAGTG CCCAGAAAAC GCACACACCC
CTGCGGCCTT CCTTATACTA AGTACAGTCA TTCTTTTCAC GGGTCTTTTG CGTGTGTGGG
```

```
                                                    0.78
      2890        2900        2910 -35    2920        2930 -10   2940 -3
GATTTGCAGG GCAATACCAA GTCGATTGTG ATCACCCCGG CTTTTCCCGA TAAATTCTTT
CTAAACGTCC CGTTATGGTT CAGCTAACAC TAGTGGGGCC GAAAAGGGCT ATTTAAGAAA

      2950  0.67  2960 -10    2970        2980        2990        3000
ATCCGCTTCA TCAAGATGAA AATAACGCGC CAGCTGAAGT TCATCGGGTT CGCCAGTGAA
TAGGCGAAGT AGTTCTACTT TTATTGCGCG GTCGACTTCA AGTAGCCCAA GCGGTCACTT

      3010        3020        3030        3040        3050        3060
CCTGCCATAA CTCTCAACCT GCTCAGTGGT CAAAAAATCA ACGGGCATAT CGGCCTCCCT
GGACGGTATT GAGAGTTGGA CGAGTCACCA GTTTTTTAGT TGCCCGTATA GCCGGAGGGA

      3070    -35 3080  0.85      3090     -35 3100 -10   3110   0.74  3120
GCCTGACGGC TTTTTTAACA CAACTGCAAC CGTTCGAAAT ATTATAAATT ATCAGACATA
CGGACTGCCG AAAAAATTGT GTTGACGTTG GCAAGCTTTA TAATATTTAA TAGTCTGTAT
                                          -10     0.81  tnpA gene
                                                        promoter

-10   3130        3140  -35   3150  0.71    3160    -10 3170        3180
GTAAAACGGC TTCGTTTGAG TGTCCATTAA ATCGTCATTT TGGCATAATA GACACATCGT
CATTTTGCCG AAGCAAACTC ACAGGTAATT TAGCAGTAAA ACCGTATTAT CTGTGTAGCA
-35

      3190        3200
GTCTGATATT CGATTTAAGG
CAGACTATAA GCTAAATTCC
```

Possible promoters (-35 and -10 regions) on the 'antisense' (non-coding) strand of this part of the transposon are shown. Also indicated is the tnp A gene promoter, on the coding strand.

Underlined figures represent the strength of each promoter as compared to a 'consensus' promoter, of which the strength is taken to be 1.00.

This 'consensus' promoter is shown below and is compared with the putative promoter for the possible antisense RNA species described in the text.



DNA sequencing gels of part of Tn1.