Transposon-Encoded Site-Specific Recombination

by

Lorraine S. Symington

A dissertation submitted for the degree of

Doctor of Philosophy at the

University of Glasgow

September 1982

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

The work contained in this thesis has never been submitted to this or any other university for a degree, and is original unless otherwise stated.

dedicated to Thomas

TABLE OF CONTENTS

Acknowledgements

I would like to thank the following:

Professor David Sherratt, my supervisor, for his patience and help during my three years of research.

Dr. Russell Thompson for advice with cloning, footprinting and DNA sequencing experiments.

To all members of the Genetics department for helpful discussions, advice and sympathy during the trials and errors of research, and for keeping me cool during working hours with the odd handful of ice in appropriate places at the most inopportune moments! A special thankyou to Jean for patiently proofreading this thesis; also to Mary, who keeps the high-spirited inmates of the sixth floor under control, and provides a constant supply of sterile solutions and buffers for all to use.

The media ladies who helped with the smooth transition from Brighton to Glasgow, kept my bacteria well-fed, glassware clean and never complained of the pile of broken glass beneath my bench.

Members of the Microbial Genetics Group in Brighton for their help during my initial six months of research before moving to Glasgow.

My family and friends for their constant support and sanity maintenance (of sorts!).

Finally, to the SRC for providing financial support.

Abstract

The closely related transposable elements, Tn*3* and γδ, share significant DNA sequence homology and functional organization. These elements have been shown to encode interchangeable resolvase proteins to mediate resolution of obligatory transpositional cointegrates. There is no apparent complementation between these elements for the *tnp*A gene function.

To analyze the mechanism of resolvase-mediated recombination, the Tn*3*-encoded *tnp*R gene was cloned into a high expression plasmid vector. This allowed large amounts of resolvase to be synthesized, thus aiding purification of the protein. The purified protein was subsequently used for biochemical analysis of the resolution reaction and for DNase I footprinting experiments.

A number of small plasmid substrates were constructed containing two *res* sites in direct or inverted orientation. *In vitro* resolution reactions were assayed by gel electrophoresis to detect the formation of interlocked circles (catenates) which appeared to be the major reaction product. The resolution reaction requires only resolvase and a supercoiled substrate containing two directly repeated sites, under the appropriate ionic conditions; the reaction is independent of host factors or an external energy supply. Catenates are always formed during resolution suggesting that this may be a direct consequence of the reaction mechanism.

Substrates containing varying lengths of DNA separating *res* sites have differing efficiencies of resolution *in vitro*, the reaction proceeds with greatest efficiency when the length of DNA between

res sites is least. This data, in conjunction with the observed preference for two directly repeated *res* sites in *cis*, has led to the proposal of a "tracking" model for resolvase-mediated recombination.

Resolvase has been shown to bind specifically and nonspecifically to DNA. Specific binding, resulting from tight association of resolvase with the target site, was investigated by sequence protection from DNase I. This revealed three binding sites for resolvase within the *res* region. The sequences of the protected sites conform to the consensus sequence proposed for other regulatory DNA binding proteins.

Abbreviations

1. For phenotype

Ap^r resistant to ampicillin

Cm^r resistant to chloramphenicol

rif^r resistant to rifampicin

Sm^r resistant to streptomycin

Su^r resistant to sulphonamide

Tc^r resistant to tetracycline

Tp^r resistant to trimethoprim

Rec^+ recombination proficient

Rec^- recombination deficient

Mob^+ transmissible by F or R388

Tra^+ self-transmissible

bom^+ basis of mobility (contains transfer origin)

Iel^+ immune to colicin El

Ika^+ immune to colicin K

Cka^+ colicin K producing

2. Chemicals

NB nutrient broth medium

EtBr ethidium bromide

DNA deoxyribonucleic acid

RNA ribonucleic acid

o/c covalently closed relaxed DNA

s/c supercoiled DNA

ATP adenosine triphosphate

APS ammonium persulphate

DTT dithiothreitol

EDTA sodium ethylene diamine tetra-acetate

GTP guanine triphosphate

SDS sodium dodecyl sulphate

TEMED tetra-methyl ethylene diamine

Tris tris (hydroxymethyl) methlyamine

3. For measures

l litres

ml millilitres (10^{-3} litres)

µl microlitres (10^{-6} litres)

g grams (or g force, when referring to centrifugation)

mg milligrams (10^{-3} grams)

µg micrograms (10^{-6} grams)

w/v weight per volume

v/v volume per volume

V volts

mA milliamps

W watts

cm centimetre

M moles per litre (molar)

mM millimolar

^{o}C degrees centigrade

sec second

min minute

bp base pair

Kbp kilobasepairs (10^{3} bases)

Mdal megadaltons (10^{6} daltons)

CHAPTER 1

GENERAL INTRODUCTION

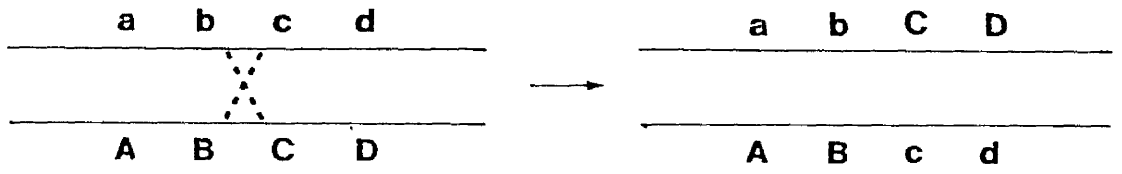
Introduction

The experiments described in this thesis were designed to study the recombination events mediated by transposable elements. Transposable elements are DNA sequences which integrate at many loci in genomes ; their insertion, and other rearrangements which they generate, has led to speculation over their possible significance in genome evolution and development. The molecular mechanisms involved in propagation of these elements have become one of the focal points of study in molecular biology over recent years and form the basis of the work reported here.

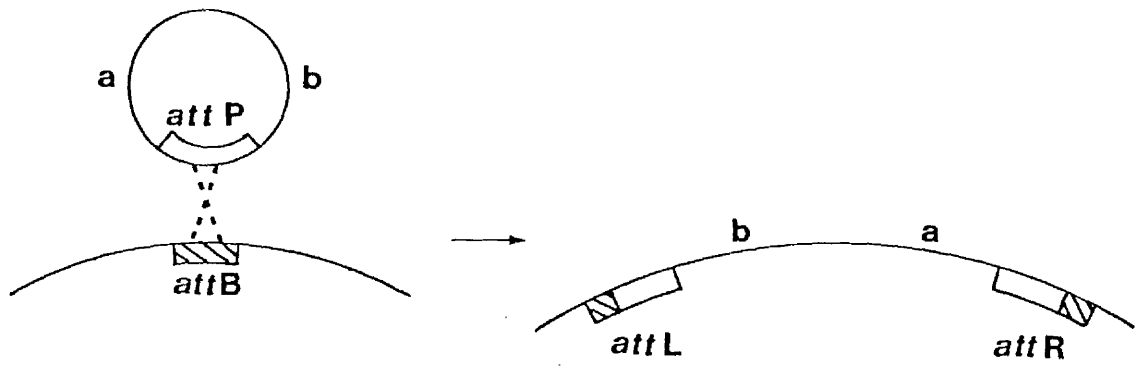
Historically, single base changes, which often occur as mistakes during replication (Radman *et al.*, 1979), were assumed to be the major cause of mutations; the substrate for Darwinian natural selection. However, genetic rearrangements such as inversion, duplication and deletion of DNA sequences also contribute to the generation of diversity within genomes. Recently, there has been speculation that transposable elements play a role in generating such rearrangements. In addition, their ability to fuse together unrelated DNA sequences, suggests that they may enable the formation of new genes from segments of pre-existing ones. Such rearrangements, together with other mutational processes, would be instrumental in generating the genetic variation for evolution to occur.

Any mechanism which places DNA in new positions, thus ensuring continuous movement of genetic material, lends greatly to genome plasticity and the generation of diversity. The major force involved in generating all rearrangements is genetic recombination. Histor-

a General genetic recombination



b Site-specific recombination



c Transposition

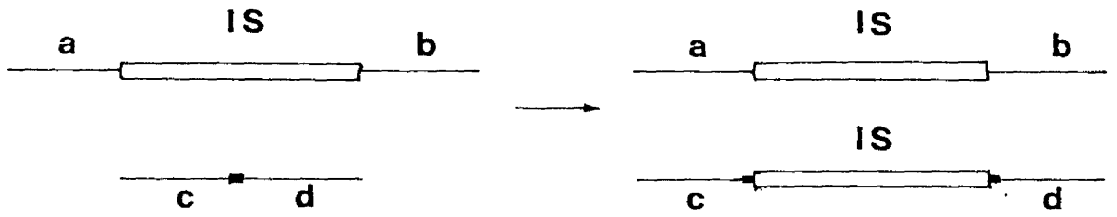


FIGURE 1.1 Diagram to illustrate the major recombination systems in *E.coli*

ically, the processes of recombination have been divided into two classes:

(a) Homologous recombination.

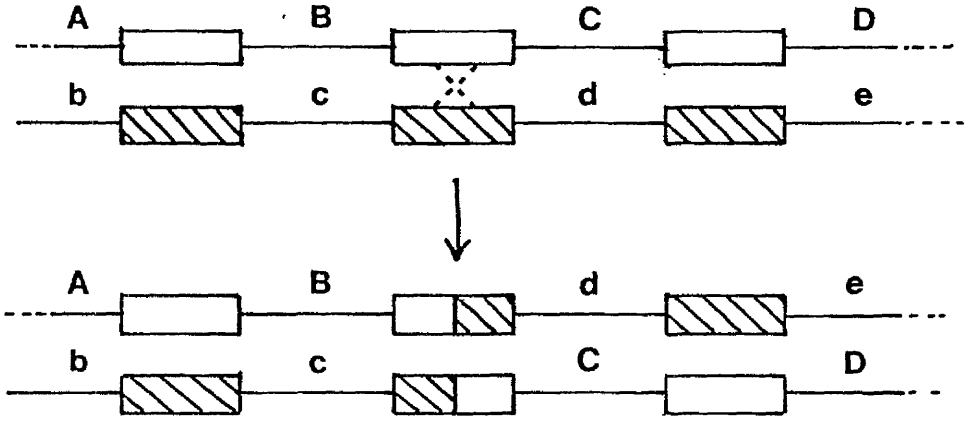
These events occur over extended regions of homology, at least several hundred base pairs, and are mediated by the general genetic recombination system; in *E.coli* this has been named the *recA*-dependent pathway of recombination. This results in reassortment of genetic markers between homologous chromosomes, but does not normally result in genetic rearrangements.

(b) Non-homologous recombination.

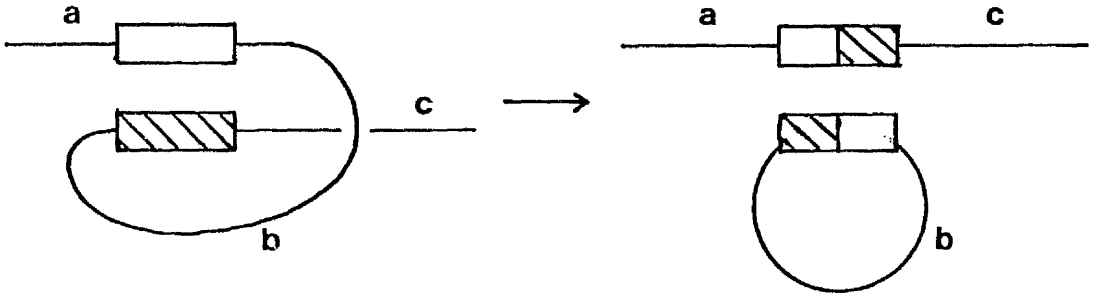
These events occur between sequences of little of no homology and are usually independent of the general genetic recombination system. The proteins which mediate these events are oftenencoded by genes adjacent to the site at which they act. This class has been subdivided into replicative and non-replicative events. The non-replicative events are typified by bacteriophage lambda integration which occurs by a reciprocal break/join mechanism with no concomitant DNA synthesis. Duplicative events are represented by transposition; the movement of transposable elements to new sites results in no loss of the element from it s original site, thus DNA synthesis accompanies transposition (see figure 1.1)

These processes of recombination have been arbitrarily divided into classes but should not be considered as independent events. Most genetic rearrangements are a consequence of a combination of these processes, eg. transposition results in duplication of the element, thus providing portable regions of homology on which the general genetic recombination system can act to generate deletions, inversions and duplications of DNA (figure 1.2).

A.



B.



C.

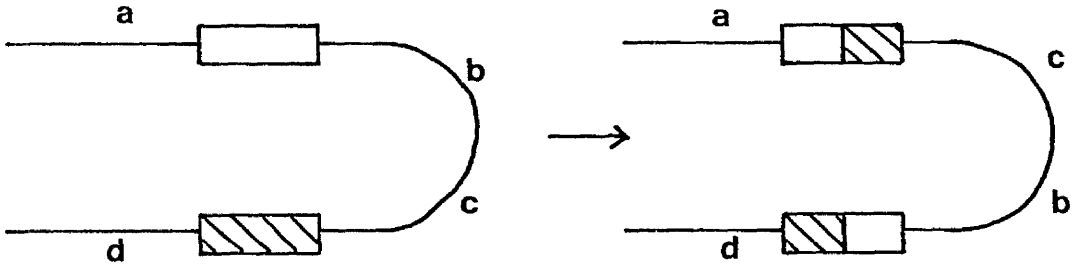


FIGURE 1.2 Homologous recombination between duplicated sequences

A. If homologous chromosomes contain repeated sequences separating genes, or parts of genes, then misalignment of the DNA helices may result in duplication or deletion of genes, or parts of genes.

B. Intramolecular homologous recombination can also occur when a chromosome contains repeated sequences. If the sequences are directly repeated then a deletion will result. If the deleted sequence has no vegetative origin of replication it will be lost unless it integrates back into the chromosome by a second recombination event. This could be at the original position or at another site where there is a copy of the element.

C. If the repeated sequences are present in inverted orientation within the chromosome, homologous recombination between the sequences will generate an inversion of the intervening DNA.

Classical genetic crosses in which markers reassort in Mendelian fashion exemplify homologous recombination; much of this data has been obtained from simple crosses in fungi or *E.coli*. In fungi all products of meiosis are recovered in tetrads, analysis of which shows which homologous chromosomes paired in the zygote and exchanged genetic material. These studies have implicated formation of heteroduplex regions as intermediates in recombination (Radding,1978). This led Holliday (1964) to propose a model for symmetrical genetic recombination which involves breakage and reunion of strands to form a heteroduplex joint (figure 1.3). However, this model does not explain the observed nonreciprocality of many recombination events. To overcome this problem Meselson and Radding (1975) proposed a modification of the initial model so that nonreciprocal progeny would result (figure 1.4). All the models proposed for the mechanism of recombination demand a number of proteins which can nick and/or ligate DNA molecules to initiate the recombination event and resolve intermediates.

The enzymology of recombination has been mainly confined to *E.coli*. Clark has isolated a number of mutants with varying degrees of recombinational proficiency; of these the *recA* mutation has the most severe effects (Clark,1973;1974). The *recA* mutation confers extreme UV sensitivity on the host, reduces cell viability, and shows deficiency in recombination of incoming linear DNA by conjugation or transduction. The *recA* gene product has been directly implicated in heteroduplex formation *in vivo* (Holloman and Radding,1976) and also has a specific protease activity intimately involved in the SOS pathway of inducible repair (see Little and Mount,1982 for a review).

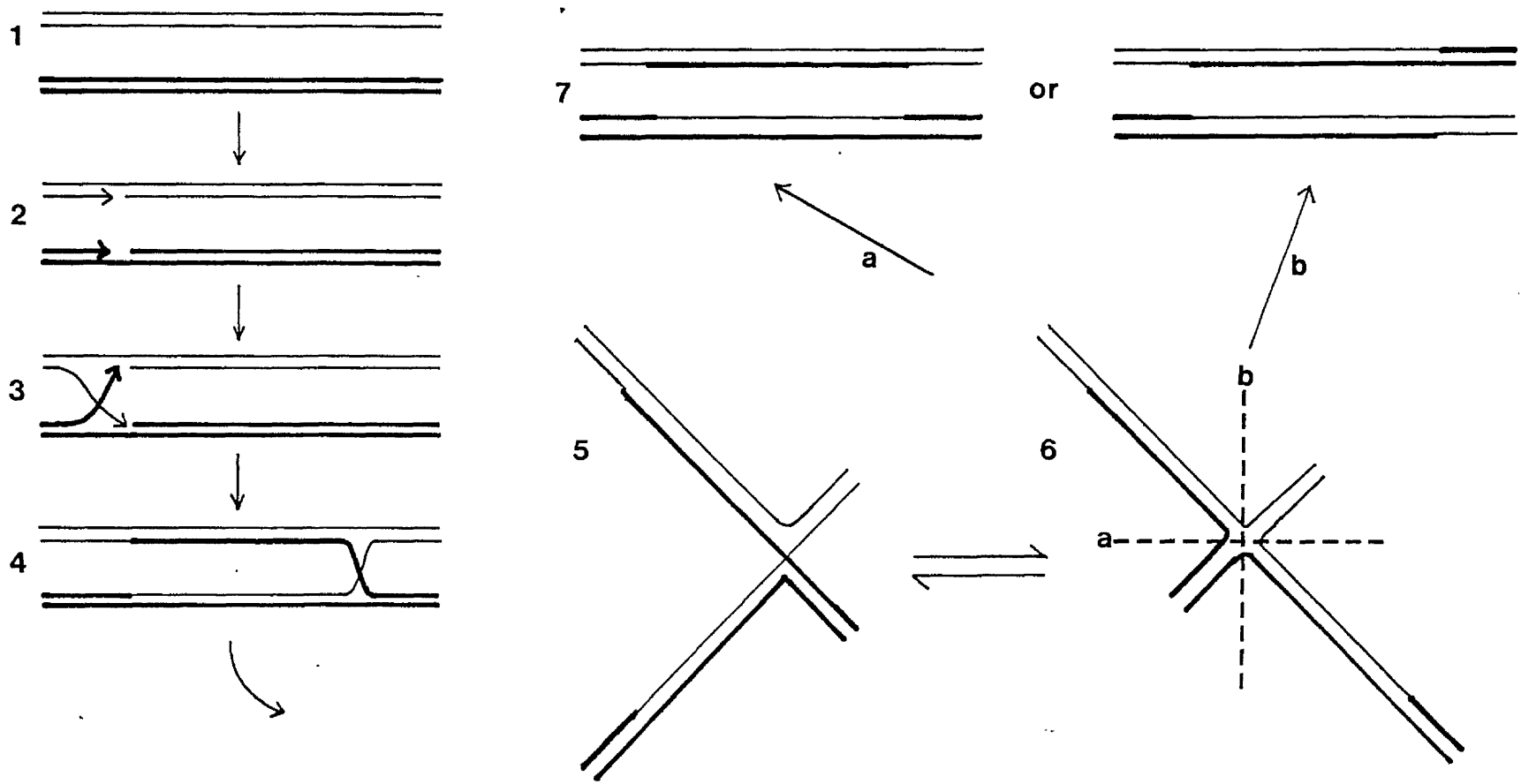

1
2
3
4
7
or
a
b
5
6
b
a

FIGURE 1.3 Holliday recombination model

1. Homologous DNA helices are aligned.

2. In each, a single-stranded nick is made in both positive (or negative) strands.

3. The free ends created dissociate from their complementary strands and instead associate with the complementary strands in the homologous double helices. The reciprocal strand invasion may be stabilized by ligation of ends.

4. The crossover point is not static and may move in either direction by rotary diffusion of the double helices, this results in heteroduplex formation.

5,6. Isomerization of the "Holliday intermediate" into another planar form makes visualization of the outcome easier (Sigal and Alberts, 1972).

7. Cutting on different axes releases recombinant molecules in which, on either side of the potentially heterozygous region, the parental alleles are either conserved in their original linkage or are reciprocally exchanged.

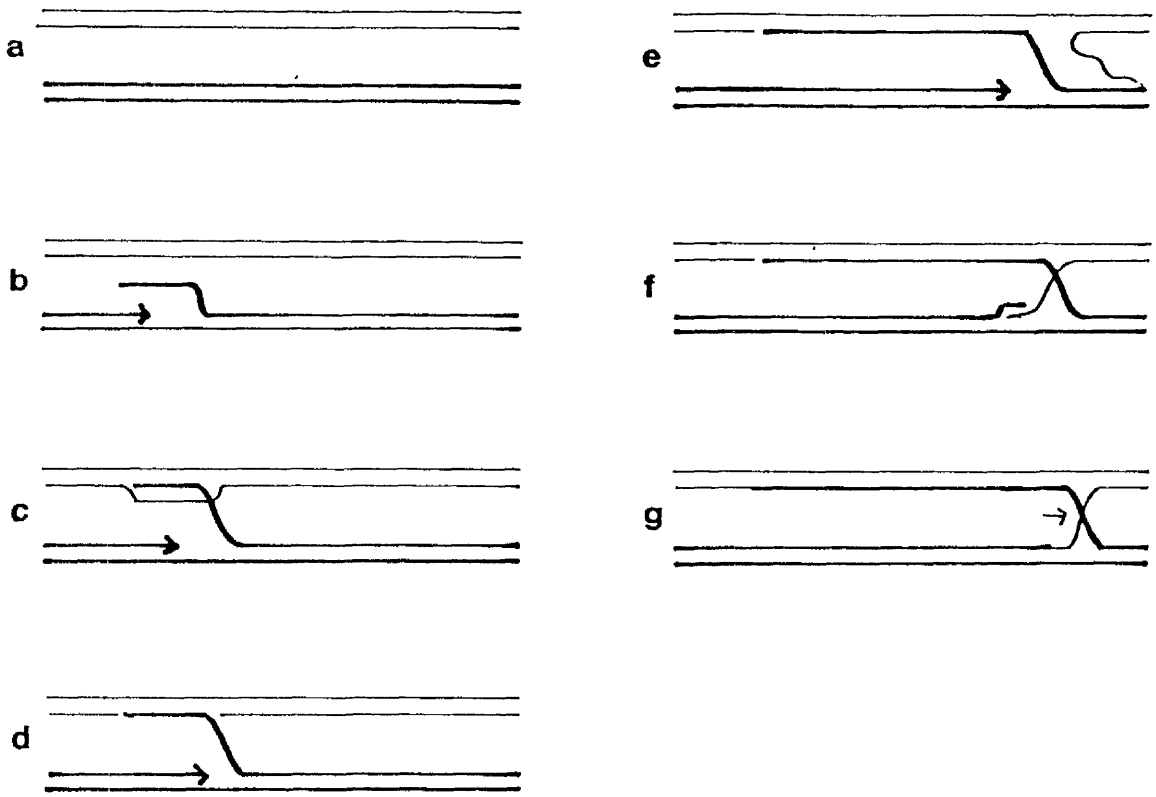


FIGURE 1.4 Meselson-Radding recombination model

a. Homologous DNA helices are aligned.

b. A single-stranded break in one molecule becomes the site of strand displacement by DNA polymerase.

c. The displaced strand invades the homologous double helix, forming a D-loop.

d.The unpaired DNA in the D-loop is removed by nucleases.

e. Branch migration is achieved by the action of DNA polymerase displacing the strand thereby forming a long heteroduplex tract in only one of the participating duplexes.

f. The strand displaced by the invading strand may itself base pair onto the complementary strand of the homologous helix thus forming a symmetrical heteroduplex joint.

g. Ends of strands are ligated to form a stable Holliday intermediate which may then be subject to the same processes of branch migration and isomerization as previously described.

The fundamental role of the *recA* gene product in many cellular processes has prompted several independent groups to clone the *recA* gene and purify it s product with the intent of studying it s action *in vitro* (Weinstock *et al.*,1979; Shibata *et al.*,1979b). The protein has a molecular weight of 37,800 dal, cleaves the lambda cI and *lexA* repressor proteins *in vitro* (Roberts *et al.*,1979; Little *et al.*,1980), has ATPase activity and also, more importantly, participates in a number of events associated with recombination.

Key features of recombination models are recognition and alignment of homologous DNA molecules followed by heteroduplex formation. There is substantial evidence that purified RecA protein promotes homologous pairing and heteroduplex formation *in vitro*. The protein binds most avidly to single-stranded DNA, in the presence of ATP, to form stable complexes. In the presence of double-stranded DNA unwinding of the duplex occurs whether the bound single strand is heterologous or homologous (Shibata *et al.*,1979a;b). It has been proposed that the initial unwinding promoted by the single-stranded DNA/RecA protein complex may be in search of homology (Shibata *et al.*, 1979a). When a region of the unwound duplex homologous to the single strand is found then pairing takes place to form a heteroduplex joint containing the new strand rather than rewinding of the original helix. The displaced strand from the duplex can be visualized as a D-loop by electron microscopy or assayed by retention on a nitrocellulose filter (McEntee *et al.*,1979; Shibata *et al.*,1979a;b;Das-Gupta *et al.*,1980). The single-stranded DNA molecule which stimulates homologous pairing may be circular or linear, suggesting that a free end is not always required, but a free end is required for stable heteroduplex formation. Stable heteroduplexes can be formed from

intact duplex molecules and single-stranded circular molecules in the presence of RecA protein and topoisomerase I (Cunningham *et al.*, 1981). The heteroduplex joint formation promoted by RecA protein can extend for several thousand base pairs due to branch migration of the crossover point; this reaction appears to be driven by ATP and is polar (Cox and Lehman,1981; Kahn *et al.*,1981; West *et al.*, 1981b). The heteroduplex always extends from 3' → 5' direction, displacing a strand from the duplex with a 5' terminus. Branch migration is halted by extensive mismatching of base pairs (more than 25%), but may pass through some mismatches; this observation may be important when considering gene conversion (DasGupta and Radding,1982). The heteroduplex can also pass through regions containing pyrimidine dimers, though this reaction proceeds at $1/50^{th}$ the normal speed (Livneh and Lehman,1982). Insertion of short DNA sequences (about 700bp) also halts branch migration (DasGupta and Radding,1982). These experiments highlight the importance of the RecA protein in the initial events of homologous pairing and heteroduplex formation leading to genetic recombination.

The major recombination pathway in *E.coli* involves the products of both the RecA gene and the RecBC gene (Clark,1974). The protein encoded by the RecBC gene has been purified; it is an ATP-dependent exonuclease, named *exoV*. *In vitro*, *exoV* binds to single-stranded DNA, can unwind and rewind duplex DNA, has endonuclease activity and degrades free ends to oligonucleotides (Telander-Muskavitch and Linn, 1980). One of the more interesting observations made with *exoV* is it's ability, under certain conditions, to bind at the ends of a linear duplex, track along the DNA in the presence of ATP, unwinding and rewinding the helix. As the rate of unwinding exceeds that of rewinding, movement of *exoV* along the duplex results in the formation

of single-stranded loops; these are visualized as double loops or loops and tails in the electron microscope (Taylor and Smith,1980a; b; Telander-Muskavitch and Linn,1980). The single-stranded loops thus formed could be instrumental in the initial synapsis event. This data does, however, conflict with genetic evidence of a role for *exoV* late in recombination (Birge and Low,1974). One cannot preclude the possibility that *exoV* may have roles both late and early in recombination depending on the substrate.

Another protein which has been directly implicated during recombination is the T4 gene 49-encoded endonuclease , *endoVII*. This has been shown to symmetrically cleave artificially constructed Holliday intermediates *in vitro* in a manner which would resolve recombinational intermediates into progeny molecules (Mizuuchi *et al.*, 1982). The *in vitro* data confirms genetic evidence for a role for this gene product late in recombination (Kemper and Janz,1976).

In contrast to general genetic recombination, site-specific recombination is independent of *recA* function. The proteins which mediate these events are often encoded by genes adjacent to one of the recombination sites. The reaction is usually by a reciprocal mechanism of breakage and reunion of strands without concomitant DNA synthesis or extensive heteroduplex joint formation. The prototype model for site-specific recombination is integration/excision of bacteriophage lambda with the *E.coli* chromosome (see figure 1.1; Campbell,1962). Extensive studies *in vivo* and *in vitro* have identified a 240bp phage attachment site, *attP* and a 25bp host site, *attB* (Landy and Ross,1977; Hsu *et al.*, 1980; Mizuuchi *et al.*, 1981). These sites contain a common core sequence of 15bp within which the crossover occurs; the crossover point is unique to each strand and

is staggered by 7bp (Nash,pers.comm.)

Nash has developed an *in vitro* system to study the molecular mechanism of this recombination event and as a means of identifying and purifying the proteins required (Nash,1975). The integration reaction requires the presence of the phage-encoded Int protein, plus the products of the *himA* and *hip* host genes which together form the integration host factor,IHF (Nash and Robertson,1981). These proteins have been purified to near homogeneity and shown to bind specifically to *att* sites (Kikuchi and Nash,1978; Ross *et al.*,1979; Hsu *et al.*, 1980; Mizuuchi *et al.*,1981; Nash and Robertson,1981; Craig and Nash, pers.comm.) The Int protein has also been shown to have the property of breaking and resealing DNA under certain conditions, which has led to it s classification as a type I topoisomerase (Kikuchi and Nash,1979b; Wang *et al.*, 1980).

Enzymes which catalyze the conversion of a particular DNA topoisomer into another,by removal or addition of superhelical turns, have been named topoisomerases. The number of supertwists in a covalently closed circular DNA molecule is always an integer, this value has been called the linking number. Topoisomerases change the linking number by cutting and rejoining DNA strands in a concerted manner. Type II topoisomerases, such as DNA gyrase, change the linking number in steps of two by mediating a double-stranded cleavage in the helix, passing an intact segment of DNA through the cleaved strands, followed by ligation of the strands. Type I enzymes have a similar reaction mechanism, but induce single-stranded cleavage thus changing the linking number in steps of one (see Cozzarelli,1980 and Gellert,1981 for reviews).

The apparent lack of specificity for *att* sites shown by Int topoisomerase is difficult to reconcile with results from DNA binding experiments. This may be attributable to the high Int concentrations required for topoisomerase activity, which may result in nonspecific binding, or could be due to recognition of a short sequence which could conceivably be widely distributed.

The preferred substrate for *in vitro* recombination is s/c DNA, though under appropriate ionic conditions o/c substrates are converted to products; this reaction is very inefficient (Pollock and Abremski, 1979; Pollock and Nash,1980). For the intermolecular reaction only *attP* is required within a s/c molecule, *attB* can be located within a relaxed circular molecule or linear DNA molecule (Mizuuchi and Mizuuchi,1979). The superhelicity of participating DNA molecules appears to be highly conserved during the reaction, suggesting that *att* sites are paired and held in place by recombination proteins during strand exchange. Nash has proposed a four-stranded intermediate structure for recombination, similar to the model proposed by McGavin (1971), in which one strand from each parent is simultaneously cleaved, followed by limited rotation around uncut strands which brings cleaved strands adjacent to new partners for the ligation step. This process would have to be repeated on the remaining two parental strands for full recombination; as yet no intermediate structure has been detected in which the reaction has only occurred on two strands. As Int exhibits type I topoisomerase activity this protein is proposed to mediate the cleavage/ligation reactions during recombination, whereas IHF may play an important role in alignment of strands prior to recombination.

In addition to Int and IHF the excision reaction requires the

presence of the Xis protein, the product of the phage-encoded *xis* gene. The substrates for this reaction are the hybrid sites, *attR* and *attL*, formed by recombination between *attP* and *attB*. The differing requirements for the integration and excision reactions ensures that after infection the phage is either stably integrated, until induced, or enters the lytic cycle.

Lambda integrative recombination represents the best characterized model for site-specific recombination, due largely to the use of an *in vitro* system. Although many of the other site-specific recombination systems recognized in prokaryotes and eukaryotes have yet to be subjected to extensive biochemical analysis, the limited genetic data suggests differences in substrate preference, efficiency of recombination and dependence on host-encoded proteins. These differences may be significant when thinking in terms of the biological roles of the systems and the reaction mechanisms involved.

The two closely related systems, *Salmonella* phase variation and G-loop flipping in the bacteriophages Mu and P1, involve site-specific inversion of a DNA segment which regulates expression of adjacent genes. In *Salmonella* two genes, H1 and H2, encode the major flagellar structural proteins. These two genes map apart on the *Salmonella* genome; cells express either one of these two genes depending on the orientation of a segment of DNA adjacent to the H2 structural gene. In one orientation the H2 gene is expressed as well as a repressor protein which prevents transcription of the H1 structural gene (Simon *et al.*, 1980). A *trans*-acting function, *hin*, maps within the invertible region; this acts on 14bp inverted repeat sequences bounding the 970bp inversion segment to mediate recombinational gene switching. The *hin* protein has no apparent activity on directly

repeated sites. An analagous system is encoded by bacteriophage Mu; in this case the expression of cell surface proteins, which determine the host range of the phage, are controlled by inversion of a DNA segment, designated the G loop (Bukhari and Ambrosio,1978; Kamp *et al.*,1978). Inversion of this DNA sequence is dependent on the *gin* gene product, which acts at 14bp inverted repeats flanking the G loop (Kamp *et al.*,1979). The Hin and Gin proteins have been shown to have interchangeable functions (Kamp and Kahmann,1981). These two systems are active only on inverted repeats, recombine at low frequency, and appear to require only the *hin* or *gin* gene products.

Site-specific recombination has only recently been recognized in eukaryotes; limited analysis has shown the presence of both inversion and deletion systems dependent on short sequence homologies. The 2μ circle plasmid in *Saccharomyces cerevisiae* encodes a site-specific recombination system; the protein which mediates this event, FLP, and the sites of inversion, are both located on the plasmid. Host-encoded factors may also be involved. The function of DNA sequence flipping in this case is unknown (Broach *et al.*, 1982). The production of antibody genes is known to occur through several recombination steps from germline to somatic cells. These events involve short sequence homologies as substrates for as yet unidentified proteins leading to maturation of antibody genes (Tonegawa *et al.*, 1981; Leder,1982).

The systems described in this section involve recombination between short, specific,homologous DNA sequences. A number of other recombination events involving short DNA sequences are recognized in *E.coli* ; an example of this is in the *lacI* structural gene where several mutation hotspots have been identified. The formation of

spontaneous mutations in the *lacI* gene is mainly attributable to short deletions. These occur at short DNA sequence homologies (5-15bp long) in *recA*$^+$ or *recA*$^-$ cells (Farabaugh *et al.*,1978; Albertini *et al.*,1982). Also included in this category is precise excision of transposable elements which occurs at low frequency by recombination between short sequence duplications (3-12bp long) flanking the integrated element (Foster *et al.*,1981; Kleckner,1981). This process is independent of the *recA* gene and may require the presence of the host-encoded *himA* gene. These events involve short sequence duplications, but do not require specific DNA sequences.

Movement of genetic elements represents another kind of recombination event; they insert at new loci with no apparent homology to the element. Such events have gross phenotypic consequences leading to insertional mutations, deletions, inversions and duplications. If these elements contain regulatory signals then the possibility exists for switching on or off the expression of adjacent genes. It was precisely this observation which identified "controlling elements" in maize (McClintock,1952;1957). McClintock's results suggested that "controlling elements" could move from loci to loci affecting the expression of many diferent genes, thus exerting clear phenotypic effects.

It was not until the late 1960's that many polar mutations in the *gal* and *lac* operons in *E.coli* were realized to be due to insertions of DNA (Malamy,1966; Adhya and Shapiro,1969). On further analysis, by physical means, it was realized that many of these insertions were of distinct size classes, ranging from 700 to 1500 bp (Starlinger and Saedler,1972). It became apparent that "insertion sequences" (IS elements) were of non-permuted length and moved as discrete

units into many sites. In the early 1970's many of the antibiotic resistance genes, known to be spread at high frequency, were identified as transposable elements (Hedges and Jacob,1974). This led the way to studying these elements in more detail as their easily selected phenotypic markers were more readily followed through populations than the simple insertion elements.

All transposon mediated rearrangements occur in the absence of *recA* function, suggesting that the elements themselves encode determinants for their own transposition. Deletion analysis has indicated the presence of at least one element-encoded protein required for transposition. Host functions are almost certainly required also. The advent of rapid DNA sequencing techniques has revealed the complete nucleotide sequence of many of the small elements and their integration sites. The most striking feature revealed by sequence data is the ubiquity of inverted repeat sequences at the ends of elements. The ends of elements are essential for transposition. All transposable elements mediate a small duplication of the target sequence during integration; these small direct repeats flank the integrated element. The repeated sequence ranges from 3-12bp and is characteristic for each element; 5bp and 9bp duplications are the most common (Grindley,1978; Johnsrud *et al.*, 1978; Kleckner,1979; Ohtsubo *et al.*,1979). The duplication of target site DNA is not required for subsequenttransposition (Kleckner,1979).

There is strong evidence that when a transposable element inserts at a new site, an intact copy of the element remains in the original position. However, this observation cannot be used as conclusive evidence for replication during transposition as there

is always more than one chromosome present during the cell cycle, therefore one cannot rule out simple excision/reintegration. All elements appear to mediate replicon fusion, or cointegration. The donor replicon interacts with a second genome, lacking the element, to generate a structure containing both replicons joined by directly repeated copies of the element (Toussaint and Faelen,1973; Gill *et al.*,1978; Arthur and Sherratt,1979). This observation directly implicates replication in the transposition process. Replication is specific to the element and the small target sequence which is presumably duplicated as part of the transposition event.

Comparison of sequence homologies, structural and functional organization, and proposed mechanisms has divided the prokaryotic transposable elements into three distinct groups (table 1.1; see Kleckner,1981 for further details).

The class I elements consist of the simple insertion sequences, which code for only one or two structural genes involved in their transposition and regulation, and composite elements which consist of accessory determinants flanked by IS elements. IS*1-5* and IS*102* are independent elements, but IS*10*, IS*50*, and IS*903* are found only in composite elements, where they flank antibiotic resistance genes. The IS modules of composite elements mediate the transposition events and subsequent rearrangements caused by these elements. Mutations within IS elements lead to a transposition deficient phenotype, the mutant element can be complemented to transpose, but the frequency with which this occurs is greatly reduced. This is thought to be due to a high preference for *cis* action by the transposase proteins (Joyce and Grindley,1981; Kleckner,1982), which have often been compared, mechanistically, to the ϕX174 *cisA* protein (Arthur and

CLASS 1A. IS-like elements: individual modules

Element	Length (bp)	Target repeat	Terminal IR	Possible coding region- no. amino acids	Source
IS*1*	768	9	18/23	70,90	F,R100 *
IS*2*	1327	5	32/41	315	*E.coli* K12
IS*3*	1300	3 or 4	32/38	n.t.	*E.coli* K12,F
IS*4*	1426	11 or 12	16/18	422,131	*E.coli* K12
IS*5*	1195	4	15/16	338,326, 118,108	*E.coli* K12
IS*10*-R	1329	9	17/22	402,62,54	Tn*10* in R100
IS*50*-R	1531	9	8/9	480,476	Tn*5* in JR67
IS*903*R,L	1050	9	18/18	370,114	Tn*903* in R6
IS*102*	≃1000	9	18/18	n.t.	pSC101^{ts}

* present on the chromosome of *E.coli* K12 and many other enterobacteria

CLASS 1B. IS-like elements: composite elements

Element	Length (bp)	Module at ends	Orientation of flanking module	Associated det.
Tn*681*	2.1	IS*1*	inverted	heat-stable toxin
Tn*2571*	23	IS*1*	direct	Cm^r,*fus*^{r},Sm^r,Su^r,Hg^r
Tn*9*	2.5	IS*1*	direct	Cm^r
Tn*2350*	10.5	IS*1*	direct	Km^r
Tn*10*	9.3	IS*10*	inverted	Tc^r
Tn*5*	5.7	IS*50*	inverted	Km^r
Tn*903*	3.1	IS*903*	inverted	Km^r

TABLE 1.1 Classification of transposable elements (from Kleckner, 1981)

CLASS II. Tn*A* family

Element	Length (Kbp)	Target repeat	Terminal IR	Associated det.	Structure
Tn*3*	4.957	5bp	38/38	Ap^r	|end.*tnp*A.*tnp*R.Ap^r.end|
Tn*1*,Tn*2*,	5	5bp		Ap^r	
Tn*801*,Tn*802*					
Tn*901*,Tn*902*					
Tn*401*,Tn*1701*					
Tn*2601*,*2602*					
γδ	5.8	5bp	36/37	?	|end.*tnp*A.*tnp*R. ? end|
IS*101*	0.209	5bp	31/37	none	sites, no functions?
Tn*501*	8.2	5bp	35/38	Hg^r	|end.Hg^r.*tnp*R.*tnp*A.end|
Tn*1721*,*1771*	11.4	5bp	35/38	Tc^r	|end.*tnp*R.*tnp*A.end ....Tc^r.part *tnp*A.end|
Tn*551*	5.3	5bp	35/35	ery^r	
Tn*951*	16.5	5bp	n.k.	lactose utiliz.	

TABLE 1.1 (contd.) The class III elements consist of the transposable bacteriophages,eg. Mu and D108

Sherratt,1979; Kleckner,1981). The transposition frequencies of composite elements are generally low and show length dependence (Chandler *et al.*,1982; Kleckner,1982). This may be due to the length of DNA which has to be replicated, thus increasing the chances of nuclease attack on exposed single-stranded regions, or due to instability of the transposase complex. Class I elements may generate cointegrates or proceed directly to transposition end-products obviating the replicon fusion step; the frequency for each type of event is characteristic for each element. Cointegrates, when formed, are stable in Rec$^-$ hosts, suggesting that these elements do not encode a system to reduce cointegrates to the normal transposition end-products. The host-encoded homologous recombination system can resolve transpositional cointegrates, albeit inefficiently, thus causing deletions between directly repeated elements. The deleted region carries a single copy of the element and may integrate by recombination into a region of the genome containing a homologous copy of the element. Intramolecular transposition results in deletion or inversion of intervening DNA (see figures 1.6 and 1.7 later). Inverted repeated copies of an IS element surrounding a unique sequence allow the entire region to transpose, thus forming a composite element. Homologous recombination across the inverted repeated modules of composite elements merely results in inversion of the accessory determinant in relation to outside markers. Composite elements flanked by directly repeated modules are always in danger of excision by recombination across the IS elements, though of course these elements also have the potential to amplify accessory determinants by unequal exchange (see figure 1.2).

Analysis of integration sites indicates some insertion specificity, particularly for the Tn*10* element. Tn*10* has "hot sites" for insertion,

a concensus sequence, GTCNAGC, has been identified by analysis of preferred integration sites (Halling and Kleckner,1982). Precise excision is rare and presumably occurs by recombination across short direct duplications of target site DNA flanking the integrated element (Foster *et al.*,1981). Precise excision may be affected by some host mutations, eg. *himA*, which has already been identifed as one of the components of IHF required for lambda integrative recombination. The *himA* gene has also been shown to regulate the expression of several lambda genes involved in establishment of lysogeny (Miller, 1981). The HimA protein could play a direct role in precise excision, or indirectly by affecting the expression of other genes (Kleckner,1982). The *ferA* gene, present on the F factor, appears to elevate the frequency of precise excision and also increases the rate of general genetic recombination between plasmids (Doherty *et al.*,1982) . This may be due to a stimulation of recombination proteins when single-stranded DNA is present in cells, ie. during transfer (Hopkins, unpubl. results).

There are no apparent sequence homologies between class I elements, or interaction between transposition determinants. Several duplicate 9bp on insertion, but there are many variations. There are no strong homologies between inverted repeat sequences of elements within this group. These elements may have evolved independently or have diverged considerably from a common ancester.

The class II elements have been classified according to their strong similarities in mechanistic and functional properties, as well as sequence homologies, which together suggest a common evolutionary origin. All elements generate 5bp duplications at the target site on integration, have 35-40bp inverted repeats, which share some

homology between elements, encode transposition functions and, usually, accessory determinants (see table 1.1). Of these the ampicillin resistant element, Tn*3*, is the best characterized. The structure and functions of this element have been revealed by DNA sequencing and analysis of a series of deletion and insertion mutants (Heffron *et al.*, 1977; 1978; 1979; Arthur and Sherratt, 1979; Sherratt *et al.*, 1981a). As for class I elements, both ends of the element are required in *cis* for transposition as well as the product of the *tnp*A gene, transposase, which is thought to act specifically at the ends of the element. Cointegrates appear to be obligatory during the transposition process, but, unlike class I elements, these are unstable in *recA*$^-$ hosts (Arthur and Sherratt, 1979; Kitts *et al.*, 1982a;b). Resolution of cointegrates is dependent on the product of the *tnp*R gene, resolvase, and occurs by recombination at a specific site, *res*, located within the element. These elements appear to have evolved a site-specific recombination system to efficiently resolve transpositional cointegrates.

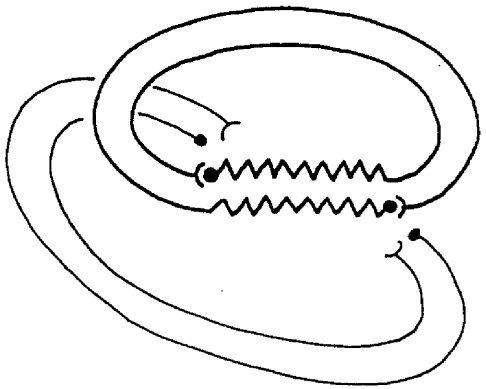
The precise excision of these elements has not been extensively studied, though it appears to occur very infrequently. There is no apparent specificity for insertion sites; integration generally occurs into AT-rich regions (Tu and Cohen, 1980).

Of the class III elements, bacteriophage Mu has been studied in most detail. Mu undergoes transposition at high frequency in the host, causing all the genetic rearrangements characteristic of a transposable element. Replicative transposition is the normal mechanism for vegetative propagation during the lytic cycle. The sequences required for transposition, demonstrated by deleting the central region of the phage to form "mini-Mus", are located within 1Kb of

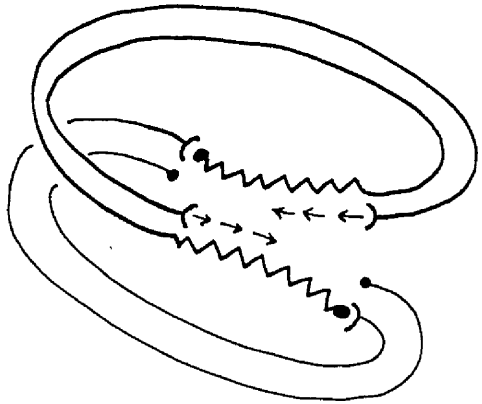
the lefthand end and 0.15Kb of the righthand end (Coehlo *et al.*, 1981). Only the ends of the element and two phage-encoded products are required for transposition. The A gene is required for all transposition processes; A mutants are defective for Mu integration, replication, deletion and cointegrate formation. B mutants retain 1-10% of wild type activity. The A gene product is intimately involved in replication of the phage, and is required in stoichiometric amounts. There is also evidence that the A gene product is unstable (Pato and Reich,1982).

The apparent replicative nature of transposition has prompted a number of models to explain these observations (Grindley and Sherratt, 1979; Arthur and Sherratt,1979; Shapiro,1979). All the models demand that the transposase protein recognizes the ends of the element. Single-stranded nicks are made at either the 5' or 3' ends of the element, followed by covalent attachment of the cleaved element strands to transposase, in analogy with the ϕX174 *cisA* protein and conjugal mobilization (Eisenberg *et al.*, 1977; Warren *et al.*, 1978). A staggered cut is made at the target site; the length of this stagger determines the size of the duplication made and is characteristic for each element. The ends of the element are then ligated to the staggered break at the target site (see figure 1.5; Arthur and Sherratt,1979; Shapiro,1979). If the 5' termini are used for the initial cleavage event, then the exposed 3' hydroxyl ends provide primers for DNA replication. This series of events results in replicon fusion; the cointegrate can then be resolved by recombination across the directly repeated copies of the element to generate the normal transposition end-products. If intramolecular transposition occurs by this pathway, inversions or deletions result depending on which strands at the target site are cleaved and ligated

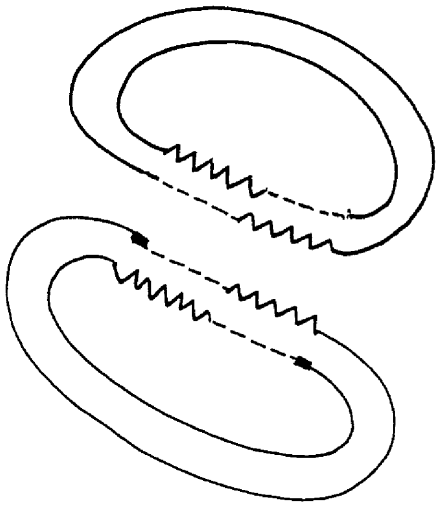
a



b



d



c

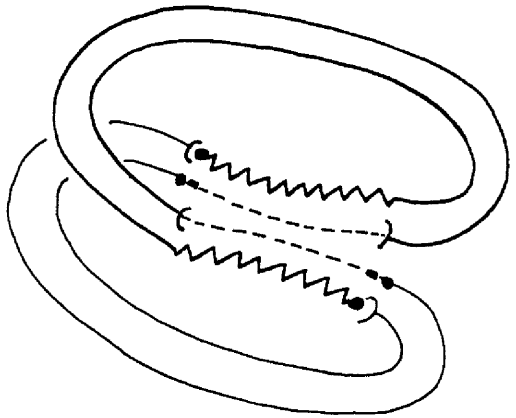


FIGURE 1.5 Symmetrical model for DNA transposition

a. Single stranded cuts are made at both ends of the transposable element (5' or 3') followed by covalent attachment of the transposase protein to the exposed termini. A staggered cut is made at the target site.

b. The ends of the element are ligated to the ends of the staggered cut at the target site, presumably using the energy conserved in the phosphodiester bond between the protein and DNA. Replication can proceed from the exposed 3' hydroxyl groups to duplicate the entire element and the small stagger at the target site.

c. Semiconservative replication has generated a new transposable element, plus a direct duplication at the target site, resulting in cointegrate formation.

d. Site-specific recombination between the two transposable elements generates the normal transposition end-products, ie. the donor replicon intact and the recipient replicon carrying a copy of the element flanked by short,direct repeats.

●— represent 5' termini

)— represent 3' hydroxyl groups

FIGURE 1.6 Consequences of intramolecular transposition

to the element (see figure 1.6).

Although cointegrates appear to be obligatory intermediates during transposition of class II elements, many elements appear to transpose directly to end-products, alleviating the necessity for a cointegrated intermediate. To explain how direct transposition may occur, asymmetric models have been proposed in which only one strand of the element is initially cleaved/ligated to the target site followed by element specific replication in a rolling circle mode (Galas and Chandler,1981; Harshey and Bukhari,1981). The outcome of these events is direct transposition or cointegration depending on which strand is used for the second cleavage/ligation step (see figure 1.7). For direct transposition the transposase protein must be able to recognize and cleave both 3' and 5' ends of the element, a step which I find difficult to reconcile. The introduction of *in vitro* systems should shed some light on the molecular mechanisms involved in transposition.

In eukaryotes transposable elements were first studied in maize by McClintock (1952;1957); they have since been identified and characterized in a number of organisms including *Saccharomyces cerevisiae* and *Drosophila melanogaster*. The *copia*-like family of *Drosophila* elements, including *copia*, 412, 297 and *mdg*, and the Ty elements of yeast have many similarities. They are generally 5-7Kb in length, have terminal direct repeats of several hundred bp with small (about 10bp) inverted repeats at the extremities, are present at about 30 copies per genome and have the ability to transpose randomly throughout genomes (Cameron *et al.*,1979; Finnegan, 1981). The *copia* and Ty1 elements generate 5bp duplications on insertion (Farabaugh and Fink,1980; Rubin *et al.*,1981). The 338bp terminal

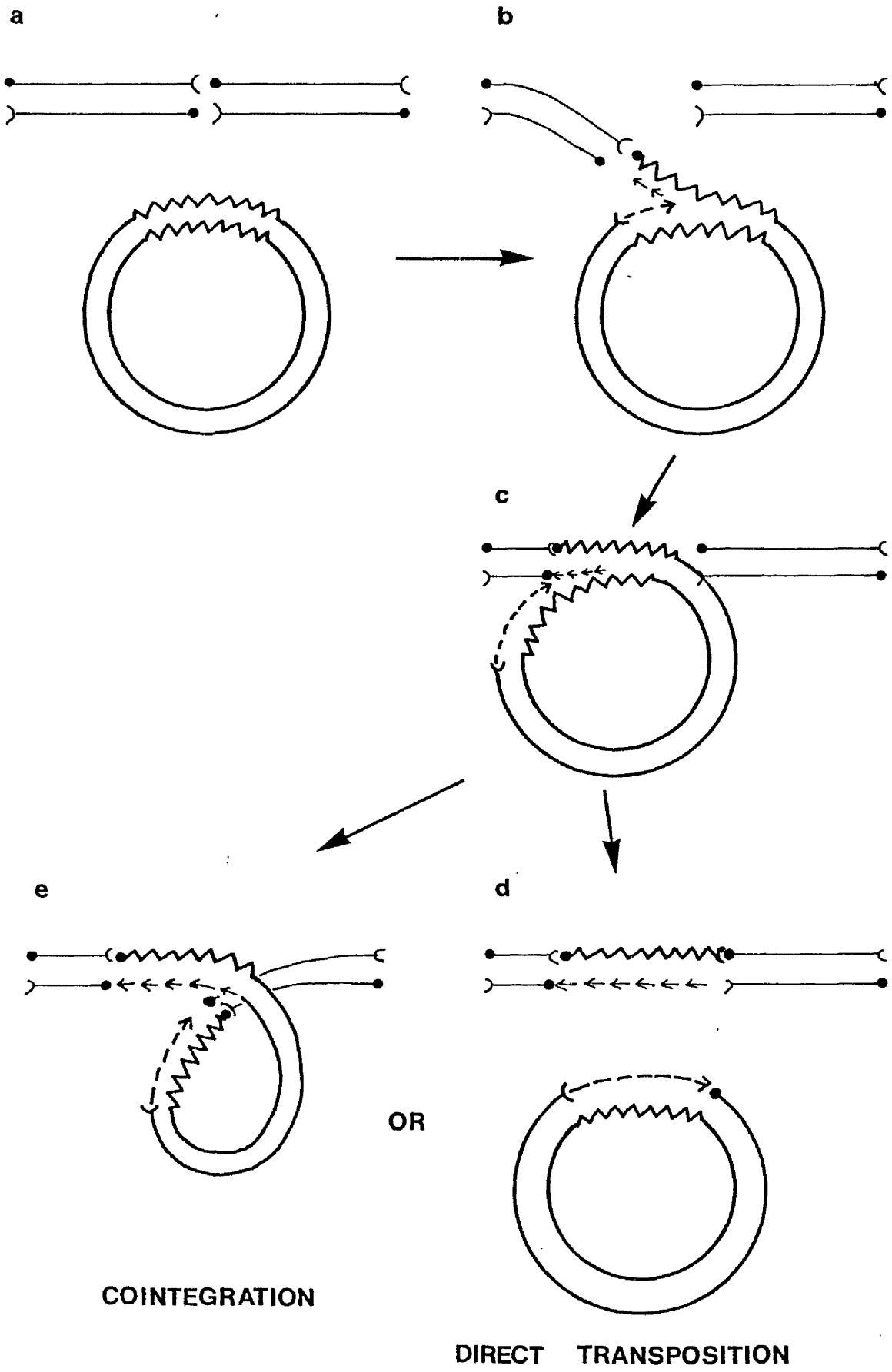

a
b
c
d
e
OR
COINTEGRATION
DIRECT TRANSPOSITION

FIGURE 1.7 Asymmetric model for DNA transposition

a. The transposase protein recognizes and brings together the ends of the element and the small target sequence. The target site undergoes a staggered cut.

b. One strand of the transposable element, shown as the 5' end, is cleaved and ligated to the target site. Replication of the element is initiated at one of it s ends from the exposed 3' hydroxyl group; complementary strand synthesis proceeds from the replication fork. The free target strand is held in place by the transposase/replication complex.

c. Rolling circle DNA replication proceeds through the entire element and small stagger at the target site.

d. Replication terminates when the other end of the element passes through the replication complex. If the 3' end of the element, opposite to the end ligated to the target site, is cleaved/ligated to the free 5' end at the target site direct transposition results.

e. If the transposase/replication complex cleaves the 5' end of the other element strand and ligates it to the exposed 3' end at the target site then cointegration results.

direct repeats of Tyl, designated δ, are found alone in yeast genomes; this is probably due to homologous recombination excising the internal region of the element rather than independent transposition of the δ sequence (Fink *et al.*, 1981). Surprisingly the direct repeat sequences of *copia* are rarely, if ever, found independent of the main body of the element (Rubin *et al.*, 1981). There is no direct evidence that eukaryotic elements replicate during transposition, though presumably they do, due to their wide dispersal throughout genomes. In some cell lines they may represent 10% of the genomic DNA. The *copia* element has recently been compared, structurally, with retroviruses (Finnegan, 1981).

Retroviruses are a family of RNA animal viruses which replicate through a DNA intermediate. On infection the linear RNA molecule is reverse transcribed; the DNA intermediate circularizes, then integrates into the host genome, generating 4bp duplications on insertion (Shimotohno and Temin, 1981). The proviral DNA contains long terminal repeat sequences (LTR) that flank the main body of the element. Each LTR contains small inverted repeats at it s termini; the LTR sequences have been shown to share some sequence homology with the *copia* element. It has been suggested that retroviruses have evolved from transposable elements, although integrated retroviruses have not been observed to transpose (Shoemaker *et al.*, 1981; Shimotohno and Temin, 1981).

The structural similarities between the *copia* and Tyl elements, and retroviruses has led to speculation that a mechanism similar to the retroviral replicative pathway may be used by eukaryotic elements as a means of transposition. Extrachromosomal circular copies of *copia* have been found in cultured *Drosophila* cells; these could of

course arise by homologous recombination across the directly repeated flanking sequences of the *copia* element rather than as replicative transposition intermediates (Flavell and Ish-Horowicz, 1981).

Transposable elements obviously play an important role in linking nonhomologous DNA. They may also mediate rearrangements by creating portable regions of homology on which the general genetic recombination system can act, resulting in deletions, inversions and duplications (see figure 1.2). Studies on the rearrangements mediated by transposable elements have revealed the fluidity of genome structure, but little is known of the mechanisms by which they occur. The introduction of *in vitro* systems to study the transposition process, in conjunction with genetic analysis, should reveal the precise molecular mechanisms involved and the proteins required for transposition and integration of viruses.

Our group at Glasgow University is currently analyzing transposition of the class II elements , Tn*3* and Tn*501*. These two elements both encode site-specific recombination systems to resolve transpositional cointegrates. The work in this thesis describes a detailed investigation into the Tn*3* encoded site-specific recombination system by genetic and biochemical techniques.

CHAPTER 2

MATERIALS AND METHODS

Name	Relevant markers	Source
AB1157	*arg, his, leu, pro, thr, lac, gal, ara, rpsL* (=Sm^r)	D.J.Sherratt
AB2463	*arg, his. leu, pro, thr, lac, gal, ara, rpsL* (=Sm^r), *recA*-13	A.J.Clark
JC9239	*arg, his, leu, pro, thr, lac, gal, ara, rpsL* (=Sm^r), *recF*-143	A.J.Clark
JC5466	*his, trp, recA*-56	F.C.Cannon (Arthur and Sherratt,1979)
DS825	*minA, minB, rpsL* (=Sm^r)	D.J.Sherratt
CSH52	*ara*, Δ*(lac, pro), strA, thi recA*	CSH strain collection (Miller,1972)
DS916	JC5466 rif^r isolate	D.J.Sherratt
AA411	Jc5466, R388	A.Arthur
N4830	λCI^{857} containing strain	R.Reed (Reed,1981b)
LS226	JC5466, R388::pRR12	L.Symington, (Chapter 3)
LS415	N4830, pLS213 - *tnp*R over-producing strain	L.Symington, (Chapter 4)

TABLE 2.1 Bacterial strains

1. Organisms

All bacterial strains, listed in table 2.1, are derivatives of *Escherichia coli* K-12. The plasmids used are shown in table 2.2.

2. Chemicals and Enzymes

All were obtained from BDH, Oxoid, Difco, Sigma or BRL except the following:

Chemical	Source
Sephadex G-75, G-150	Pharmacia Fine Chemicals
CM-Sepharose	Pharmacia Fine Chemicals
Pd/Pt wire	Agar acids
Piperidine	Fleska
Radiochemicals	Amersham (courtesy of Virology department)

3. Basic Media

L-broth: 10g tryptone, 5g yeast extract, 5g NaCl, 1g glucose; made up to 1 litre with distilled water, adjust to pH 7.0 with NaOH.

Nutrient agar: 25g Oxoid No.2 NB, 12.5g agar; made up to 1 litre in distilled water.

Iso-sensitest agar: 23.4g IS medium, 12.5g agar; made up to 1 litre in distilled water.

Davis-Mingioli Salts (X4): 28g K_2HPO_4, 8g KH_2PO_4, 4g $(NH_4)_2SO_4$, 1g Na_3citrate, 0.4g $MgSO_4.7H_2O$; made up to 1 litre with distilled water. For minimal agar 100ml were added to 300ml molten 2% agar, plus the appropriate supplements as necessary.

Phage buffer: 7g Na_2HPO_4, 3g KH_2PO_4, 5g NaCl, 0.25g $MgSO_4.7H_2O$, 0.05g $CaCl_2$; made up to 1 litre with distilled water.

Supplements: Where required growth supplements were added at the following concentrations- amino acids at 40µg/ml, thymine at 20µg/ml, thiamine(B1)

(a) Conjugative plasmids

Name	Derivation	Phenotype	Source
F'*lac,pro*	naturally occurring, IncF1	Lac^+,Pro^+,Tra^+	D.Leach (Miller,1972)
R388	naturally occurring, IncW	Su^r,Tp^r,Tra^+	M.Robinson (Datta and Hedges,1972)

(b) Non-conjugative plasmids

Name	Size	Derivation	Phenotype	Source
pMB9	5.5	mini ColE1-*tet*	Iel^+,Tc^r	G.Warren (Bolivar *et al.*,1977)
pMB9::Tn*103*	9.5	insertion of Tn*103* into pMB9	Iel^+,Tc^r*	L.Symington
pMB9::Tn*1031*	8.9	*Pst*I deletion derivative of pMB9::Tn*103*	Iel^+,Tc^r*	A.Lamond (Kitts *et al.*,1982a)
RSF1050	7.8	pMB8::Tn*3*	Iel^+,Ap^r	S.Falkow (Heffron *et al.*,1977)
RSF1341	6.3	deletion derivative of RSF1050	Iel^+,Ap^r	S.Falkow (Heffron *et al.*,1977)
RSF1365	6.5	deletion derivative of RSF1050	Iel^+,Ap^r	S.Falkow (Heffron *et al.*,1977)
pBR322	4.36	engineered	Ap^r,Tc^r	D.J.Sherratt (Bolivar *et al.*,1977)

pOX14	10.06	insertion of γδ into pBR322	Ap^r,Tc^r	R.Reed (M.Guyer,1978)
pRR1	6.86	*Xho*I deletion derivative of pOX14 - *tnp*A^-,*tnp*R^-,*res*$^-$	Ap^r,Tc^r	R.Reed (Kitts *et al.*,1982b)
pRR12	10.06	linker insertion into pOX14-*tnp*A^+,*tnp*R^-	Ap^r,Tc^r	R.Reed (Kitts *et al.*,1982b)
pRR17	10.06	linker insertion into pOX14-*tnp*A^-,*tnp*R^+	Ap^r,Tc^r	R.Reed (Kitts *et al.*,1982b)
pλ8	5.4	insertion of 1100bp P_L containing fragment from λ into pBR322	Ap^r	R.Reed (Reed,1981b)
pPAK329	4.64	insertion of 282bp *res* containing fragment into pBR322	Ap^r,Tc^r	P.Kitts
pLS138	4.92	insertion of 282bp *res* containing fragment into pPAK329	Ap^r,Tc^r	L.Symington (Chapter 5)
pLS139	4.99	insertion of 357bp *res* containing fragment into pPAK329	Ap^r	L.Symington (Chapter 5)
pLS140	4.99	as above-opposite orientation	Ap^r	L.Symington (Chapter 5)

Name	Size	Derivation	Phenotype	Source
pLS213	5.2	insertion of *tnp*R into pλ8	Ap^r	L.Symington (Chapter 4)
pACYC184	4.0	engineered from p15a	Cm^r,Tc^r	F.Cannon (Chang and Cohen,1978)
pLS109 (pACYC184::γδ)	9.7	insertion of γδ into pACYC184	Cm^r	L.Symington (Chapter 3)
pLS110	8.0	engineered insertion of Tn*103* into pACYC184	Cm^r,Tc^r*	L.Symington (Chapter 4)
pLS129	8.0	insertion of Tn*103* into pACYC184	Cm^r,Tc^r	L.Symington (Chapter 5)
pLS134	6.0	constructed from pLS129 and pMB9::Tn *1031*	Cm^r	L.Symington (Chapter 5)
pLS135	4.0	resolution product of pLS134	Cm^r	L.Symington (Chapter 5)
pLS137	10.0	constructed from pLS129 and pMB9::Tn *1031*	Cm^r,Tc^r	L.Symington (Chapter 5)
pLS142	9.7	insertion of RR12 γδ into pACYC184	Cm^r,Tc^r	L.Symington (Chapter 3)
pPAK317	4.28	insertion of 282bp *res* containing fragment into pACYC184	Tc^r	P.Kitts
pAA131	20.3	RSF103::pACYC184::Tn*103* cointegrate	Cm^r,Sm^r	A.Arthur (Arthur and Sherratt,1979)

pACYC184::Tn3	8.9	insertion of Tn 3 into pACYC184	Cm^r Tc^r	A. Arthur
RSF103	12.3	RSF1010::Tn*103*	Sm^r (Mob^-?)	S.Falkow (Heffron *et al.*,1977)
pDS4101	12.2	ColK::Tn*1*	Ap^r,Cka^+,Ika^+, Mob^+,bom^+	D.J.Sherratt (Tacon *et al.*,1981)
pDS4153	8.1	*Hae*II scramble of pDS4101	Ap^r,Cka^+,Ika^+, Mob^+,bom^+	D.J.Sherratt (Tacon *et al.*,1981)

TABLE 2.2 Bacterial plasmids

*denotes Tc^r is repressed in the presence of the Tn*3* *tnp*R gene product

at 2µg/ml.

Antibiotics: Ampicillin, streptomycin, trimethoprim, kanamycin and rifampicin were used at 50µg/ml; chloramphenicol at 25µg/ml and tetracycline at 10µg/ml. All antibiotics except trimethoprim were used in nutrient agar. Trimethoprim was used in minimal agar or, more usually, in iso-sensitest agar (similar composition to nutrient, but lacks thymine). Streptomycin selection for the plasmid RSF1010 was at 20µg/ml in minimal agar.

4. Buffers

E: 0.04M Tris, 0.02M NaAc, 0.001M EDTA, pH8.2 with acetic acid; usually kept as 10X concentrate.

TBE(X10): 108g Tris, 55g Boric acid, 9.3g EDTA; made up to 1 litre with distilled water. The pH should be about 8.3.

SRB: 15.5g Tris, 72.05g glycine, 0.1% SDS; made up to 5 litres in distilled water

SDS-PAGE lower buffer(X4): 1.5M Tris-HCl, 0.4% SDS, pH 8.8.

SDS-PAGE upper buffer(X4): 0·5M Tris-HCl, 0.4% SDS, pH 6.8.

TE: 10mM Tris-HCl, 1mM EDTA, pH7.5

TEN: 10mM Tris-HCl, 1mM EDTA, 50mM NaCl, pH 7.5

FSB: 10% Ficoll(w/v), 0.5% SDS(Serva)(w/v), 0.06% bromophenol blue (w/v), 0.06% orange G (w/v) made up with buffer E.

SCCLB: 2.5% Ficoll(w/v), 1.25%SDS(w/v), 0.015% bromophenol blue(w/v), 0.015% orange G (w/v); made up with buffer E.

PFSB: 10% glycerol(v/v), 0.01% bromophenol blue (w/v), 5% mercapto-ethanol(v/v), 3% SDS(w/v), 0.625M Tris, pH 8.0.

Elution buffer: 500mM NH_4Ac, 1mM EDTA, 0.1% SDS(w/v)

Lytic Mix: 2% Triton X-100, 0.05M Tris, pH 8.0, 0.06M EDTA, pH 8.0; made up with distilled water.

STET: 8% sucrose(w/v), 5% Triton X-100(w/v), 50mM EDTA, 50mM Tris-

HCl, pH 8.0.

Low salt restriction buffer-LSRB(X10): 100mM Tris-HCl, 100mM $MgCl_2$, 10mM DTT, pH 7.4. Used for restriction with *Taq*I.

Medium salt restriction buffer-MSRB(X10): 100mM Tris-HCl, 100mM $MgCl_2$, 10mM DTT, 500mM NaCl, pH7.4. Used for restriction with *Bam*HI, *Bgl*II, *Cla*I, *Hinc*II, *Hind*III, *Sau*3AI, and *Pst*I.

High salt restriction buffer-HSRB(X10): 500mM Tris-HCl, 100mM $MgCl_2$, 1M NaCl, pH7.4. Used for restriction with *Eco*RI, and *Sal*I.

NTE(X10): 500mM Tris-HCl, 50mM $MgCl_2$, 100mM mercaptoethanol, pH7.9.

Kinase buffer(X10): 500mM Tris-HCl, 100mM $MgCl_2$, 50mM DTT, pH 7.6.

5. Growth conditions

Liquid cultures were grown in L-broth at 37°C with aeration unless otherwise stated. "Stationary phase cultures" were small liquid cultures, usually 2ml, grown without shaking at 37°C overnight.

Plates contained 25ml agar with the appropriate supplements and were incubated at 37°C for 16-18 hr unless otherwise stated.All dilutions prior to plating were made in phage buffer.

6. Purification of plasmid DNA

(a) Cleared lysates

Cleared lysates were usually prepared from 100ml cultures; for cultures of 10ml or one litre the volumes of reagents were scaled up or down accordingly. Cells were harvested by centrifugation at 12,000g, 4°C for 10min. The cell pellet was resuspended in 3.3ml cold 25% sucrose/ 0.25M Tris-HCl, pH 8.0. 0.67ml of fresh lysozyme solution, made up at 10mg/ml in 0.25M Tris-HCl, pH 8.0, were added and the cell suspension swirled frequently, on ice, for 10min. 1.3ml of 0.25M EDTA, pH 8.0 were added and swirled again, while on ice, for 5min. 5.3ml of lytic

mix were added gently to the cell suspension; the lysate was swirled gently to ensure thorough mixing. The cells were judged to have lysed when the mixture was clear and viscous; this usually took about five min. The lysate was centrifuged at 43,000g, 4°C for 20min to pellet the chromosomal DNA and membranous material. The supernatant, containing mainly plasmid DNA, was carefully decanted - this is referred to as the cleared lysate. The cleared lysate was purified further by phenol/isopropanol extraction or by CsCl/EtBr equilibrium centrifugation.

(b) Phenol/isopropanol extraction

The cleared lysate was mixed with an equal volume of freshly distilled phenol, saturated with 1M Tris-HCl, pH 8.0. The phases were resolved by centrifugation at 5,000g for 5min. The upper aqueous phase was phenol extracted twice more. Residual phenol was removed by ether extraction. An equal volume of diethyl ether was added, mixed well, the upper phase(ether) was discarded and the process repeated until the lower(aqueous) phase was clear. Residual ether was removed by blowing compressed air over the surface of the solution. 0.1 volume of 3M NaAc was added, then 0.54 volume of isopropanol. This was mixed well and left at room temperature for 45min to precipitate nucleic acids. The solution was centrifuged at 12,000g, 15°C for 25min; the pellet obtained was washed with 70% ethanol/TE then dried. The dried pellet was resuspended in 500µl TE and stored in airtight plastic tubes at 4°C.

(c) CsCl/EtBr equilibrium centrifugation

For DNA of greater purity, cleared lysates were run through CsCl/EtBr gradients. For each tube 5.0g CsCl, 4.83ml cleared lysate, 0.33ml EtBr(3mg/ml), 0.1ml 0.2M Na_2HPO_4 were mixed and placed in Beckman quick-seal tubes. The density of the mixture was checked(it

should be about 1.58g/cc), the tubes were filled with liquid parafin and heat sealed. They were centrifuged in a Ti50 or Ti70 rotor in a Beckman L-2 ultracentrifuge at 48Krpm for 16hr at 15°C. Two bands were formed within the gradient; the lower, denser one contains covalently closed DNA. This band was removed by inserting a syringe needle through the side of the tube and slowly drawing off the band laterally. EtBr was removed by repeated extraction with butan-1-ol. The solution was diluted with three volumes of water, then precipitated by addition of two volumes of ethanol at -20°C for one hr. The DNA was collected by centrifugation at 32,000g, 4°C for 25min. The pellet was washed with 70% ethanol/TE, dried, then resuspended in 500µl TE.

(d) STET DNA Purification

For small scale DNA isolation a quick method was used to provide DNA suitable for transformation or restriction.A clump of cells, which had been "patched"onto selective plates, were resuspended in 100µl of STET buffer in plastic eppendorf tubes and briefly vortexed. 10µl of freshly made lysozyme(10mg/ml) were added and the lysate mixed well. The tubes were placed in a boiling water bath for two min, then spun in an Eppendorf microfuge for 10min. The supernatant was removed and nucleic acids precipitated from it by adding an equal volume of isopropanol and leaving at -20°C for 15min. DNA (and RNA) were collected by centrifugation in the Eppendorf microfuge for five min. The pellet was resupended in 20µl of TE buffer; 5µl were sufficient for each restriction.

(e) Birnboim/Doly DNA Purification

The basic principle of this method is alkaline denaturation of linear DNA molecules. Cells are lysed with NaOH/SDS; chromosomal DNA is denatured, but not covalently closed DNA. On neutralization

the mass of chromosomal DNA renatures and aggregates, proteins complex with SDS; when centrifuged the supernatant contains only covalently closed DNA and low molecular weight RNA (Birnboim and Doly, 1979). Small overnight cultures, usually 1-3ml, were harvested by centrifugation in the Eppendorf microfuge for one min; the supernatant was removed with a drawn out pasteur pipette. The cell pellet was resuspended in 100µl 50mM glucose, 25mM Tris-HCl, pH 8.0, 10mM EDTA, then incubated at 22°C for five min. 200µl 0.2M NaOH/1% SDS were added, gently mixed, and left on ice for five min. To neutralize, 150µl precooled 5M KOAc, pH 4.8(3M KAc pH'd with acetic acid) were added, mixed gently and left on ice for five min. The crude lysate was spun in the Eppendorf for one min, the supernatant was removed and to it added two volumes of ethanol.This was mixed well and left to precipitate on dry ice/ IMS for 10min. The precipitate was collected by centrifugation in the Eppendorf microfuge for seven min; the supernatant was discarded. The pellet was washed with 500µl 70% ethanol/TE, mixed well, then precipitated as before. The last step was repeated before finally drying the pellet and resuspending in 20µl TE. 5µl were used for each restriction.

(f) Single colony cleared lysates(referred to as SCCL)

This technique provides a quick method for analyzing the total DNA content of a particular clone. The single colonies (isolated from transformation or conjugation) were "patched" onto selective plates. Clumps of cells, picked up using sterile toothpicks, were resuspended in 150µl of SCCLB. The lysates were spun in an Eppendorf or Sarsted microfuge for 15min; the supernatant, about 50µl, was loaded directly onto agarose gels for visualization of the DNA content.

7. DNA Electrophoresis through gels

Vertical gel kits held two glass plates measuring 16X16cm and separated by perspex spacers of 3mm(1mm for acrylamide gels). The gel liquid was poured between the two plates; 3mm thick teflon combs (1mm thick perspex for acrylamide gels) with 10 or 15 teeth were pushed into the top displacing the gel material to produce sample pockets, when removed from the set gel. Electrophoresis was from top to bottom with the cathode at the top.

(a) Agarose gels

Sigma agarose (typeII, medium EEO) and buffer E were used for electrophoresis unless otherwise stated. Agarose gels for SCCL and detection of *in vitro* resolution were made up at 0.8%(w/v); for restriction gels 1% agarose was usually used. The agarose/buffer E mixture was heated to 100°C until completely molten, then cooled to 46°C for gel pouring. DNA samples were mixed with 0.25 volume FSB; 20µl were usually loaded per pocket. The gel was run with buffer at anode and cathode at between 20 and 120V until the blue dye marker reached the bottom, unless otherwise stated. The relative separation between bands increases when gels are run at low voltage for longer times, but bands tend to be more diffuse. Restriction gels were usually run at 100V for 4hr, whereas gels analyzing large plasmids, such as R388 derivatives, were run overnight at 20V. The gels were stained in gel running buffer containing 0.5µg/ml EtBr for 30min. Gels were viewed on a 260nm transilluminator and photographed using a 35mm camera (Ilford HP5 film) with a red filter.

(b) Polyacrylamide gels

Optimal resolution of small DNA fragments was achieved by analyzing restriction digests of plasmid DNA on polyacrylamide gels. 10% (w/v) acrylamide, made up in TBE, was used to analyze small fragments of 10-300bp; 5% (w/v) acrylamide was used to assay for larger fragments, 100-1000bp. Acrylamide solutions (kept as 20% acrylamide, 1% bis-acrylamide stock) were polymerized by the addition of 0.36ml 10% (v/v) TEMED and 0.7ml 10% (w/v) APS in 60ml total volume. Gels were run with TBE buffer at anode and cathode at 30mA until the yellow dye reached the bottom of the gel. The gel was removed from the running apparatus, separated from the glass plates, then stained and photographed as previously described.

(c) Sequencing polyacrylamide gels

High resolution sequencing gels resolve DNA strands of n and n+1 nucleotides over a range of one to several hundred base pairs.The common gel sequencing techniques are versions of methods described by Peacock and Dingman (1967) and Maniatis *et al.* (1975). For the experiments described in this thesis only 8% sequencing gels were employed. 8% gels contained 7.6% (w/v) acrylamide, 0.4%(w/v) bis-acrylamide, 50% (w/v) urea (8.3M), 100mM Tris-borate, pH 8.3, 2mM EDTA, 0.07% (w/v) APS, and 0.7%(w/v) TEMED as catalyst. This solution was injected between two clean, siliconized glass plates, measuring 200X400mm separated by 0.3mm plasticard spacers. The gel material was displaced from the top of the gel by a 0.3mm thick plasticard comb to form the sample pockets. The gel was installed in the vertical running apparatus; buffer reservoirs were filled with TBE. The comb was gently eased from the top of the gel and sample pockets were immediately flushed out with buffer to remove unpolymerised acrylamide, which slides into the pockets from the top of the gel. Gels were

pre-run at 40W (1600V) for 30min, then disconnected prior to loading samples. Samples, usually 3µl, were loaded with a glass capillary tube evenly across the pocket. The power supply was reconnected and the gel run at 40W until the bromophenol dye marker had nearly reached the bottom. The gel was disconnected. One glass plate was lifted from the gel, a sheet of "clingfilm" was spread over the gel, smoothed down and secured with tape on the underlying glass plate. This was placed over a sheet of X-ray film inside a light-tight aluminium box. Exposing the gel to preflashed film, juxtaposed with an intensifying screen at $-70^{o}C$ produced an image much faster and was generally employed.

8. Interpretation of gel data

The distance migrated is related to the size of a molecule of a given conformation; small molecules run the furthest. The distance migrated by a given DNA molecule is dependent on it's conformation; supercoils generally run fastest and open circles slowest, with linear molecules between the two. Above a certain size all linears run at the same rate, therefore large plasmid linears run faster than corresponding supercoils. Large linear fragments of chromosomal DNA appear as a single band.

For restricted DNA molecules, the relationship:

$$\text{distance migrated} = K.\log(\text{size}) + c \quad \text{holds}$$

If a DNA species is restricted to give fragments of known size,e.g. λ or pBR322, then the sizes of fragments from other DNA molecules may be determined from a graph of distance against log(size).

9. Elution of DNA fragments from gels

(a) Electroelution

Fragments of DNA or DNA of unique conformation were separated by electrophoresis through acrylamide or agarose gels. The stained gel was placed on the long-wave transilluminator, the relevant bands were excised with a scalpel and placed in a dialysis sac with 100µl of TBE or E buffer. The dialysis sac was subjected to electrophoresis at 50V for 2hr. The current was reversed,to remove any DNA from the sides of the dialysis sac,for 5min. The contents of sac were removed, agarose discarded and DNA precipitated by the addition of 0.1 vol. 3M NaAc plus two vol. ethanol. This was placed at -20°C for 30min, then centrifuged in the Eppendorf microfuge for 7min. The DNA pellet was washed, dried and resuspended in TE, ready for use to restrict, ligate etc.

(b) Crush-soak-precipitate method

This method was used exclusively for extracting labelled DNA fragments from acrylamide gel slices. The DNA-containing band was mashed with a siliconized glass rod in 200µl of elution buffer in a 1ml eppendorf tube. The mixture was then incubated at 45°C for 16hr. Acrylamide was filtered from the solution by pouring through a glass wool plug in a punctured 0.5ml eppendorf tube. Acrylamide pieces were retained on the surface of the glass wool; this was washed through with an additional 200µl of elution buffer. DNA from the filtered solution was precipitated by the addition of 2 vol. ethanol, then left at -20°C for 30min. The precipitate was collected by centrifugation for 7min in the Eppendorf microfuge, washed, dried and resuspended in TE.

10. DNA Manipulations *in vitro*

(a) Restriction of plasmid DNA

Restriction of plasmid DNA was performed in 0.5ml polypropylene eppendorf tubes. The final reaction volume was 20µl, containing:

0.5µg plasmid DNA

2 µl 10X appropriate restriction buffer

1 unit enzyme

gelatin/water(at 0.1mg/ml) to 20µl total volume

Complete digestion was usually achieved in two hr at 37°C. For DNA isolated by the mini-prep. or phenol/ isopropanol extraction methods, 1µl of 1mg/ml ribonuclease was usually added as RNA species tend to obscure the small DNA fragments. Digests were analyzed by electrophoresis through agarose or polyacrylamide gels.

(b) Alkaline phosphatase treatment

Bacterial alkaline phosphatase catalyzes the removal of phosphate groups from the 5' terminus of DNA and RNA fragments. The resulting 5' hydroxyl terminus may be used as a substrate for polynucleotide kinase, or as a means of preventing vector recircularization during ligation. When no direct selection is available for cloning DNA fragments, phosphatasing the vector provides a means of enriching for clones containing inserts. Restricted vector DNA was precipitated with NaAc/ethanol; the resulting pellet was dried and resuspended in 50µl of 10mM Tris-HCl, pH 8. 20 units of bacterial alkaline phosphatase were added and the reaction mixture incubated at 65°C for one hr. To remove the enzyme, the solution was phenol extracted twice, ether extracted twice, then precipitated using NaAc/ethanol. The DNA pellet was washed, dried and resuspended in the appropriate buffer for ligation or kinase treatment.

(c) Ligation of restriction fragments

Endonucleases used to digest plasmid DNA were destroyed either by heating to 65°C for 5 min or by phenol/ether extraction of the reaction mixture. DNA was precipitated with NaAc/ethanol then resuspended in the following buffer:

66mM Tris-HCl, pH 7.5

10mM $MgCl_2$

10mM mercaptoethanol

1mM EDTA

0.4mM ATP

0.1 unit ligase/μg DNA

Ligation mixtures were incubated at 14°C for 16-18hr. These were then diluted in TE for use in transformation. Prior to some ligations, individual DNA fragments were isolated from digests by separating on 1% (w/v) low melting point agarose horizontal gels. The relevant bands were excised from the gel, melted at 65°C, then 2μl were added to the ligation reaction mixture. When thoroughly mixed and cooled the concentration of agarose was reduced to below that required to set and did not appear to inhibit the reaction significantly.

(d) DNA polymerase end-labelling

For DNA sequencing and footprinting reactions, a DNA fragment with one labelled terminus is required. Enzymes which produce a staggered cleavage with a 3' recessed hydroxyl group provide a substrate for DNA polymerase I. Only one labelled $[\alpha\text{-}^{32}P]$ triphosphate is required to be polymerized adjacent to the recessed 3' hydroxyl group for sequencing and footprinting reactions. The labelled triphosphate is chosen in reference to the sequence cleaved by the particular restriction enzyme. Restricted plasmid DNA was precipitated and dried, the following were added:

5µl $[\alpha-^{32}P]$ XTP (10mCi/ml)

5µl NTE(X10)

0.5µl *E.coli* DNA polymerase I large fragment (1 unit)

39.5µl distilled water

The reaction was incubated at 14°C for one hr.

(e) T4 polynucleotide kinase end-labelling

T4 polynucleotide kinase catalyzes the transfer of the γ-phosphate from ATP to the 5' hydroxyl terminus of DNA, RNA and mononucleotides. As a 5' hydroxyl end is required, the fragment must be dephosphorylated with alkaline phosphatase prior to labelling. Following phosphatase treatment the purified fragment was precipitated and dried, the following were added:

10µl $[\gamma-^{32}P]$ ATP (10mCi/ml)

5µl kinase buffer (X10)

0.5µl T4 polynucleotide kinase (1 unit)

34.5µl distilled water

The reaction was incubated at 14°C for one hr.

(e) Chemical sequencing reactions

Chemical sequencing involves three consecutive steps: modification of the base, removal of the modified base from it's sugar, and strand scission at that sugar. The end-labelled fragment is treated such that each strand should be cleaved only once; each fragment has one common end, the other is variable in length. An array of fragments is produced, representative of cleavage at each position held by a particular base. By performing four separate reactions each using a chemical which cleaves only one of the four bases, separating the bands on a sequencing gel, then analyzing the ladder of bands produced, it is possible to read the sequence

from an autoradiogram of the gel. For the experiments described in this thesis only the purine-specific cleavage reaction was performed; the following protocol was employed:
2µl of ^{32}P labelled DNA (about 50 cps on the mini-monitor) were dried down under vacuum. 15µl DPu mix (0.1g diphenylamine; 5ml 66% HCOOH+1mM EDTA) were added and incubated at 22°C for 8min. The reaction was stopped with 45µl water and 500µl diethyl ether and mixed well. When the phases had resolved the ether was removed; the ether extraction was repeated twice more. The solution was frozen by placing in a dry ice/IMS bath, then lyophilized. 100µl of 1.0M piperidine were added, then incubated at 90°C to induce strand scission. The solution was frozen then lyophilized. To the dried pellet 10µl of distilled water were added, mixed well, then frozen and lyophilized. This step was repeated once more. The dried DNA pellet was resuspended in 6µl of formamide-containing marker dyes (0.05% orange G (w/v); 0.05% bromophenol blue (w/v); 0.05% xylene cyanol (w/v) in 98% formamide). The sample was heated to 90°C for one min, to separate DNA strands then quickly chilled. 3µl of the sample were loaded per gel pocket.

11. Transformation

A fresh overnight culture of the recipient strain in L-broth was diluted 1:40 in 20ml of L-broth. The culture was grown at 37°C, with shaking, for about 90min, or until there were about $2X10^8$ cells/ml. The time required to reach this stage is strain dependent; $RecA^+$ strains grow much faster than $RecA^-$ derivatives. Cells were pelleted at 12,000g for five min, resuspended in 10ml cold 50mM $CaCl_2$ and left in an ice/water bath for 15-20min. The cells were pelleted again, resuspended in 1ml of cold 50mM $CaCl_2$ and left in an ice/water bath for at least 15min. The competence for DNA uptake

by $RecA^+$ cells increases for up to 24hr on ice at this stage (Dagart and Ehrlich, 1979). Mix 0.2ml of the cell suspension with up to 0.1ml of DNA solution (DNA was usually diluted in TE buffer - 0.1µg of DNA is ample for each transformation). The DNA/cell mixture was mixed well, then incubated in an ice/water bath for 20min; at 37°C for 7min, then ice/water again for 30min. 0.8ml of fresh, pre-warmed L-broth were added to each tube and incubated at 37°C, shaking, for 30-120min to allow expression of antibiotic resistance. All antibiotic selections required expression times of at least 90min, except ampicillin, which needs less than 30min for expression. 100µl aliquots were plated onto selective agar and incubated overnight at 37°C.

12. Bacterial Conjugation

(a) Liquid matings (for F factor crosses)

Fresh overnight cultures of recipient and donor strains were diluted 1:20 into 5ml L-broth. The cultures were grown at 37°C, with shaking, until the donors were in early log. phase ($2X10^8$cells/ml) and the recipients had reached late log. phase ($2X10^9$cells/ml). The two cultures were mixed together and left for 1hr at 37°C, without shaking, then diluted in phage buffer before plating onto selective media.

(b) Plate matings (routinely used for R388)

R388 produces short sex pili, thus high cell to cell contact is required for conjugation to take place. This was achieved by growing cultures of donors and recipients to the same densities as used for liquid matings, mixing the cultures, then concentrating by centrifugation. The cell pellet was resuspended in 1ml L-broth and poured over the surface of a well-dried nutrient agar plate. After

one hr the mating mixture was removed from the plate by resuspending in phage buffer, then diluted and plated accordingly.

13. SDS-Polyacrylamide gel electrophoresis (for proteins)

SDS-polyacrylamide slab gels were routinely used to assay whole cell lysates for resolvase and also to assay column fractions during purification of resolvase. All the gels used were "stacking" gels containing 12.5%(w/v) acrylamide in the lower separating gel and 4.5% (w/v) acrylamide in the upper stacking gel. Gels were prepared from a stock solution containing 30% (w/v) acrylamide and 0.8% (w/v) N,N'-bis-methylene acrylamide. 0.03% (w/v) APS and 0.025% (v/v) TEMED were added prior to pouring gels. The gel apparatus consisted of two glass plates of the same dimensions as used for DNA gels, separated by 1.5mm perspex spacers. Buffer and stock acrylamide were mixed, then freshly made APS and TEMED were added prior to pouring. 0.1% SDS solution was sprayed over the top of the lower gel to ensure an even surface. When the lower gel had set the SDS solution was poured off. The components of the top gel were mixed and poured over the lower gel. A 12 space perspex comb was inserted into the top of the gel and left for a further 30min to set. Top and bottom reservoirs of the vertical gel apparatus were filled with SRB. The comb was removed and sample pockets were flushed out with buffer to remove any unpolymerized acrylamide. Electrophoresis was carried out at 30-40mA until the blue dye marker reached the bottom of the separating gel. Gels were stained for one hr, with slight agitation, in 0.2% coomassie blue in fix (50% methanol; 40% water and 10% acetic acid). The stain was poured off, then washed several times in fix until destained. The gel was placed in 10% acetic acid for the final wash to enlarge the gel to it s original size, then photographed. For photography Panatomic-X film was used in a 35mm

camera with an orange or yellow lens filter.

The molecular weight marker proteins used to calibrate gels were purchased from Bio-Rad, and contained the following proteins at 0.1µg/µl:

Phosphorylase B	93,000 Mdal
BSA	68,000
Ovalbumin	43,000
Carbonic anhydrase	30,000
Soybean lectin inhibitor	21,000
Lysozyme	14,300

14. Electron Microscopy

DNA molecules from *in vitro* resolution reaction mixtures were visualized by electron microscopy. The copper support grids used were coated with a film of 3.5% (w/v) parlodion in amylacetate. These were left to dry at room temperature, overnight, in a dessicator. The hypophase for DNA spreading consisted of 16ml of formamide, 0.8ml Tris-HCl/EDTA (1M/0.1M , pH 7.5) made up to 80ml with distilled water. This solution was poured into a petri dish. A clean microscope slide, which had been soaked in chromic acid, washed with water and ethanol flamed before use, was used as a ramp by placing half in the petri dish of hypophase solution and resting against it s side. The hyperphase consisted of 10µl Tris/EDTA (as above), 20µl DNA solution, 50µl formamide, 10µl cytochrome c (1mg/ml in 0.1M Tris/0.01M EDTA, pH 8.5) and 20µl distilled water. 25µl of the hyperphase were run down the ramp onto the hypophase using a Hamilton syringe and microcapillary tube. Areas of the hyperphase were picked up on coated grids by touching gently on the surface of the solution around the ramp. Grids were stained with

uranyl acetate (0.05mM uranyl acetate, 0.05mM HCl in 90% ethanol) for 30sec, then washed in isopentane for 30sec. The grids were left at room temperature, overnight, to dry, then shadowed with Pd/Pt in a rotary shadower. The grids were viewed in a Jeol JEM-100S electron microscope and photographed at 20,000X magnification.

CHAPTER 3

INTERACTIONS BETWEEN γδ AND Tn*3*

Introduction

Many of the genes encoding antibiotic resistance in multiply drug resistant bacteria can reside on transposable elements. For the purpose of studying the transposition process, these elements are easier to work with than the simple insertion elements as their transposition can be monitored by the phenotypic markers, which they carry. Of these, the ampicillin resistant transposons, Tn*1/3*, have been the subject of intense investigation to gain some insight into the mechanism of transposition.

The problem has been approached by constructing a number of deletion and insertion mutations using *in vitro* techniques (Heffron *et al.*,1977;1978); these could then be assayed for transposition. Mutants with an apparent transposition$^-$ phenotype were also tested in the presence of a complementing element to identify sites and proteins involved in transposition. These procedures identified the inverted repeats of Tn*3* as essential sites for transposition. Internal mutations fell into two complementation groups affecting transposition. Mutations in one group map within the first 3,000bp of Tn*3*, extending from the left inverted repeat through more than half of the element. These mutants express a transposition deficient phenotype, but may be complemented *in trans* to transpose (Heffron *et al.*,1977;1978). Fine structure complementation analysis,sequencing and SDS-PAGE have indicated the presence of a single polypeptide of 120,000 dal. from this region. This protein appears to be the only element-encoded product required for transposition and has been named *tnp*A, or transposase (Chou *et al.*,1979b; Gill *et al.*,1979; Heffron *et al.*,1979); host-encoded factors are almost certainly required also.

The second complementation group extends over a short region, about 450bp, between *tnp*A and the structural gene for Ap^r, *bla*. Mutations within this region lead to an increased frequency of transposition and thus identified this gene, *tnp*R, as a repressor of transposition (Gill *et al.*,1978;1979; Chou *et al.*,1979b). Analysis of some of these mutants by SDS-PAGE indicated that *tnp*R is autoregulated.(Chou *et al.*,1979a; Dougan *et al.*,1979).

A third site required in *cis* was identified by a subset of deletions which on transposition formed aberrant products. These consisted of the entire donor plasmid inserted into the recipient genome; the inserted replicon was shown to be flanked by directly repeated copies of the element (Gill *et al.*,1978). This provided direct evidence for the replicative nature of transposition and formed the basis for many transposition models. The fused replicons, or cointegrates, were considered to be intermediates in the transposition process. The site deleted was presumed to be required for resolution of this structure by recombination between the direct repeats of Tn*3* to generate the normal transposition end products. This site will be referred to as *res* (Reed,1981a).

Most of the assay systems used to measure transposition rely on insertion of the element into a conjugative plasmid; the frequency of transfer of Ap^r, encoded by the element, may be directly compared with the transfer frequency of the conjugative plasmid. When the F factor was used in this assay system *tnp*R^-, *res*$^+$ rarely formed cointegrates. However, when the conjugative plasmid R388 was used as a recipient for transposition, *tnp*R^-,*res*$^+$elements resulted exclusively in replicon fusion in $RecA^-$ cells (Sherratt *et al.*,1981a;Kitts *et al.*,1982a). This suggested that the *tnp*R gene product, or resolvase,

plays a direct role in the resolution of transpositional cointegrates by recombination at *res* sites. The conflicting data obtained by using the F factor or R388 in the assay system suggested to us that the F factor may itself code for a recombination protein capable of acting at *res* to resolve transpositional cointegrates.

The F factor plasmid has been associated with a number of recombination events , many of which are RecA-independent. Three different sequences have been identified on F, which serve as attachment sites between F and the chromosome in the formation of Hfr strains. Hfr strains formed by recombination between the γδ sequence on F and the chromosome are known to be unstable and revert at high frequency by a RecA-independent event between directly repeated copies of the γδ element (Davidson *et al.*,1974). Davidson has proposed that F encodes a recombination protein, which mediates this event. Electron micrographs of F plasmid DNA have shown the presence of hairpin loop structures in the γδ region. The stem structure is short, between 20 and 50bp,and suggests that the γδ element has inverted repeats at it s ends (Broker *et al.*,1977).

More recently γδ has been shown to mediate another kind of recombination event, transposition. This was originally detected by the ability of the F factor to transfer the small Mob^- plasmid, pBR322, during conjugation, albeit at low frequency. Analysis of the Tc^rAp^r transconjugants revealed the presence of a plasmid, in addition to F, of 10Kb - equivalent to the size of pBR322 plus one copy of the γδ element (5.7Kb). Restriction analysis and electron microscopy confirmed that the new plasmid species was indeed pBR322 carrying a copy of γδ (Guyer,1978).

These insertions of γδ into pBR322 have subsequently been analysed in more detail by DNA sequencing to determine the precise insertion points and the sequence of the ends of the γδ element (Reed *et al.*,1979). Analysis of several independent insertions indicated that a 5bp duplication of pBR322 DNA had occurred at the integration point, adjacent to the γδ element. The sequencing data obtained from these experiments also revealed that the γδ element has 35bp perfect inverted repeats at it s ends, with significant homology to the ends of Tn*3*. These observations have led to the classification of γδ within the"Tn*3* family" of transposable elements (Calos and Miller,1980; Kleckner,1981). The strong similarities between these elements prompted an investigation into the possible interaction of transposition functions.

Results

1. Resolution of Tn*1/3* transpositional cointegrates by γδ

The small, broad host range plasmid RSF1010 confers resistance to Su and Sm (Heffron *et al.*,1975). This plasmid has been used as a recipient for transposition of the Ap^r transposon, Tn*1*, from R1*drd*. One particular insertion of Tn*1* into the Su^r structural gene was polar on Sm^r, resulting in a Su^sSm^s phenotype. Selection for Sm^r to relieve polarity led to a spontaneous deletion removing parts of the *tnp*R and *bla* genes of Tn*1* (Heffron *et al.*,1977). The plasmid carrying this deletion, named RSF1103, was used for complementation of *tnp*A with various pMB8::Tn*3* mutants (Heffron *et al.*,1977;1978). The Tn*103* element transposes at high frequency due to relief of *tnp*A repression by *tnp*R (Sherratt *et al.*,1981a). Insertion of RSF1103 into the Tc^r structural gene of pACYC184 results in replicon fusion between the two plasmids to generate a 20Kb cointegrate plasmid containing directly

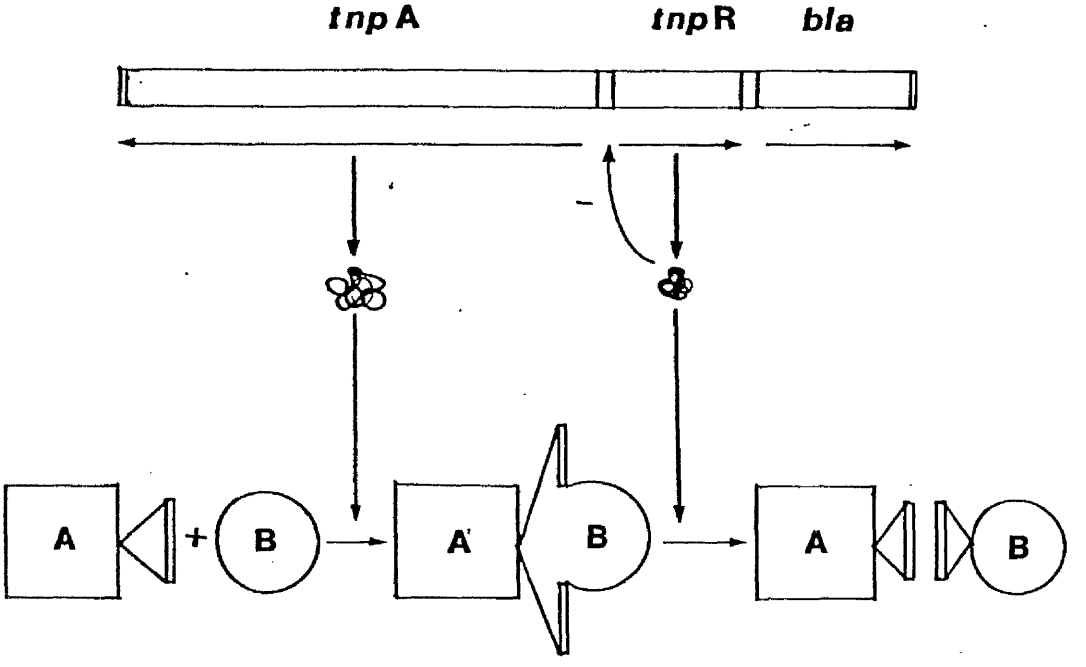


FIGURE 3.1 A general scheme of the transposition process for Tn*3*-like elements

The *tnp*A gene product mediates the formation of an obligatory transpositional cointegrate, which is then resolved by the *tnp*R gene product to generate the normal transposition end products. The *tnp*R gene product also controls the expression of the *tnp*A and *tnp*R genes at the transcriptional level.

repeated copies of Tn*103*; this was designated pAA131. pAA131 was stably maintained in Rec$^-$ hosts; in RecA$^+$ hosts the cointegrate was slowly recombined to form the normal transposition end products. It was also noted that this plasmid could be stably maintained in a RecA$^+$, RecF$^-$ strain. This suggested to us that the RecF pathway for recombination may be required for plasmid intramolecular recombination, though it does not appear to be required for integration of linear DNA by conjugation or transduction (Clark,1973; Gillen *et al.*,1981; James *et al.*,1982). pAA131 is efficiently recombined in the presence of a complementing *tnp*R$^+$ transposon to generate the normal transposition end products, i.e. RSF103 and pACYC184::Tn*103* (figure 3.1) (Arthur and Sherratt,1979). It was observed that in the presence of *tnp*R the Sm^r of RSF103 was repressed, conferring sensitivity to this antibiotic on strains containing these plasmids. This is presumably due to transcriptional readthrough from the *tnp*A gene promoter into the Sm^r structural gene, regulated by resolvase. The presence of a functional *tnp*R gene product can therefore be monitored by resolution of pAA131 and by repression of Sm^r (for details see A.Arthur,1981). A similar, but more effective system has been devised for studying transcription from *tnp*A and *tnp*R by cloning the β-galactosidase structural gene into the *tnp*A and *tnp*R genes, then monitoring the levels of this enzyme under various conditions (Chou *et al.*,1979a;b; Cohen *et al.*,1979).

To test for the presence of a function analagous to resolvase on the F factor, strain CSH52 containing the F prime, F'*lac pro*, was mated into strain JC9239 containing pAA131. Selection was made on minimal/lactose/Cm plates plus the appropriate supplements for the recipient strain. Exconjugants were patched onto selective plates to test for Sm^s and for analysis by SCCL. Most colonies had retained

pAA131 pAA131, F'*lac,pro*

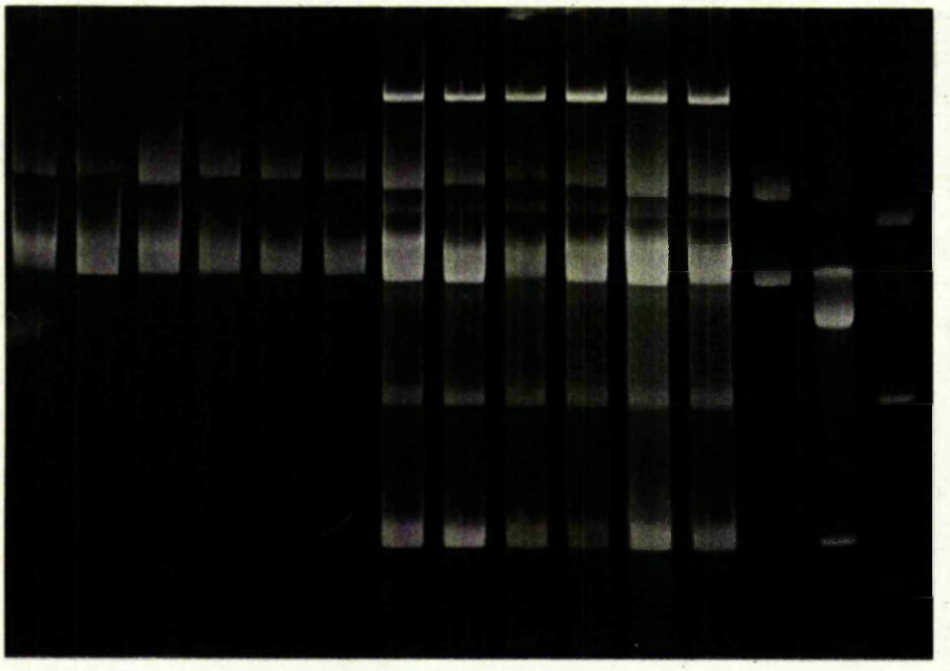


FIGURE 3.2 Resolution of pAA131 by F'*lac,pro*

Exconjugants from the cross between CSH52,F'*lac,pro* and JC9239, pAA131 were analyzed by SCCL to screen for resolution of pAA131. Although all the exconjugants contain the normal resolution products, pAA131 is still visible suggesting that resolution is not complete.

0.8% agarose gel run at 25V.

partial resistance to Sm, but growth was very poor; this indicated that repression was not complete.

The DNA content of exconjugants indicated four plasmid species; the F' factor, pAA131, plus the products of pAA131 resolution (figure 3.2). As pAA131 was still visible in all the colonies examined this suggested that resolution is not complete in the presence of the F' factor. This may be due to the fact that the copy number of F is stringently controlled (only 2-3 copies /cell); if the apparent resolvase activity present on F is autoregulated, as found for Tn*1/3*, then there may be very little free resolvase available in the cell. As pAA131 is a multicopy-number plasmid there may not be enough resolvase present to bind to all of the available *res* sites; or, alternatively, the recombination activity found on F may be inefficient at resolving Tn*1/3* cointegrates.

The pAA131 resolution products were identified by comparison with DNA size markers and by the ability of purified DNA from exconjugants to transform the markers for Cm^r and Sm^r independently. Analysis of single colonies from the recipient strain in the cross, which had not received the F factor showed no resolution, as detected by these methods.

The F factor is known to carry several different insertion elements, any of which could possibly be implicated in the observed recombination event (Davidson *et al.*,1974). To further test the role of the γδ element as a source of complementing recombination functions, an insertion of γδ into pBR322, termed pOX14, was used (Guyer,1978). Insertion mutants of pOX14, containing *Eco*RI linkers inserted into the *tnp*R and *tnp*A genes, have been constructed by Reed; these plasmids

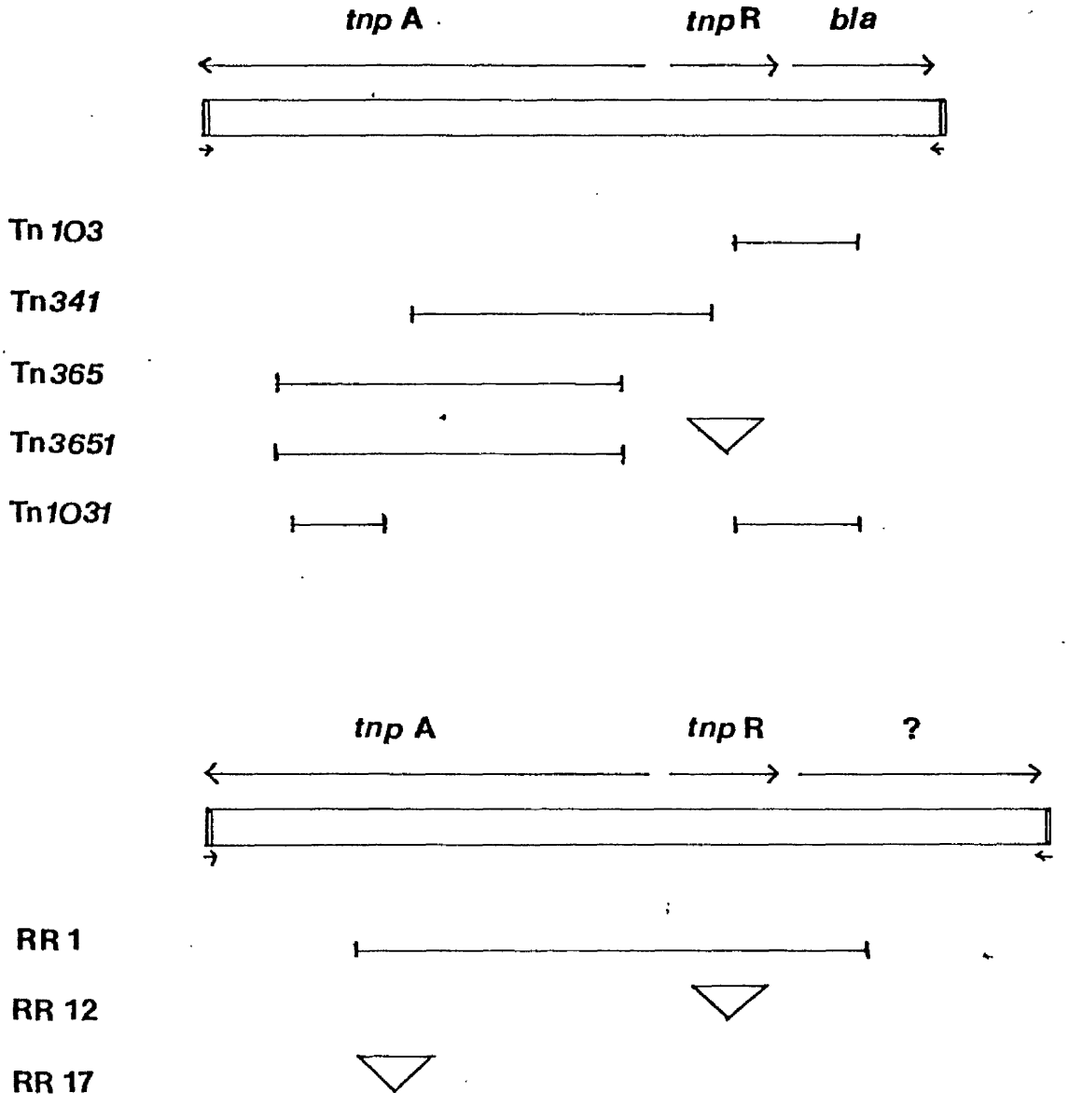


FIGURE 3.3 Maps of Tn3 and γδ indicating some deletion and insertion mutants

Lines represent deletion of a DNA segment, triangles denote insertion of a DNA segment.

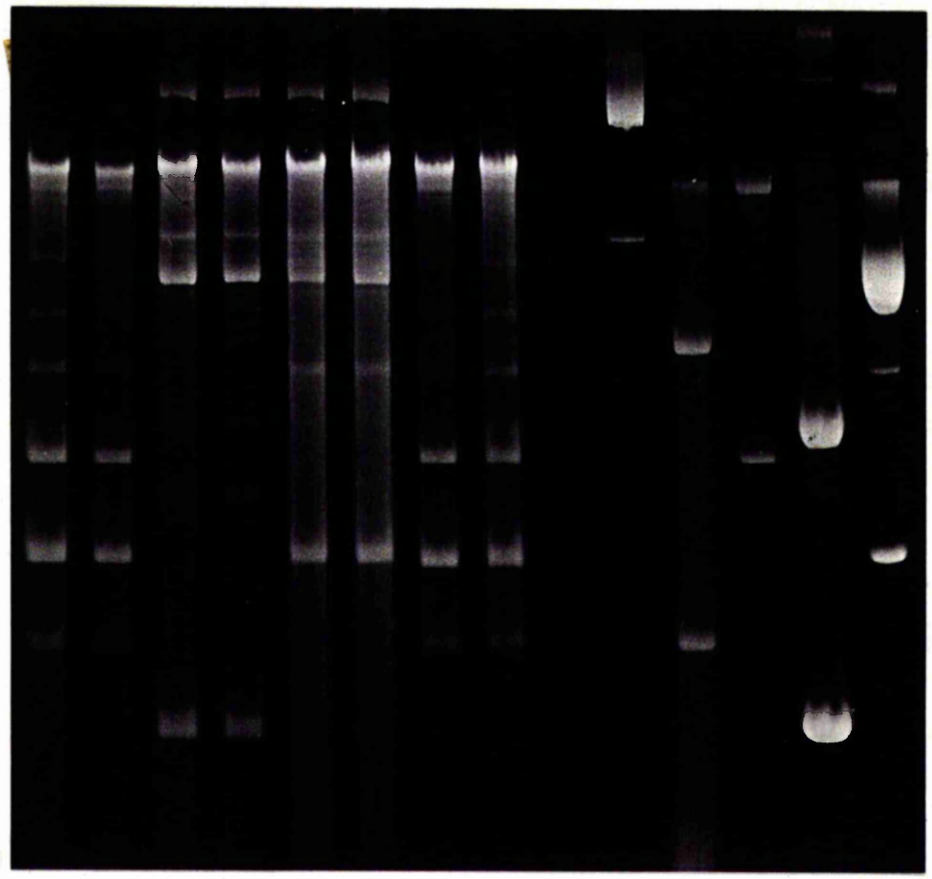


FIGURE 3.4 Resolution of pAA131 by derivatives of pOX14

Ap^r,Tc^r,Cm^r transformants were analyzed by SCCL to screen for resolution. Those derivatives coding for a functional *tnp*R gene product, pOX14 and pRR17 are able to resolve pAA131.

0.8% agarose gel run at 25V.

have been named RR12 and RR17, respectively (figure 3.3). A *tnp*A$^-$ *tnp*R$^-$ *res*$^-$ mutant has also been constructed by deletion of 3.2Kb of DNA between the two *Xho*I sites of γδ to form RR1 (Kitts *et al.*,1982b).

These mutants were tested for resolvase complementation by transforming into a pAA131-containing strain. Transformant colonies resistant to Ap,Tc and Cm were replica plated onto Sm. RR1 and RR12 (*tnp*R$^-$ elements) transformants maintained resistance to Sm, whereas pOX14 and RR17 transformants were completely sensitive to Sm. Transformants were also analysed by SCCL; pAA131 was efficiently resolved by plasmids pOX14 and RR17, but not by RR1 or RR12 (figure 3.4). These results suggested that γδ provides a function analagous to resolvase for resolution of Tn*1/3* cointegrates and repression of the *tnp*A promoter. It is assumed that the incomplete resolution observed, when F'*lac pro* was used to complement,was due to a copy number effect.

2. Resolution of γδ cointegrates by Tn*1/3*

A transpositional cointegrate containing directly repeated copies of γδ was constructed by transposition of RR12 (*tnp*R$^-$ derivative of γδ) into R388, followed by conjugation. Tp^r,Ap^r,Tc^r exconjugants were analysed by SCCL and shown to contain one large plasmid, larger than R388, and with an increased copy number, typical of a pMB1 replicon. This plasmid was stably maintained in RecA$^-$ hosts and subsequently transferred Tp^r,Ap^r,Tc^r at 100% compared with Tp^r transfer. The cointegrate containing strain, designated LS226 was independently transformed with each of the following DNAs: pACYC184, pACYC184::Tn*3*, pDS4153 (*tnp*A$^-$*tnp*R$^+$ derivative of Tn*1* in ColK) and pACYC184::γδ.

Analysis of the DNA content of individual transformants indicated resolution of the cointegrate into R388::γδ and RR12 by *tnp*R$^+$ derivatives of Tn*1/3*/γδ(figure 3.5). This data confirms the prediction that γδ and Tn*1/3* have complementing resolvase functions.

3. Interaction between Tn*1/3* and γδ transposase functions

The data presented above, and since by McCormick *et al.* (1981) and Casabadan *et al.* (1982), has shown complementation for resolution functions between γδ and Tn*1/3*. There has, however, been no evidence to suggest that the transposase proteins of the two elements interact. It has been reported that γδ, but not Tn*3*, can promote replicon fusion between the IS*101* element present on pSC101 (Ravetch *et al.*, 1979) and plasmid ColE1, suggesting that γδ complements transposition of IS*101* (Miller and Cohen, 1980). The ends of IS*101*, Tn*3* and γδ share great sequence homology (figure 3.6). Between Tn*3* and γδ there 28/35 identical bp; the two ends of IS*101* are not perfect inverted repeats, they show 22 or 29/35bp matches with Tn*3*. The comparison between γδ and IS*101* reveals 29 or 34/35 identical bp. The transposase proteins of Tn*3* and γδ share less than 30% homology, but as the sequences which they are thought to recognize, i.e. the inverted repeats of the elements, are highly homologous, it would be interesting to see if any interaction could be detected.

A transposition assay was devised to measure low frequencies of transposition of *tnp*A$^-$,*tnp*R$^-$,*res*$^-$ derivatives of Tn*3* or γδ into R388 in the presence of complementing elements. The *tnp*A$^+$ complementing plasmids used were pACYC184::Tn*103* or pACYC184::γδ (*tnp*R$^-$ derivative). The plasmids to be tested for transposition are both Mob$^-$, therefore the only way that these can be transferred during conjugation is by formation of a cointegrate with R388. Transpositional cointegrates

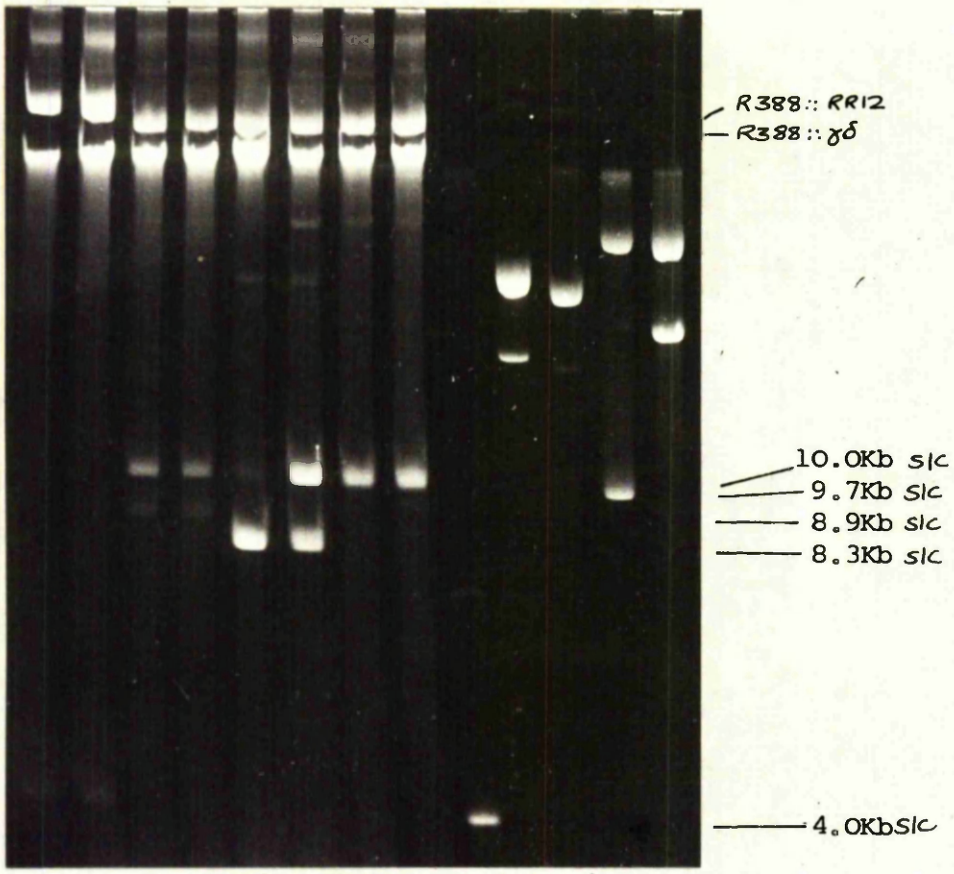


FIGURE 3.5 Resolution of LS226 by Tn*1/3* derivatives

The DNA content of transformants was analyzed by SCCL to screen for resolution. Plasmids containing a functional *tnp*R gene resolve pLS226 to the normal transposition end-products. (pLS226 ≡ R388::RR12)

0.8% agarose gel run at 25V

Tn*3* - both ends		GGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAG
γδ	γ end	GGGGTTTGAGGGCCAATGGAACGAAAACGTACGTT*TA T*
	δ end	GGGGTTTGAGGGCCAATGGAACGAAAACGTACGTT*AAG*
pSC101	end 1	GGGGTTTGAGGTCCAACCGTACGAAAACGTACGGTAAG
	end 2	GGGGTCTGAGGGCCAATGGAACGAAAACGTACGTTAGT

FIGURE 3.6 A comparison of the inverted repeat sequences of Tn*3*, γδ and pSC101.

The solid lines beneath sequences indicate homology to Tn*3*.

Strain	No. Tp^r, rif^r exconjugants*	No. Tp^r, rif^r, Ap^r exconjugants*	%age Cm^r	Frequency RSF1341/ RR1 transfer
(a)	8.2×10^8	4.0×10^1	0	4.8×10^{-8}
(b)	1.2×10^8	1.0×10^6	< 0.1	8.3×10^{-3}
(c)	2.6×10^8	1.0×10^2	> 90	3.8×10^{-7}
(d)	5.2×10^8	2.4×10^1	0	4.6×10^{-8}
(e)	2.3×10^8	1.1×10^4	< 0.1	4.7×10^{-5}
(f)	3.8×10^8	8.6×10^4	< 0.1	2.2×10^{-4}

TABLE 3.1 Transposition frequencies of RSF1341/ RR1 complemented with either Tn*3* or γδ transposase proteins

* indicates number of recipient cells receiving R388/ml.

thus formed transfer the antibiotic resistance carried by the small plasmid, Ap^r, with Tp^r carried by R388. The frequency of Tp,Ap transfer can be directly compared with the frequency of Tp transfer to give a measure of transposition.

The following strains were constructed by transforming the appropriate DNAs (maps of elements are shown in figure 3.3) into strain AA411, which contains R388:

(a) R388/RSF1341 (Tn3-*tnpA*⁻, *tnpR*⁻)

(b) R388/RSF1341/pACYC184::Tn*103* (*tnpR*⁻, *tnpA*⁺)

(c) R388/RSF1341/pACYC184::γδ(*tnp*R⁻)

(d) R388/pRR1 (γδ-*tnpA*⁻, *tnpR*⁻)

(e) R388/pRR1/pACYC184::Tn*103*

(f) R388/pRR1/pACYC184::γδ(*tnp*R⁻)

Transformants were analysed by SCCL to check for the presence of plasmids, already indicated by the antibiotic resistance of the strains (figure 3.7). Representative clones for each strain were grown at 30°C, the optimal temperature for transposition (Kretschmer and Cohen,1979), and used as donors for plate matings with DS916 (rif^r) as the recipient strain. Serial dilutions from each cross were plated onto isosensitest media containing rif,Tp to measure the frequency of R388 transfer, and rif,Tp,Ap to measure the frequency of RSF1341/pRR1 transfer. The apparent transposition frequencies obtained have been tabulated (table 3.1). Exconjugants from crosses (a) and (d), on analysis by SCCL, were shown to be donor mutations to rif^r, this effectively reduces their frequency of transfer to less than indicated. Colonies from crosses (b),(c),(e) and (f) all contained stable cointegrate plasmids (figure 3.8). The co-integrates from cross (c) had probably arisen by complex trans-position events in the donor strain involving all three plasmids

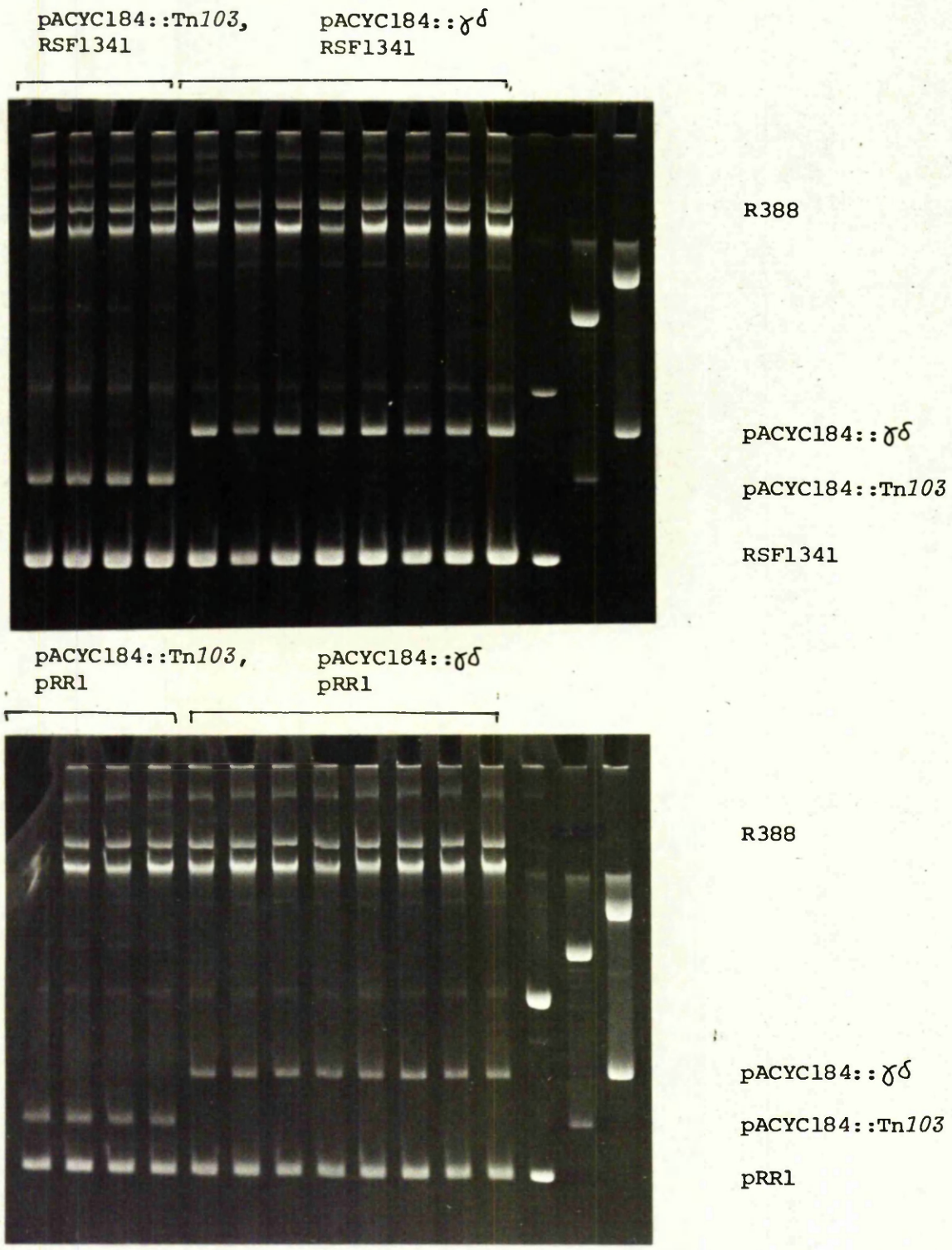


FIGURE 3.7 Analysis of plasmid DNA content of strains used for transposition assay

Tp^r,Ap^r,Cm^r transformants were analyzed by SCCL to ensure the presence of the plasmids for the complementation assay.

0.8% agarose gel run at 25V.

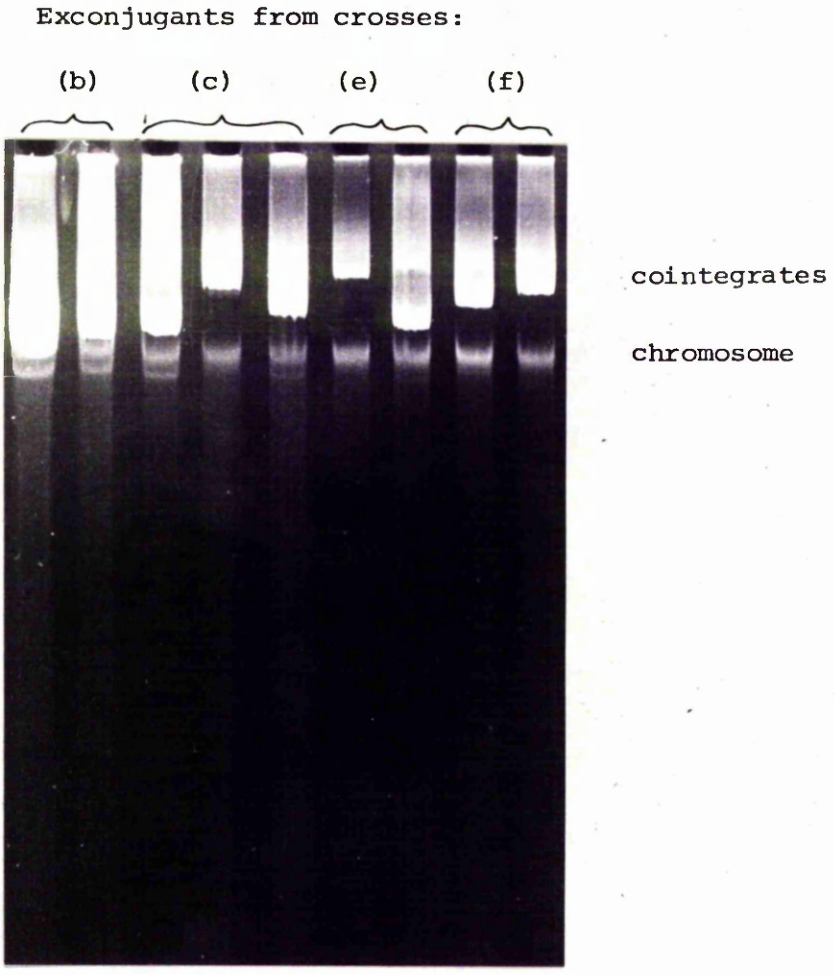


FIGURE 3.8 Analysis of exconjugants from matings

The DNA content of exconjugants from the transposition assay was analysed by SCCL. Cointegrate plasmids, identified by their size and copy number, are present in exconjugants from crosses (b),(c),(e) and (f).

0.8% agarose gel run at 25V

as these carried antibiotic resistance markers for Tp,Cm and Ap. It is likely that pACYC184::γδ had transposed into RSF1341, which had then inserted into R388. Exconjugants which were Cm^S from this cross proved to be donor mutations. As no transpositions involving just RSF1341 had occurred, it was concluded that γδ does not complement transposition of Tn*3*.

Cointegrate plasmids from crosses (b),(e) and (f) were all Cm^S and stable in the presence of *tnp*R^+ elements (data not shown), indicating that RSF1341/pRR1 (both *res*$^-$ derivatives) had transposed into R388. The cointegrates from the control crosses, (b) and (f), were as expected and provide a comparison of the transposition frequencies between the two elements. This indicates that transposition of Tn*1/3* is 40-fold higher than that observed for γδ. The low transposition frequency of γδ could result from less efficient transcription/translation signals, less efficient transposase protein or instability of the transposase protein.

The most surprising result was obtained from cross (e), where it appeared that Tn*1* could complement transposition of γδ, only five-fold down on the frequency observed using γδ (to complement transposition of γδ). This anomalous result could, however be due to an aberrant transposition event involving one end of γδ and the inverted repeat of Tn*3* present on pBR322, at the end of the *bla* gene (Bolivar *et al.*,1977). A scheme showing how this aberrant event could take place between pRR1 and R388 is indicated (figure 3.9). This would involve transposition of the entire region between the right inverted repeat of Tn*3* and the left inverted repeat of γδ. Such an event would result in duplication of the Ap^r and Tc^r genes; therefore the *Cla*I restriction site, present at the

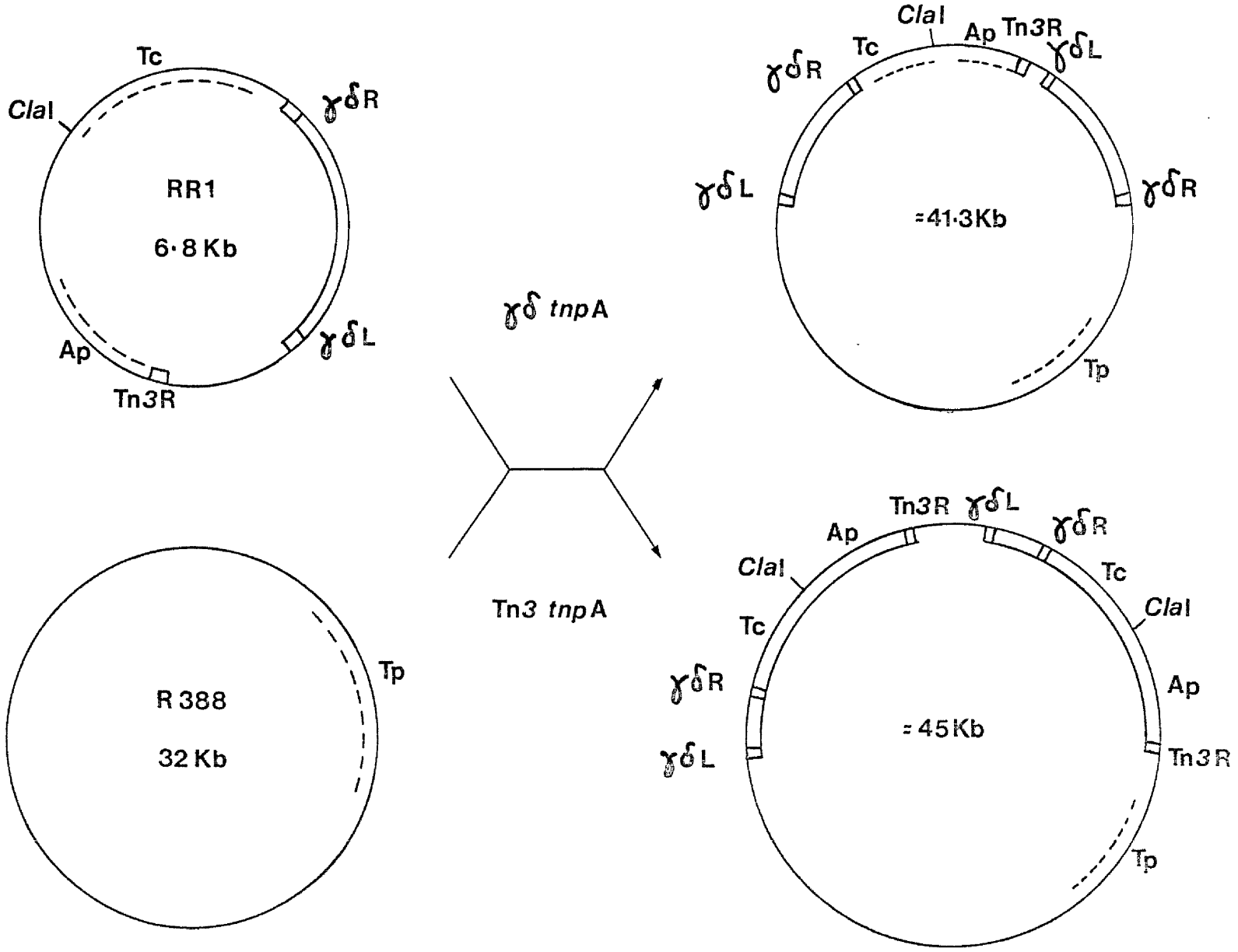

Tc
ClaI
γδR
RR1
6·8 Kb
γδL
Ap
Tn3R
γδ tnpA
Tn3 tnpA
ClaI
Ap
Tn3R
γδL
Tc
γδR
γδL
≈41·3Kb
γδR
Tp
R 388
32 Kb
Tp
Tn3R
γδL
Ap
γδR
ClaI
Tc
Tc
ClaI
γδR
Ap
≈45Kb
γδL
Tn3R
Tp

FIGURE 3.9 A scheme to show the formation of normal and aberrant transpositional cointegrates from pRR1

Transposon sequences are represented by double lines and antibiotic resistance genes by dotted lines. The positions of *Cla*I restriction sites are indicated. R and L denote the righthand and lefthand inverted repeats of the named elements, respectively.

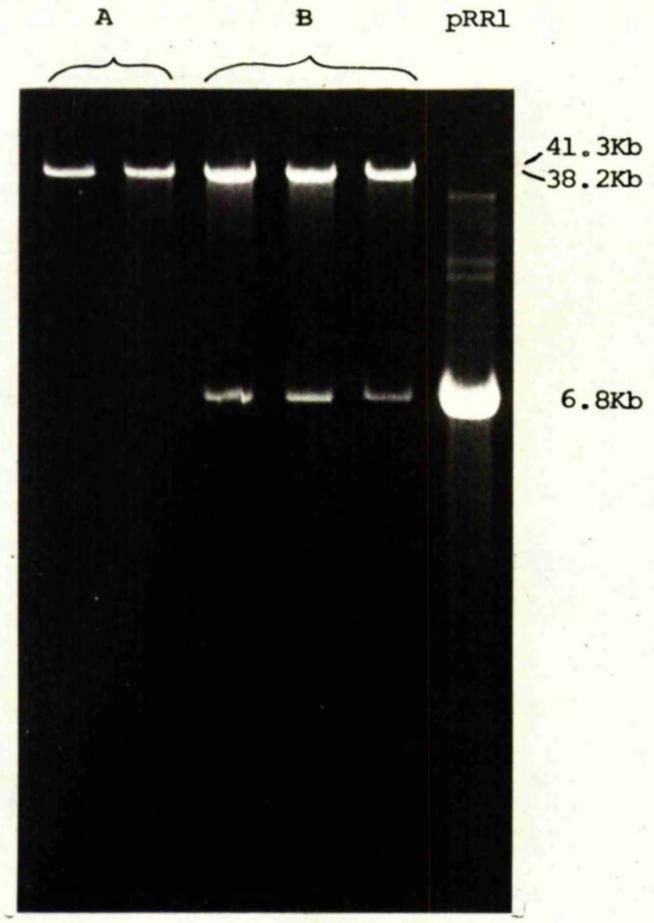


FIGURE 3.10 Restriction analysis of exconjugants

A - exconjugants from cross (f) restricted with *Cla*I

B - exconjugants from cross (e) restricted with *Cla*I

Of the exconjugants from cross (e) analyzed, all contained two *Cla*I restriction sites. This indicated that an aberrant transposition event had occurred involving one end of γδ and the Tn*3* inverted repeat present in pBR322 as shown in figure 3.9.

0.8% agarose gel run at 25V.

beginning of the Tc^r gene, would also be duplicated in the cointegrate. The normal transposition event results in duplication of just the γδ element and therefore cointegrates contain only one *Cla*I site (figure 3.9).

To test this possibility, several independent clones from crosses (e) and (f) were chosen (for comparison) and their DNA isolated. Restriction of purified DNA from cross (e) clones revealed the presence of two *Cla*I fragments of the sizes predicted; clones from cross (f) had only one *Cla*I site (figure 3.10). These results confirm the hypothesis that Tn*1/3* transposase can complement transposition of the hybrid element using one end of Tn*3* plus one end of γδ, in preference to both ends of γδ. As only a few clones were checked the possibility cannot be ruled out that a single γδ element can transpose using Tn*1/3* transposase or that the γδ transposase can complement transposition of the hybrid element. These experiments should really be repeated using the pRR1 element on a different replicon, such as pMB8 - to directly compare with RSF1341, before concluding definitively whether Tn*1/3* transposase can complement transposition of γδ.

Discussion

Genetic and biochemical evidence accumulated over recent years has suggested strong similarities between the Tn*1/3*elements and γδ. Sequence data has shown highly conserved regions between these two elements, in particular the inverted repeats, intercistronic region between the *tnp*A and *tnp*R genes, and the primary structure of the resolvase proteins. The inverted repeats have 28/35bp sequence homology in which there is a run of 12 identical

bp (figure 3.6). They also share the property of duplicating 5bp of target DNA during integration (Ohtsubo *etal.*, 1979; Reed *et al.*, 1979).

Genetic analysis has indicated that both elements appear to go through an obligatory cointegrate intermediate stage during transposition. Formation of the transpositional cointegrate is absolutely dependent on the *tnp*A gene product, or transposase (Kostriken *et al.*, 1981; Reed, 1981a; Kitts *et al.*, 1982a). Resolution of the cointegrated replicon occurs by a very efficient site-specific recombination event, mediated by the *tnp*R gene product, resolvase, at a defined site, *res* (Reed, 1981a).

It was observed by two independent groups in America that cointegrates formed by transposition of *tnp*R$^-$*res*$^+$ elements were rarely detected in the presence of F factor plasmids. These results conflicted with those obtained using R388 in the assay system, where cointegrates were exclusively formed by *tnp*R$^-$ elements (Gill *et al.*, 1978; Sherratt *et al.*, 1981a). This suggested to us that the F factor carried a function analagous to resolvase, which could reduce Tn*1/3* cointegrates to the normal transposition end products. The data presented in this chapter substantiate this hypothesis.

Artificial cointegrates have been constructed by cloning the *res* region from Tn*3* into a $\gamma\delta$ containing plasmid, and vice versa. Resolution of such plasmids yields products with hybrid crossover sites. By comparing the DNA sequences of the hybrid sites with those of the wild-type sequences it has been possible to identify the recombination point within a 19bp sequence, which is homologous between the two elements. (Kostriken *et al.*, 1981; Reed, 1981a). This

sequence is very AT rich, containing 16/19 AT residues. The 19bp sequence contains presumptive overlapping Pribnow boxes for both the *tnp*A and *tnp*R genes. This suggests that the dual roles of resolvase in resolution and repression may be mediated in a single step by specific binding within this region.

The high degree of sequence and function conservation between these two elements suggests a common evolutionary origin. Of the two gene products encoded by these elements, which are involved in their transposition, the resolvase proteins exhibit complementation, but apparently not the transposase proteins. This is not really surprising as the transposase proteins share less than 30% sequence homology, whereas the resolvase proteins are 80% homologous. Most of the models which have been formulated to explain the mechanism of transposition invoke a specific interaction between the transposase protein and the ends of the element (Arthur and Sherratt,1979; Shapiro,1979;Hershey and Bukhari,1981). The transposase proteins must be highly discriminating to distinguish between the inverted repeats of these two elements.

The results presented in section 3 suggest that Tn*1* transposase can recognize one inverted repeat from Tn*3* and one from $\gamma\delta$ to mediate an aberrant transposition event, albeit at reduced frequency when compared with Tn*3* transposition (about 200-fold down). The entire region between the right inverted repeat of Tn*3*, present at the end of the Ap^r gene, and the left inverted repeat of $\gamma\delta$ has been shown to transpose, including the Ap^r and Tc^r genes. It is possible that one inverted repeat may be more important than the other in recognition or during the transposition process. This possibility is at present being investigated by

another member of our group at Glasgow. The hybrid elements have been constructed and their transposition properties will be studied using either Tn*1/3* or γδ transposase proteins. These kinds of experiments should show if there is any functional asymmetry between ends. It may also be of interest to induce single site mutations within the ends of the elements by directed site-specific mutagenesis *in vitro*, to determine which bases are the most important for transposition.

The presence of a transposon-encoded site-specific recombination system appears to be unique to the Tn*3*-like family of elements. The IS elements, which appear to transpose either by direct transposition or cointegration, rely on the host-encoded general genetic recombination system to resolve cointegrates, when formed. Why have an element-encoded recombination system? The reason for this may be found by studying the biology of the two groups of elements. The IS elements are located mainly on the chromosome; most of the recombination events they are involved in are intramolecular. Current models for transposition suggest that intramolecular transposition results in deletion or inversion of intervening DNA according to which strands at the target site are cleaved/ligated during the initial stage of transposition (figure 1.)(Arthur and Sherratt,1979; Shapiro,1979).The only recombination event required for such rearrangements to occur is the initial break/join reaction mediated by the transposase protein; a resolution step is not necessary. It has been shown for Tn*3* that deletion formation, by intramolecular transposition, is independent of the *tnp*R gene function (Bishop, pers.comm.). One could argue that elements, which transpose preferentially by direct transposition, or intramolecularly, would not require their own recombination

system. The Tn*3*-like elements occur most frequently on plasmids, often encoding accessory determinants such as antibiotic resistance. Their efficient spread through populations is underlined by the rapid increase of multiple drug resistance in bacteria. It has been shown that the host-encoded Rec system resolves transpositional cointegrates very poorly (Arthur and Sherratt,1979). Elements which always use the cointegrate pathway would require a RecA-independent, transposon-encoded system for resolution to ensure their efficient spread through populations.

Transposons usually encode accessory determinants, which provide their hosts with a selective advantage under appropriate conditions; one could argue that there is strong natural selection to maintain these elements in populations. But what of IS elements? These provide no direct selective advantage to their host, but their ubiquitous occurrence in a diverse range of organisms pinpoints their success. One may argue that these elements confer some selective advantage on their hosts by mediating genetic rearrangements, which could increase the adaptivity of the host and hasten the evolution process. They also provide portable regions of homology at which the host-encoded general genetic recombination proteins may act to generate deletions, duplications and inversions. Transposition of these elements is usually tightly regulated; if not, the gross rearrangements which could occur may be suicidal to the host and thus to the element contained within it. Elements which have the property of replicating by horizontal transmission, as well as vertically to daughter progeny during vegetative propagation have often been called "selfish DNA elements" on the basis that they have no phenotypic advantage, they just replicate. For further discussion of these issues see Doolittle and Sapienza

(1980;1981) and, Orgel and Crick (1980).

The simple IS elements and the Tn*3*-like family of transposable elements appear to have fairly distinct properties, suggesting that the two classes of elements may have evolved independently, or, if they are derived from a common ancestor, have diverged considerably. Their transposase proteins generally duplicate 9bp or 5bp of target DNA during integration, respectively; this may be indicative of different mechanisms and/or different specificities during the initial break/rejoin event. The greatest difference is, however, the presence of a site-specific recombination protein encoded by the Tn*3*-like elements. Both classes of elements have evolved efficient systems for their propagation; the IS elements by undergoing direct transposition,thereby alleviating the necessity for a cointegrated intermediate, and the Tn*3* elements by using site-specific recombination to resolve obligatory cointegrates.

Site-specific recombination has been recognized in a variety of systems as an efficient means of generating rearrangements,in some cases to control gene expressinn, e.g. inversion of the G-loop in phage Mu to control production of host-range proteins (Bukhari and Ambrosio,1977). Of these the lambda integration system has been well characterized using genetic and biochemical techniques. The Tn*3* resolvase system has been investigated in detail,genetically, by several members of our group at Glasgow. To discover the exact requirements, and shed some light on the mechanism of the resolvase-mediated recombination event, I decided to tackle the problem using a biochemical approach. The aim of the experiments in the following chapters was to purify the resolvase protein and study it s action in an *in vitro* system.

CHAPTER 4

CLONING AND PURIFICATION OF RESOLVASE

Introduction

One of the most successful approaches used to study a particular system in detail has been to reconstruct it *in vitro*, complementing the genetic analysis. This has proved most fruitful in the study of DNA replication and, more recently, for various recombination systems. Temperature sensitive mutations had identified many gene products involved in DNA replication, but these could only be classified as initiation or elongation factors; their precise function and mechanism were more difficult to determine. Using plasmid or phage substrates, a detailed biochemical analysis has been made of the replication apparatus. Principally, replication defective mutants were used to provide crude cell extracts for use in complementation assays *in vitro*; this allowed purification of various replication proteins, which were then reconstituted to partial or complete systems (Alberts and Sternglanz.1977; Kornberg,1980). These techniques have uncovered many proteins involved in DNA replication, most of which interact to form large complexes.

The use of *in vitro* systems to study recombination is a fairly recent innovation due mainly to the lack of a suitable assay. Work has, however, progressed rapidly in the understanding of phage lambda integrative recombination. This is mediated by the phage-encoded protein, Int, and acts specifically at a site located on the phage, *att*P, and on the host, *att*B (see Nash,1981 for a review). The use of a cell-free system has aided the purification of the proteins involved and has prompted a model for the recombination mechanism. The host-encoded general genetic recombination system has also been probed *in vitro*. The work of C.Radding and P.Howard-Flanders at Yale University has revealed many of the recombinational properties of *recA* protein (Radding,1981; West *et al.*,1981). Studies using purified *recBC* protein,

or *exoV*, *in vitro* have also led to a proposed mechanism of action of this protein during recombination (Muskavitch and Linn,I980; Taylor and Smith,I980a,b). The role of host- encoded recombination proteins in plasmid recombination is at present being investigated *in vivo* and *in vitro* (Kolodner,I980 and pers. comm.).

All of these systems rely on an assay system, which clearly distinguishes between substrate and product molecules. Once a suitable assay has been devised it may be possible to isolate the particular proteins involved from crude cellular extracts. This may be achieved by using a strain which produces the protein in fairly large quantities, or by cloning the gene into a high expression vector, thus allowing large amounts of the protein to be synthesized.

The experiments described in this chapter were initiated with the intention of studying resolvase-mediated site-specific recombination *in vitro*. Production of the *tnp*R protein, or resolvase, is known to be autoregulated (Chou *et al.*I979a); thus even on a high copy number plasmid very little free resolvase is present in the cell. To produce sufficient quantities for purification it has proved necessary to separate the *tnp*R gene from it s natural promoter and place it under the control of a different promoter, thereby releasing it from repression. One of the problems encountered in cloning genes under high expression promoters is the possibility that the protein may be lethal to the cell in high concentrations. This may be overcome by using a promoter which is regulated under certain conditions. Many of the commonly used expression vectors are derived from pBR322 (Bolivar *et al.*,I977), and contain an insertion from phage lambda carrying the leftward promoter, P_L. This promoter is usually controlled by the lambda repressor protein,CI. A temperature sensitive CI repressor has been isolated, CI^{857}, which allows

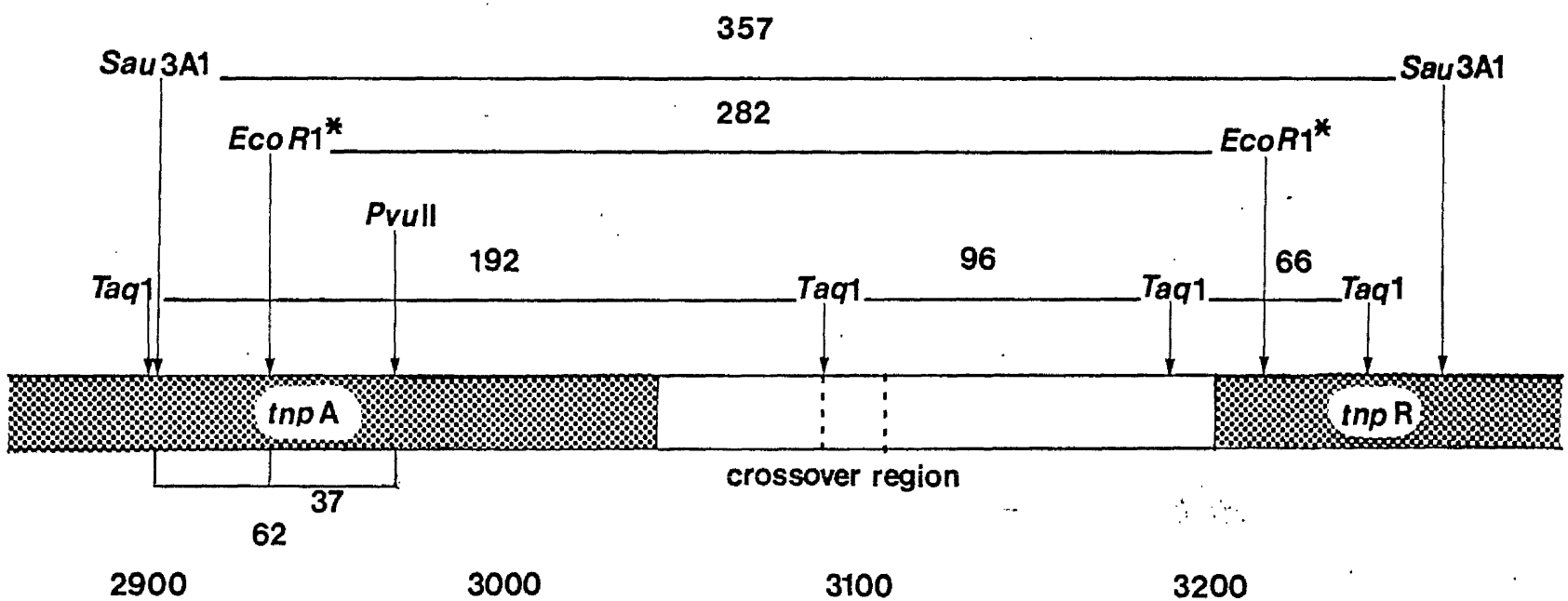


FIGURE 4.1 Restriction map of the Tn*3* intercistronic region

The locations and distances between restriction sites are indicated. The crossover region, defined by Reed (1981a) and Kostriken *et al.*, (1981), is depicted within the intercistronic region; this 19bp sequence also contains the presumptive Pribnow boxes for both *tnp*A and *tnp*R genes (Pribnow, 1975).

uncontrolled transcription from P_L at 42°C (Franklin,I97I; Bernard and Helinski,I979). The plasmid vector, pλ8, has an IIOObp *Bam*HI/*Bgl*II fragment from lambda, containing the P_L promoter and the N gene, inserted into the *Bam*HI site of pBR322 (Reed,I98Ib). This plasmid has been used as a vector for cloning the Tn*3* *tnp*R gene to produce sufficient quantities for purification.

The strategy for cloning has been to first isolate the gene from it s natural promoter by cloning a partial *Taq*I digest into pACYCI84, then to insert the isolated *tnp*R-containing fragment into pλ8 so that transcription reads through from P_L.

Results

I. Isolation of the *tnp*R gene

The nucleotide sequence of Tn*3* reveals that *Taq*I is the only enzyme, which cleaves between the putative Pribnow box and the beginning of the *tnp*R structural gene. As *Taq*I recognizes a 4bp sequence it does have several sites within Tn*3*, two of which are in the structural gene for *tnp*R and two in the intercistronic region between *tnp*R and *tnp*A (see figure 4.I). A partial digest of Tn*3* should produce some fragments carrying *bla* and extending leftwards through *tnp*R into the regulatory region.

The initial stage of the cloning strategy involved partially digesting a Tn*3*-containing plasmid with *Taq*I to generate a *bla*, *tnp*R fragment, using Ap^r as a selective marker for cloning. These fragments were cloned into the single *Cla*I site of pACYCI84; reaction mixtures containing ligated DNA were transformed into DS825. Transformant clones resistant to Cm and Ap were analyzed by SCCL to screen for plasmids of about 6kb; i.e. pACYCI84 containing a 2Kb insert.(fig.4.2). Plasmids of about this size

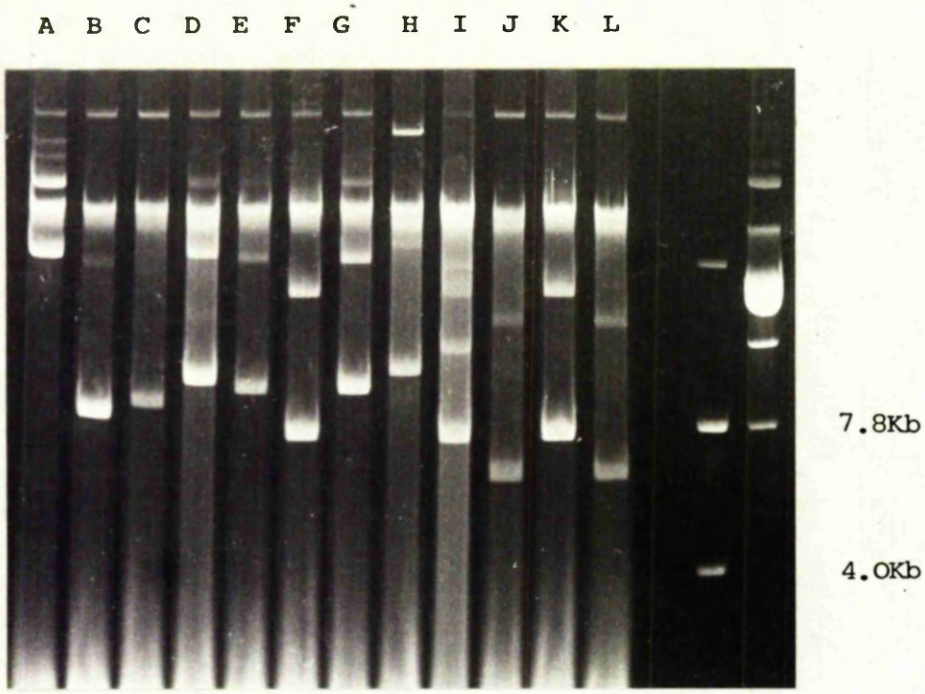


FIGURE 4.2 SCCL of Cm^r,Ap^r transformants from cloning stage 1

Plasmids which were within the 6-8Kb size range were purified for further analysis by *Taq*I restriction.

0.8% agarose gel run at 25V.

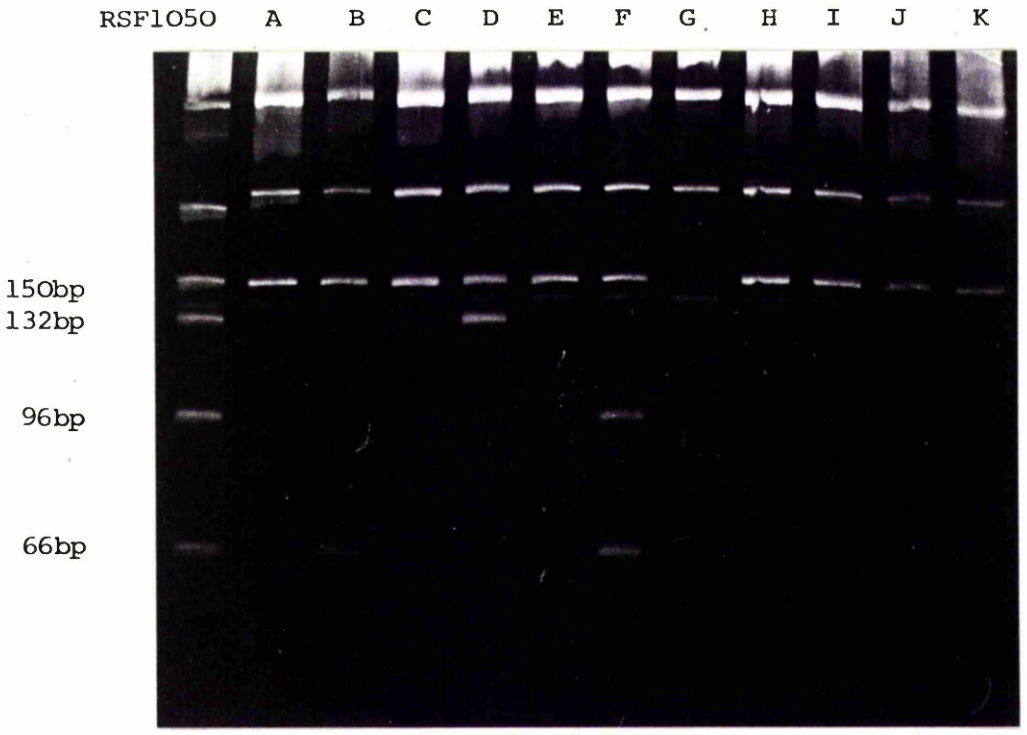


FIGURE 4.3 Restriction analysis of prospective *tnp*R-containing clones

The prospective clones were restricted with *Taq*I; the first track represents the fragments produced from a wild-type Tn*3* element. Plasmids represented in tracks B and K contain the 66bp and 150bp fragments required for the *tnp*R structural gene, but lack the 96bp promoter-containing fragment.

10% polyacrylamide gel run at 30mA.

were purified and analyzed further by *Taq*I restriction; this indicated which fragments had been cloned. To allow detection of small DNA fragments (less than IOObp), digests were run on IO% acrylamide gels (figure 4.3). The gel indicated that two of the clones analyzed contained the I5Obp and 66bp fragments of *tnp*R, but lacked the 96bp fragment containing the putative promoter sequence. It was assumed that the order of these fragments was the same as in wild-type Tn*3* and one of these clones, named pLS2O4, was used for the subsequent cloning stage.

Plasmid pLS2O4 has no apparent repressor or resolvase activity, when analyzed *in vivo*; this was expected as the gene had been isolated from it s promoter and transcriptional studies of the region around the *Cla*I site suggest that there are no other promoters present, which could read through the insert (Stüber and Bujard,I98I). Quite fortuitously the junction sequences generated by insertion of the *Taq*I fragment were *Cla*I recognition sites. This provided a means of lifting the *tnp*R and *bla* genes from pLS2O4 on a single *Cla*I fragment.

2. Insertion of the *tnp*R gene into pλ8

The pλ8 plasmid is unstable in non-lysogenic *E.coli* strains; this effect is presumably due to uncontrolled transcription interfering with plasmid functions, probably replication (Bernard and Helinski,I979). To overcome this problem all manipulations were performed in either a wild-type λ lysogen or in the strain N483O, which contains the CI^{857} repressor gene, but has most other λ genes deleted (Reed,I98Ib). For cloning purposes there is a *Cla*I site downstream from the P_L promoter, five triplets before the end of the N gene. There are two other *Cla*I sites in pλ8, one other within the IIOObp cloned λ fragment and the other at the beginning of the Tc^r gene of pBR322. There is also a single *Pst*I site located within the Ap^r structural gene; this is also present in the Ap^r

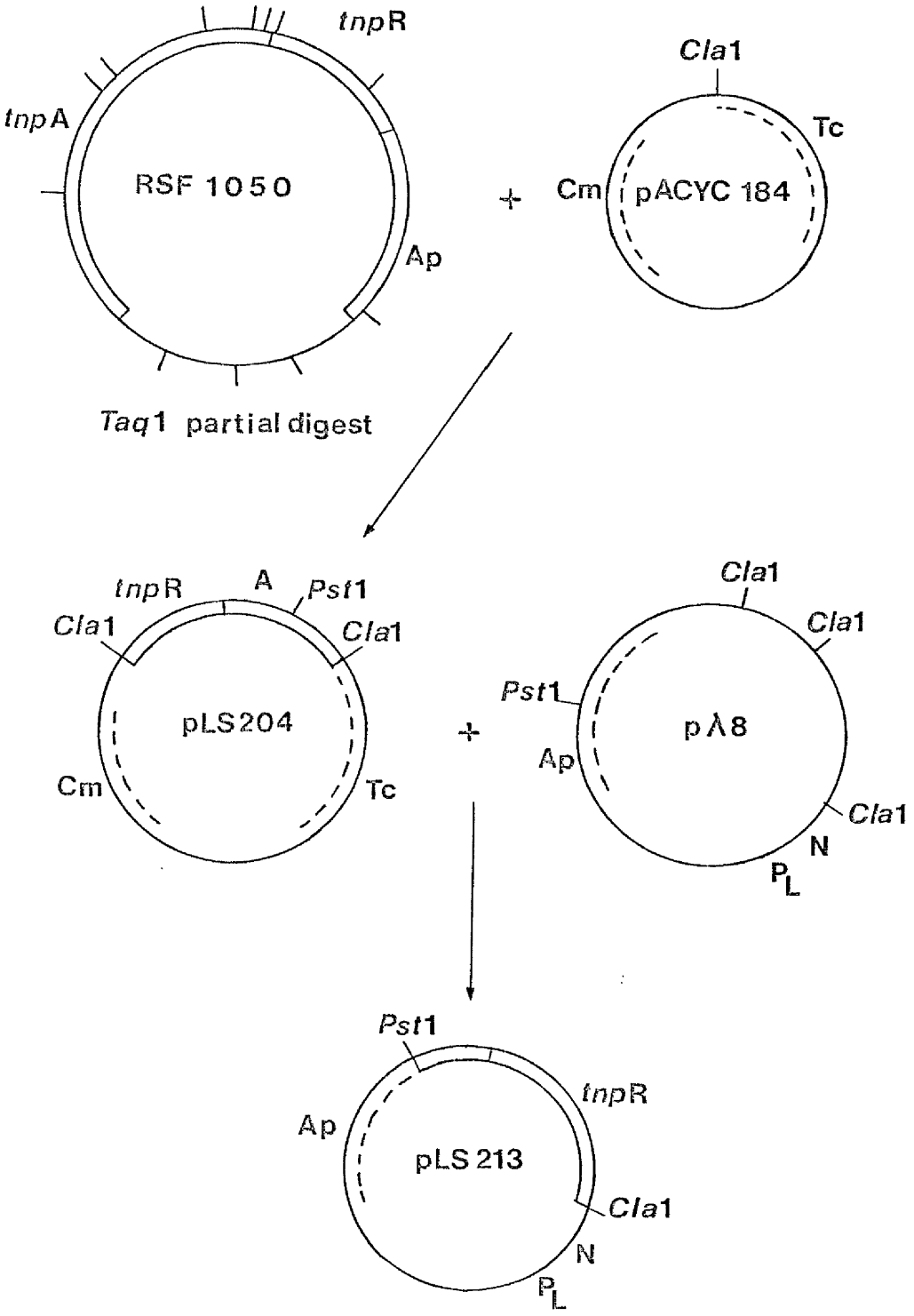

tnpR
tnpA
RSF 1050
Ap
+
Cla1
Tc
Cm
pACYC 184
Taq1 partial digest
tnpR
A
Pst1
Cla1
Cla1
pLS204
Cm
Tc
+
Cla1
Cla1
Pst1
pλ8
Ap
Cla1
N
PL
Pst1
tnpR
Ap
pLS 213
Cla1
N
PL

FIGURE 4.4 Strategy for cloning the *tnp*R gene from Tn*3* into pλ8

A *Taq*I partial digest of RSF1050 was mixed and annealed to *Cla*I digested pACYC184 DNA. Of the Cm^r,Ap^r clones obtained, some contained the complete structural gene for *tnp*R, but lacked the promoter region. One of these, named pLS204, was then digested with *Cla*I and *Pst*I and ligated to *Cla*I/*Pst*I restricted pλ8 DNA. Two major size classes were formed; the larger of these contained the pλ8 replicator fragment plus the *tnp*R-containing fragment from pLS204.

gene of pLS204 as the *bla* gene of pBR322 was originally derived from Tn*3* (Bolivar *et al.*, 1977). Cleavage of pλ8 with *Cla*I and *Pst*I yields four fragments, the largest of which carries *oriV*, P_L, most of the N gene and the terminal end of the *bla* gene. Cleavage of pLS204 with *Cla*I and *Pst*I yields a 1.3Kb fragment containing *tnp*R and the proximal end of the *bla* gene. When these two digests are mixed and annealed, the simplest way of regenerating the *bla* gene is by ligation of the major pλ8 fragment to it's original adjacent fragment, or by ligating to the *tnp*R,*bla* fragment from pLS204 (figure 4.4). Such an experiment was performed and ligation mixtures were transformed into strain N4830; selection was made for Ap^r colonies at 30°C, the permissive temperature. Transformants were analyzed by SCCL to screen for plasmids of the predicted size, about 5Kb. The two major size classes obtained were of 4.5 and 5.0Kb, of these the larger size class were assumed to have inserts containing *tnp*R(figure 4.5). Clone pLS213(track L) was analyzed further by restriction of purified DNA with *Cla*I and *Pst*I and compared with similar digests of pλ8 and pLS204 (figure 4.6). The digests clearly show that pLS213 has the structure predicted.

3. Detection of a functional *tnp*R gene product

The physical structure of pLS213, as determined by restriction analysis, suggests that the *tnp*R gene had been cloned, including the ribosome binding site 8bp upstream from the ATG start codon (Shine and Dalgarno,1974). For the gene to be expressed it is required that the N* gene is terminated before the beginning of the *tnp*R message. If there is no stop codon in frame with the N* gene, then a fusion polypeptide would probably result. However, analysis of the sequences involved reveals the presence of a stop codon, UAA, just after the *Cla*I site; therefore a functional protein should be produced.

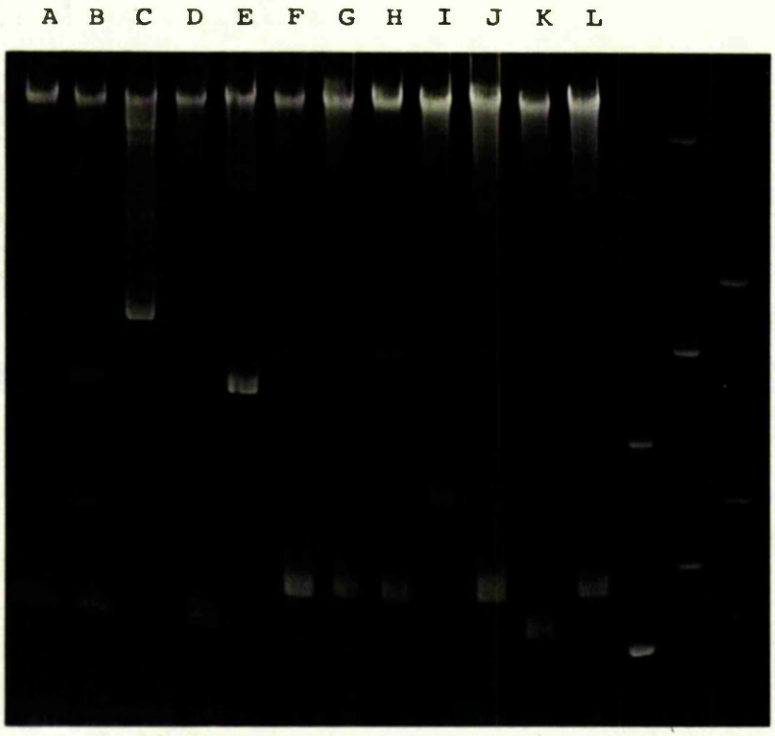


FIGURE 4.5 SCCL of Ap^r transformants from cloning stage 2

Tracks F,G,H,J and L represent plasmids of the size predicted for pLS213(5.0Kb). The other major plasmid size class, represented in tracks A,B,D and K were probably due to ligation of the 3.7+ 0.77Kb *Cla*I/*Pst*I fragments of pλ8.

0.8% agarose gel run at 100V.

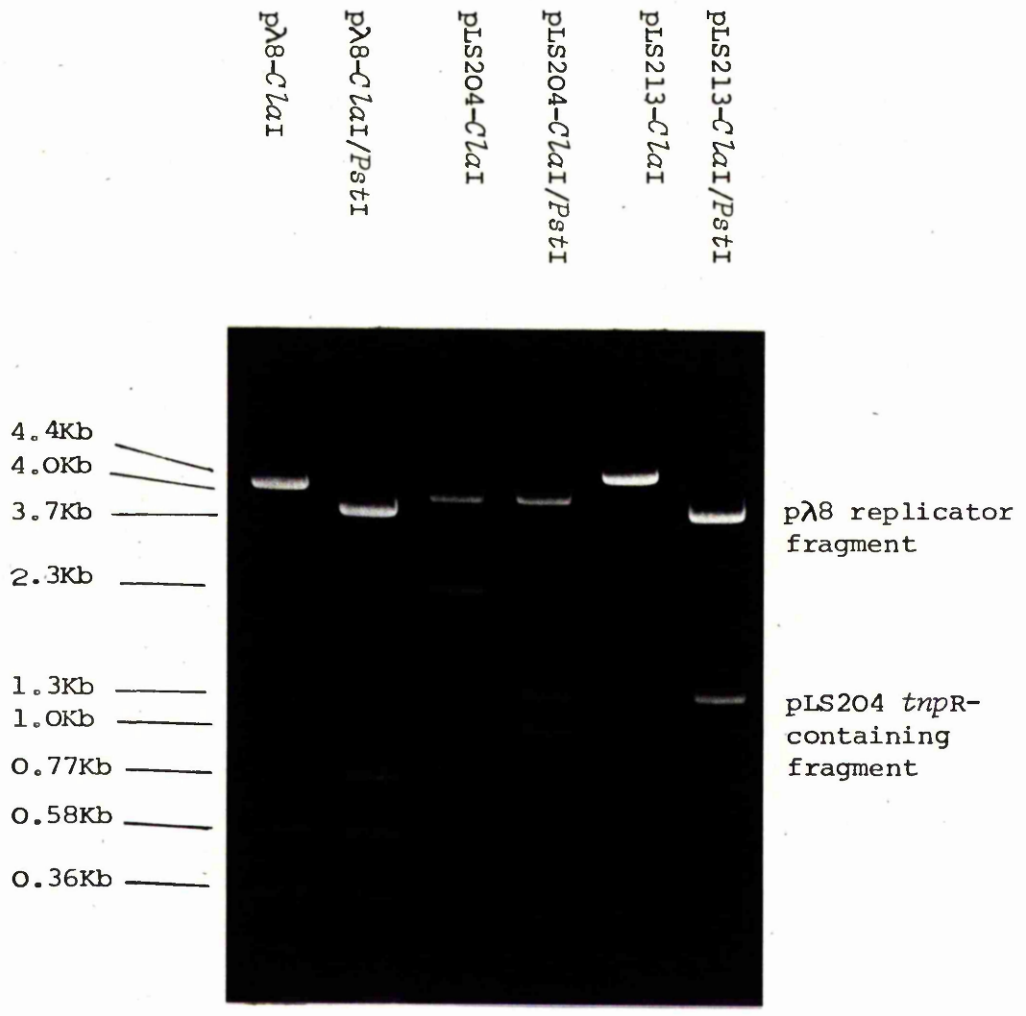


FIGURE 4.6 Restriction digests of pλ8, pLS204 and pLS213

Purified DNA was digested with *Cla*I *or* *Cla*I/*Pst*I. The 3.7Kb pλ8 replicator fragment and the 1.3Kb *tnp*R-containing fragment are both present in pLS213.

0.8% agarose gel run at 100V.

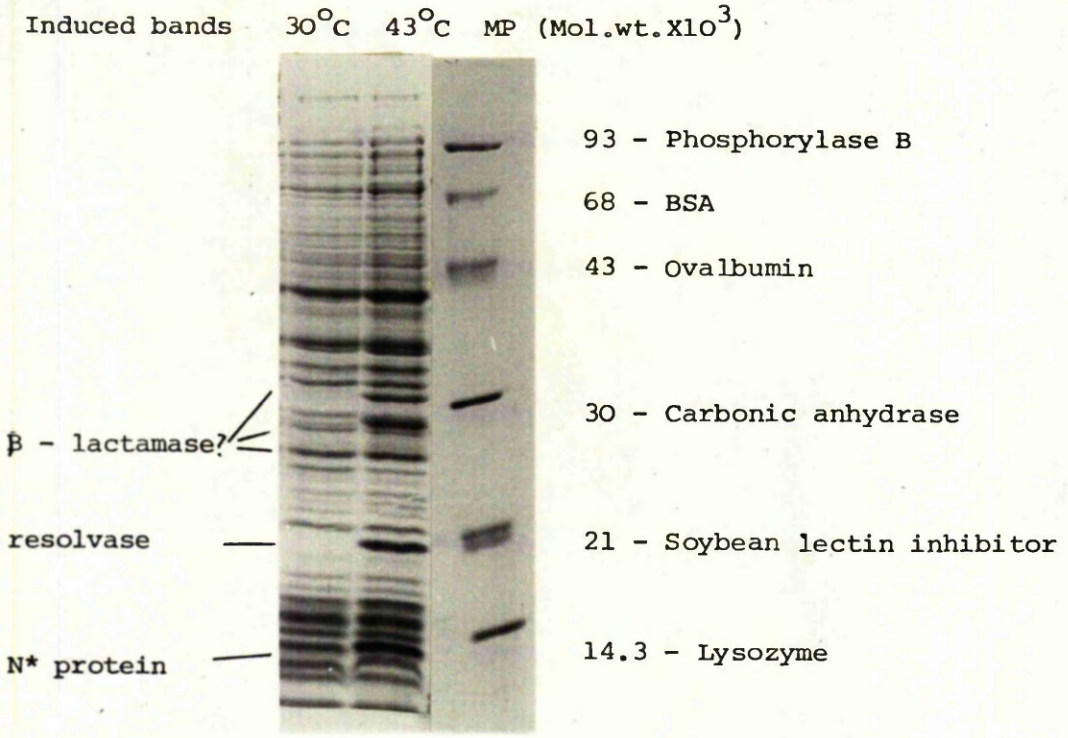


FIGURE 4.7 SDS-PAGE of total cell lysates from uninduced and induced LS415 cells

Five bands increase in intensity on heat induction as shown. 12.5%,4.5% stacking gel run at 35mA

To check that resolvase was in fact being produced in large amounts from pλ8 it was necessary to assay the total cellular proteins present, by SDS-PAGE. The strain containing pLS2I3, named LS4I5, was grown in 4ml L-broth at 30°C to 5XIO^{8}cells/ml with constant shaking. 2ml of this culture were transferred to 43°C for one hour to induce transcription from P_L. The samples of induced and non-induced cultures were centrifuged at I2,000g for 2min. to pellet the cells. The cells were resuspended in IOOμl of PFSB and boiled for 2min. The total protein content was analyzed by SDS-PAGE.(fig.4.7). There appear to be five protein bands which increase in intensity when the cells are induced. By comparison with size markers one of these has a molecular weight of 2IK, corresponding to the size of resolvase. The smallest of the induced proteins, about I2K, is presumably the N^* protein. The remaining three induced bands are due to β-lactamase and it s precursors; this suggests that the Ap^r gene is also under some control by P_L (Dougan *et al.*,I979). A rough estimate of the amount of resolvase produced can be made by measuring the density of stained bands on a gel showing induced and non-induced cell samples. The peaks on the densitometer trace representing the protein profile were cut out and weighed. The protein peak corresponding to resolvase was compared with the total cellular protein of induced cells to give a value of 3.8%. It was concluded from this data that the *tnp*R gene had been successfully cloned under P_L control; induction of transcription elevates the level of resolvase to approximately 4% of the total cellular proteins.

As the *tnp*R gene had undergone several manipulations during cloning it was considered necessary to check that the protein produced was still functional *in vivo* before proceeding with the purification. Two assays were used to test for the presence of a functional protein: repression of transcription from *tnp*A, and screen for resolution.

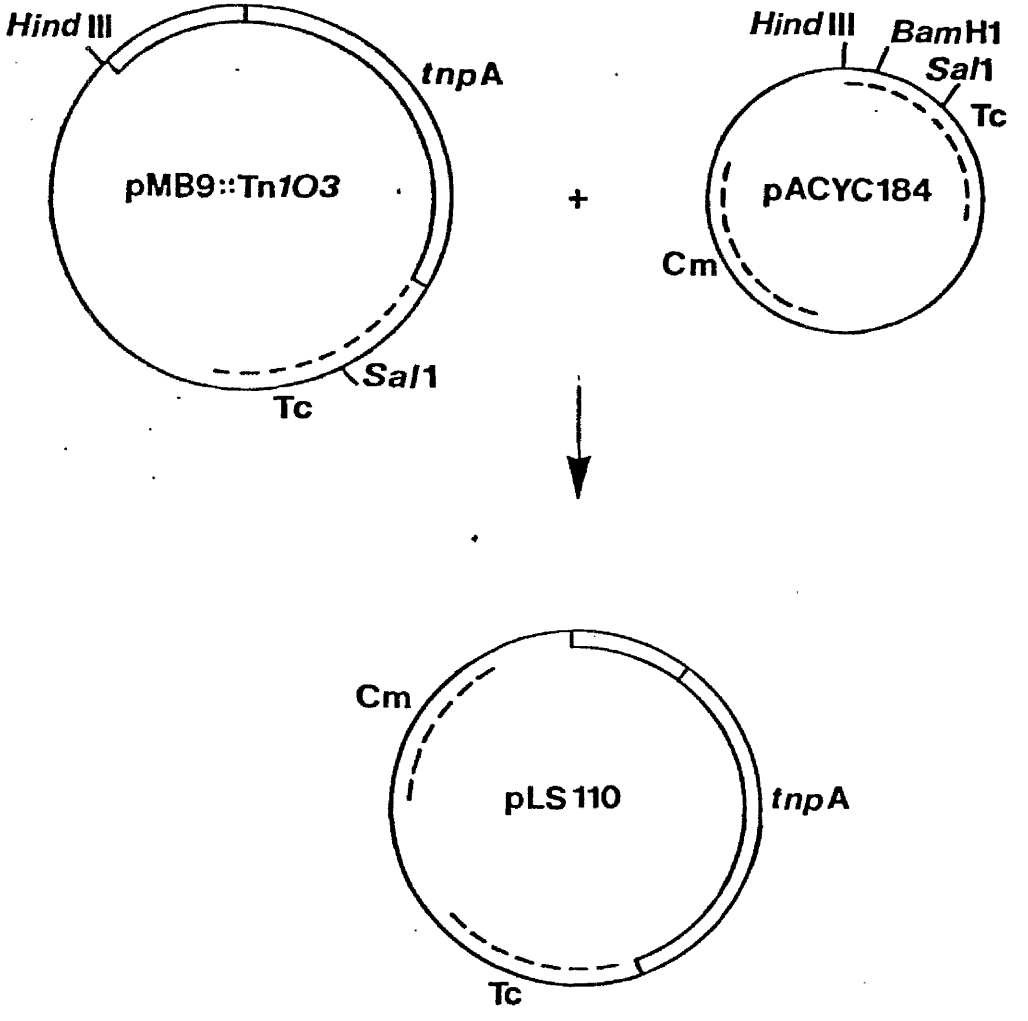


FIGURE 4.8 Strategy for construction of pLS110

pMB9::Tn*103* DNA was digested with *Hind*III/*Sal*I to produce a fragment carrying the entire Tn*103* element and part of the Tc^r gene. This was ligated to *Hind*III/*Bam*HI/*Sal*I restricted pACYC184 DNA. The major size class of Cm^r,Tc^r transformants was of the size predicted for pLS110. *In vivo* analysis indicated that the Tc^r structural gene of this plasmid was under *tnp*R control as expected.

(i) The repression assay makes use of a fortuitous insertion of the transposon Tn*I03* (*tnp*A$^+$,*tnp*R$^-$) into the promoter region of the Tc^r gene of pMB9. It was observed that in the presence of resolvase this plasmid was completely Tc^s, indicating that expression of Tc^r gene is directed from the *tnp*A promoter. The plasmid pMB9 is in the same compatibility group as pBR322 and therefore cannot be maintained within the same cell as pLS2I3. As pACYCI84 is in a different compatibility group, the Tn*I03* insertion was transferred,by cloning, from pMB9 to pACYCI84 as these have analogous Tc^{r}genes (figure 4.8). These two plasmids have been used to select resolvase amber mutations (Kitts *et al.*,I982b) and could potentially be used to isolate *res* site mutants.

It is assumed that even at 30^{o}C there is a small amount of resolvase present in strain LS4I5 due to low level of readthrough from P_L in the presence of CI^{857}. Strain LS4I5 was transformed with pLSIIO, containing the repressible Tc^r gene; transformant clones resistant to Cm and Ap were subsequently checked for Tc^r by replica plating. All Cm^r colonies were Tc^s indicating that the cloned gene produces functional resolvase to repress transcription from the *tnp*A promoter.

(ii) To check for resolution activity a transpositional cointegrate derived from Tn*I03*, pAAI3I, was transformed into strain LS4I5. Transformants resistant to Ap and Cm were analyzed by SCCL and shown to contain only the resolution products of pAAI3I and pLS2I3 (data not shown). Together these data indicated that a functional *tnp*R gene had been cloned and that resolvase could be produced in large amounts by heat induction.

4. Purification of resolvase (Reed, 1981b)

Fresh overnight cultures of LS4I5 in L-broth were diluted I:20 into two litres of L-broth, with the addition of 50µg/ml ampicillin to maintain selection for the plasmid.Cultures were grown to about 5XIO^8 cells/ml

at 30°C. Flasks were transferred to 43°C to induce transcription from P_L. Culture flasks containing at least 500ml of liquid would normally take 20-30 min. to heat to 43°C; an hour after this point the cells were harvested.

Cells were harvested by spinning at 9,820g for 5min. Pellets were resuspended in 20ml of 100mM NaCl, 20mM Tris, pH7.5. The cells were again pelleted, then resuspended in 10ml of 10% sucrose/TEM buffer (TEM: 20mM Tris, pH7.5; 1mM EDTA; 1mM mercaptoethanol). The cell suspension was sonicated in four 30 sec. bursts using a Dawe Soniprobe type 7532B. The crude lysate was centrifuged at 12,000g for 10min. to remove unlysed cells and cell debris. The supernatant was carefully decanted and to it added polymin P (10% w/v in TEM), dropwise, to a final concentration of 0.5%. This was left on ice for 10min. to allow polymin P to aggregate with DNA/protein complexes; these were then precipitated by centrifugation at 12,000g for 10min. The supernatant was discarded; the pellet was resuspended in 0.3M NaCl/TEM and left on ice for 10min. At this concentration of NaCl most proteins dissociate from DNA, but most of the resolvase remains bound. The DNA pellet was collected by centrifugation as before, then resuspended in 10ml of 1M NaCl/TEM and left on ice for 10min. to remove resolvase from the DNA/polymin P complex. The DNA/polymin P pellet was removed by centrifugation.

The resulting supernatant was loaded directly onto a 100X2.5cm. Sephadex G-75 (medium grade) gel filtration column, equilibrated with 1M NaCl/TEM. Fractions were collected using an LKB 7000 Ultrorac fraction collector and scanned using an LKB 8300 Uvicord II, set at 280nm during collection. The uvicord trace, though not an accurate measurement for optical density, clearly shows where the void volume is eluted

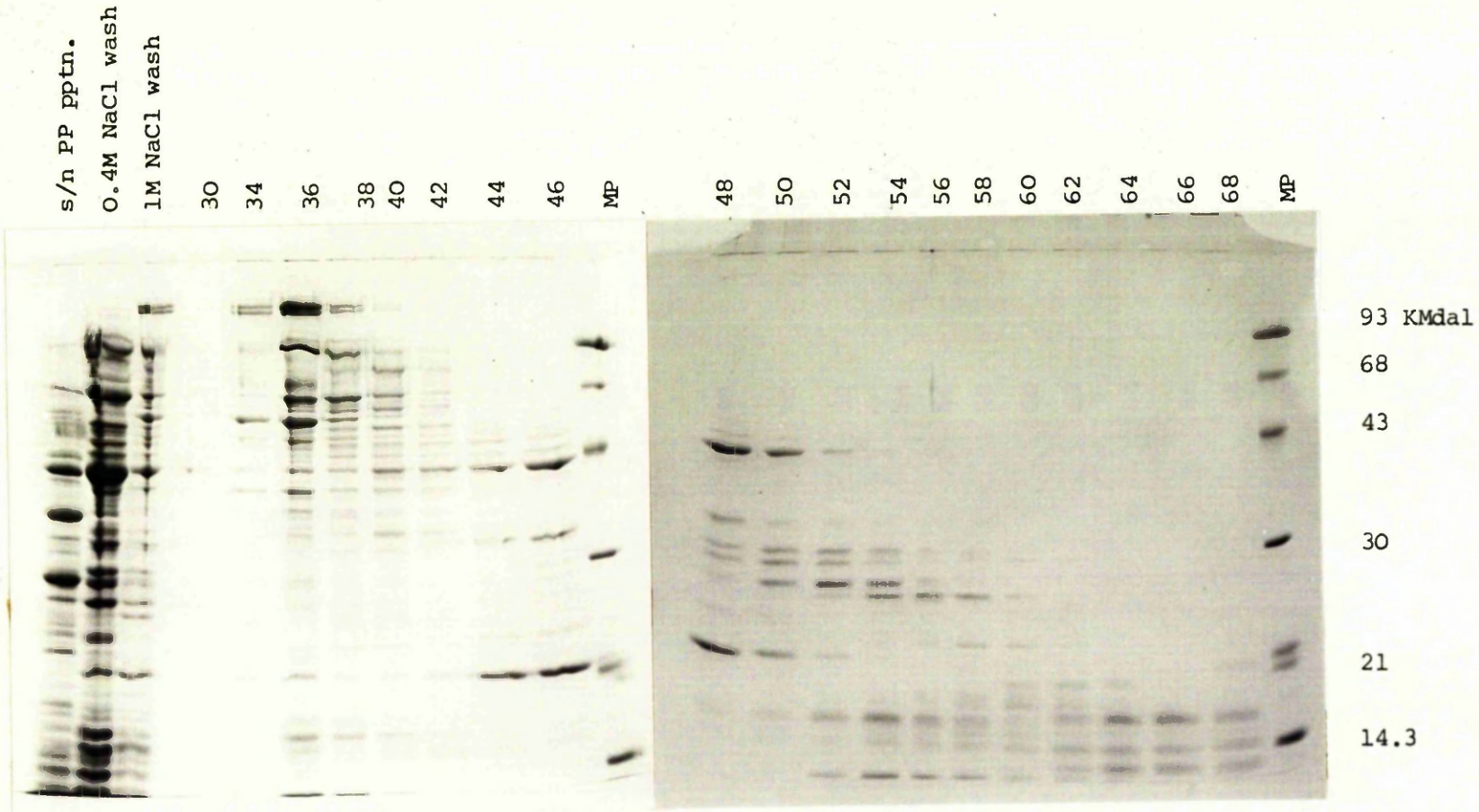


FIGURE 4.9 SDS-PAGE assay of fractions from G-75 gel filtration column

Fractions 43-49 were pooled and dialysed as described in text.

12.5%,4.5% stacking gel run at 40mA.

(greater than 80,000 dal. for Sephadex G-75), and fractions containing major protein peaks. Fractions were analyzed further by SDS-PAGE to determine those containing resolvase. The gels indicate resolvase-containing fractions and it's elution point in relation to other proteins (figure 4.9). Resolvase appears to be eluted from the column at several different positions suggesting that it is present in multimeric forms under the conditions used, i.e. IM NaCl/TEM. Some resolvase is eluted during the major large protein peak, representative of proteins of about 80,000 dal.; it appears again in fractions, which elute at a similar size to proteins of 40,000 dal. Finally, a small amount is eluted at the end of the protein elution profile with other small proteins of about 20,000 dal. This data suggests that under the conditions used for the column, the protein exists in equilibrium between monomers, dimers and higher forms, with dimers predominating.

monomers $\rightleftharpoons$ DIMERS $\rightleftharpoons$ tetramers?

It would be interesting to know which way the equilibrium between forms may move, if at all, under physiological conditions, estimated at 150-250mM sodium equivalents (Pollock and Abremski, 1979).

Fractions containing resolvase were pooled and dialyzed three times against 7M urea/TEM. The pooled, dialyzed protein was then applied to a 10ml CM-Sepharose cation exchange column, equilibrated with 7M urea/TEM. After loading, one bed volume of running buffer (7M urea/TEM) was passed through the column. A salt gradient of 0→200mM NaCl was applied to the column and collected as above. Fractions were assayed by optical density and gel electrophoresis. Resolvase gives very low O.D. readings at 280nm due to the lack of tryptophan residues in the protein, but this assay does give some indication of which fractions to analyze by

SDS-PAGE. All fractions from the column were measured for conductivity; by comparison with standards this gives the molarity of each fraction, which can then be represented graphically along with O.D. measurements (figure 4.IO). Two main peaks were indicated by O.D. readings; these fractions were assayed by SDS-PAGE to determine which contained resolvase (figure 4.II).

Resolvase-containing fractions were pooled and dialyzed against 7M urea/IM NaCl/TEM, then twice against IM NaCl/TEM. At this stage half of the sample was removed from the dialysis sac and stored at $4^{o}C$; the remainder was dialyzed further against IM NaCl/TEM/50% glycerol(v/v) and stored at $-20^{o}C$. After several months the samples were compared by SDS-PAGE; the sample stored at $4^{o}C$ contained far less resolvase than the $-20^{o}C$ sample (even after taking into account the glycerol concentration effect), and had lost all activity *in vitro*. Henceforth all resolvase samples were stored in 50% glycerol at $-20^{o}C$; samples stored in this way retained activity 8-9 months after preparation.

To determine the predominant multimeric forms of resolvase, samples of purified protein were applied to a I50ml Sephadex G-I50 gel filtration column calibrated with markers of known molecular weight. By comparing the resolvase elution point with that of transferrin(86,000),BSA(68,000), ovalbumin(43,000), and lysozyme(I4,300), one should be able to determine it's molecular weight in IM NaCl (used for storage and column running buffer). The results from these experiments indicated that resolvase is eluted over the same fraction range as ovalbumin, and is therefore present in dimeric form (43,600 dal.).

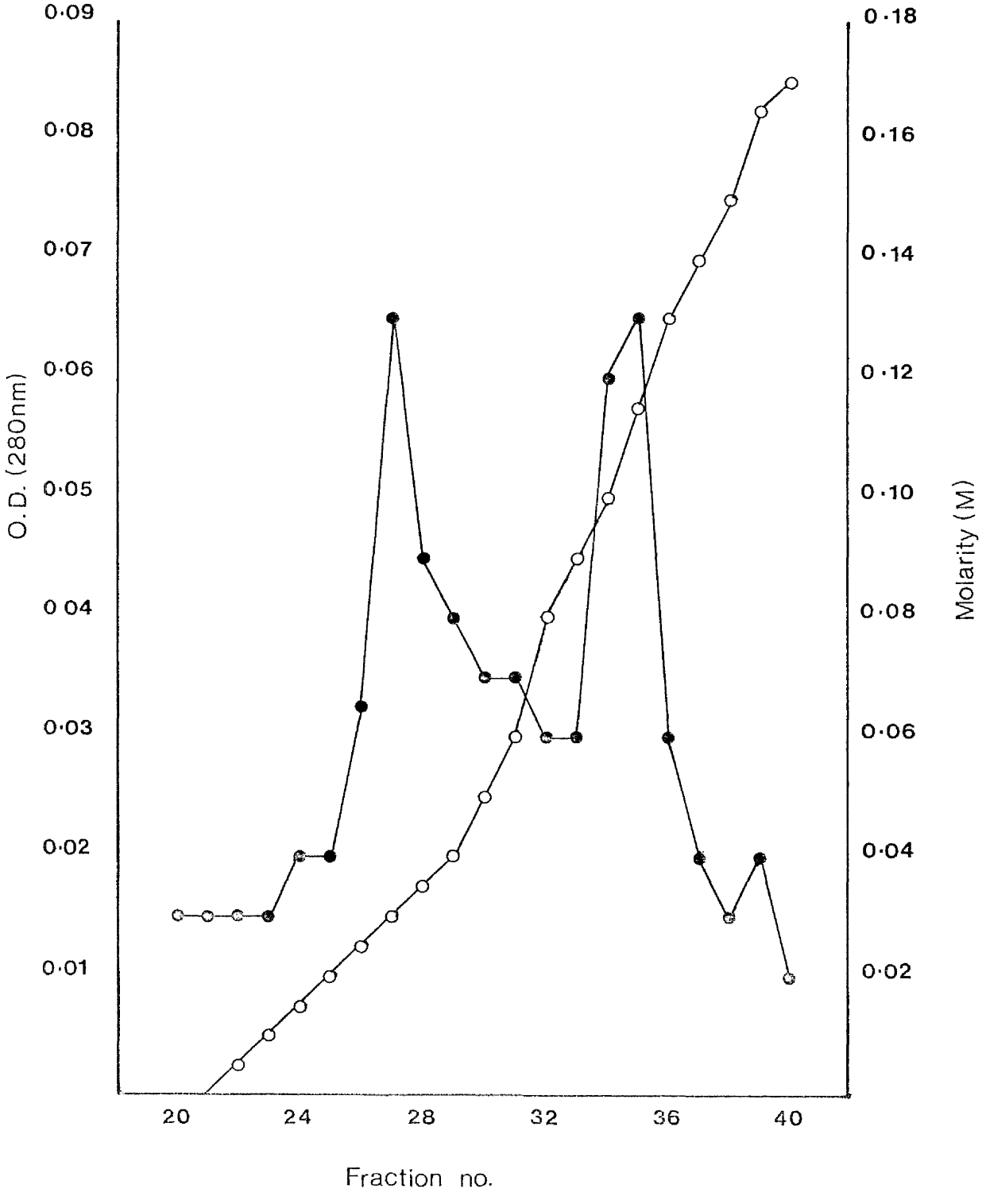
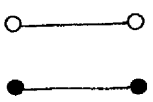

0·09
0·08
0·07
0·06
0·05
0 04
0·03
0·02
0·01
O.D. (280nm)
0·18
0·16
0·14
0·12
0·10
0·08
0·06
0·04
0·02
Molarity (M)
20
24
28
32
36
40
Fraction no.

FIGURE 4.10 Graph showing the optical density and molarity of fractions from the CM-Sepharose ion-exchange column



represents molarity

represents optical density

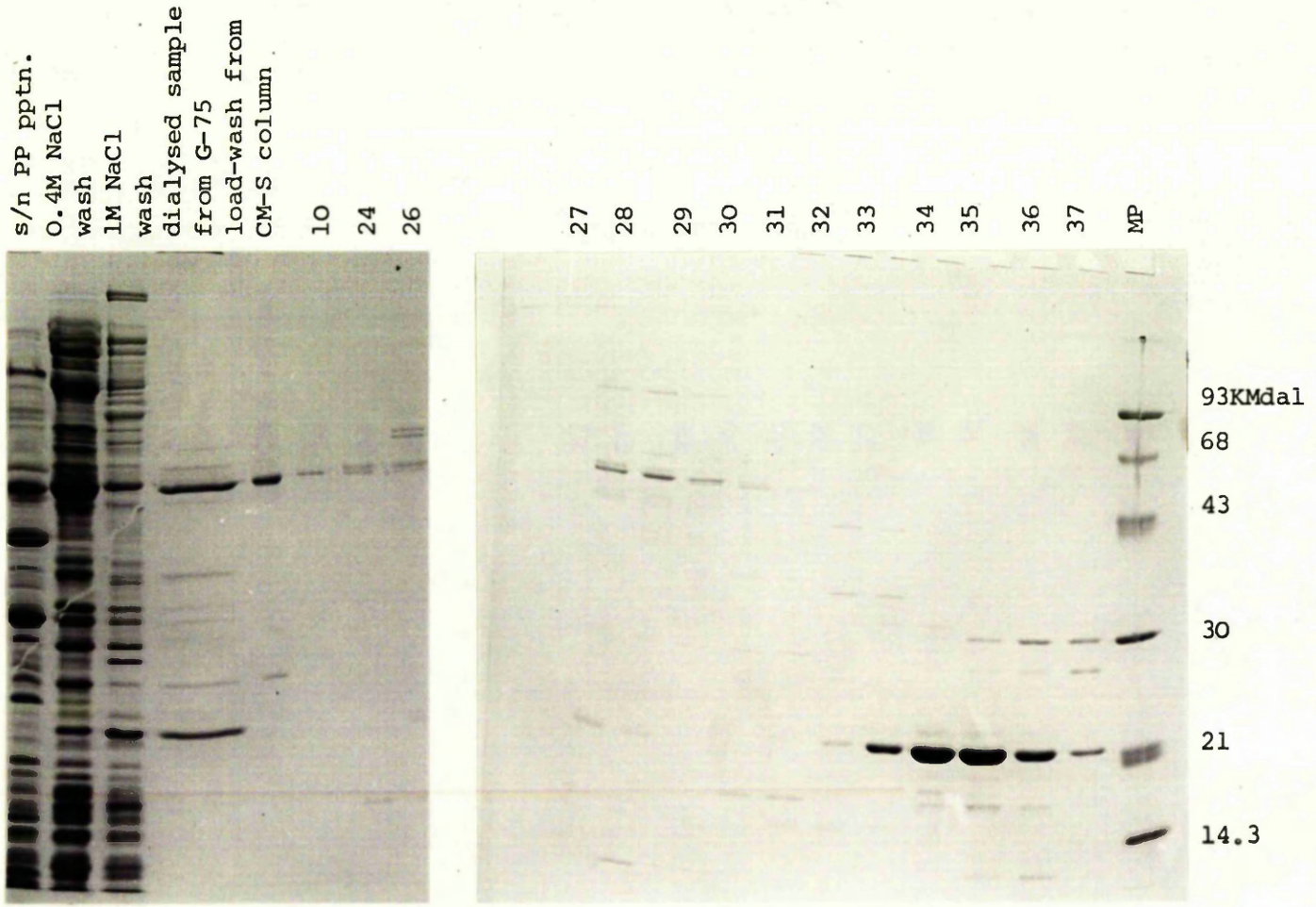


FIGURE 4.11 SDS-PAGE assay of fractions from CM-Sepharose ion-exchange column

Fractions 33-37 were pooled and dialysed (details in text).

12.5%,4.5% stacking gel run at 40mA

Discussion

The aim of the experiments in this chapter was to clone the Tn*3* *tnp*R gene into a high expression vector, and subsequently purify it s product, resolvase. The basic strategy used was similar to that employed by R.Reed for isolation of the γδ resolvase protein; the vector, pλ8, and strain containing CI^{857}, N4830, used for cloning were kindly provided by R.Reed (Reed, 1981b).

Although previous data had shown that there are many similarities between the two resolvase proteins from Tn*3* and γδ, it was not known how their properties would compare during purification and *in vitro*. I decided to study the Tn*3* resolvase protein in order to compare it s activity *in vitro* with other site-specific recombination proteins and, more importantly, to attempt to deduce it s recombination mechanism.

One of the interesting features uncovered during purification of resolvase is the apparent molecular weight of the protein when eluted from gel filtration columns. Under the conditions used, i.e. 1M NaCl/ TEM, the protein appears to be eluted predominantly as dimers, some in higher forms and very little as monomers. It has not been possible to deduce the form of the protein under physiological conditions due to insolubility problems in low salt solutions (at 100mM NaCl the purified protein precipitates from solution). In cells containing wild-type Tn*3* elements the levels of resolvase are very low, barely distinguishable even in minicells when analyzed by SDS-PAGE (Dougan *et al.*,1979). At this concentration the protein appears to be soluble and is presumably bound to DNA most of the time. Consequently there would be very little free resolvase present in the cell. It would be interesting to know

if resolvase binds to DNA as a monomer, dimer or in higher forms.

It is assumed that during purification of resolvase the protein is not irreversibly affected by the high urea concentration used to solubilize the protein for running on the ion exchange column. During this stage of the purification the dialysis steps, column running, and subsequent gel assays were performed as quickly as possible to ensure that the protein was not exposed to urea for any longer than necessary. Analar grade urea was always used and solutions freshly made at each stage of the purification; the conductivity of urea solutions was routinely checked. It may be of significance that the final resolvase preparation was kept in urea solutions for a much shorter time than had other preparations, and appears to exert higher activity in *in vitro* resolution experiments.

The results presented in the following chapters suggest that resolvase, prepared by these methods, has retained specific binding and recombination activities.

CHAPTER 5

CONSTRUCTION OF THE *IN VITRO* RESOLUTION SYSTEM

Introduction

Genetic analysis has indicated that for resolvase mediated site-specific recombination to occur, resolvase and a replicon containing two directly repeated copies of *res* are required.(Kitts *et al.*,1982a). As yet, no host factors have been directly implicated in this event. Resolvase and a suitable substrate are essential components for construction of an *in vitro* system to study resolution in detail. Using this basis it should be possible to determine any other factors, which may be required. Purification of resolvase has been described; the second requirement is a suitable substrate. The substrate is required to have two *res* sites in direct orientation and on resolution must be easily assayed to distinguish between substrate and product molecules. The most commonly used and convenient assay for recombination is restriction cleavage of reacted DNA. This assay relies on the substrate and product molecules producing fragments of different sizes, thus enabling clear distinction between substrate and products on gel electrophoresis (Mizuuchi and Nash,1976; Reed,1981b).

A series of small plasmid substrates has been constructed, using *in vitro* techniques, which may be assayed for resolution by gel electrophoresis, transformation or electron microscopy. The plasmid substrates thus constructed have been used with purified resolvase to answer some of the following questions relating to the mechanism of the reaction:

(a) Does the system require any host-encoded products?

(b) What are the basic ionic conditions for optimal reaction?

(c) Is an external energy source required?

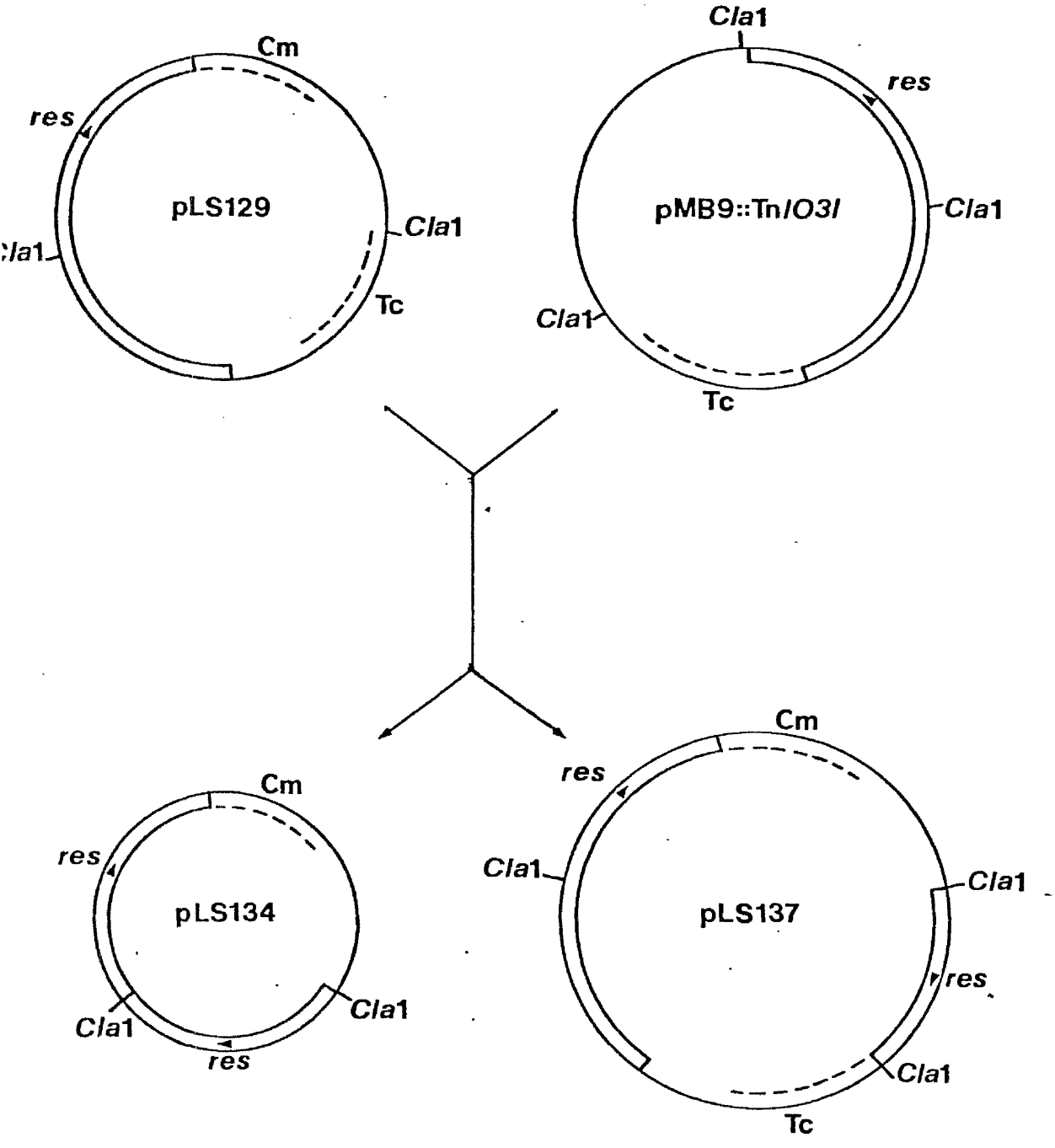


FIGURE 5.1 Strategy for construction of pLS134 and pLS137

(for details of manipulations see text)

(d) Is the reaction reciprocal?

(e) Is the substrate required to be supercoiled?

(f) Can resolvase join molecules by intermolecular recombination at *res* sites on two separate molecules?

(g) Does resolvase have any activity on inverted *res* sites?

(h) Does resolvase conserve superhelicity during the reaction?

(i) Does resolvase exhibit any topoisomerase activity?

(j) Does resolvase catenate products - if so, is this an intrinsic part of the reaction?

(k) Is resolvase catalytic?

(l) Does resolvase induce double-stranded or single-stranded cleavage at *res* sites?

Results

1. Construction of pLS134 and pLS137

Substrates were constructed using the techniques of restriction, ligation, transformation and SCCL as described in Materials and Methods. These substrates were both derived from pLS129, an insertion of Tn*103* into pACYC184. On *Cla*I restriction two 4Kb fragments are produced; one carrying the Cm^r gene, *oriV* and *res*, the other, the Tc^r gene and most of the *tnp*A gene. The source of the second *res* site was pMB9::Tn*1031* (Kitts *et al.*,1982a), in which *res* is carried on a 2Kb *Cla*I fragment (see figure 5.1). pLS129 was partially digested with *Cla*I to produce a mixture of unit length and 4Kb fragments. This digest was treated with bacterial alkaline phosphatase (to help prevent vector recircularization) then electrophoresed through a 1% low melting point agarose gel to separate fragments. Gel bands corresponding to 8Kb and 4Kb fragments were excised and used in ligation with *Cla*I digested pMB9::Tn*1031* DNA. The ligation

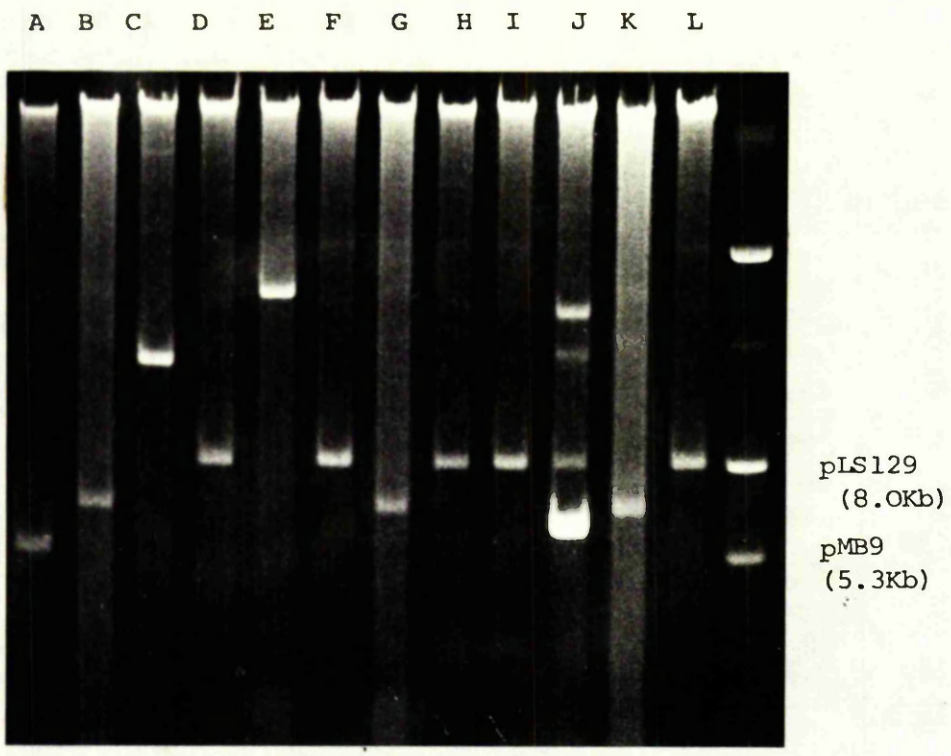


FIGURE 5.2 Analysis of clones by SCCL

Clones A and C were of the size expected for pLS134 and pLS137; these were chosen for further analysis.

0.8% agarose gel run at 100V.

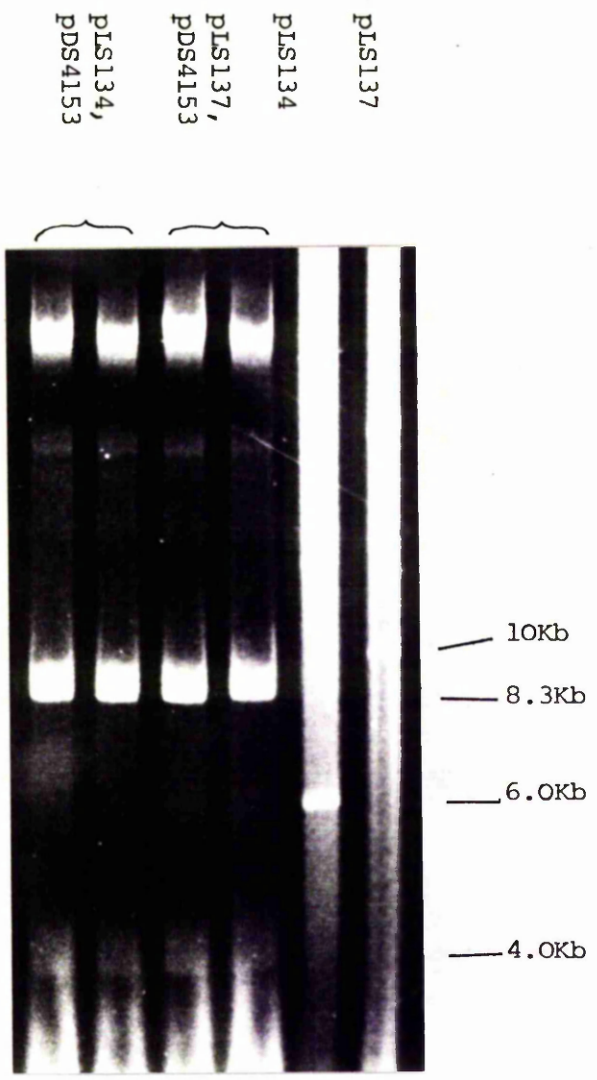


FIGURE 5.3 Resolution of pLS134 and pLS137 *in vivo*

Clones of the size expected for substrates (A and C from previous gel) were transformed into a strain containing pDS4153 (ColK:: Tn*1* *tnp*A$^-$, *tnp*R$^+$ derivative). Analysis of Cm^r,Ap^r transformants by SCCL revealed the presence of the expected resolution product, pLS135(4Kb). The reciprocal resolution product lacks *oriV* and is therefore not seen *in vivo*.

0.8% agarose gel run at 100V.

reaction mixture was diluted and transformed into AB2463. Transformants resistant to Cm or Cm,Tc were analyzed by SCCL to detect insertions by size (figure 5.2). Plasmids of predicted sizes were purified by the STET method and checked for resolution by transforming into a *tnp*R$^+$ plasmid strain (figure 5.3). The plasmids designated pLS134 and pLS137 were shown to resolve *in vivo* to produce a smaller plasmid of 4Kb, which carries *oriV* and Cm^r. The reciprocal product of resolution for both substrates carries no *oriV* and is therefore lost *in vivo*. The product of *in vivo* resolution was purified for use as a size marker for *in vitro* experiments.

pLS137 has been used mainly for transformation experiments; as the Tc^r gene is flanked by *res* sites, resolution yields a Cm^r,Tc^s plasmid. The resolution product carrying the Tc^r gene has no *oriV* and is thus not detected *in vivo*. On transformation of *in vitro* reaction mixtures the proportion of Cm^r transformants, which retain resistance to Tc provides a measure of resolution.The plasmid pLS134 has been used for most of the experiments to characterize the *in vitro* system, using electrophoresis as the assay.

2. Construction of pLS138

Analysis of the DNA sequence surrounding the *res* region of Tn*3* reveals the presence of two *Eco*RI* cleavage sites, one within the beginning of the *tnp*A structural gene and the other in the *tnp*R gene (see figure 4.1). The *Eco*RI* sites are located 282bp apart. This fragment has been cloned into the *Eco*RI site of pACYC184, rendering the plasmid Cm^s (Kitts,1982). Both junction sequences of this plasmid are now *Eco*RI cleavage sites enabling the 282bp *res*-containing region to be lifted from the plasmid, named pPAK317, on a single *Eco*RI fragment. This fragment was subsequently purified

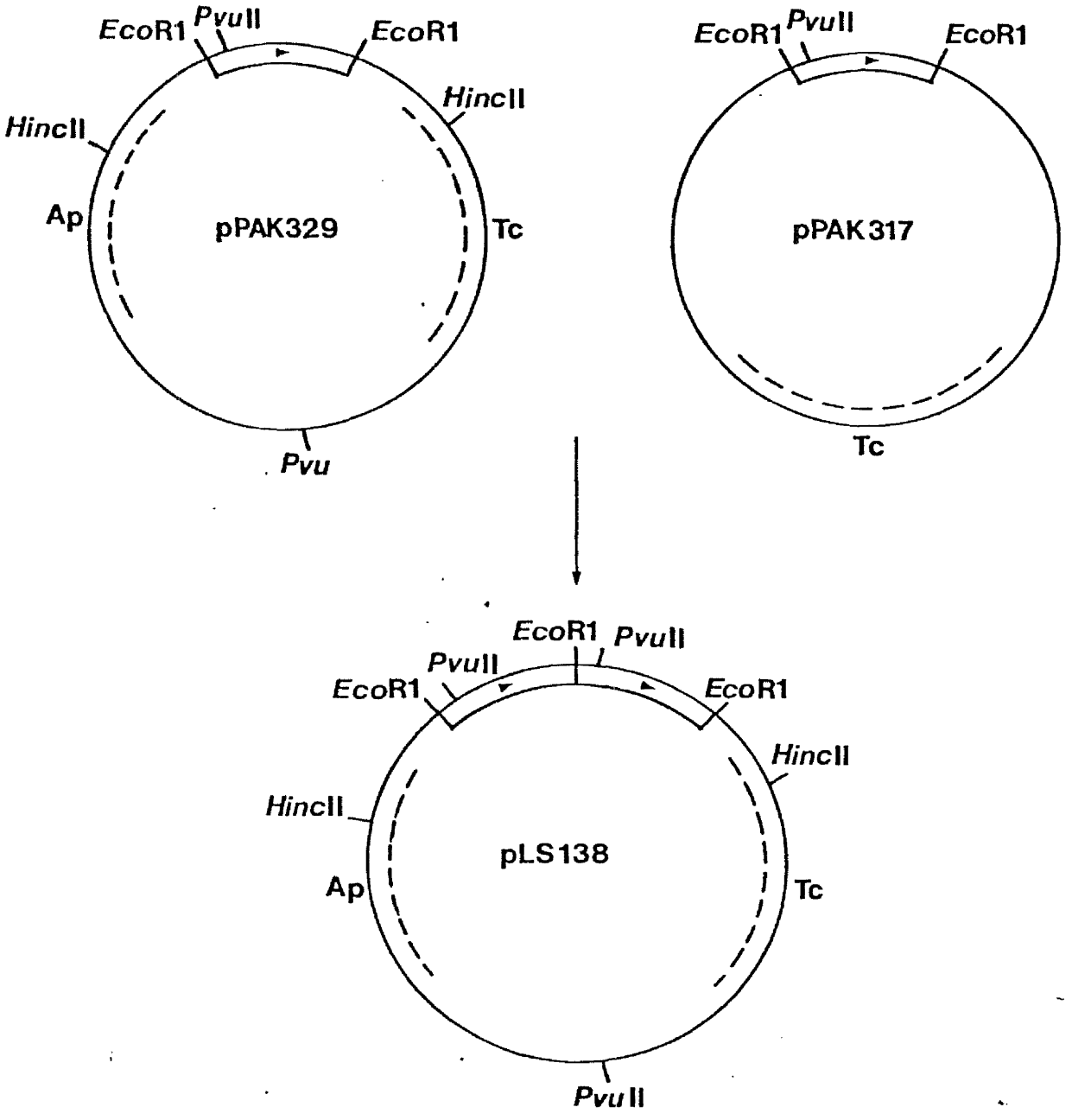


FIGURE 5.4 Strategy for construction of pLS138

(see text for details)

and cloned into the single *Eco*RI site of pBR322 to produce pPAK329 (Kitts,1982).

The plasmid pLS138, was constructed by cloning a second copy of the 282bp *res*-containing fragment into pPAK329 (figure 5.4). The 282bp fragment was purified by cleaving pPAK317 with *Eco*RI followed by electrophoresis through a 5% polyacrylamide gel; the band corresponding in size to 282bp was excised and the DNA eluted by the electrophoresis method. pPAK329 was partially cleaved with *Eco*RI to produce linear fragments; the digest was treated with bacterial alkaline phophatase, then annealed to the purified 282bp fragment. Ligated DNA was transformed into AB2463; transformants resistant to Ap,Tc were selected and analyzed by SCCL (figure 5.5). Clones E-H represent plasmids slightly larger than pPAK329; clone B is of the size expected for dimers of pPAK329. It was assumed that all clones containing inserts would have directly repeated copies of *res* due to the observed instability of palindromic DNA (Collins, 1981; Warren,pers.comm.). The orientation of *res* sites was confirmed by restriction analysis with *Pvu*II, which cuts asymmetrically within the 282bp fragment (figure 5.6). The structures of pBR322, pLS138 and pPAK329 have been compared by restriction with *Hinc*II and *Pvu*II; the increase in size of the smallest *Hinc*II fragment of pBR322 indicates insertion of 564bp or 282bp, respectively. Double digests with *Hinc*II and *Pvu*II orientate the 282bp fragment(s) in pPAK329 and pLS138 as indicated (figure 5.4).

Of major importance was the ability of pLS138 to resolve; this was initially tested by transforming pACYC184::Tn*3* into the pLS138-containing strain. Transformant colonies resistant to Cm were analyzed by SCCL (data not shown). This indicated the presence of

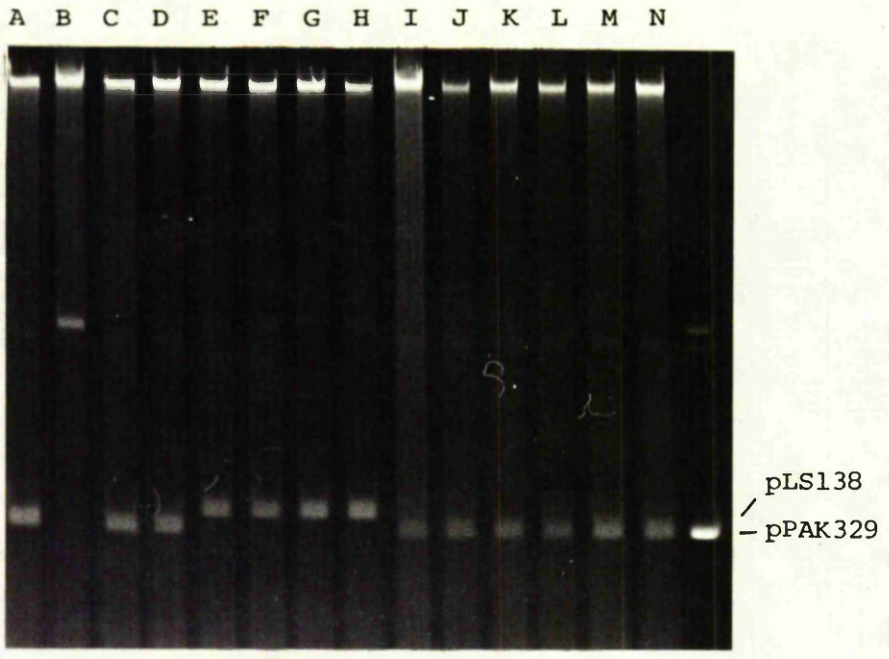


FIGURE 5.5 Analysis of clones by SCCL

The plasmids represented by tracks E-H are slightly larger than pPAK329, presumably due to insertion of the 282bp *res*-containing fragment. One of these was chosen for further analysis.

0.8% agarose gel run at 25V.

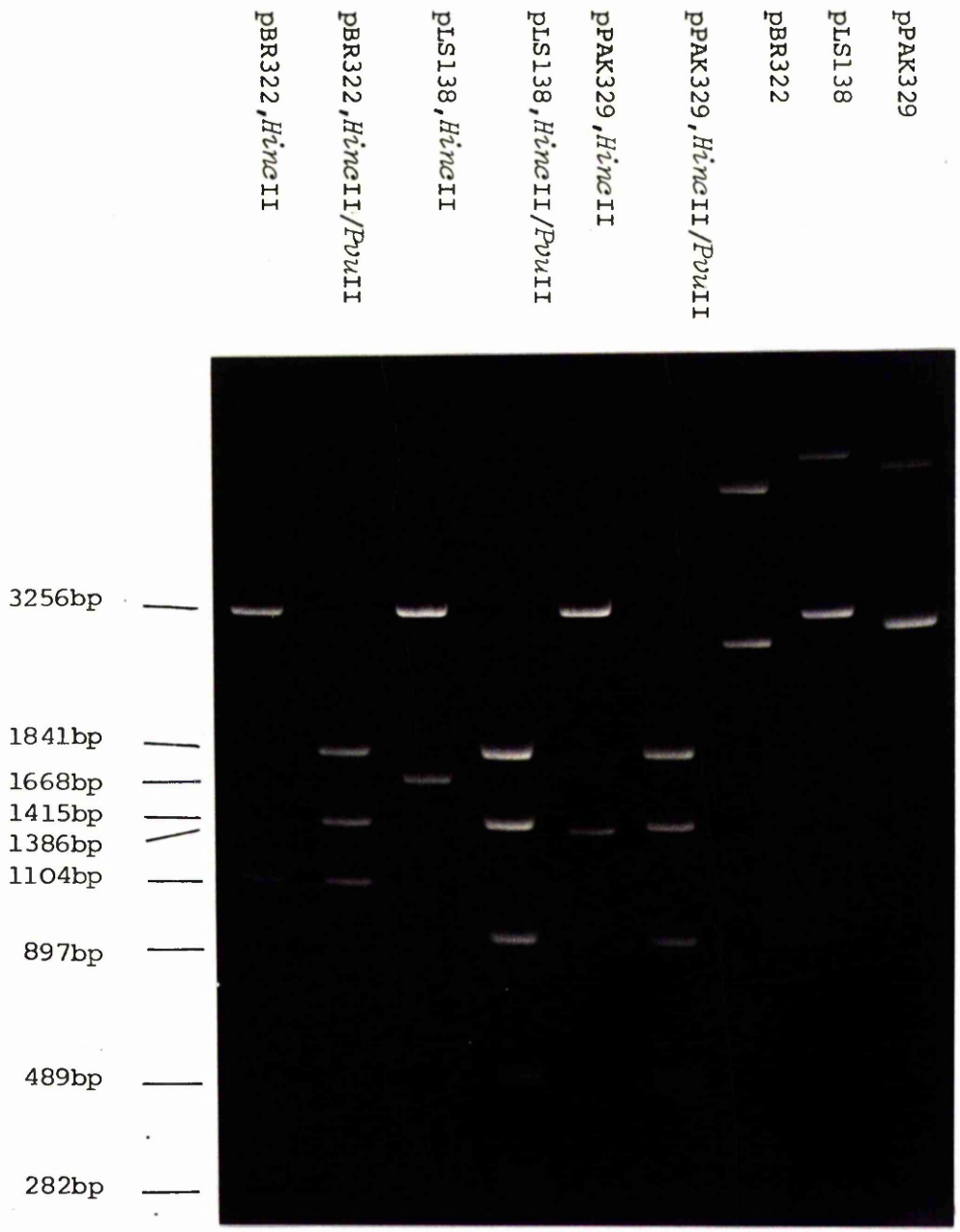


FIGURE 5.6 Comparative restriction digests of pBR322, pLS138 and pPAK329

The smaller *Hinc*II fragment of pBR322 contains the *Eco*RI site into which the 282bp *res*-containing fragment(s) have been cloned; this is indicated by the increase in size of this fragment in pLS138 and pPAK329. The *Pvu*II digests confirm the orientation of the inserts in pLS138 and pPAK329.

1% agarose gel run at 100V

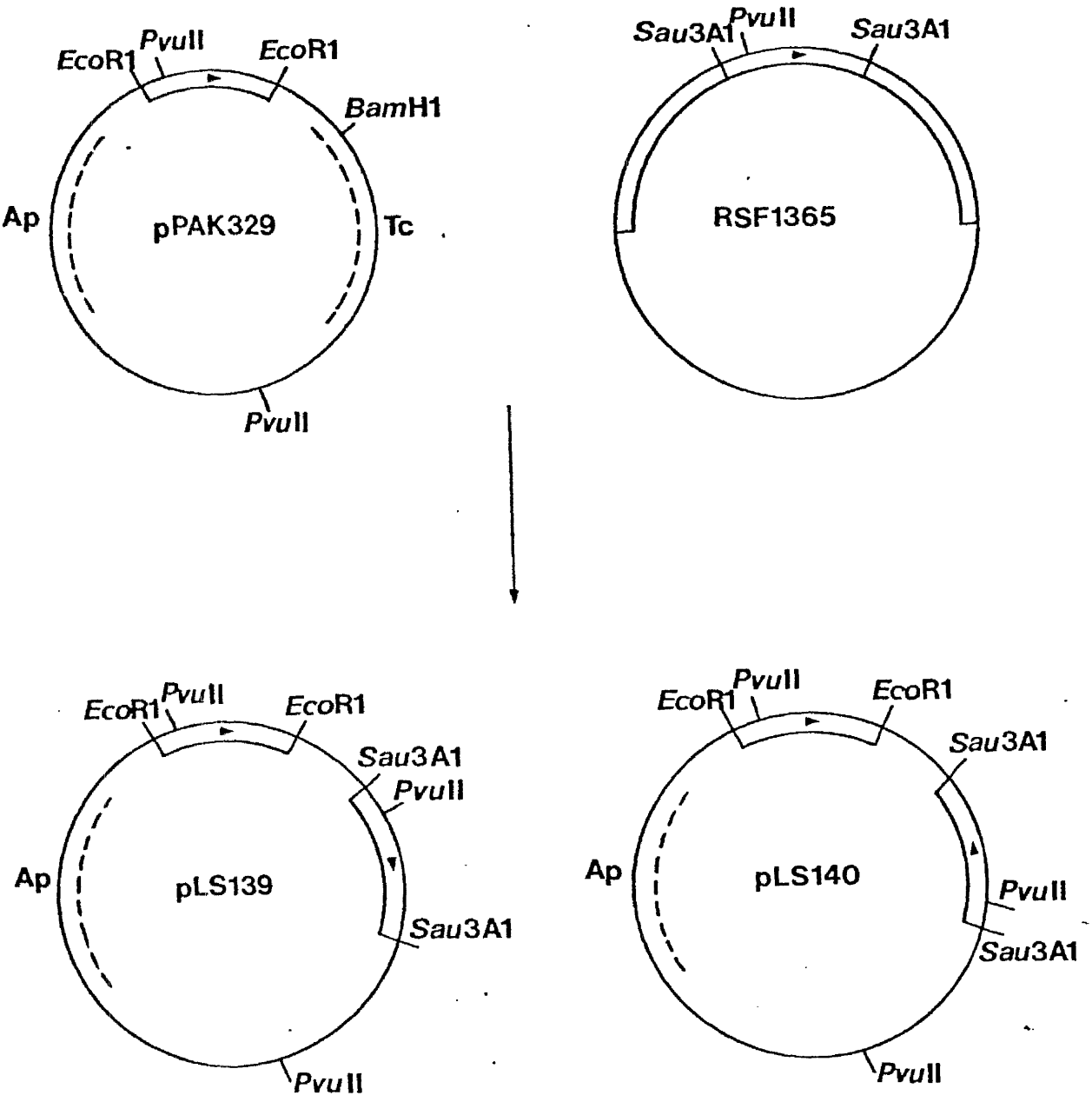


FIGURE 5.7 Strategy for construction of pLS139 and pLS140

(for details refer to text)

only two plasmids, corresponding in size to pPAK329 and pACYC184::Tn*3*, suggesting that the 282bp fragment contains sufficient sequence for resolution and that the short distance between sites does not interfere with the reaction.

3. Construction of pLS139 and pLS140

These plasmids were also constructed using pPAK329 as a vector. The *res* site of Tn*3* is contained within a 357bp *Sau* 3AI fragment (figure 4.1); this fragment contains sufficient sequence for resolution *in vivo* (Reed,1981a), and also contains the regulatory region for transposition functions (Chou *et al.*,1979b). This fragment was purified from RSF1365, a deletion derivative of Tn *3* (a map of Tn*365* is shown in figure 3.3). *Sau*3AI digested RSF1365 DNA was electrophoresed through a 5% polyacrylamide gel to separate fragments. Fragments were identified by comparison with pBR322 *Sau*3AI digested DNA, for which sizes are known (Sutcliffe,1978). The band corresponding to the 357bp *res*-containing fragment was excised from the gel and DNA eluted by the electrophoresis method. There is a single *Bam*HI site located within the Tc^r structural gene of pPAK329; inserts into this site render the plasmid Tc^s. The purified 357bp fragment was annealed to *Bam*HI digested pPAK329 DNA (figure 5.7). Analysis of the sequences at either end of the fragment and at the *Bam*HI site indicates that the junction sequences will not be recognized by *Bam*HI, therefore all inserts into this site will be resistant to *Bam*HI cleavage. Recircularized vectors will be sensitive to *Bam*HI cleavage. As the transformation frequency of linear DNA is several fold lower than circular DNA, digestion with *Bam*HI prior to transforming should enrich for clones containing inserts. Ligated DNA was precipitated with 66% ethanol/0.3M NaAC, resuspended in MSRB and restricted with *Bam*HI before transformation into AB2463. Transformants resistant

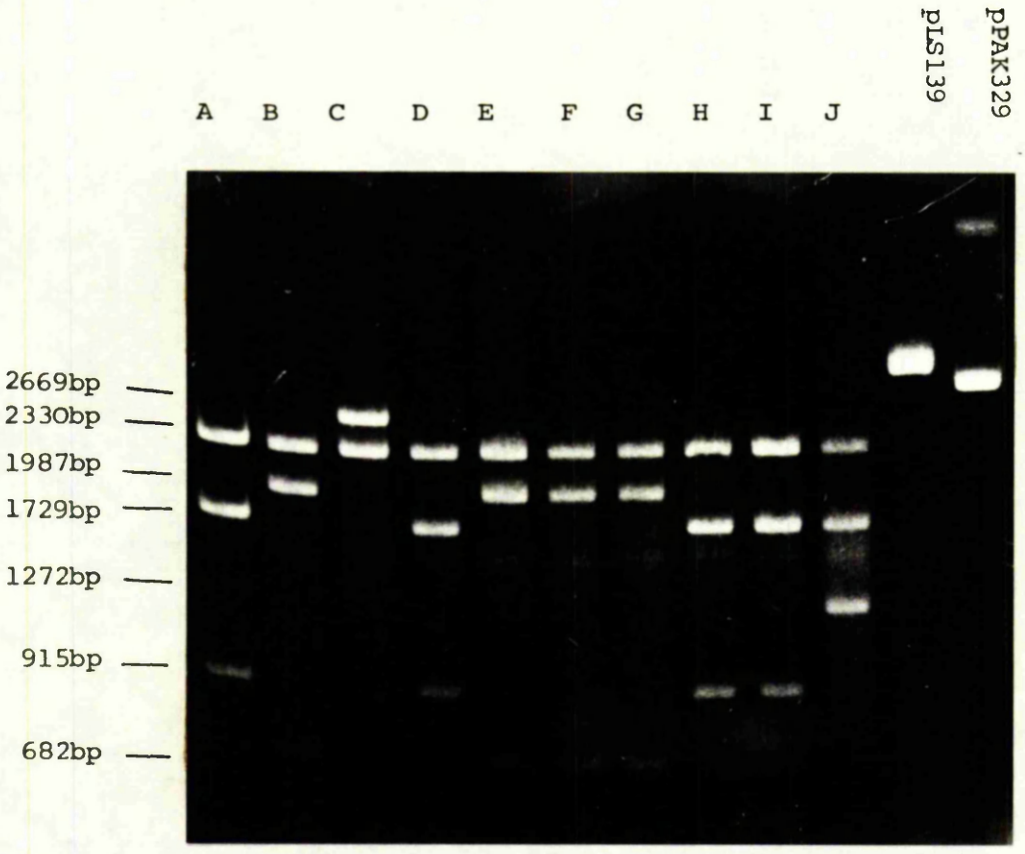


FIGURE 5.8 Restriction of Ap^r, Tc^s clones with *Pvu*II

The size of the two smallest fragments produced by *Pvu*II digestion indicates the orientation of the inserted fragment. Clones A and B are representative of insertions in both orientations, inverted and directly repeated *res* sites. Clones C and J have lost the *Pvu*II site from the inserted fragment. Clone J, as suspected, carries two copies of the 357bp fragment.

1% agarose gel run at 120V.

to Ap were checked for Tc^S; 5% of the Ap^r clones had retained resistance to Tc, presumably due to unrestricted vectors or transformants from linearized vectors. SCCL of Ap^r, Tc^r transformants indicated 18/20 plasmids larger than pPAK329; of these one was slightly larger than the rest, probably due to insertion of two copies of the 357bp fragment. The remaining two clones were smaller than pPAK329; these could have arisen by transformation of linear molecules, which were then subject to exonucleases *in vivo* to generate deletions of varying length.

Clones containing inserts should have three *Pvu*II sites, the distance between sites indicating the orientation of inserted fragments. DNA was purified from ten clones by the Birnboim/Doly method and restricted with *Pvu*II (figure 5.8). Clones A and B are representative of inserts in both orientations and were subsequently purified for *in vitro* experiments. Clone J, which was presumed to have two copies of the 357bp fragment, has lost one *Pvu*II site, but the fragment sizes are consistent with an additional 357bp. Clone C also appears to have lost a *Pvu*II site; for both of these clones the *Pvu*II site which has been lost is from inserted DNA, not from the vector. This may be due to the method used to purify the 357bp fragment which involves excising a band from an ethidium bromide stained gel on a long wavelength transilluminator.

4. Construction of the *in vitro* resolution system - basic requirements

The reaction conditions used for the initial experiments were based on those devised by Reed for the $\gamma\delta$ resolvase *in vitro* system (Reed,1981b):

20mM Tris-HCl, pH 7.5

10mM $MgCl_2$

50mM NaCl

1mM DTT

0.5µg DNA

0.2µg resolvase, in 20µl total volume

When pLS134 was used in a resolution reaction, under the conditions described above, it was observed that a new band was present in reacted samples, migrating slightly faster than s/c pLS134, on gel electrophoresis. In analogy with the γδ and Int systems, this could be due to the formation of s/c catenated product molecules (Nash *et al.*, 1977; Reed, 1981b). To test this hypothesis reacted DNA was restricted with *Eco*RI or *Hind*III; these two enzymes each have a single site on pLS134 (figure 5.9A). Unreacted material migrates as unit length molecules of 6Kb, whereas catenated molecules would be separated to produce one linear molecule of 4Kb or 2Kb and one circle of 2Kb or 4Kb, depending on which enzyme was used for cleavage. The results obtained by restricting resolution reactions indicated that the DNA species migrating slightly ahead of s/c substrate was due to product molecules. The catenated product band is not always clearly distinguishable on gels, especially those run quickly; for most experiments reaction mixtures were divided so that half of each sample could be assayed by restriction also.

To determine the optimal buffer pH for resolution, a range of Tris-HCl buffers were substituted into the standard reaction conditions. The optimal pH was found to be between 7.6 and 8.2; however very little difference was noticeable across this range (figure 5.10). The addition of KCl up to 50mM does not appear to affect the reaction, nor does 5mM spermidine, but spermidine

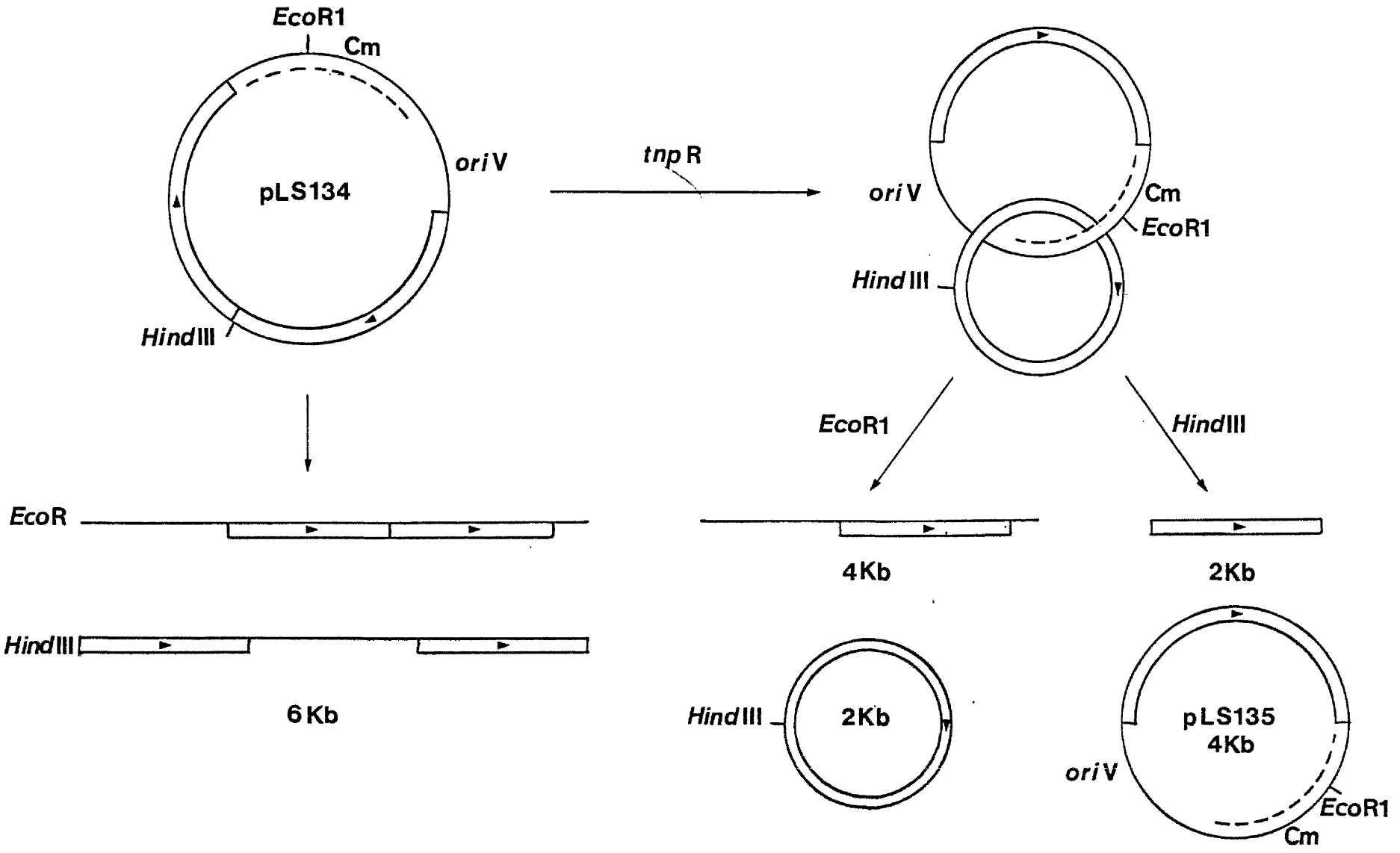

EcoR1
Cm
pLS134
oriV
HindIII
tnpR
oriV
Cm
EcoR1
HindIII
EcoR1
HindIII
EcoR
4Kb
2Kb
HindIII
6Kb
HindIII
2Kb
pLS135
4Kb
oriV
EcoR1
Cm

FIGURE 5.9A Restriction assay for pLS134 resolution

If the products of resolvase mediated recombination are catenates then it should be possible to unlink the circles by restriction enzyme cleavage. Using an enzyme which has a single site on the substrate, unreacted material will migrate as 6Kb unit length molecules, whereas catenates will be separated into 4Kb or 2Kb unit length molecules plus 2Kb or 4Kb circles, depending on the enzyme used. The product circle which remains intact retains the superhelicity that was present in the catenate, thus giving some indication of the conservative nature of the reaction.

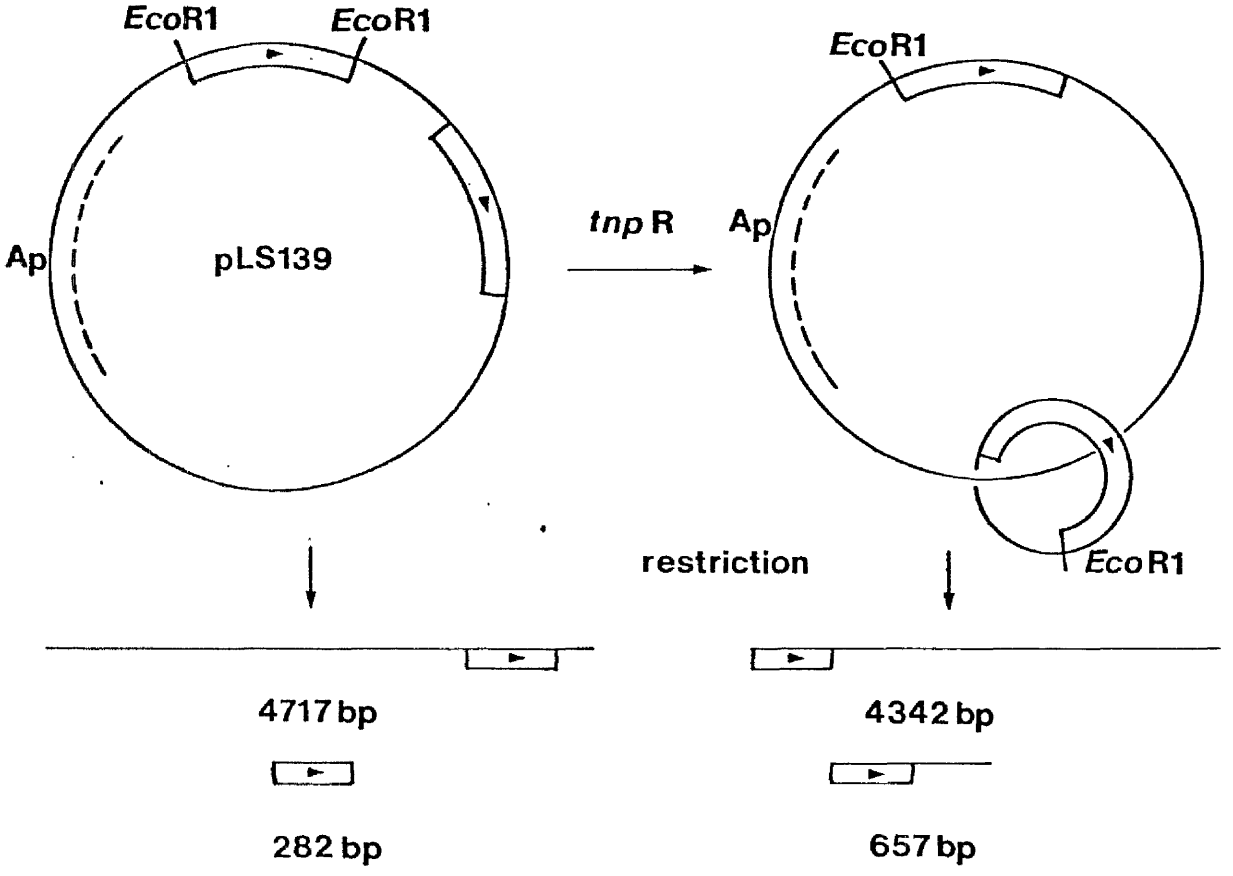


FIGURE 5.9B Restriction assay for pLS139 resolution

Cleavage of pLS139 resolution reactions with *Eco*RI produces fragments which may be clearly distinguished as unreacted substrate or product molecules on gel electrophoresis. In fact any enzyme which has a single site on pLS139 may be used for the restriction assay as it will separate catenated products.

In a similar manner inversion of DNA between *res* sites on pLS140 may be assayed by *Eco*RI digestion as the two orientations will produce fragments of the same size as pLS139 reacted and unreacted molecules.

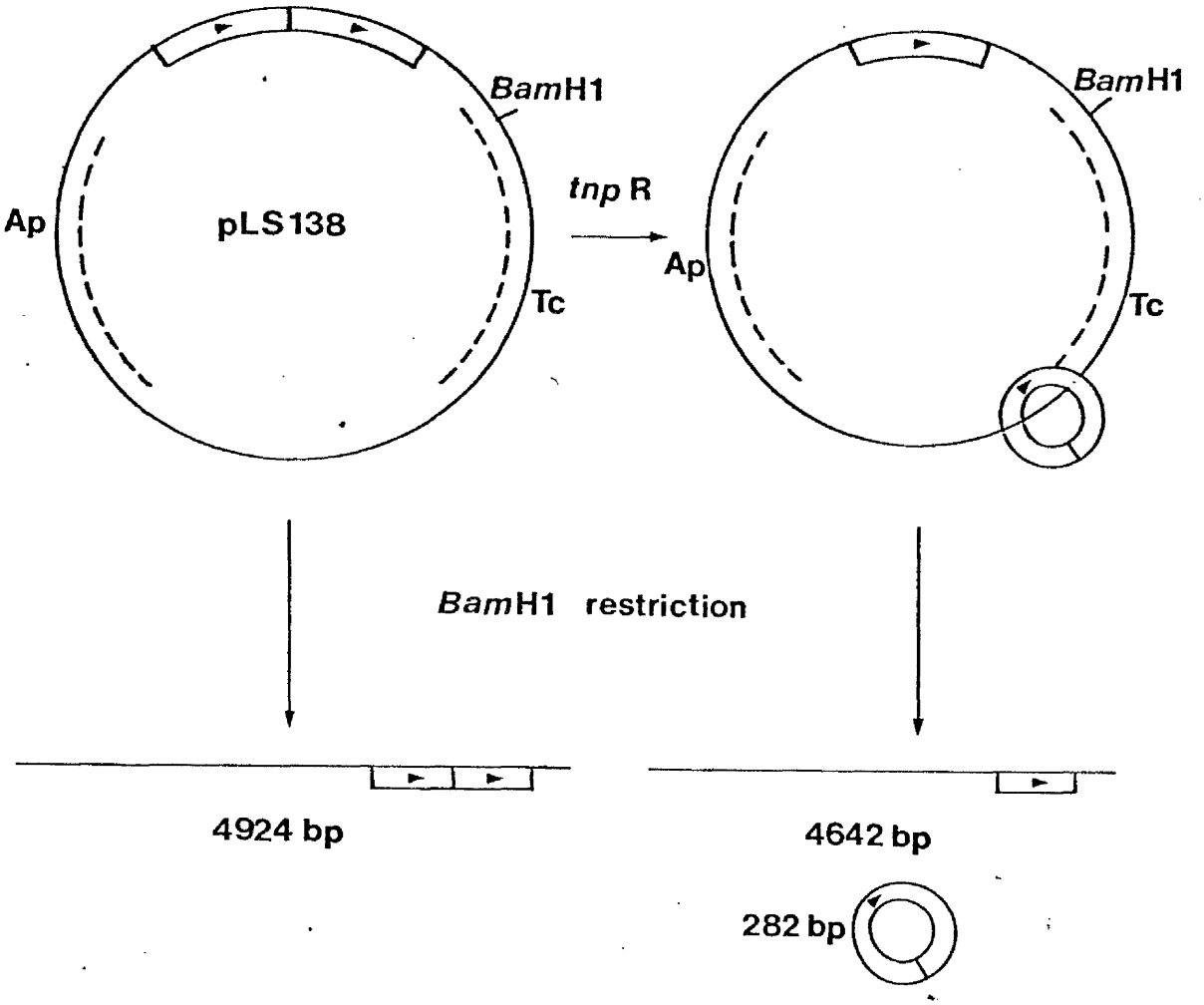


FIGURE 5.9C Restriction assay for pLS138 resolution

Any restriction enzyme which has a single site located on pLS138 will separate catenates thus clearly distinguishing between substrate and product molecules on gel electrophoresis.

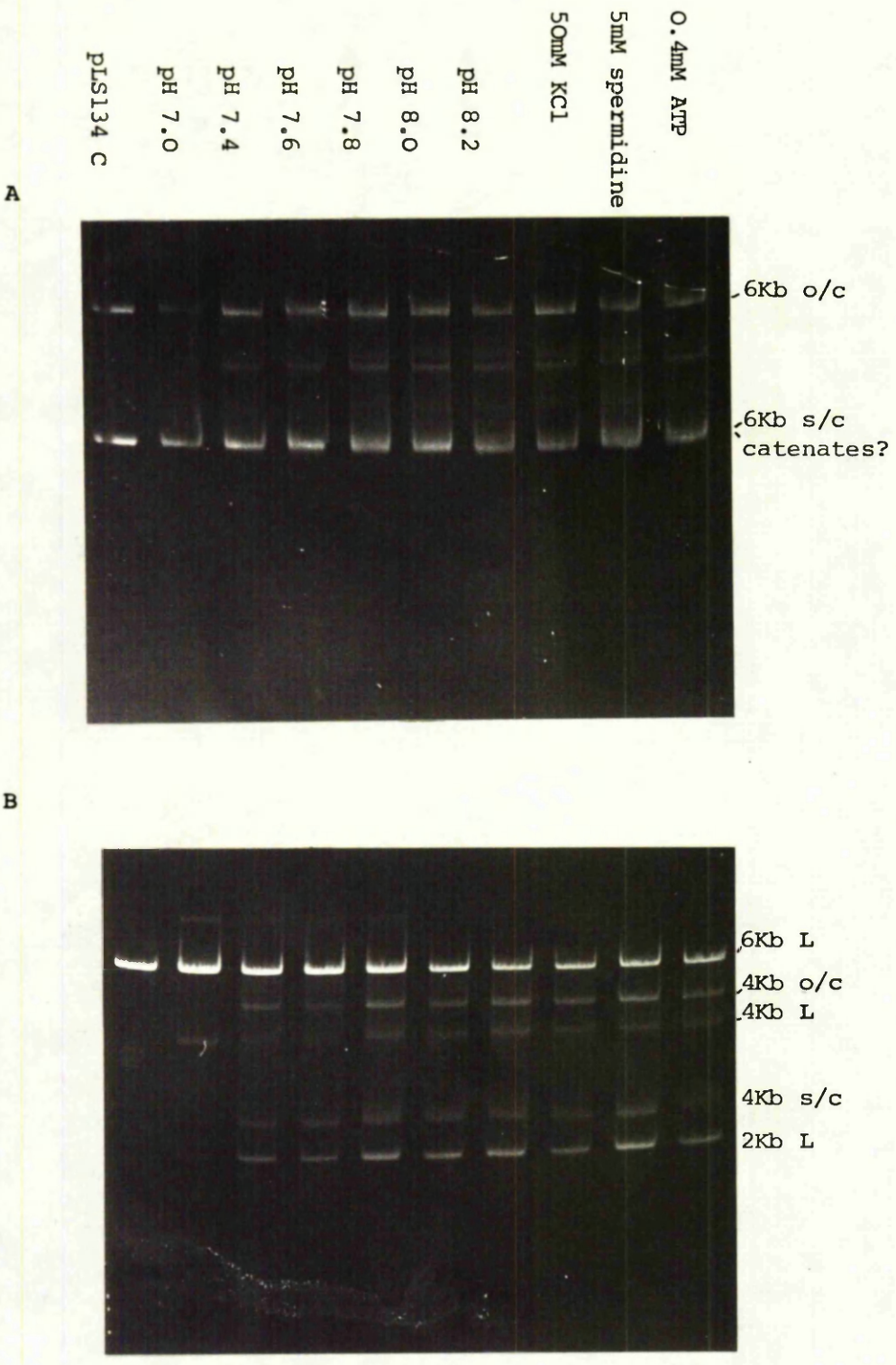


FIGURE 5.10 Determination of optimal conditions for resolution

A. Unrestricted samples, incubated for 1hr with 0.5 g of resolvase prep.1. S/c catenated products are not distinguishable from s/c substrate, but several other new bands are visible migrating between s/c and o/c.

B. Samples as above, restricted with *Hind*III.

0.8% agarose run at 25V

concentrations in excess of 5mM inhibit the reaction. Addition of 0.4mM ATP has no apparent affect on resolution activity. Increasing the NaCl concentration up to 250mM does not appear to diminish the reaction significantly, and it has not been possible to reduce the concentration of NaCl to less than 30mM due to the high salt storage buffer used for resolvase. Three independent samples of resolvase have been purified, each with varying degrees of resolution activity. Preparation 3, purified during the latter part of experimental work, appears to be the most concentrated and has by far the greatest activity *in vitro*. In the following experiments, for each gel shown, the resolvase preparation used is indicated in the figure legend. Tris-HCl buffers of pH8.0 have subsequently been used for *in vitro* resolution reactions; the other conditions in the standard reaction mix have been unchanged.

5. Determination of the minimal amount of resolvase required for resolution *in vitro*

The use of predetermined concentrations of substrate and protein allows calculation of the ratio of these two components in the cell-free system. This ratio has been determined for Int-mediated recombination and for γδ resolvase; as the minimal amount of protein required is high it has been suggested that these proteins may not be catalytic (Nash and Robertson,1981; Reed,1981b). The ratio of Int monomers per recombinant is 20-40 and for γδ resolvase about 50. However, it has not been possible to determine how much of the protein used in these reactions is in fact in an active state.

To determine the minimum amount of protein required in the Tn*3* resolution system, standard reaction conditions were used with the addition of increasing amounts of resolvase. Resolvase

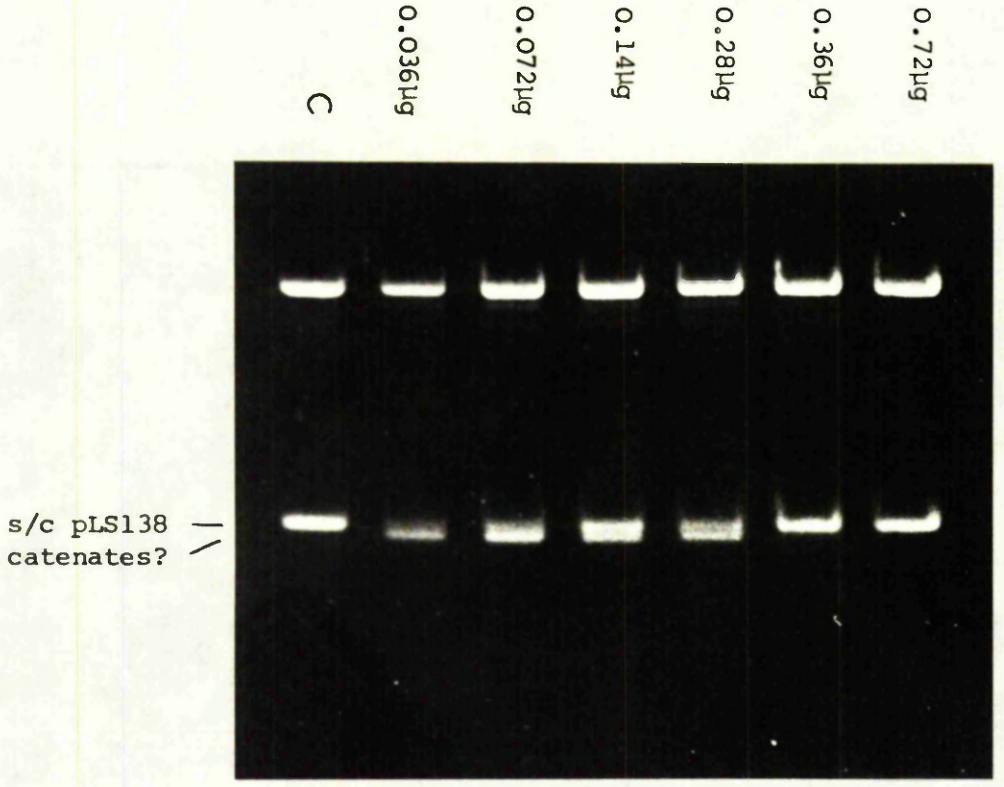


FIGURE 5.11 Determination of optimal resolvase concentration for resolution

Optimal resolution, judged by formation of s/c catenated product band, is observed with 0.036-0.072μg of resolvase (prep. 3) after 30min incubation.

0.8% agarose gel run at 25V.

concentrations were estimated by comparing known volumes of the protein preparation with purchased molecular weight markers by SDS-PAGE. The stained gel was scanned with a microdensitometer, the resulting protein peaks were cut out from the scan trace, weighed and compared. It should be stressed that values obtained provide a very rough estimate due to inaccuracies of the microdensitometer trace, possible differential staining with coomassie blue and the assumption that bought molecular weight markers really are at the concentration stated.

Protein preparation 3 was used for this experiment as it appears to be more active than the other samples. The protein concentration,as determined using the microdensitometer, is 0.72 mg/ml. Dilutions of this resolvase preparation in 0.7M NaCl/TMD were used in a series of reactions with pLS138 as substrate DNA. Incubations were for 30min at 37^{o}C, reactions were analyzed by gel electrophoresis to indicate presence of catenated products (figure 5.11). The gel indicates that at the lowest concentration tested, about 0.036μg resolvase, almost complete conversion of s/c substrate to s/c catenated products has occurred. Based on the assumption that the reaction contains 0.5μg DNA and 0.036μg resolvase, the following calculation can be made:

21,300g resolvase = $6X10^{23}$ molecules

2.13μg resolvase = $6X10^{13}$ molecules

0.036μg resolvase = $1X10^{12}$ molecules

$3.2X10^{6}$g pLS138 = $6X10^{23}$ molecules

3.2μg pLS138 = $6X10^{11}$ molecules

0.5μg pLS138 = $1X10^{11}$ molecules

Therefore under standard reaction conditions the ratio of protein monomers to DNA molecules is 10:1. Due to error range of the method used to determine the protein concentration this ratio could be 2-5 fold out. If this ratio is correct, it would suggest that the resolvase preparation is fairly active. It is of significance that this preparation was kept in urea for only a short time compared with other samples suggesting that, as expected, urea probably has a detrimental effect on the protein's activity.

Comparison of reactions containing varying amounts of resolvase indicates that as the resolvase concentration is increased there is a reduction in the conversion of s/c substrate to catenated products. This inhibitory effect has also been observed with γδ resolvase (Reed,pers.comm.). The reasons for this are unclear; it could be due to excess resolvase binding non-specifically to DNA and interfering with the reaction in some way, for example, by changing the superhelicity of the molecule. Concentrations below those used in these experiments result in very little,if any, resolution.

6.Stoichiometry of the resolvase-mediated reaction

The experiments described above suggest that five monomers of resolvase are required per *res* site. This figure is quite high for a catalytic protein, which has been the basis for suggesting that Int and γδ resolvase are non-catalytic. I do not believe that the high concentrations required for these proteins to act in site-specific recombination necessarily reflect their stoichiometric nature. There is evidence to suggest that Int is a type I topoisomerase (Kikuchi and Nash,1979a;Wang *et al.*,1980); topoisomerases are known to be catalytic. γδ resolvase has been shown to create

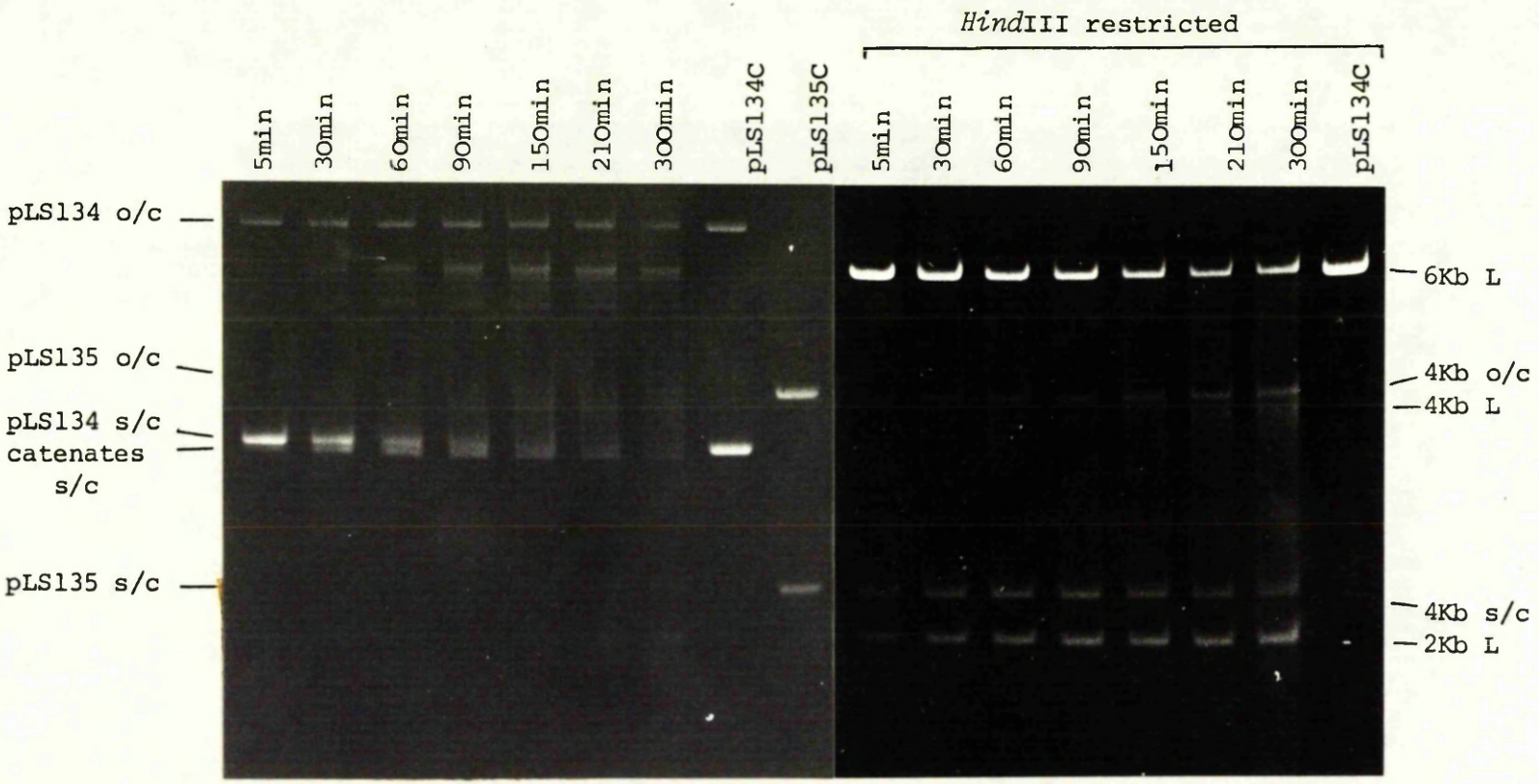


FIGURE 5.12 Time course of pLS134 resolution

The majority of reaction products are formed within the first 30min of the reaction time. These appear as s/c catenated products, initially, but during the latter part of the time course free reaction products are also visible, presumably due to non-specific nicking activity.

0.8% agarose run at 25V.

double-stranded cleavage at the *res* site on substrates containing two copies of *res* in direct repeat (Reed and Grindley,1981); the protein then becomes covalently attached to the recessed 5' terminus produced. This mechanism is reminiscent of DNA-gyrase. The breakage and subsequent formation of phosphodiester bonds would not be expected to change the active site of the protein, thus it should be catalytic.

Time course experiments using standard reaction conditions suggest that the reaction occurs fairly quickly. The majority of products are formed within the initial 20min of the reaction and then little conversion occurs over several hours (figure 5.12). The tail off of the reaction could be due to use of all the available substrate, use of all of the available resolvase, or instability of resolvase. These possibilities can be easily tested. Incubation of a reaction mix containing resolvase, but lacking DNA, for one hour followed by addition of substrate results in normal conversion of substrate to products. This suggests that resolvase is stable under the conditions used. If extra substrate DNA is added after the "plateau point", this is not converted to products; this suggests that resolvase is the limiting factor in the reaction. The simple kinetics of the reaction may be studied by adding competitor DNA to time course experiments. A series of experiments was constructed based on the substrate pLS139 and using the minimal amount of resolvase required for resolution. It was observed that if an equal amount of pACYC184::Tn*3* (one *res* site) was added at the same time as pLS139, the amount of pLS139 converted to products was approximately half of that converted in the absence of competitor (figure5.13). This is expected if both species of DNA titrate out all of the available resolvase equally. A similar

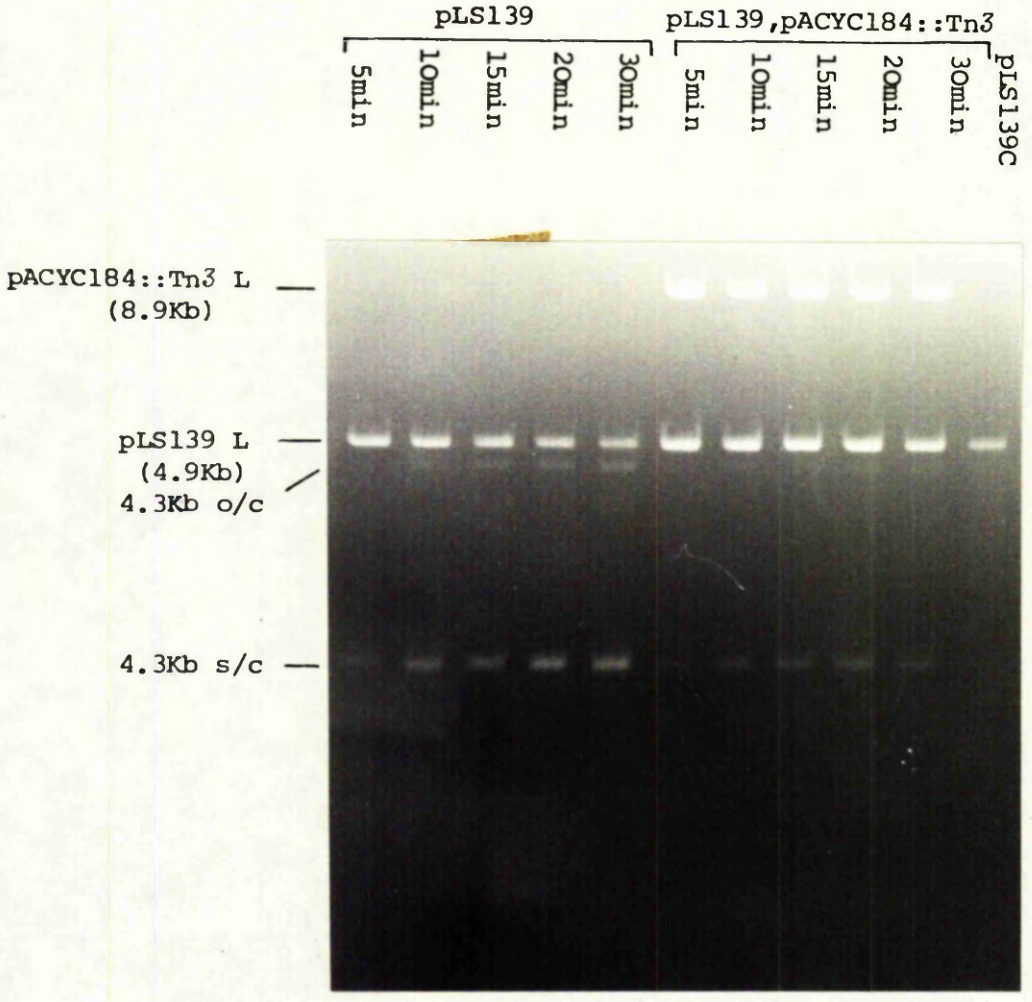


FIGURE 5.13 Effect of adding competitor DNA to pLS139 resolution reactions

The samples have been restricted with *Hind*III to clearly distinguish products from substrate. *Hind*III has one site on pLS139 and one site on pACYC184::Tn*3* . The resolution products from pLS139 consist of 4.32 and 0.65Kb circles; the smaller of these carries the *Hind*III restriction site. Thus restriction of reacted pLS139 with *Hind*III yields a 4.32Kb circle and 657bp linear fragment.

0.8% agarose gel run at 25V (resolvase prep. 3).

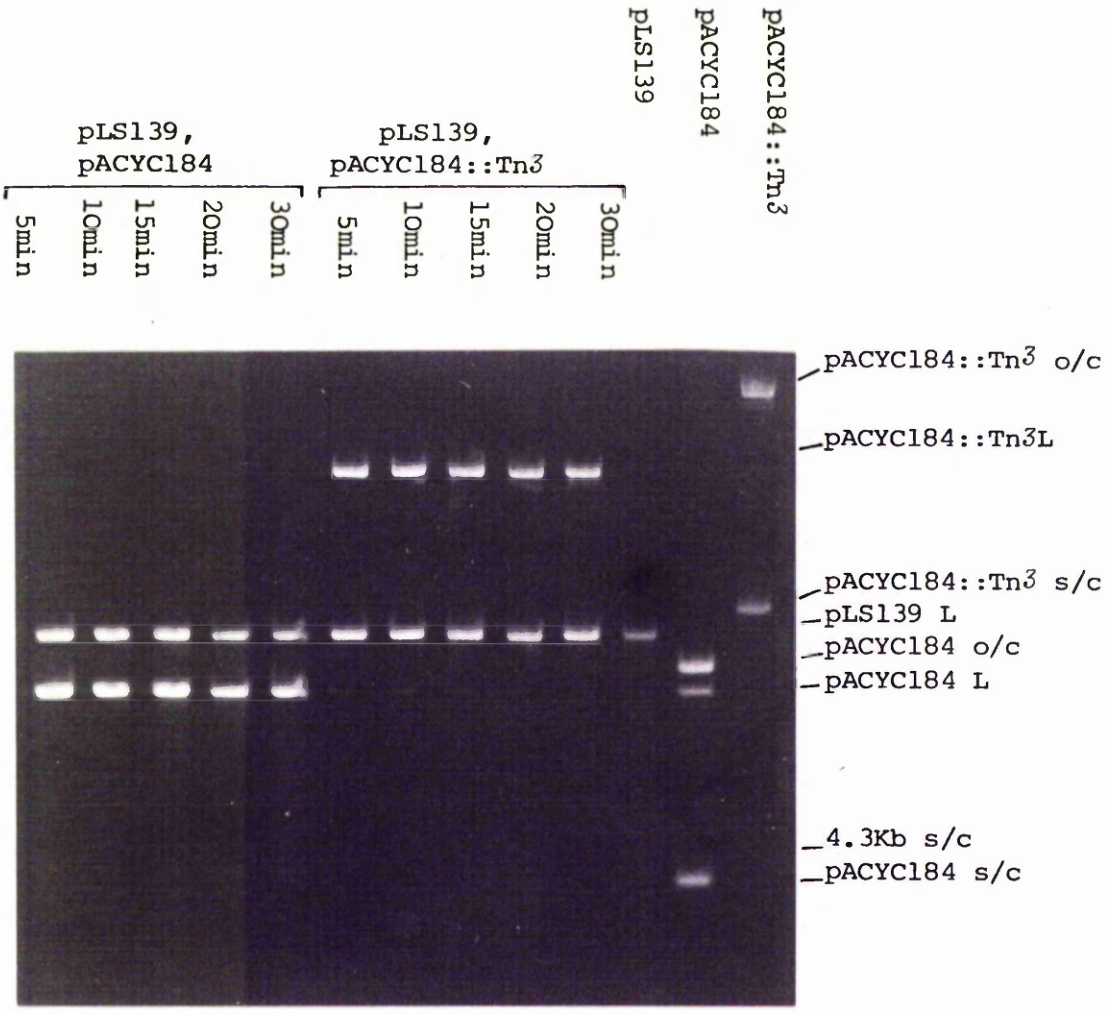


FIGURE 5.14 Effect of competitor DNA on pLS139 resolution

When pACYC184 (no *res* sites) is incubated with resolvase prior to addition of pLS139, very little conversion of pLS139 to products is observed. If the competitor contains a single *res* site (pACYC 184::Tn*3*), then no resolution of pLS139 is detectable.

0.8% agarose gel run at 25V (resolvase prep. 3).

experiment was performed, but this time pLS139 was not added to the reaction mix until 15min after pACYC184::Tn*3* had been added. Over the same time course, starting after addition of pLS139, there was no apparent conversion of substrate to products (figure 5.14). This indicated that pACYC184::Tn*3* had titrated out all of the resolvase leaving none available for pLS139 resolution, suggesting that once bound to a DNA molecule resolvase remains firmly attached and is unavailable to additional DNA added after this point. To test the specificity of resolvase binding the above experiment was repeated using pACYC184 as the competitor DNA. The reaction mixture was incubated for 15min before the addition of pLS139, then the kinetics followed as previously. It can be seen that only a small proportion of pLS139 is converted to products (figure 5.14).This may be due to non-specific binding of resolvase to pACYC184, but as this plasmid contains no *res* sites the binding may not be as tight and therefore some resolvase will dissociate from pACYC184 to bind specifically to pLS139.

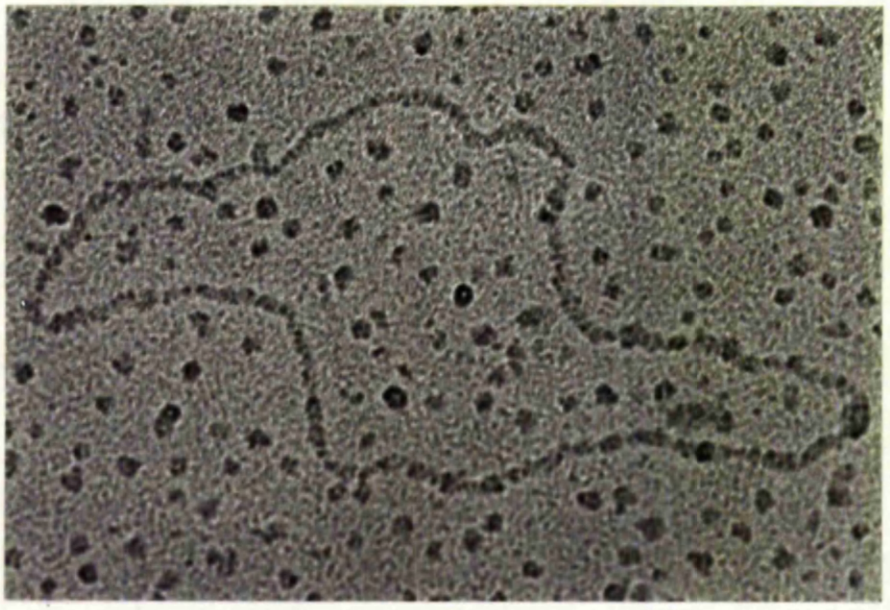
One of the problems in trying to determine whether resolvase is catalytic is the fact that once resolution has occurred, the protein probably remains tightly bound to the *res* sites present on the product molecules. When tightly bound at *res* the protein does not readily dissociate, thus it is difficult to tell whether it can then mediate a second resolution event. The protein may be considered to be catalytic in that the active site is probably not irreversibly altered during the reaction, but may also be considered to be stoichiometric in that once resolution is completed resolvase remains firmly bound at *res* and is therefore not available for further resolution. In a natural situation if the transposon should transpose again, then resolvase is already sitting

on one *res* site ready to resolve the new cointegrate.

It was noticed that throughout prolonged time course reactions several new bands appeared on gels; this was more noticable with resolvase preparation 1. The band corresponding to linear substrate molecules increases in intensity, indicative of nicking throughout the reaction. This nicking does not appear to be specific as subsequent restriction would otherwise yield products of characteristic size. Non-specific nicking would presumably cause a reduction in the amount of s/c substrate available for resolution. Reaction products, which may originally be s/c catenates, on nicking would produce a variety of products, eg. one s/c linked to one o/c, two linked o/c and free reaction products, s/c and o/c. These forms are probably represented by the bands of new mobility which are formed and increase in intensity, during the reaction (see figure 5.12). These molecules have not been examined further though it would be possible to extract the DNA from individual bands to examine by electron microscopy. The nicking activity observed with resolvase preparation 1 could be due to a minor protein contaminant or due to resolvase itself acting with altered specificity.

A commonly used assay for detecting recombination intermediates is electron microscopy. This provides a direct visualization of participating DNA molecules (Nash *et al.*,1977; Potter and Dressler,1979; Shibata *et al.*,1979; Cunningham *et al.*,1980; Kolodner,1980; Fishel *et al.*,1981; Mizuuchi *et al.*,1982). The band migrating slightly ahead of s/c pLS134 was eluted from a gel, by the electrophoresis method, and used for electron microscopy. Although this DNA had been exposed to ethidium bromide

(a)



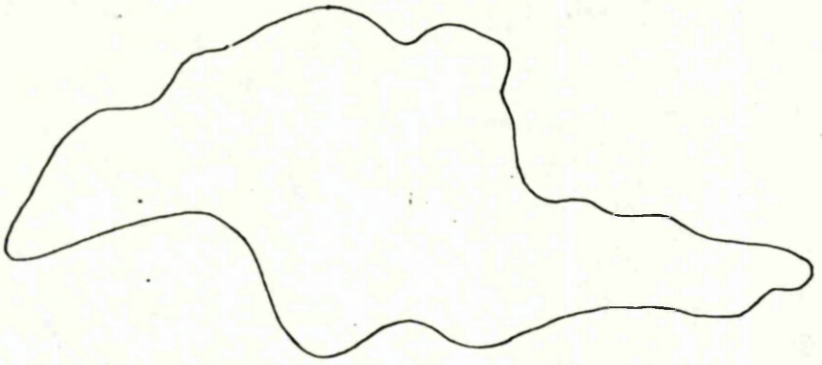


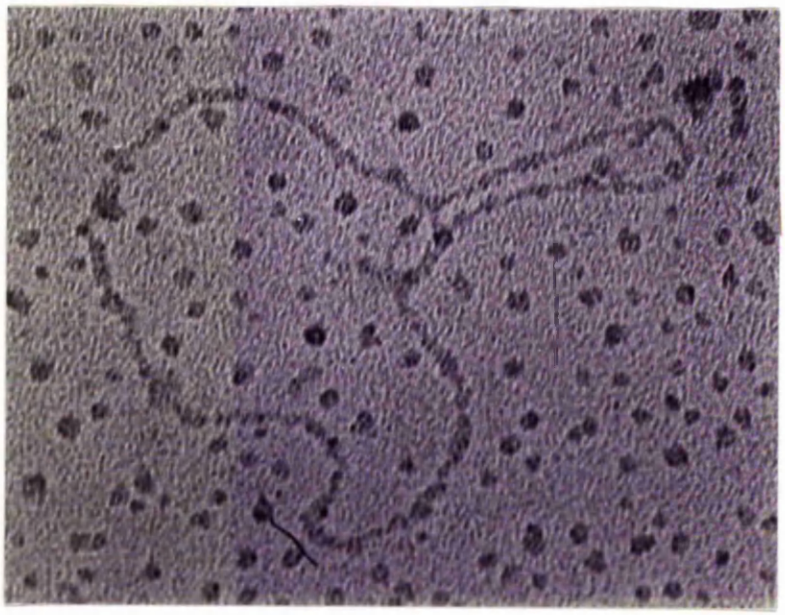
FIGURE 5.15 Electron micrographs of reacted pLS134 DNA

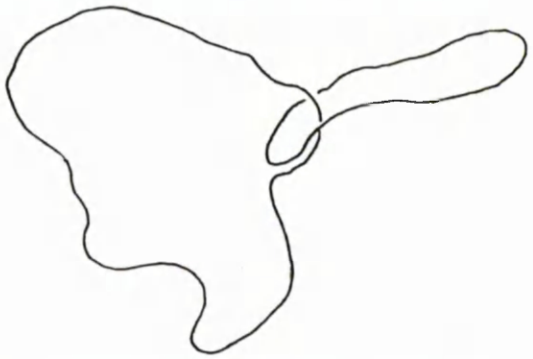
(a) Relaxed substrate molecule

(b)-(e) Apparent catenated molecules, with graphic interpretation

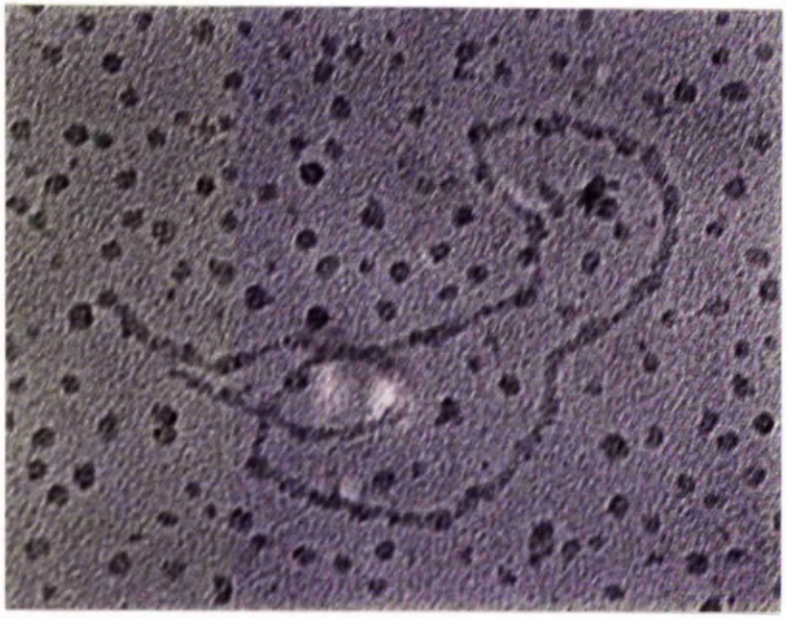
(photographed at 20,000 X magnification)

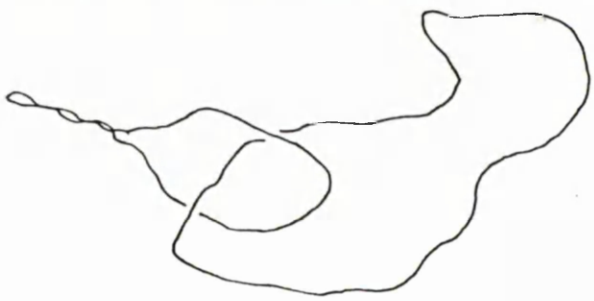
(b)



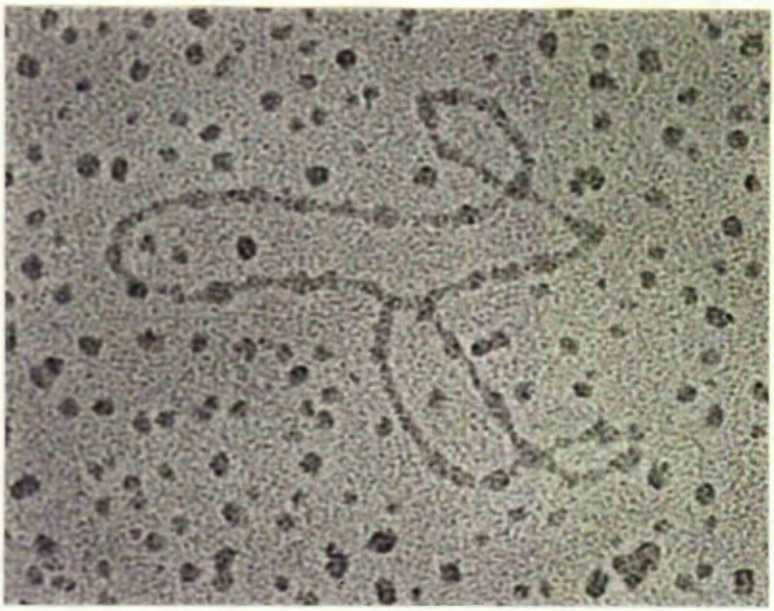


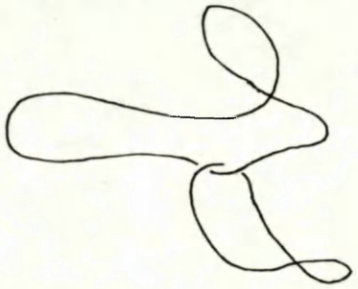
(c)



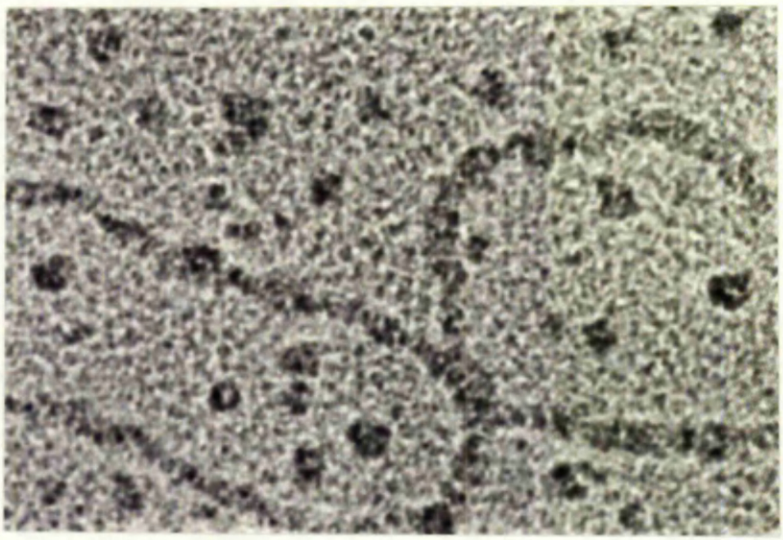


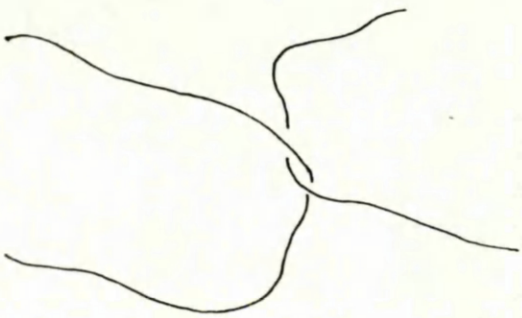
(d)





(e)





staining and u.v. light most of the molecules on grids were still s/c and therefore difficult to distinguish. Some molecules were, however, relaxed and appeared to be catenated; examples of electron micrographs are shown (figure 5.15).

7. Topological considerations

It has been reported that a s/c substrate is an absolute requirement for γδ resolvase activity *in vitro* (Reed,1981b). The lambda Int system also works with greatest efficiency on a s/c substrate, though under conditions of low ionic strength, some reaction is observed using a relaxed substrate (Mizuuchi and Nash ,1976; Kikuchi and Nash,1979a;Pollock and Abremski,1979). The data presented in this chapter is consistent with these findings. When catenated products are formed there is a corresponding decrease in the s/c substrate band; there appears to be no decrease in the o/c band (figure 5.16). It seems unlikely that s/c are converted to o/c, then to s/c catenated products, therefore a s/c substrate would appear to be a requirement for resolution. A preparation of purified o/c molecules was used for *in vitro* resolution reactions. There was no detectable conversion to products, which would presumably be visible on gels as free reaction products or simple catenates. The addition of ATP could not compensate for the lack of superhelicity (data not shown). As the reaction is dependent on superhelicity and requires no external energy source, it is tempting to speculate that the energy contained within supercoiling is used in some way to "drive" the reaction.

When catenated reaction products are cleaved with an enzyme which cuts only one of the two circles, the other remains in a

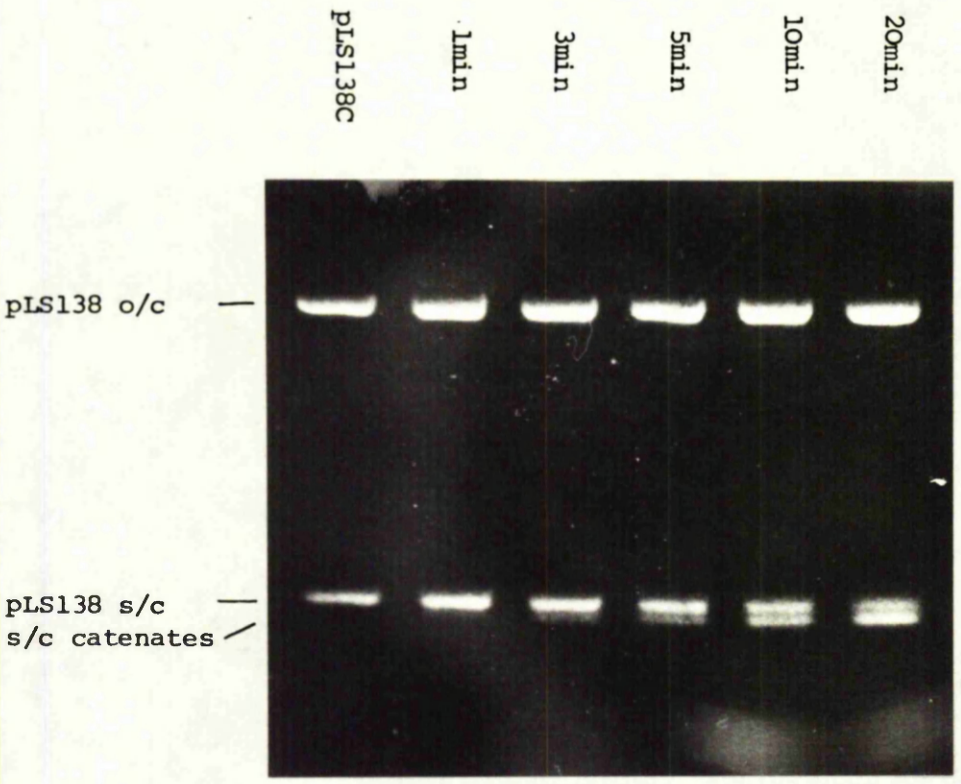


FIGURE 5.16 Time course of pLS138 resolution

The s/c substrate band decreases throughout the reaction with a corresponding increase in the band migrating ahead of the s/c substrate, presumably this is due to formation of s/c catenated products. There is no appaernt change in the o/c substrate band. 0.8% agarose gel run at 25V, resolvase prep.3.

highly supercoiled form. Some o/c are present and increase throughout the reaction, presumably due to nicking activity, but when initially formed the products maintain about 70% s/c (see figures 5.12 and 5.13). This suggests that the recombination reaction is very conservative with little loss of superhelicity. During the cleavage and ligation of strands to new partners the strands must be held in some way, presumably by proteins, so that free rotation of strands can not occur. Experiments by Reed and Grindley (1981) suggest that γδ resolvase makes a double stranded cleavage at *res* sites and attaches covalently to the recessed 5' terminus. Resolvase may act in a manner analagous to type II topoisomerases: simultaneous cleavage of both strands at *res* sites, conservation of the phosphodiester bond energy by covalently bonding to the protein followed by ligation of the strands to new partners. This would result in no significant loss of superhelicity, though the strand rotation required to align strands for ligation to new partners may result in a slight decrease in superhelicity, maybe changing the linking number by one.

During the ligation step one might expect that the original DNA strands could be rejoined (maybe following strand passage as for topoisomerases) or ligated to new partners (resolution) depending on alignment of the strands and specificity of the protein. If resolvase could mediate the strand passage event, like gyrase, then one would expect that it could change the linking number of a particular substrate (see Gellert, 1981 for a review of DNA topoisomerases). Reed and Grindley have been unable to detect any topoisomerase activity associated with γδ resolvase (pers. comm.). However the three independent Tn*3* resolvase preparations exhibit topoisomerase activity to varying degrees. This activity

is visualized on agarose gels by the formation of a ladder of topoisomers between the s/c and o/c substrate DNA bands (figure 5.17). This property of resolvase has not been characterized as it appears to be very inconsistent and may depend on critical ionic conditions in the reaction. Of course, the possibility cannot be ruled out that it could be due to minor protein contaminants present in resolvase preparations rather than intrinsic to resolvase itself. The harsh conditions to which the protein is exposed during it s purification may affect the specificity of the protein in some way. The reaction conditions required for resolvase topoisomerase activity appear to mirror the conditions for resolution over the same pH range and ionic strength. On cleavage with *Eco*RI or *Hind*III the circular products of pLS134 resolution were also present as a mixture of topoisomers. This suggested that resolution works on molecules which contain varying degrees of superhelicity, ie.they do not have to be highly supercoiled, or that the topoisomerase activity works on the products which have only one *res* site (figure 5.17B).

One would expect resolvase topoisomerase activity to change the linking number of a particular topoisomer by two steps. As a naturally derived s/c DNA preparation contains a mixture of topoisomers , the resolvase topoisomerase always appears to change the linking number of substrates in steps of one (see figure 5.17). To determine the linking number change induced by resolvase one would have to prepare DNA of a unique linking number. Another feature expected of resolvase topoisomerase is specificity to molecules containing *res* sites. Reed and Grindley (1981) have shown that resolvase cleavage only occurs on molecules with directly repeated copies of *res*. One would therefore expect

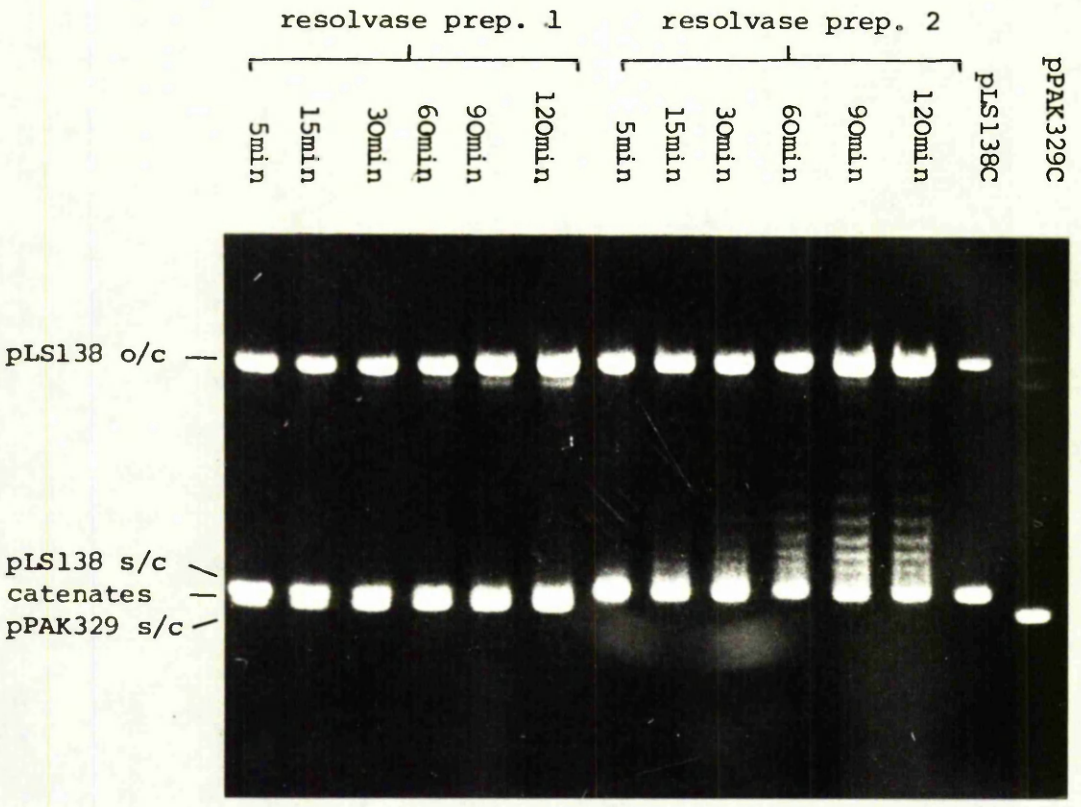


FIGURE 5.17A Comparison between resolvase preparation 1 and 2 showing topoisomerase activity

This gel compares the resolution and topoisomerase activities of preps.1 and 2. It can be clearly seen that prep.1 exhibits predominantly resolution activity, indicated by the formation of catenates. Resolvase prep.2 shows greatest topoisomerase activity; some of the topoisomers produced may be catenated products, but as these samples have not been restricted these are indistinguishable from substrate molecules.

0.8% agarose gel run at 25V.

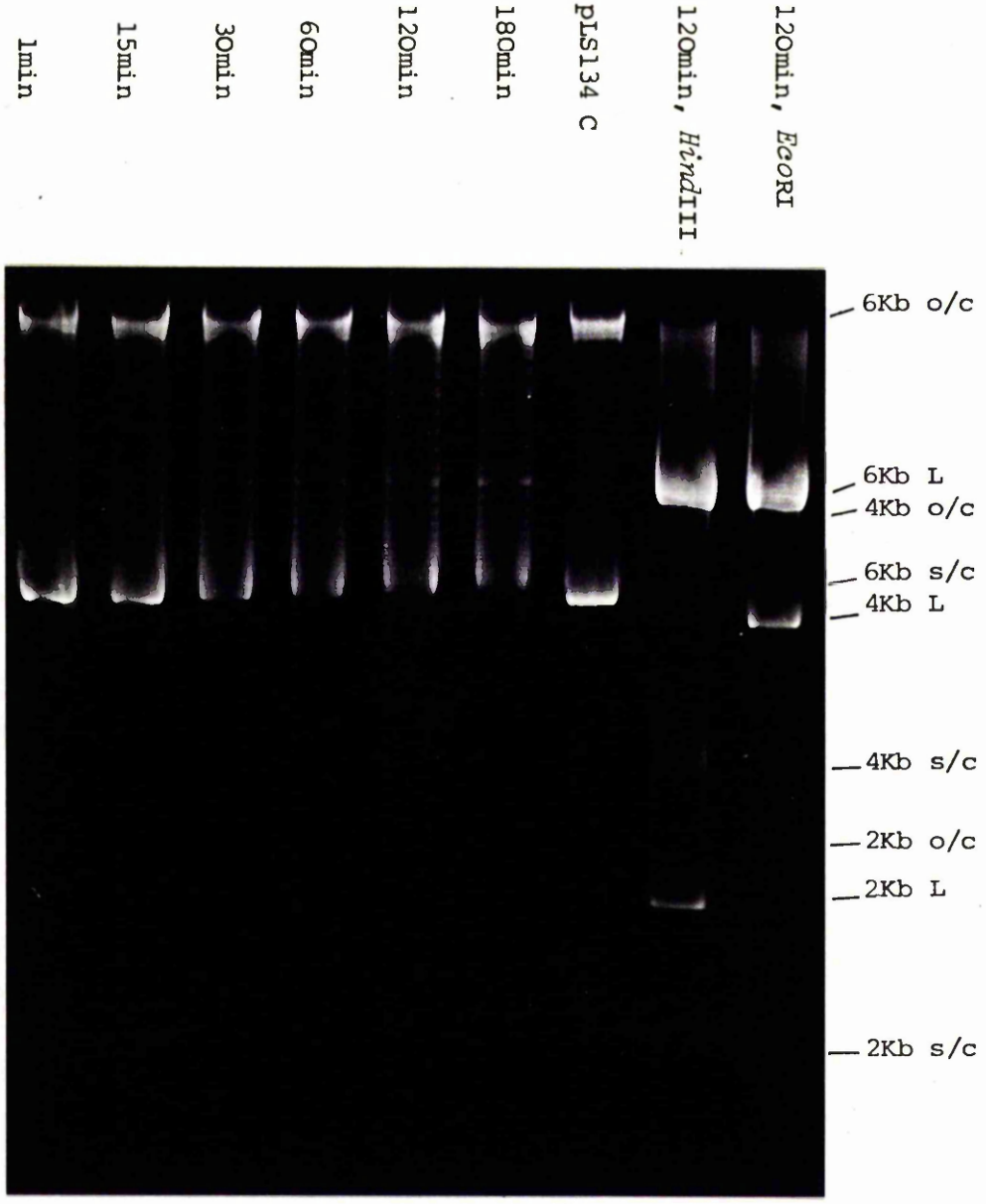


FIGURE 5.17B Topoisomerase activity exhibited by resolvase

A time course reaction of pLS134 with resolvase prep.2. It can be seen that a range of topoisomers are formed very quickly; these do not change significantly throughout the time course. The restricted samples indicate that reaction products are present in various topoisomeric forms also.

0.8% agarose gel run at 25V.

the topoisomerase activity to be exerted only on such substrates. However, the topoisomerase activity,which was detected with Tn*3* resolvase, appeared to sometimes occur on molecules with no *res* sites.

The lambda Int protein has been shown to exhibit topoisomerase activity (Kikuchi and Nash,1979b). This activity has been characterized and shown to change the linking number of a purified topoisomer in steps of one. The surprising feature of the reaction is that it is not specific to *att* sites, even though purified Int has been shown to bind specifically to *att* sites in DNA protection experiments (Mizuuchi *et al.*,1981). Int has been classified as a type I topoisomerase as it appears to cause single stranded cleavage/ligation during topoisomerization (Wang *et al.*,1980). The observed single stranded cleavage/ligation activity of Int has been a major force in formulating a model for the recombination mechanism (Nash *et al.*,1981).

8. Catenation - an intrinsic property of resolvase?

It appears that catenated plasmids are the major reaction products. The appearance of free reaction products, only observed after prolonged incubation with resolvase preparation 1, is presumably due to non-specific nicking activity. Is catenation a consequence of the reaction mechanism? If the formation of catenated products is due to tangling of different parts of the s/c substrate circle prior to strand exchange, then one would expect that relaxed circles could be converted to either free products or catenates (Abremski and Nash,1980). As resolvase does not appear to resolve relaxed DNA this approach to the problem has been fruitless. However, one might expect that reducing

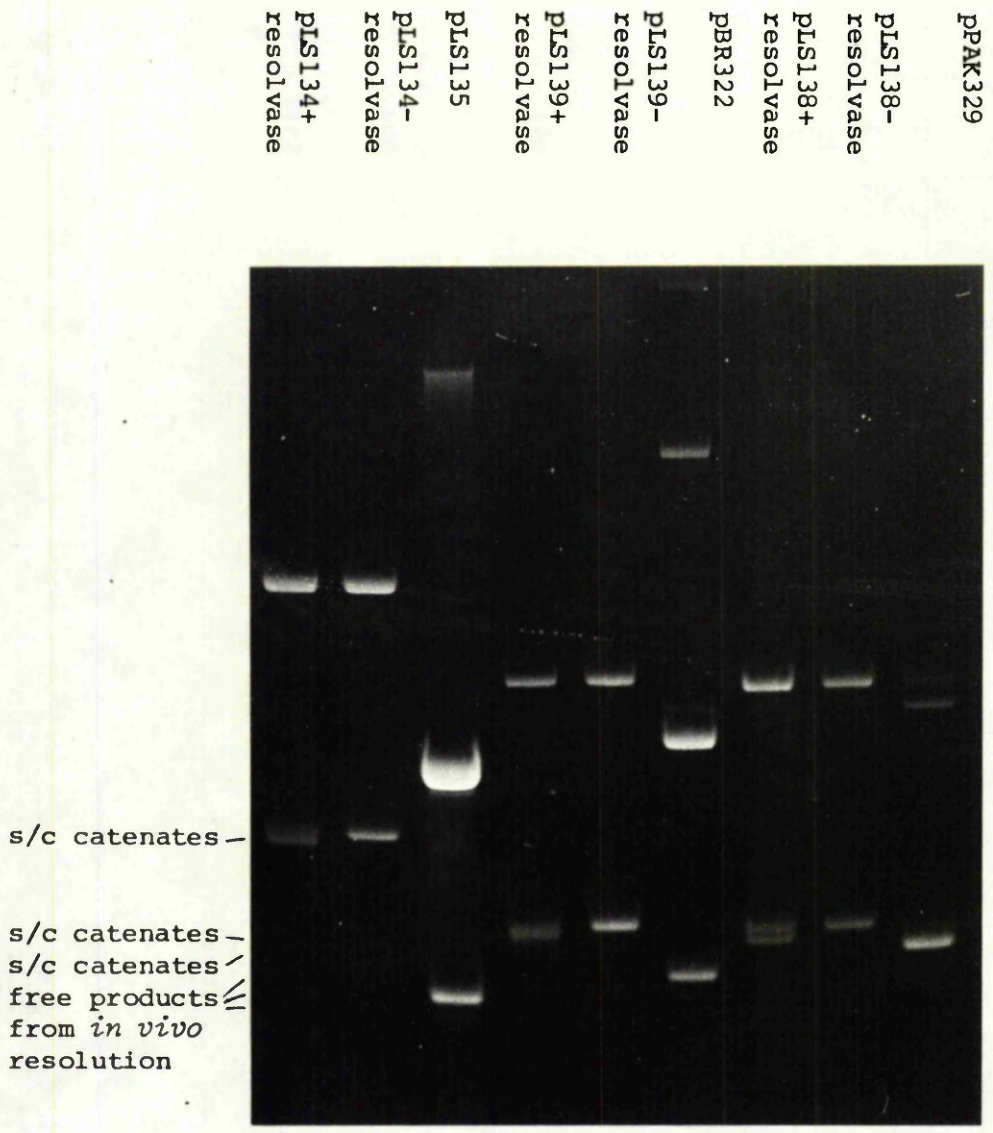


FIGURE 5.18 Catenate formation by pLS134, pLS139 and pLS138

It can be clearly seen that the band migrating slightly ahead of s/c substrate is distinguishable from free product molecules for all three substrates.

0.8% agarose gel run at 25V. Samples incubated with resolvase prep. 3 for 30min.

the length of DNA between *res* sites would decrease the possibility of supercoils between *res* sites and therefore the proportion of catenated products. This hypothesis has been tested by comparing the formation of catenates from substrates pLS134, pLS139 and pLS138, which have 2Kb, 657bp and 282bp separating *res* sites respectively (figure 5.18). The results clearly show that all three substrates are converted to catenated products; no free reaction products are detectable.

9. Requirement for magnesium ions

If Mg^{++} is omitted from the reaction mixture then no resolution is detected. It has been shown that γδ resolvase cleaves at *res* sites in the absence of Mg^{++}, but only when two *res* sites are directly repeated within a molecule (Reed and Grindley,1981). This has allowed alignment of the precise cleavage point within the 19bp crossover region. As hybrid sites have been generated *in vivo* between Tn*3* and γδ it is expected that the cleavage points by both proteins are the same (Kostriken *et al.*, 1981; Reed,1981a).

Incubation of plasmids pLS134 and pLS139 with resolvase, in the absence of Mg^{++}, for 16hr yielded new bands which had not been previously observed in the presence of Mg^{++}. These bands were only detected using resolvase preparation 3 ; the other samples did not show this activity (figure 5.19). Substrate molecules containing only one *res* site or with two inverted *res* sites were not susceptible to cleavage. This suggests that site-specific cleavage at *res* is intermediate in the reaction. Sequencing of the fragments produced should reveal the exact cleavage point. Reed and Grindley have evidence that the protein

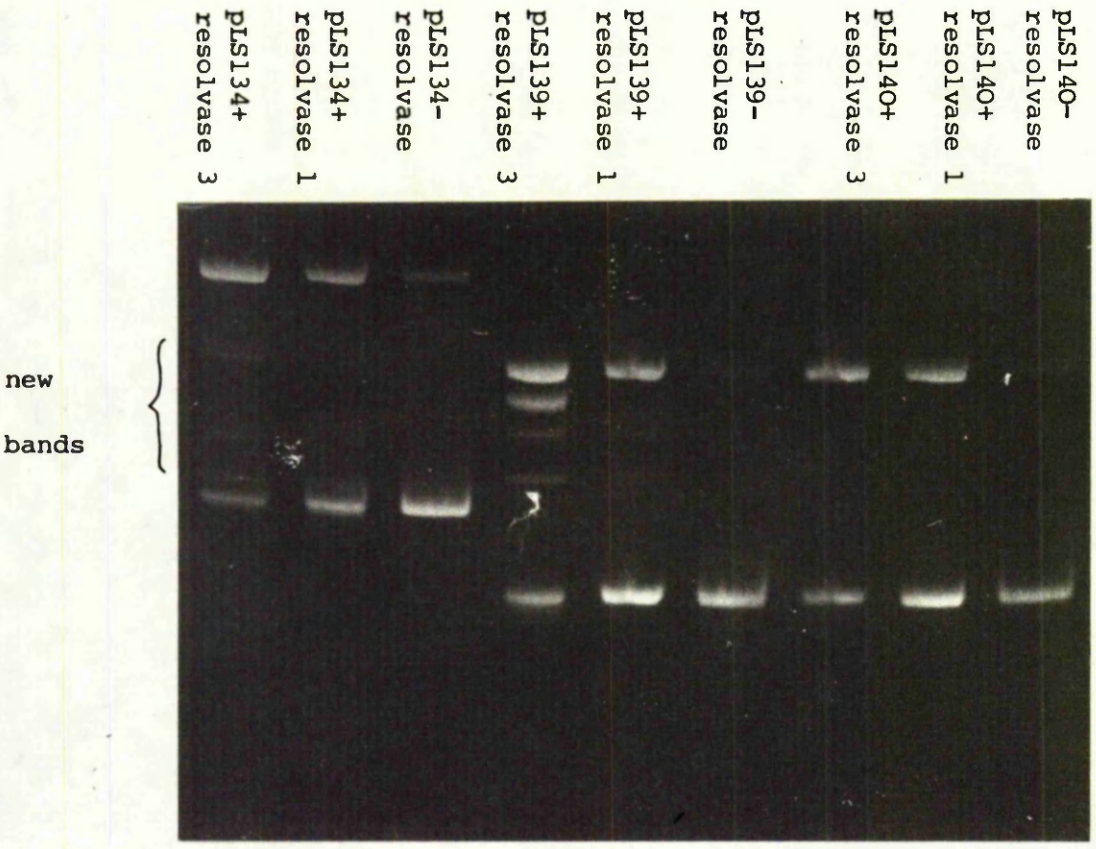


FIGURE 5.19 Incubation of substrates in the absence of Mg^{++}

With resolvase prep.3 new bands are formed when substrates pLS134 and pLS139 are incubated for 16hr in the absence of Mg^{++}. The substrate with inverted *res* sites, pLS140, is unaffected under these conditions.

0.8% agarose gel run at 100V.

attaches covalently to the exposed 5' terminus of the staggered cut (1981). It would be interesting to determine which amino acid forms the covalent bond to the DNA backbone. DNA gyrase forms a phosphodiester bond between a tyrosine residue on subunit *gyr*A and the cleaved DNA 5' terminus (Sugino *et al.*, 1980; Tse *et al.*, 1980). There are two tyrosine residues present in the resolvase molecule, located within each end of the coding region. The most highly conserved region between the γδ and Tn*3* resolvase proteins is at the amino terminus; due to the observed similarities between the proteins one might expect that the active site would be within the most highly conserved region. It is tempting to speculate that the tyrosine residue at the amino terminus may be part of the active site and participate in formation of a phosphodiester bond with DNA at the *res* site.

10. Requirement for *res* sites in *cis* for resolvase activity

The substrates used throughout this chapter contain two *res* sites directly repeated within the same molecule; resolvase works very efficiently on this arrangement. The plasmids pLS135 and pPAK329 contain only one *res* site; one would expect that if resolvase could join these molecules by recombination across *res* sites then the products of the reaction would be dimers, or catenates if these were then immediately resolved. Using various reaction conditions no recombination activity could be detected (data not shown). It has been shown that addition of up to 5mM spermidine may assist catenation by topoisomerases due to aggregation of DNA molecules (Krasnow and Cozzarelli, 1982). The addition of spermidine to reactions containing pLS135 or pPAK329 led to no detectable formation of dimers or catenates. The apparent lack of joining activity by resolvase

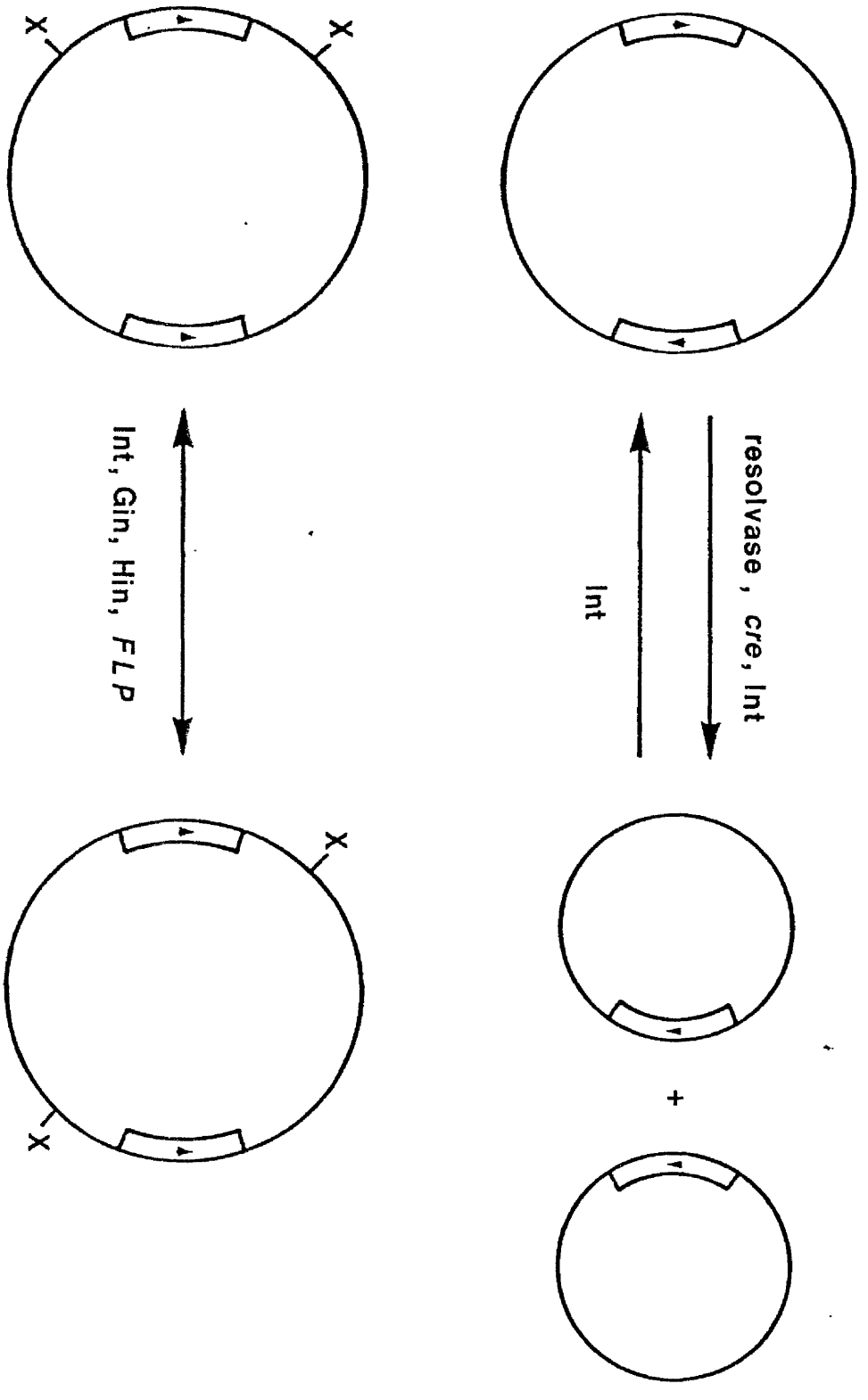
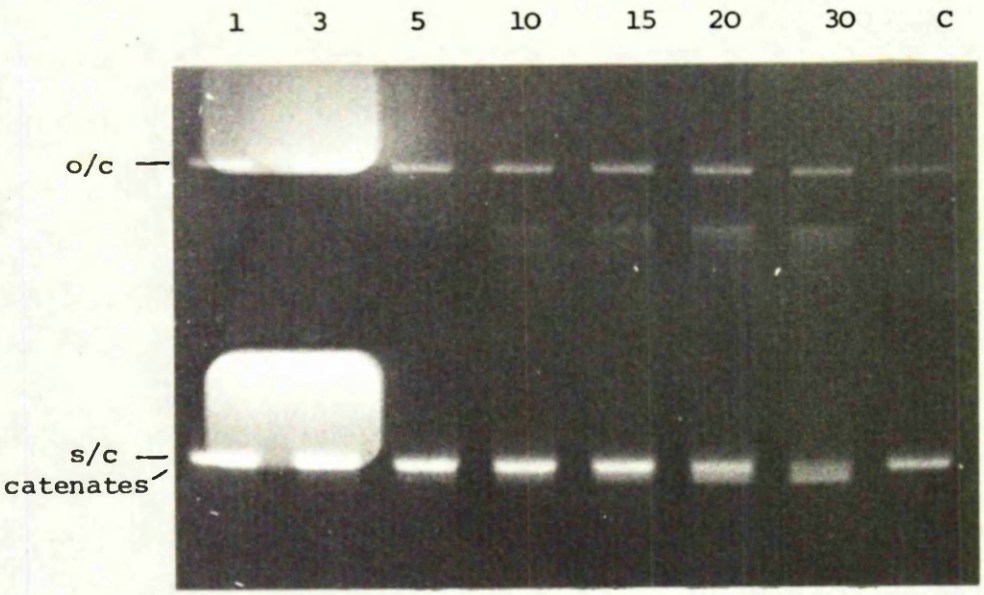


FIGURE 5.20 A comparison of the preferred substrates for various site-specific recombination proteins

is in direct contrast to the lambda system in which the major activity of Int is to mediate integration by joining replicons. (see figure 5.20). This probably reflects different mechanisms involved in these systems. One of these differences may be in the way the two recognition sites are aligned for the recombination event. The two sites may be brought into close proximity by random movement of the supercoiled substrate molecule. If protein molecules are attached to both of the recombination sites then one may imagine that protein to protein interaction may stabilize the two sites in a position close enough together for recombination to occur (Sherratt *et al.*, 1981b). Such a model would satisfactorily explain the data obtained from the Int system, but is incompatible with the observations made with resolvase. Due to the constraints imposed by DNA structure one would expect that reducing the length of DNA between *res* sites would decrease the chances of random alignment of sites, thus reducing the efficiency of the reaction. The plasmids, which had been constructed to study catenation, pLS134, pLS139 and pLS138 have 2Kb, 657bp and 282bp separating *res* sites respectively. While studying resolution of these plasmids over a short time course, it was observed that pLS138 was converted to catenated products most rapidly (figure 5.21). These results suggest that resolvase has an active mechanism for seeking *res* sites within a molecule rather than relying on random alignment. This hypothesis is discussed in detail in the following chapter.

A


1
3
5
10
15
20
30
C
o/c
s/c
catenates

B

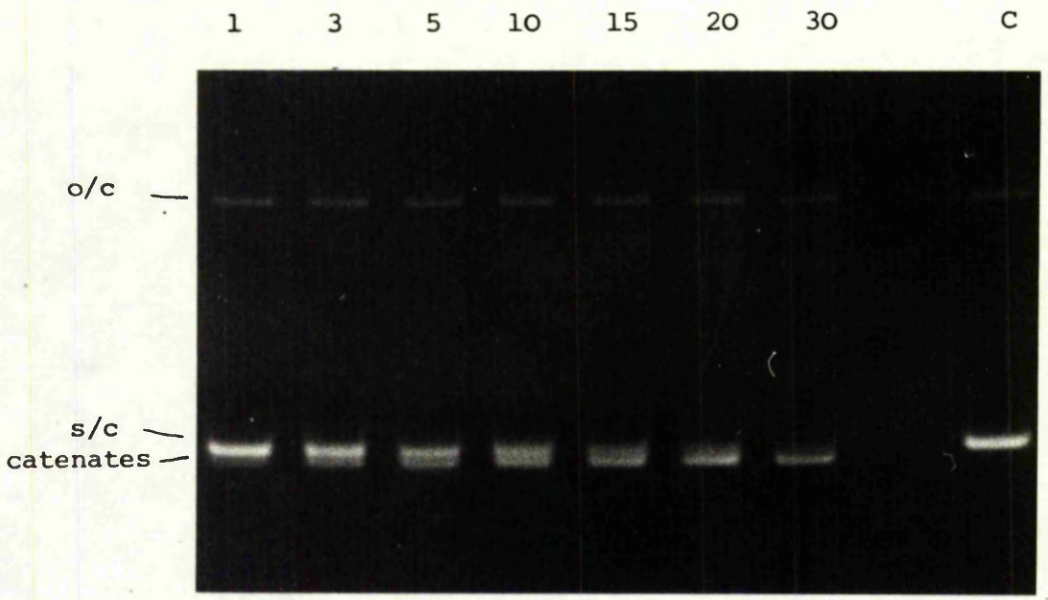

1
3
5
10
15
20
30
C
o/c
s/c
catenates

FIGURE 5.21 A comparison of the efficiency of catenate formation between pLS134 and pLS138

A. Time course of pLS134 resolution

Within the first 15min of the reaction less than 10% of pLS134 substrate is converted into catenated products. Over the following 15min conversion to catenates is almost complete.

B. Time course of pLS138 resolution

For pLS138 more than 50% of s/c substrate has been converted to catenates within 15min; by 30min conversion is almost complete. Both substrates reach complete conversion at about 30min, but for pLS138 the products appear much faster initially, within 1min. This suggests that an active mechanism is operating to bring *res* sites together for recombination, favouring sites close together.

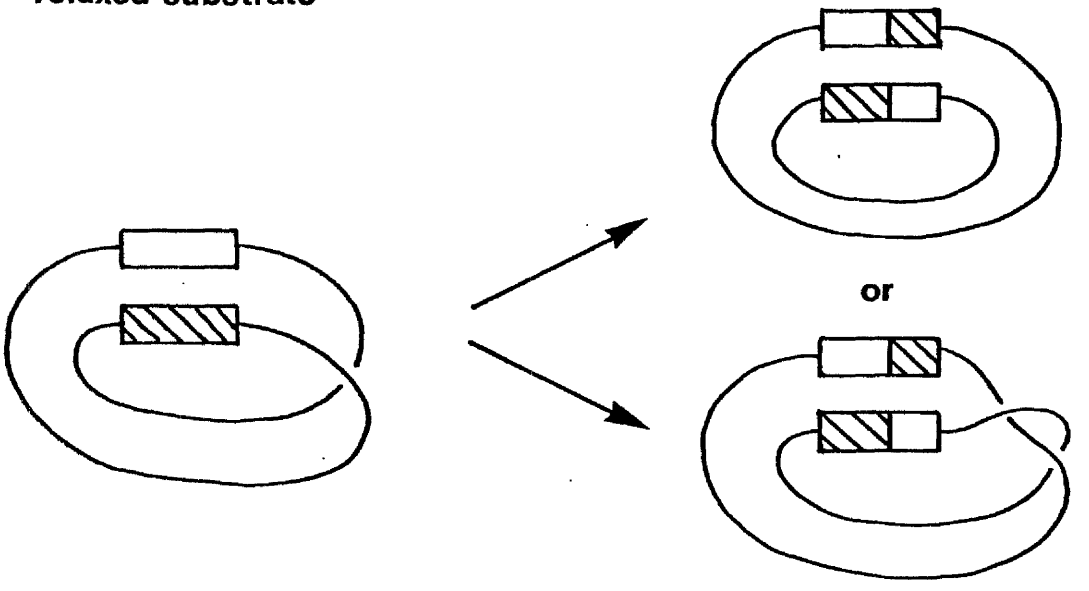
Discussion

The data presented in this chapter suggests that the only requirements for resolution *in vitro* are a s/c substrate, resolvase, buffer, Mg^{++} and NaCl. No host factors appear to be required for the reaction. The absence of high energy cofactors and triphophates indicates that there is no concomitant DNA synthesis during the recombination event or requirement for an external energy supply. Evidence provided by the formation of catenated products suggests that the reaction is reciprocal, i.e. all strands broken during the recombination event are rejoined to produce continuous DNA helices.

The formation of catenates is an interesting feature of the reaction; it has also been shown for other *in vitro* site-specific recombination systems (Nash *et al.*, 1977; Reed, 1981b). Nash has proposed that recombination between specific sites on a s/c molecule always results in catenation (see figure 5.22). This is due merely to continual DNA writhing such that the DNA between sites is always s/c. If, however, the substrate is relaxed, then the reaction products should be a mixture of catenates and unlinked circles in roughly equal proportions. Int works at a low, but detectable, frequency on relaxed substrates thus allowing a comparison between s/c and o/c substrates on formation of catenates to be made (Pollock and Abremski, 1979; Pollock and Nash, 1980). Even using the relaxed substrate 85% of reaction products are catenated, Nash suggests that this is due to the low degree of writhing in a relaxed molecule and concludes that catenation is not an intrinsic property of Int.

Although I have been unable to detect recombination using a

A relaxed substrate



B supercoiled substrate

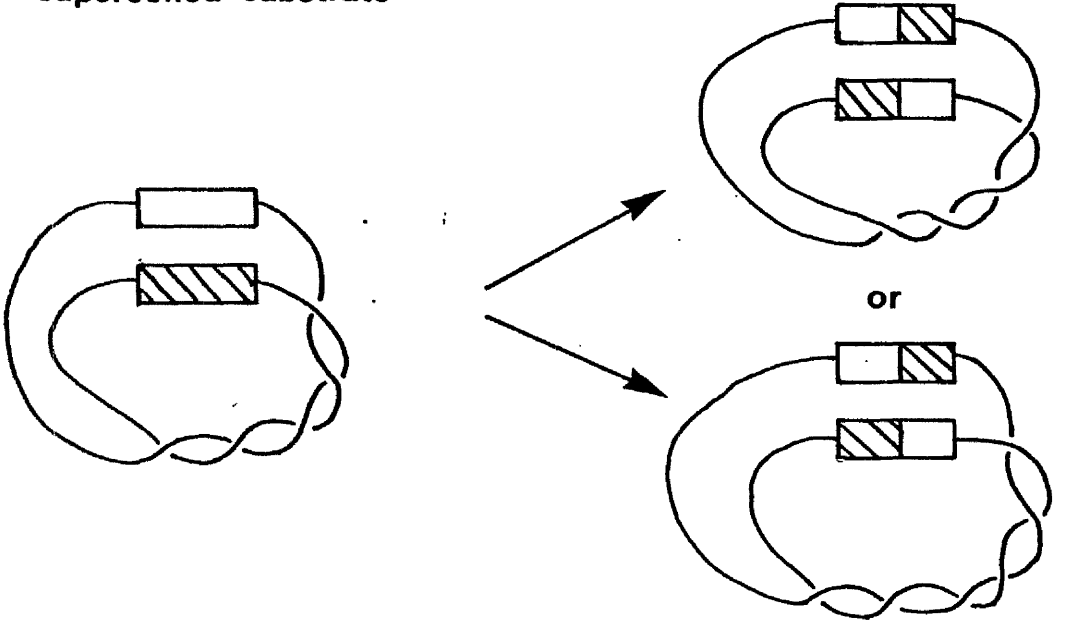


FIGURE 5.22 Comparison of catenate formation from relaxed and supercoiled substrates

If the two *res* sites are brought into alignment by a random process, then the superhelicity of the substrate should determine whether the products are catenated or unlinked circles.

A. When the substrate is relaxed one twist is enforced between the *res* sites for alignment of directly repeated *res* sites. Depending on how the strands are rotated to join to their new partners this results in either simple catenates or unlinked circles.

B. When the substrate is s/c there are several twists between *res* sites; this substrate always results in catenated products. Again, depending on strand rotation, these may gain or lose a twist. In practise the system probably results in a loss of energy, ie. to a less s/c state.

relaxed substrate, the approach used to this problem has been to construct substrates with varying lengths of DNA between *res* sites. If formation of catenates is due merely to superhelicity between *res* sites then,by shortening the length of DNA between sites the chances of this region being supercoiled are reduced, therefore the ratio of catenates to unlinked products should decrease correspondingly. The substrate pLS138 has only 282bp between *res* sites. Working on the assumption that pBR322 has about 16 superhelical turns, then there should be one twist every 270bp. As DNA is constantly writhing one would expect that the region separating *res* sites on pLS138 would lack supertwists some of the time. One would therefore predict that free reaction products would be formed at a detectable frequency. It has been clearly demonstrated that unlinked product circles are not detected using pLS138 as the substrate in a resolution reaction (see figures 5.16 and 5.18). This suggests that catenation is an intrinsic property of resolvase and may be a direct consequence of the mechanism used to seek *res* sites. Free reaction products have been detected using resolvase prep.1 after long (more than 3hr) incubation times; this is presumably due to the high degree of non-specific nicking activity associated with this preparation.

During the resolution reaction there appears to be no significant loss of superhelicity; the products retain 70-80% superhelicity. This suggests that the reaction is very conservative. There may be a slight decrease in superhelicity during the reaction as some strand rotation is required to bring new partner strands into alignment for the ligation step (see figure 5.22). This may result in the loss of one superhelical twist, effectively reducing the linking number by one. This is difficult to measure as the

substrate to product conversion involves resolution of one plasmid into two linked plasmids. One approach to this problem, currently being investigated by Nash using Int (pers.comm.), is to look at recombination between inverted repeats. For this reaction the substrate and product are both plasmids of the same size. Theoretically it should be possible to observe a change in linking number by reacting a purified topoisomer with the recombination protein and analyzing the products by gel electrophoresis. In practise this has proved more difficult as Int appears always to knot the intervening DNA when using a s/c substrate. As the reaction on o/c inverted substrates is barely detectable, Nash has tried reducing the length of DNA between *att* sites to decrease the possibility of random "flopping" of DNA over itself which may be the cause of knotting. However even when the length of DNA is reduced to 200bp, simple trefoil knots are always formed. This suggests that knotting may be a consequence of the reaction mechanism. By comparing the linking number of purified,artificially constructed,knotted substrates with the reaction products from Int-mediated recombination between inverted repeats on specific substrates, Nash has shown that there is a change of + 2 in the linking number following recombination ie,a decrease in negative superhelicity (pers.comm.). We have been unable to reproduce similar experiments with resolvase as there is no detectable recombination between inverted *res* sites *in vitro* and is very low *in vivo* (Chaing and Clowes,1979; Dyson, pers.comm.)

There is some evidence that resolvase exhibits topoisomerase activity, reflected in the ability of most preparations to produce a ladder of topoisomeric forms. This property could conceivably be due to minor protein contaminants or, if due to resolvase

itself, a modification of the protein's specificity. It may be of significance that the protein preparation showing highest topoisomerase activity was left in 7M urea for several days during it s purification; this may have affected the specificity of the protein in some way. When analysed by SDS-PAGE this preparation appears to be about 95% pure. One of the puzzling features of the topoisomerase activity is the observation that it is not specific to molecules containing *res* sites; again this could be due to altered specificity of the protein.

It has recently been shown that both type I and type II topoisomerases can catenate and decatenate DNA circles (Liu *et al.*, 1980; Kreuzer and Cozzarelli,1980; Brown and Cozzarelli,1981; Krasnow and Cozzarelli,1982). Type I enzymes require a nick in one of the participating duplex molecules, whereas type II enzymes can catenate s/c molecules. This reflects their proposed reaction mechanisms which involve strand passage through a single-stranded or double-stranded break in the DNA backbone respectively. One might expect that resolvase could decatenate reaction products if it is indeed a topoisomerase. However as the reaction products are s/c this would require type II topoisomerase activity for decatenation. Although double stranded cleavage/ligation is thought to occur during resolution this does not necessarily mean that the topoisomerase activity is also of this nature (Reed and Grindley, 1981). It has been demonstrated that catenation/decatenation reactions rely on crucial ionic conditions, and the polyamine, spermidine, is usually required (Krasnow and Cozzarelli,1982) Cozzarelli has recently shown that resolvase exhibits type I topoisomerase activity (pers.comm.);if this observation is correct it would adequately explain why s/c catenated products are not

decatenated. The proposed reaction mechanism for topoisomerases involves single or double-stranded DNA cleavage, conservation of the phosphodiester bond energy by covalent attachment of the exposed 5' terminus to a tyrosine residue on the protein, then ligation of the DNA strands following passage of an unbroken strand(s) through the DNA "gate" (Morrison *et al.*, 1980; Tse *et al.*, 1980). Enzymes which mediate a double strand breakage/reunion event have the potential to act as topoisomerases; likewise, topoisomerases have the potential to act as recombination enzymes. In fact there is evidence that DNA gyrase may be implicated in illegitimate recombination events (Ikeda *et al.*, 1981).

It has been difficult to determine whether resolvase is a catalytic or stoichiometric protein. Nash and Reed have concluded that Int and $\gamma\delta$ resolvase are stoichiometric based on the assumption that a high protein:DNA ratio is indicative of a non-catalytic protein (Nash and Robertson, 1981; Reed, 1981b). The results presented in section 6. suggest that addition of competitor DNA lacking *res* sites effectively titrates out resolvase; a competitive molecule with one *res* site is even more efficient and can titrate out all of the available resolvase (see figures 5.13 and 5.14). This suggests that resolvase binds specifically and non-specifically to any DNA which is added to a reaction mix. Once bound to a DNA molecule it remains associated with it either tightly at *res* (normal repression conditions), or more loosely when bound non-specifically (while searching for a *res* site?). It seems fairly unlikely that the active site of resolvase would be irreversibly altered by resolution, and one would therefore expect the protein to be catalytic. If, however, following resolution the protein remains tightly bound at *res* and does not readily dissociate, then it would not be available

to further substrate molecules. This is reflected in the high protein:DNA ratio observed. Although resolvase has been shown to associate with any DNA molecule, whether it contains *res* sites or not, one would expect the protein to bind specifically to *res* sites to mediate resolution and repression. To investigate the specific binding properties of resolvase in detail I decided to use the purified protein with small *res*-containing DNA fragments in a series of DNA footprinting experiments.

CHAPTER 6

IDENTIFICATION OF THE DNA BINDING SITES FOR RESOLVASE

Introduction

The specific binding of proteins to regions of DNA may be analyzed by random cleavage with non-specific nucleases. This technique has been used extensively to probe nucleosome structure, the *lac* repressor binding site and DNA-DNA gyrase interactions. Digestion of chromatin with micrococcal nuclease yields fragments of a regular size, which appear to be multiples of about 200bp. Continued digestion results in fragments of 146bp; these are known as core particles and are resistant to further digestion by micrococcal nuclease (Felsenfeld,1978). Close examination of the core particles using other nucleases, such as DNase I, has revealed that the DNA within the core is not completely protected from nuclease attack. For each nuclease used,a characteristic ladder pattern of fragments is produced, with each band spaced at roughly 10bp intervals, equivalent to the distance between each turn of the helix (Noll,1974). Fine structure analysis using DNase I on DNA fragments of defined length has shown that the sensitive sites are staggered by 2bp between complementary strands (Lutter,1977; 1979). With DNase II and *Staphylococcal* nuclease the staggers produced are 4bp and 2bp respectively (Sollner-Webb and Felsenfeld, 1977; Sollner-Webb *et al.*,1978). Liu and Wang (1978) have shown that a similar cleavage pattern is produced when DNA is bound to a solid surface, such as calcium phosphate (see figure 6.1).

These results suggested that the nucleosomal DNA is bound to the histone proteins such that only the outermost part of the helix is accessible to nuclease attack (Lutter,1977). This has led to the formulation of a model in which 146bp of DNA are wrapped around the outside of a central histone core (figure 6.2).

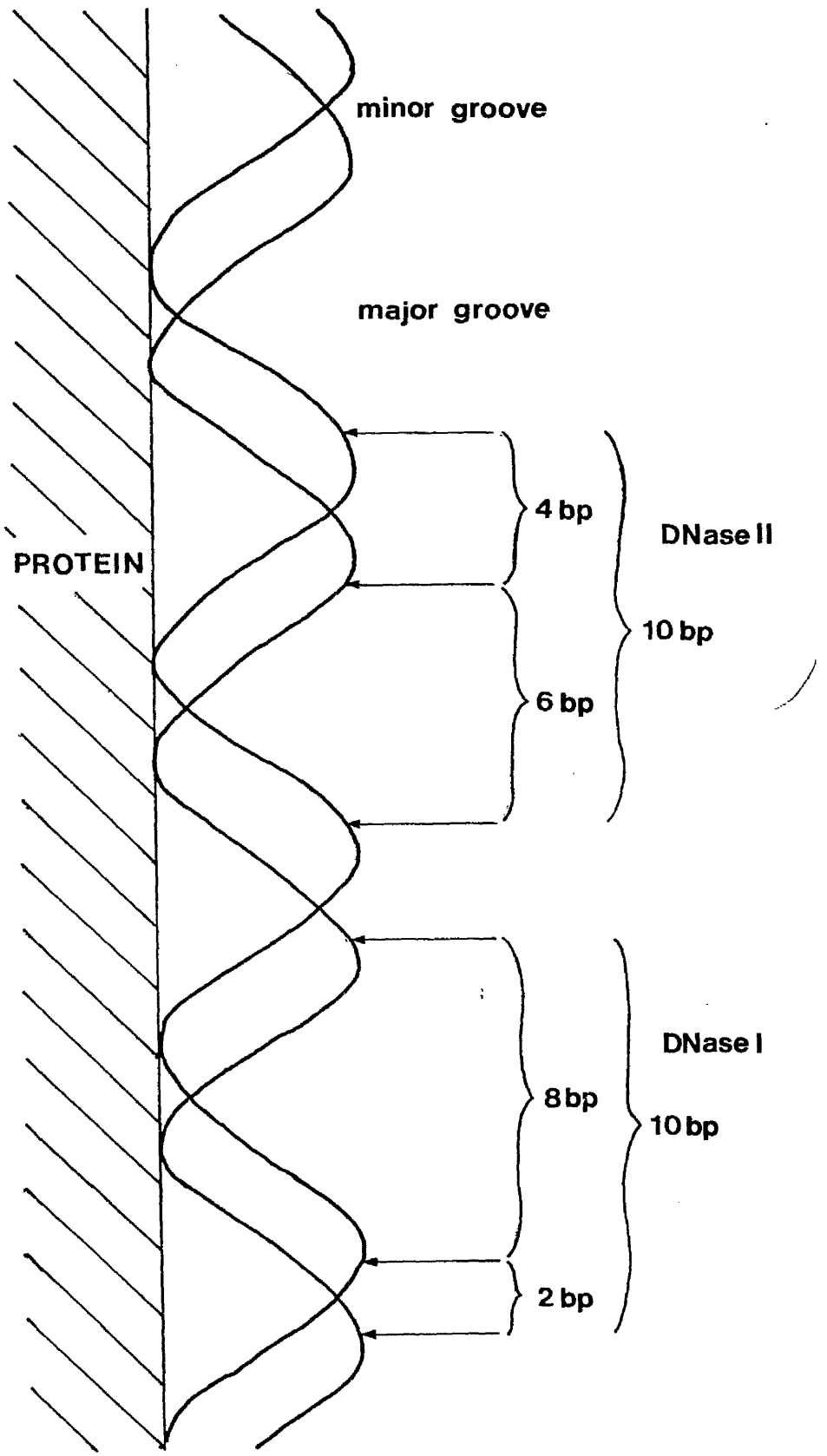

minor groove
major groove
PROTEIN
4 bp
DNase II
10 bp
6 bp
DNase I
8 bp
10 bp
2 bp

FIGURE 6.1 A model to demonstrate the nuclease cleavage sites on B form double helix bound to a solid surface

It has been proposed that nucleases have a limited angle of approach to the DNA helix, when bound on one side to a solid surface such as protein or calcium phosphate (Lutter,1977; Liu and Wang,1978). The cleavage sites for DNase I and DNase II are staggered by 2bp and 4bp, respectively, on complementary strands. The distance of 10bp between each turn of the helix (as shown in the diagram) is based on measurements of fragments produced when core particles are cleaved with nucleases; this value is less than observed when DNA is bound to a solid plane, 10.6bp (Liu and Wang,1978; Rhodes and Klug,1980)

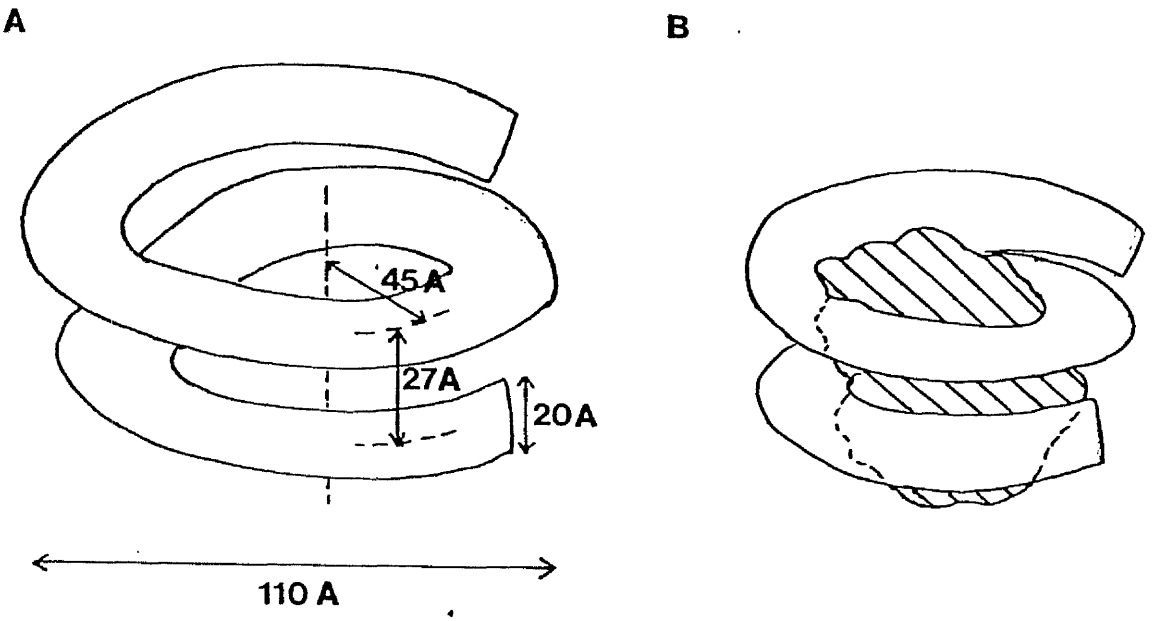


FIGURE 6.2 Structure of the nucleosome core particle

A. This diagram illustrates the dimensions of the DNA helix coil in the core particle. The overall measurements (in angstrom units) are based on electron microscopy, X-ray crystallography and nuclease sensitivity data.

B. Model of how the DNA coil is wrapped around the histone core.

The histone core is comprised of four dimeric subunits: $H2A_2$, $H2B_2$,$H3_2$, and $H4_2$; H1 is thought to be involved in separating individual nucleosome particles (Kornberg and Klug,1981). On the assumption that H2-H4 form compact globular proteins, then the core particles should have a diameter of about 80 angstroms. With DNA wrapped around the outside of the histones the diameter of each core particle should be about 110 angstroms (see figure 6.2). These predictions comply with observed data from electron micrographs of chromatin and X-ray crystallography.

The technique of DNA "footprinting" to define specific sequences which interact with known regulatory or DNA binding proteins was first published by Galas and Schmidt in 1978. The idea behind the experiments is very simple: a protein which binds tightly to DNA then protects that region of DNA from subsequent attack by nucleases. As well as identifying the specific sequences to which proteins bind, this technique may also indicate some of the topological features of the interaction; this has been clearly shown for nucleosome structure and, more recently, for the DNA-DNA gyrase interaction. When present in limiting amounts DNA gyrase binds at two specific regions on plasmid ColE1, named sites a and b. Footprinting of fragments containing these sites reveals a region of 145-155bp of DNA protected from DNase I cleavage. This protected region has been divided into a central core region at which the breakage/reunion reaction mediated by DNA gyrase takes place, and two flanking arms. The core sequence is completely protected from DNase I, but the flanking arms show a periodicity in nuclease sensitive sites, similar to that observed for nucleosomes (Liu and Wang,1978; Kirkegaard and Wang,1981; Morrison and Cozzarelli, 1981). This data suggests that the flanking sequences may be

wrapped around the outside of the DNA gyrase holoenzyme. It has been proposed that the DNA is always wrapped in the same direction, inducing polarity into the interaction (Liu and Wang,1978; Morrison and Cozzarelli,1981).

The precise extent of the DNA sequences required for resolvase mediated site-specific recombination has not been determined. A 357bp *Sau*3AI fragment spanning the *tnp*A-*tnp*R intercistronic region is known to contain sufficient sequence for resolution (Reed,1981a, Chapter 5). During the purification of resolvase the protein binds non-specifically to all DNA, suggesting that at high concentrations, represented in the overproducer strain, it binds to any DNA. However, under normal *in vivo* conditions, and in the *in vitro* system, the amounts of resolvase present are far less. Does resolvase bind specifically within the *res* region?

To analyze the specific sequences, if any, that resolvase binds to, the purified protein has been used in a series of DNA footprinting experiments. The experiments described in this chapter explain the strategies used to isolate and footprint labelled *res*-containing fragments. This data, in combination with results obtained from *in vitro* and *in vivo* resolution , has prompted the formulation of a model for resolvase mediated site-specific recombination.

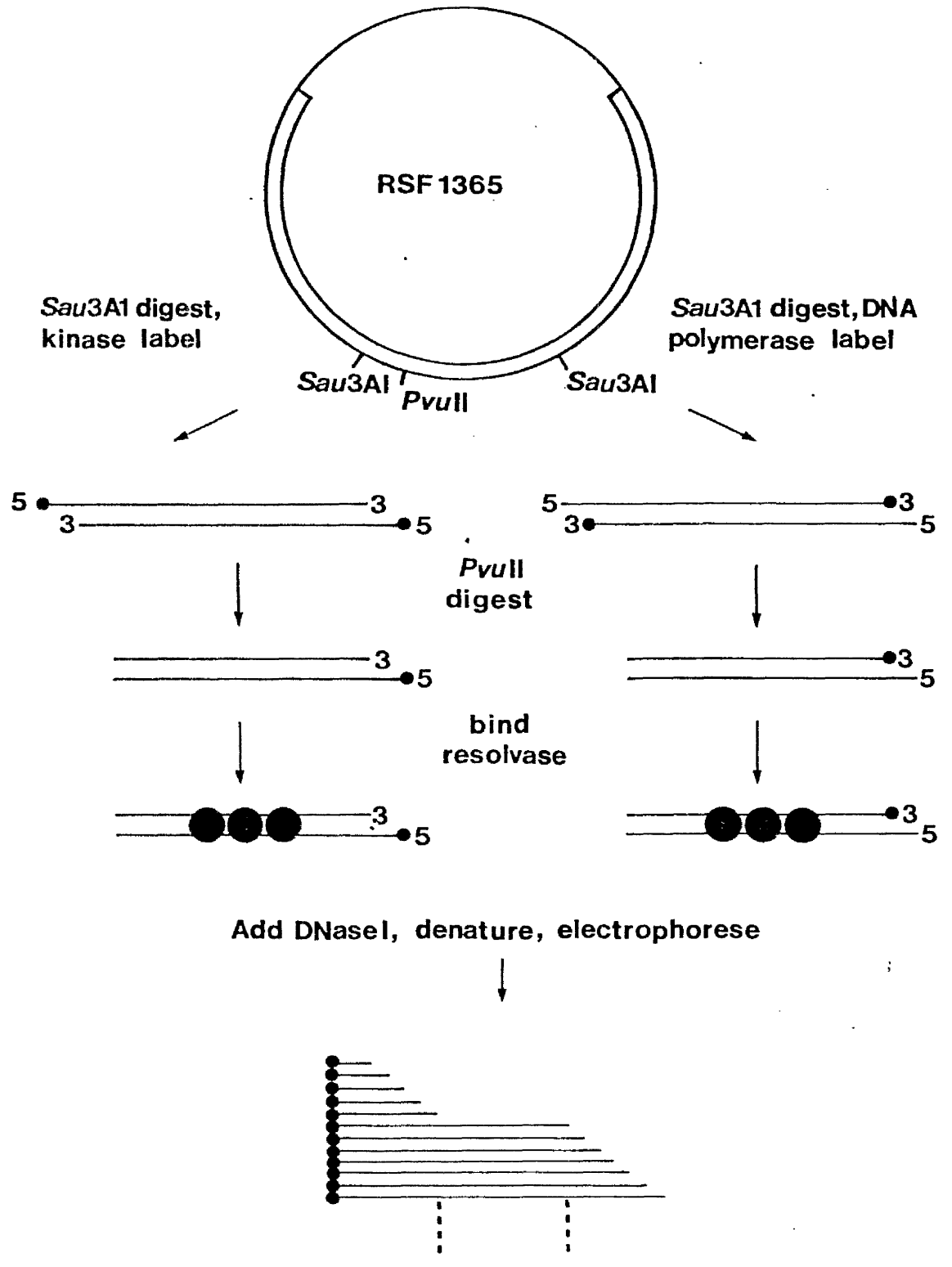


FIGURE 6.3 Strategy for footprinting the *res*-containing *Sau*3AI restriction fragment.

Results

1. Strategy for isolation of *res* containing fragments

The basic substrate for both DNA footprinting and sequencing techniques is a DNA fragment of defined length, usually less than 500bp, with the terminus of one strand labelled with ^{32}P. This may be achieved by labelling both 3' or 5' ends of a unique fragment, then separating the strands by denaturation/electrophoresis, or by cutting with a second restriction enzyme which cleaves the fragment to produce two fragments each with one labelled terminus. Provided that the second enzyme cuts asymmetrically within the fragment, these can be separated by electrophoresis followed by purification of the terminally labelled fragment from the gel material (Maxam and Gilbert,1977;1980). If the unique, terminally labelled strand is cleaved randomly along it s length, either induced by base-specific chemicals or by a non-specific nuclease, then a series of fragments of varying length should be produced. Under conditions which cut each strand once the series of fragments produced should have one common (labelled) end and the other of varying length. As long as the cleavage is random along it's length the nest of fragments produced should differ by one bp and when resolved on a sequencing gel should produce a ladder of bands, each rung differing from it s neighbours by one nucleotide. When a protein binds to a specific DNA sequence that region is protected from subsequent nuclease attack. If the protein bound fragment is subjected to nuclease cleavage prior to electrophoresis a gap is observed in the ladder of fragments where the nuclease has been unable to cleave. This gap identifies the region of DNA bound by the protein (figure 6.3).

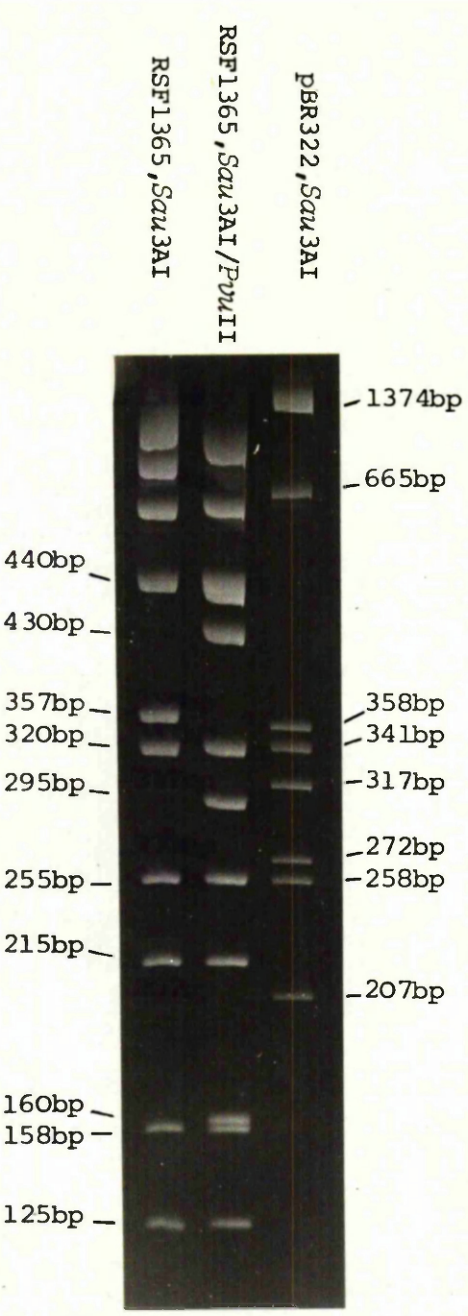


FIGURE 6.4 Restriction analysis of RSF1365

Plasmid RSF1365 was digested with *Sau*3AI or *Sau*3AI/*Pvu*II. The sizes of fragments produced were determined by comparison with *Sau*3AI digested pBR322 DNA, for which the fragment sizes are known (Sutcliffe,1978).

5% polyacrylamide gel run at 30mA.

The complete nucleotide sequence of Tn*3* has been determined (Heffron *et al.*, 1979), and from this a detailed restriction map predicted. The site at which recombination takes place has been defined within a 357bp *Sau*3AI fragment (see figure 4.1; Reed, 1981a) This fragment contains sufficient sequence for resolution *in vivo* and *in vitro*, and also contains the promoters for *tnp*A and *tnp*R (Chou *et al.*, 1979b). There is a single *Pvu*II site within this fragment located 62bp from the lefthand end (on the Tn*3* map). As *Sau*3AI is a 4bp recognition enzyme it has many sites within Tn*3*. To isolate the 357bp fragment a small deletion derivative of Tn*3*, RSF1365, was used (map shown in figure 3.3); this provides a series of fragments on *Sau*3AI digestion which may be clearly distinguished by electrophoresis. When doubly digested with *Sau*3AI and *Pvu*II, the 357bp *res*-containing fragment is cleaved to produce two fragments of 295 and 62bp; there are also bands of 430 and 160bp produced by *Pvu*II as there is a second *Pvu*II site located in the *tnp*A gene (figure 6.4). The fragments produced from RSF1365 were sized in reference to a *Sau*3AI digest of pBR322, for which the sizes are known (Sutcliffe, 1978).

2. Determination of resolvase binding sites on the 3' labelled strand

The restriction enzyme *Sau*3AI produces a 4bp staggered cut with a 5' extension; the resulting recessed 3' hydroxyl provides a suitable substrate for DNA replication by DNA polymerase I:

```
- - - - G A T C - - - -   3'
        └─────┘
- - - - C T A G - - - -   5'
```

Labelling the termini of the 357bp *Sau*3AI fragment was achieved by addition of $[\alpha^{32}P]$-GTP and DNA polymerase I (Klenow fragment) to

a total RSF1365 digest. The labelled DNA was digested with *Pvu*II followed by electrophoresis through a 5% acrylamide gel to isolate the 292bp end-labelled fragment. The DNA band was excised from the gel and the fragment eluted by the crush-soak-precipitate method. The purified fragment was subsequently used for footprinting and sequencing reactions. For each reaction sufficient labelled fragment was added to produce about 50 cps on the minimonitor; this amount of radioactivity was sufficient for an exposure time of only 16hr (under optimal conditions) to visualize bands on an autoradiograph.

The enzyme DNase I works under most conditions, whereas resolvase has specified reaction conditions. For this reason all DNase I digestions, plus or minus resolvase, were performed using conditions optimal for *in vitro* resolution. To determine the enzyme concentration required to give a ladder of bands representing the entire 295bp fragment, DNase I solutions of varying concentration were used to digest the labelled fragment. To the standard reaction conditions (20mM Tris-HCl,pH 8.0, 10mM $MgCl_2$, 1mM DTT, 50mM NaCl) were added 0.4µg of λ DNA, as competitor for DNase I, and about 30ng of labelled DNA fragment in a total volume of 50µl. Reactions were incubated for 20min at 30°C before addition of varying amounts of DNase I (10-100ng). The DNase I digests were performed for one min, then stopped by the addition of 12.5µl of 3M NH_4Ac/0.25M EDTA. DNA was precipitated from the solution by addition of two volumes of ethanol, left at -20°C for 20min then spun in the Eppendorf microfuge to pellet the DNA. The pellet was washed twice, dried, then resuspended in 6µl of formamide containing marker dyes. This solution was heated to 90°C to separate the DNA strands, then quick chilled on ice prior to gel loading. 3µl of each sample were loaded onto an 8% polyacrylamide/8.3M urea sequencing gel to analyze the

ladder of DNA fragments produced; the gel was run until the bromophenol blue dye marker was just above the bottom of the gel. The gel was exposed to preflashed X-OMAT X-ray film, in the presence of an intensifying screen, at -70°C, for 18hr. This experiment indicated the optimal DNase I concentration required to produce a ladder of bands representative of the entire 295bp fragment; for subsequent experiments 80ng of DNase I were used.

Using the conditions described above, varying amounts of resolvase were added to reactions prior to DNase I digestion. Reactions were analyzed by electrophoresis through 8% polyacrylamide sequencing gels, then autoradiographed for 19hr (figure 6.5). Clearly resolvase binds to specific sequences, indicated by the gaps in the ladder of fragments produced by DNase I. The entire protected region extends over approximately 120bp; within this region there is a short, unprotected sequence. The two separate regions have been arbitrarily designated site I and site II. It is noticeable that within the long protected site that there are sites of enhanced sensitivity to DNase I, some of which are not recognized in the absence of resolvase. As the concentration of resolvase in reactions is increased there appears to be an intermediate point which shows only partial protection (figure 6.5, track C). With 0.2µg of resolvase only the central part of the region, at the beginning of site II, appears to be protected; the flanking sequences are protected when the resolvase concentration is increased to 0.3µg. This suggests that the central sequence may be the most important during the initial recognition by the protein and is the first site to be protein bound.

To assign nucleotide positions to the sequences protected

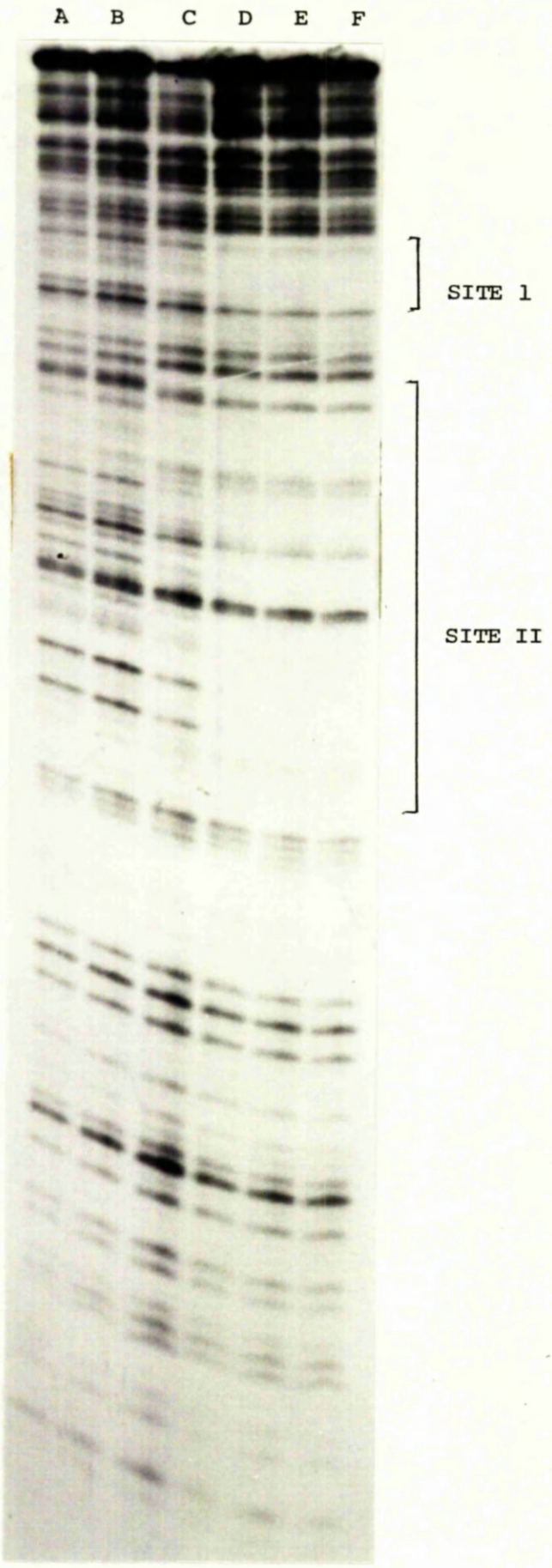

A
B
C
D
E
F
SITE 1
SITE II

FIGURE 6.5 Determination of resolvase binding to the 3' end-labelled *Sau*3AI fragment

A no resolvase

B + 0.1µg resolvase

C + 0.2µg resolvase

D + 0.3µg resolvase

E + 0.4µg resolvase

F + 0.5µg resolvase

Resolvase binds at two distinct sites within this fragment, these are separated by about 20bp of unprotected DNA. For all DNA protection experiments 80ng of DNase I were added after incubation of DNA with the specified resolvase concentration (resolvase preparation 1 was used throughout).

8% polyacrylamide/8.3M urea gel run at 40W.

by resolvase it was necessary to repeat these experiments and run them alongside chemical cleavage reactions of the labelled fragment. A purine specific cleavage reaction was performed on the labelled strand by the Maxam and Gilbert sequencing technique (1977;1980). As the complete sequence of Tn*3* is known, it was considered unnecessary to use all of the base specific cleavage reactions; the purine specific cleavage track provides sufficient information to assign sequence positions. The addition of a sequencing track has allowed determination of the DNA sequences of the sites protected by resolvase (figures 6.6 and 6.8). The region at which the crossover occurs,and for γδ resolvase site-specific cleavage,lies within site I (Kostriken *et al.*,1981; Reed,1981a; Reed and Grindley, 1981); this extends from coordinates 3089-3118 on the Tn*3* map. This site also contains putative Pribnow boxes for both *tnp*A and *tnp*R, and RNA initiation sequences. A 20bp unprotected region separates site I from the major protected region, site II. Most of the sequence from 3139-3208 is protected, but does contain DNase I sensitive sites, some of which are greatly enhanced. There is also a resolvase-dependent DNase I cleavage point 4bp into the *tnp*R gene, this region is outside site II; the significance of this is unclear. It is difficult to define exact site boundaries, particularly of site I, as resolution is poor at the top of the gel. Also it can be seen that DNase I does not cleave completely randomly, therefore there are some sites in the sequence which are resistant to DNase I in the presence or absence of resolvase. Partial protection from DNase I when using 0.2μg of resolvase was not observed in this experiment.

A summary of gel data, which shows the complete nucleotide sequence of the intercistronic region, indicating resolvase binding

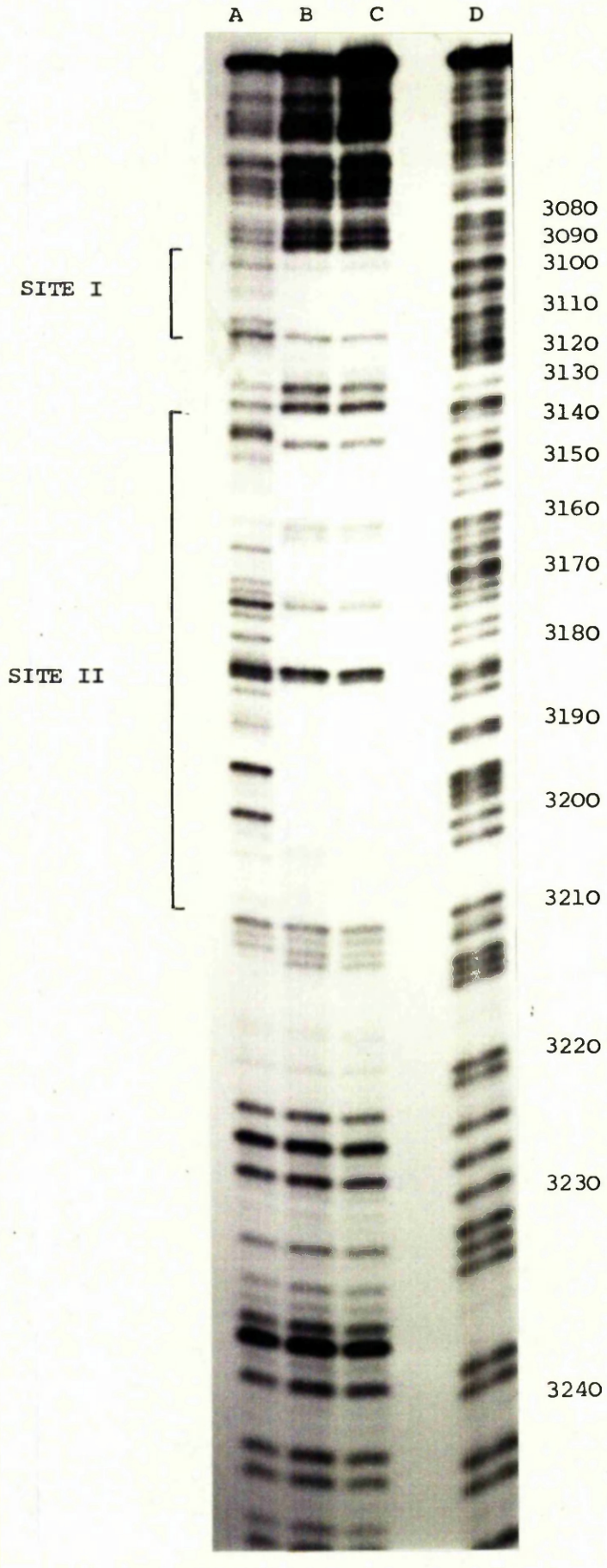

A
B
C
D
SITE I
SITE II
3080
3090
3100
3110
3120
3130
3140
3150
3160
3170
3180
3190
3200
3210
3220
3230
3240

FIGURE 6.6 Determination of resolvase binding sites on the 3' labelled *Sau*3AI fragment - assignment of nucleotide positions

A no resolvase

B + 0.2µg resolvase

C + 0.4µg resolvase

D AG sequence track

Site I extends over 29bp, from position 3089 to 3117 on the Tn*3* map. Site II is 67bp long, extending from position 3138 to 3205. The exact endpoints of these sites cannot be precisely defined as resolution at the top of the gel is poor and some sites are insensitive to DNase I cleavage in the presence or absence of resolvase.

8% polyacrylamide/8.3M urea gel run at 40W.

and DNase I enhanced cleavage sites is presented (figure 6.8). There is no clear periodicity of DNase I cleavage sites within the protected regions as observed for nucleosome structure, but these may still have some significance in terms of the topology of the resolvase-DNA interaction. As site II contains DNase I cleavage sites it suggests that there may be a looser association of resolvase with the DNA sequence compared with site I, which appears to be completely protected. The observed patterns of sensitive sites in the presence and absence of resolvase are different and thus may not be due simply to lack of protection on some strands. This change in specificity suggests a conformational change to the DNA on binding resolvase. If the enzyme can only make a limited angle of approach then some potential cutting sites may not be orientated correctly, when bound to resolvase, and hence not cleaved. It is possible that parts of site II are wrapped around the outside of resolvase exposing them to DNase I cleavage on the parts of the helix furthest from the protein. If the DNase I sensitive sites are a result of DNA wrapping around the outside of the protein then it would be expected that the complementary strand should have sensitve sites staggered by 2-3bp (Lutter,1977;Sollner-Webb *et al.*,1978). To confirm these results resolvase binding to the 5' labelled strand was investigated.

3. Determination of resolvase binding sites on the 5' end-labelled strand

Footprinting experiments, similar to those described above, were repeated using the 357bp *Sau*3AI *res*-containing fragment labelled at the 5' terminus. The fragment was initially purified from an RSF1365 *Sau*3AI digest by electrophoresis through 5% acrylamide followed by elution of DNA from the excised gel slice by the

electrophoresis method. The purified fragment was treated with bacterial alkaline phosphatase, phenol/ether extracted, then precipitated with 65% ethanol without the addition of carrier tRNA, which competitively inhibits the kinase reaction. The phosphatased fragment was labelled with $[\gamma\text{-}^{32}P]$ ATP by T4 polynucleotide kinase, then treated as previously described to obtain a 295bp fragment with one labelled end. This fragment was used, under similar conditions to those employed for footprinting the 3' labelled strand, to determine the sequences protected by resolvase (figure 6.7). The representative sequencing track run alongside the footprints is again from a purine-specific chemical cleavage reaction. As expected the sites protected on the 5' labelled strand complement the data obtained from the opposite strand. The crossover point is located within site I and is completely protected from DNase I. Within the long protected region, site II, there are again sites of DNase I cleavage; most of these are not enhanced. If these were due to a background of unbound strands then one would expect that all bands present in the minus resolvase reaction to be present at reduced intensity; as this is not observed one may deduce that they are a consequence of resolvase binding. Some of the sites are enhanced and staggered by 3-4bp compared with those on the 3' end-labelled strand as indicated (see figure 6.8). There appears to be no sequence protection beyond site I into the *tnp*A gene, though a small protected site may be undetected due to low resolution at the top of the gel. This contrasts with the *in vivo* analysis of Kostriken *et al.* (1981) which suggests that there is a site within the *tnp*A gene required for resolution.

The results from these experiments indicated that resolvase binds specifically within the intercistronic region, protecting

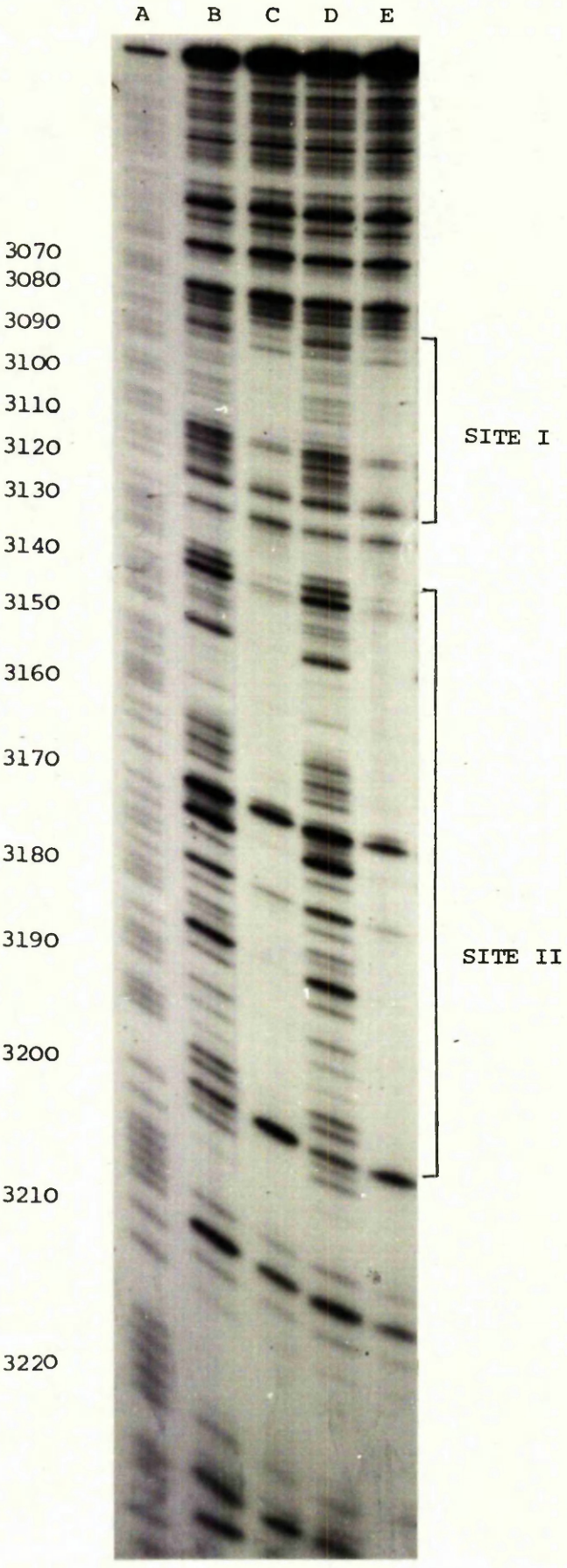

A
B
C
D
E
3070
3080
3090
3100
3110
3120
3130
3140
3150
3160
3170
3180
3190
3200
3210
3220
SITE I
SITE II

FIGURE 6.7 Determination of resolvase binding sites on the 5' labelled *Sau*3AI fragment

A AG sequence track

B no resolvase

C + 0.2µg resolvase

D no resolvase

E + 0.4µg resolvase

Site I extends over 30bp from nucleotide position 3090 to 3119 on the Tn*3* map; site II extends over 63bp from 3141 to 3204.

8% polyacrylamide/8.3 urea gel run at 40W.

FIGURE 6.8 Location of the resolvase binding sites within the *tnp*A-*tnp*R intercistronic region of Tn*3*

Solid bars indicate regions protected by resolvase from DNase I claavage; O indicate sites not cleaved by DNase I in the presence or absence of resolvase; ↓ represent sites of enhanced DNase I sensitivity in the presence of resolvase; ▼ indicate DNase I cleavage points within the protected region. The large arrowheads represent the presumptive crossover point in analogy with γδ; the presumptive Pribnow box sequences for both *tnp*A and *tnp*R are shown, as are the putative transcription start points (Rosenberg and Court,1979).

the putative Pribnow box/resolvase cleavage site from DNaseI attack; also sequences to the right of this site up to the beginning of the *tnp*R gene. The significance of resolvase binding within site II is unclear; it could be necessary for correct alignment of strands during recombination or may be important for repression of transcription. It is of some significance that a *tnp*A derepressed mutation has been isolated, which is caused by a single GC to AT base change at position 3140, the beginning of site II (Chou *et al.*, 1979b). This mutation, designated *cis*10, has unfortunately not been tested for resolution.

4. Does DNA conformation affect the resolvase binding pattern?

Although the footprinting experiments provide data on the specific interaction between resolvase and a linear DNA fragment there are always doubts that this may be an artifactual situation. Previous data has shown that resolution is absolutely dependent on a supercoiled substrate carrying two copies of *res* in direct orientation; repression requires only one *res* site. One can imagine that several of the steps involved in resolvase-mediated recombination could require an energy source, provided by the negative superhelicity of the substrate:

(i) Association of resolvase with DNA.

(ii) Specific binding to the target site; the high AT content of the *res* region may lead to "melting" in a s/c molecule, which in turn may facilitate resolvase binding within this site.

(iii) Alignment of *res* sites; if this is an active process it may require energy input.

(iv) Cleavage/ligation of strands to new partners.

The possibility remains that the resolvase/DNA interaction

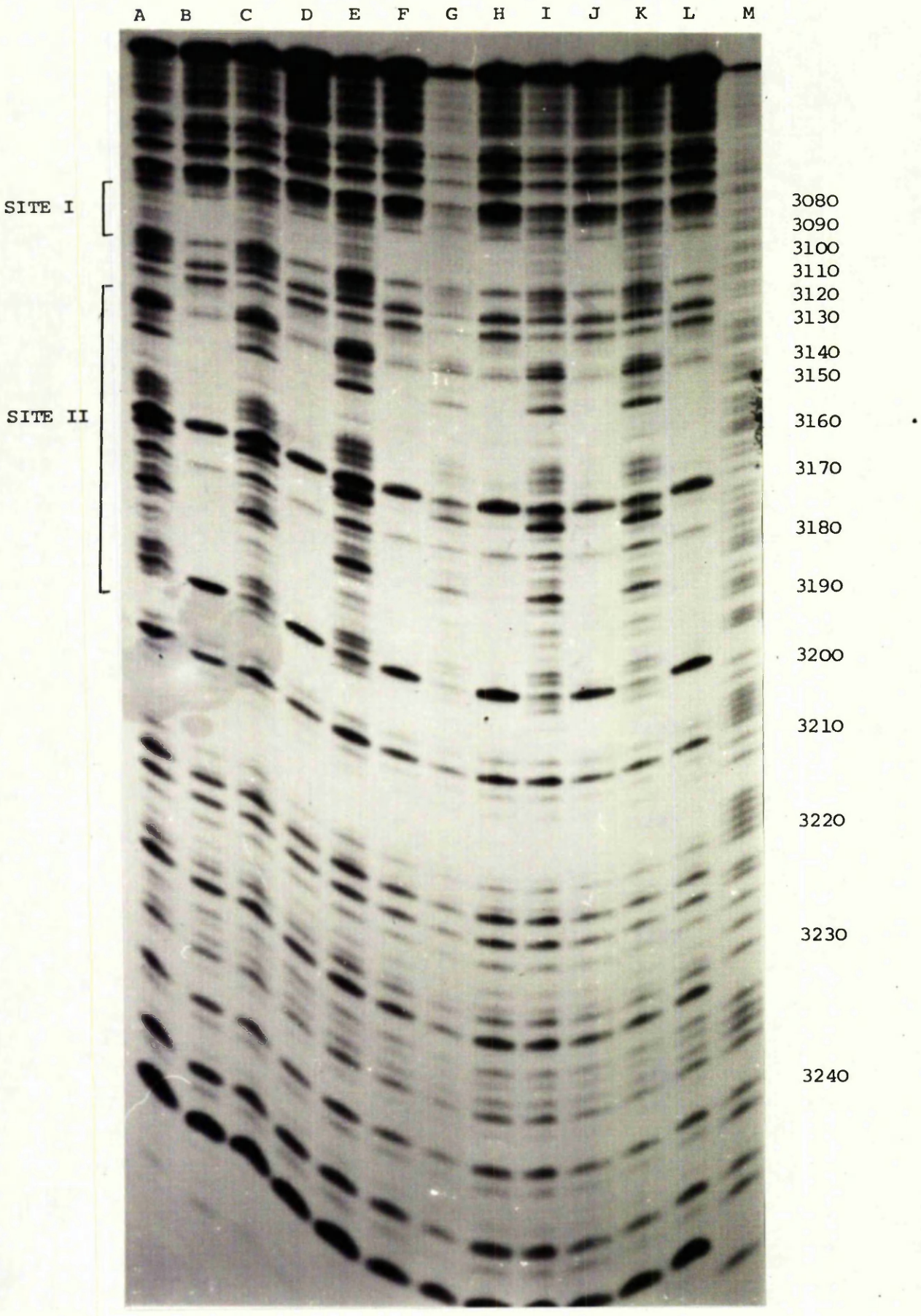

A
B
C
D
E
F
G
H
I
J
K
L
M
SITE I
SITE II
3080
3090
3100
3110
3120
3130
3140
3150
3160
3170
3180
3190
3200
3210
3220
3230
3240

FIGURE 6.9 Addition of increasing concentrations of EtBr to the 5' end-labelled *Sau*3AI fragment - affect on resolvase binding

A no EtBr, no resolvase

B no EtBr, + resolvase

C + 0.05µg EtBr, no resolvase

D + 0.05µg EtBr, + resolvase

E + 0.1µg EtBr, no resolvase

F + 0.1µg EtBr, + resolvase

G + 0.2µg EtBr, no resolvase

H + 0.2µg EtBr, + resolvase

I + 0.5µg EtBr, no resolvase

J + 0.5µg EtBr, + resolvase

K + 1.0µg EtBr, no resolvase

L + 1.0µg EtBr, + resolvase

M AG sequence track

The addition of EtBr from 0.1µg/ml to 20µg/ml does not appear to affect the specific binding pattern of resolvase to this fragment.

8% polyacrylamide/8.3M urea gel run at 40W.

may be different on a s/c molecule and the topology may also change during resolution, when the two *res* sites are brought into close contact. There is no easy way of footprinting a s/c molecule. It may, however, be possible to mimic the conformational change in DNA structure induced by superhelicity by the addition of an intercalating agent such as ethidium bromide. Ethidium bromide has been extensively used as an unwinding agent in determining the number of superhelical turns in covalently closed molecules (Keller,1975; Deleys and Jackson,1976; Espejo and Lebowitz,1976). When ethidium bromide intercalates the unwinding angle of the helix is changed by 26° (Wang,1974). If the ethidium bromide concentration is gradually increased from 0.01µg/ml to about 0.15µg/ml a s/c plasmid, such as ColE1, is unwound until it's mobility on electrophoretic gels is similar to that observed for relaxed molecules. Increasing the concentration of ethidium bromide further leads to greater unwinding; this results in the introduction of positive superhelical turns.

Based on the concentration of ethidium bromide used in these experiments,and on recommendations from D.Lilley (pers.comm.), I have tried to mimic the degree of unwinding induced by superhelicity by addition of ethidium bromide to footprinting reactions. The concentrations of ethidium bromide employed ranged from 0.05µg/ml to 1µg/ml; this was added to reaction mixtures prior to the addition of resolvase. As can be seen from the DNA footprint there appears to be no difference in the pattern of resolvase binding to the *res* region in the presence of ethidium bromide (figure 6.9). Assuming that ethidium bromide binds to linear and closed circular DNA molecules in a similar manner, one may conclude that the change in DNA conformation induced by superhelicity does not change the resolvase binding pattern.

5. Analysis of resolvase binding to the righthand end of Tn*3*

Examination of the DNA sequence within the *tnp*A-*tnp*R intercistronic region reveals a very high AT content, particularly within site I. The 19bp crossover region, defined by *in vivo* analysis, contains 16/19 AT pairs (Kostriken *et al.*, 1981; Reed, 1981a). Does resolvase bind to other AT rich sequences?

To test this possibility a different fragment was isolated from Tn*3* , encompassing the righthand end of the element. This fragment has a high AT content and, more significantly, contains a sequence highly homologous to the *res* site. (see figure 6.10). The plasmid pLS134, constructed for *in vitro* resolution experiments, was used as a source of this fragment. The *res*-like sequence may be isolated on a *Cla*I/*Bgl*I fragment (figure 6.11); this was predicted by analysis of the Tn*3* nucleotide sequence. A *Cla*I digestion of pLS134 was labelled with $[\alpha^{32}P]$-GTP using DNA polymerase I, then cut with *Bgl*I to produce a 360bp end-labelled fragment. This was purified by electrophoresis through 5% polyacrylamide followed by elution of the DNA from the gel slice by the crush-soak-precipitate method. Footprinting of the fragment indicated no protection by resolvase over the entire region, as detected by these methods (figure 6.12); as no protection was apparent a DNA sequence track was not included. These results confirm that resolvase binds specifically to the *res* region under the conditions employed.

Discussion

The technique of DNA footprinting has proved a useful tool to probe the interactions between DNA and specific DNA binding proteins. It has aided the identification of sequences involved in regulation

Tn*3* *res*

A T A T T A T A A A T T A T

Tn*3* nuc.4902-4889

T T A T C A A A A A G G A T

Hin inverted repeat

T T A T C A A A A A C C T T

Gin inverted repeat

T T A T C C A A A A C C T C

λ *att* core sequence

T T A T A A A A A A G C T G

FIGURE 6.10 A comparison between the DNA sequences of *res*, the sequence at the righthand end of Tn*3*, Hin inverted repeat, G loop inverted repeat and the λ *att* sequence

The bases which share homology with the *res* site have been underlined.

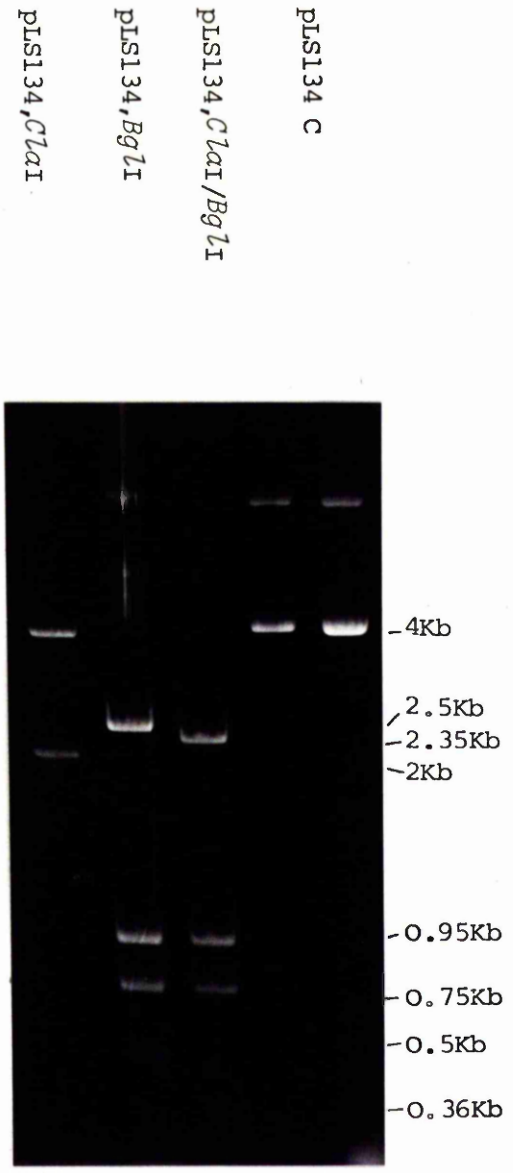


FIGURE 6.11 Restriction analysis of pLS134

Plasmid pLS134 was restricted with *Cla*I, *Bgl*I or *Cla*I/*Bgl*I as indicated. The 500bp *Bgl*I fragment is cleaved by *Cla*I to produce 360 and 135bp fragments. The 360bp fragment contains the terminus of the Ap^r gene, the sequence bearing homology to *res*, and the righthand inverted repeat.

1% agarose gel run at 100V.

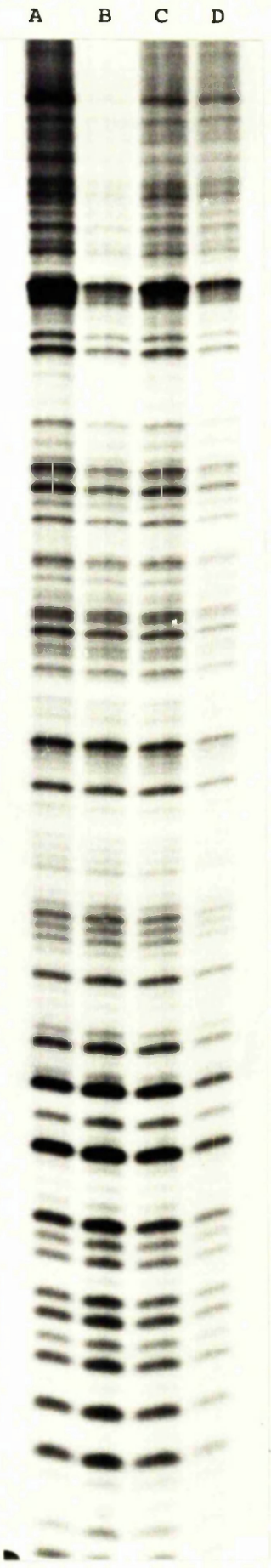

A
B
C
D

FIGURE 6.12 Analysis of resolvase binding to the righthand inverted repeat region of Tn*3*

A no resolvase

B + 0.2µg resolvase

C + 0.4µg resolvase

D + 0.8µg resolvase

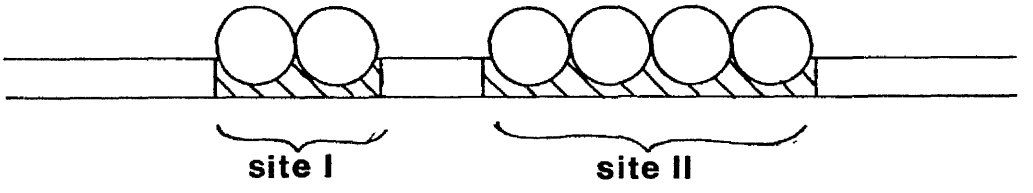
There appears to be no specific binding of resolvase to this region, though at the highest concentration of resolvase used all the bands are slightly faded; this could be due to non-specific binding by the protein. As no specific binding pattern was observed a sequence track was not included in this experiment.

8% polyacrylamide/8.3M urea gel run at 40W.

of gene expression, eg. *lac* operator/repressor, *uvrB* operator/*lexA* repressor (Galos and Schmidt,1978; Sancar *et al.*,1982), and has lent insight into the topography of DNA-protein interactions. The best documented example of this is fine structure analysis of chromatin core particles, which has been revealed by use of non-specific nucleases (Noll,1974; Lutter,1977; Sollner-Webb *et al.*,1978). More recently the interaction between DNA and DNA gyrase has been examined using DNase I (Liu and Wang,1978; Kirkegaard and Wang,1981; Morrison and Cozzarelli,1981). Sequence protection by DNA gyrase indicates that the site at which cleavage takes place is within a 40bp region tightly protected from nuclease attack, whereas the flanking arms, each of about 50bp and also protected, have bases highly susceptible to DNase I cleavage. There is a regular periodicity between DNase I sensitive sites, reminiscent of that observed for nucleosome structure. This has led to the proposal that DNA sequences in the flanking arms of the protected region are wrapped around the outside of the protein holoenzyme. This would introduce "handedness" into the interaction if the DNA always coils around the protein in a particular direction. Negative superhelicity by the sign inversion model requires that DNA gyrase fixes the polarity of the passed DNA relative to the cleaved DNA. If there were no directionality in the interaction then introduction and removal of negative supertwists would occur equally (Brown and Cozzarelli,1979). The footprinting data suggests that DNA is wrapped around the DNA gyrase holoenzyme in a right-handed coil; the polarity of this interaction is supportive evidence for the sign inversion model (Liu and Wang,1978; Morrison and Cozzarelli,1981).

The results in this chapter indicate that resolvase interacts specifically with the *res* region of Tn*3*. Resolvase protects about

A



B

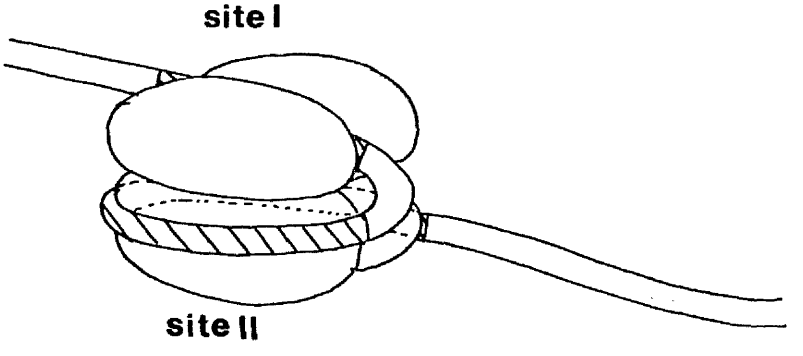


FIGURE 6.13 Models for resolvase binding to the *res* region

A. The simplest model involves attachment of resolvase monomer or dimer units to each of the sites. A model of this kind does not explain the disparity in DNase I enhanced cleavage sites observed over site II.

B. In this model the DNA *res* region is folded around a reolvase protein core structure, in analogy with the proposed nucleosome core particle model, and DNA-gyrase interaction. This model agrees with data obtained from DNase I cleavage, which demands that site I is completely protected from nuclease attack (except the extreme lefthand end), but site II is susceptible to limited cleavage. Sites of enhanced cleavage by DNase I occur throughout site II, but with no regular periodicity.

120bp from DNase I digestion, though within this region there is an unprotected region of 20bp arbitrarily dividing the entire region into two protected sites. Site I contains the proposed recombination crossover site (Kostriken *et al.*,1981; Reed,1981a) and extends over at least 29bp; the sequences to the right of site I constitute the major protected region and extend over at least 64bp. Site II may be important for repression by resolvase as the *cis*10 derepressed transposition mutant is located at the beginning of this site; it would be interesting to analyze resolvase protection to this mutant element. Preliminary data suggests that the sequences contained within the terminal 15bp of site II are required for resolution (Kitts, pers.comm.). The pattern of DNase I cleavage suggests that site I is tightly protected from nuclease attack, whereas site II may be more loosely associated with resolvase. Although the periodicity of DNase I enhanced sensitivity sites is not as clear as that found in nucleosomes, the pattern of cleavage is suggestive of a conformational change induced by resolvase which could involve wrapping of the DNA helix around a resolvase protein complex.

It is difficult to assess details of the DNA-resolvase interaction without knowing how many protein monomer units bind at each *res* site. Gel filtration of the protein suggests an equilibrium in 1M NaCl solution between monomers, dimers and tetramers, with dimers as the predominant form. However, as the NaCl concentration is decreased 20 fold for *in vitro* reactions the equilibrium may well be shifted, perhaps towards higher forms? If each resolvase molecule assumes a globular tertiary structure then the diameter of each monomer unit would be about 4nm. Resolvase binding to DNA in the simplest form may be envisaged as monomers/dimers attached to the linear *res* region (figure 6.13A). Such a model would require

6-7 protein monomers/*res* site, but does not account for the apparent disparity in DNase I cleavage patterns observed between sites I and II. An alternative model based on comparisons with nucleosome structure and the model for DNA-DNA gyrase may be envisaged; this involves folding of the DNA around a protein core structure, represented as a tetramer in the diagram (figure 6.13B). For comparison, the dimensions of the histone core are slightly larger than the predicted size of a resolvase tetramer. The length of DNA wrapped around the histone core is 146bp; the DNA associated with resolvase is about 120bp in length. The second model would agree with data obtained from DNase I cleavage of site II. Wrapping of site II partially on the outside of the protein core would account for the sites of enhanced sensitivity to DNase I within this region. Although these sites are not spaced at regular intervals of 10.6bp as expected for DNA bound to a solid surface (Liu and Wang,1978; Rhodes and Klug,1980), this could be due to conformational changes in the helix induced by attachment to resolvase. If DNase I can only make a limited angle of approach then some potential cutting sites may not be orientated correctly if there are "kinks" in the DNA helix, hence these sites may not be cleaved. The regular site periodicity assumes that DNA is in the B configuration; if this region of DNA is equilibrium between the B conformation and other structures such as Z DNA, then this may also affect the availability of DNase I cleavage sites.

The data presented shows that resolvase binds specifically to the *res* site; results from the previous chapter suggested that resolvase also binds non-specifically to any DNA molecule. How does resolvase find it's specific binding site? One can consider two possible mechanisms by which a DNA binding protein can find it's

target site. The first is a simple diffusion model; the protein diffuses freely throughout the cell and binds tightly when it finds it's specific site. 3D diffusion would be very slow as the concentration of target sites is very low compared with non-specific sites. The evidence provided in the last two chapters suggests that resolvase finds *res* sites much faster than the diffusion model would allow. During the random 3D search the protein would have to test a large fraction of non-specific sites before the target was located.

A protein which has some affinity for DNA, already shown for resolvase, in addition to it's specific binding properties, will almost certainly bind non-specifically at first due to the vast excess of non-specific sites over specific ones. It may be useful to consider the specific association as a two stage process with a non-specific complex as an intermediate step (Berg *et al.*,1981). Once bound non-specifically the protein then has to find it's specific site, though if it remains in the vicinity of DNA then the chances of finding it's target are much greater than by random diffusion . Two models have been proposed for the mechanism of seeking the target site:

(a) Dissociation-reassociation ("hopping")

Non-specific binding brings the protein into close contact with the DNA molecule. If the initial binding is non-specific the protein dissociates, but remains in the vicinity of DNA and immediately reassociates; this could be at the same or a nearby site. The protein continually dissociates and reassociates with DNA until it reaches the target where it binds tightly. This process has been likened to hopping along the DNA molecule in search of the target site, and would presumably require no external energy source (Berg *et al.*, 1981; Winter and von Hippel,1981).

(b) Sliding

The basic assumption for this model is that the protein can slide along the DNA in a 1D random walk while bound. The sliding model relies on the electrostatic binding between protein and DNA. Binding affinity is assumed to be due entirely to the entropy of dilution of the counterions displaced from DNA by binding of the protein. As the protein moves along the DNA molecule it displaces counterions from DNA "ahead" of it and the same number of counterions are replaced when it has vacated that DNA region. If the counterions are not totally"site bound" to phosphates and the positive charges on the protein are not all precisely in register with DNA phosphates at non-specific sites, then very little resistance to sliding would be expected (Winter *et al.*, 1981). Presumably once the target site has been located, the electrostatic charges on the protein register with the specific binding site, neutralizing the DNA phosphates, such that a tight association occurs.

Kinetic data obtained from detailed analysis of the *lac* operator/*lac* repressor interaction is supportive of the two stage process. The observed interaction at the target site occurs 10^4-fold faster than expected for random 3D diffusion. Sliding is proposed as the major site-seeking mechanism (Winter *et al.*, 1981).

In vitro resolution reaction experiments with resolvase preparation 3 indicate that the reaction occurs very quickly; the majority of reaction products are formed within 10min. The first stage of the reaction may involve location of DNA by resolvase, followed by a tracking mechanism (sliding, hopping, or both) to find the target site, *res*. Once firmly bound at *res*, repression of transcription from *tnp*A and *tnp*R promoters results, but for

the resolution reaction two *res* sites are required. One can envisage at least two ways in which *res* sites may be aligned for this to occur:

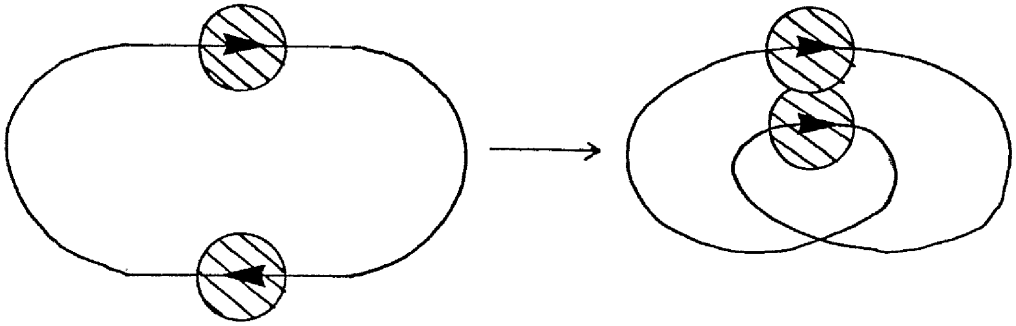
(i) Random interaction

For this model it is assumed that two resolvase "units" (a unit could be equivalent to monomers, dimers or higher forms) have located *res* sites. As the DNA is continually writhing the sites would become randomly aligned at some point; this complex may be stabilized by protein-protein interactions (see figure 6.14A). Once in this position the resolvase mediated recombination event can take place. However, random interaction is inconsistent with several observations made *in vitro*. Resolvase appears to be very inefficient at recombining *res* sites located on separate molecules; this would occur at a reasonable frequency if the random interaction model was used to align sites.(This method may well be used by the λ Int protein which works with equal efficiency in joining and separating molecules by recombination across *att* sites).As for location of target sites this would also limit the time taken for resolution as it would take longer to randomly align sites than if some active mechanism were operating. The substrate pLS138 has *res* sites only 282 bp apart; due to the constraints of DNA structure one would predict that random alignment of *res* sites would occur very infrequently; this is not compatible with the results obtained *in vitro*, which suggest that pLS138 is an efficient substrate.

(ii) Tracking

The results from *in vitro* experiments suggest that there is an active mechanism not only to locate *res* sites, but also to bring two *res* sites into close enough contact for recombination. The tracking model proposes that initially a resolvase "unit" binds

A



B

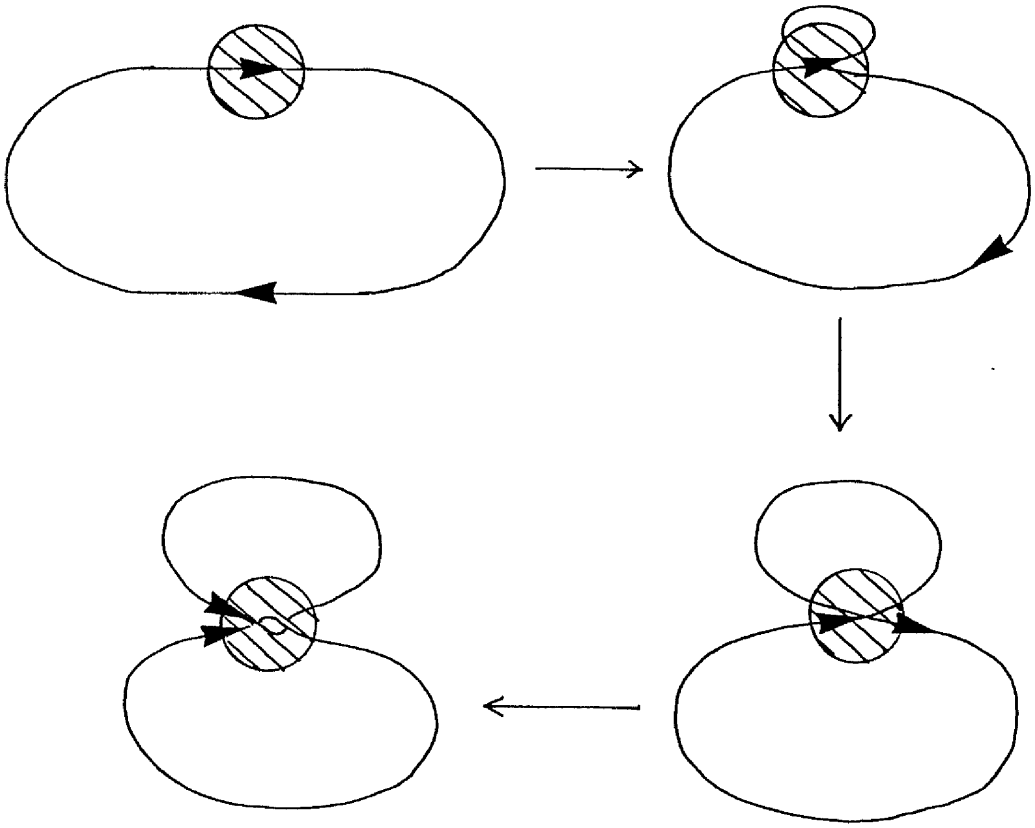


FIGURE 6.14 Models for alignment of *res* sites prior to resolution

A. Random alignment: Resolvase molecules bind to both *res* sites, these will come into close contact by random writhing of the DNA molecule. When this occurs the interaction may be stabilized by bonding between the resolvase units prior to strand exchange.

B. Tracking model: This model invokes an active mechanism for bringing the two *res* sites into close proximity. It is proposed that only one resolvase "unit"(dimer or tetramer?) is required bound at one *res* site for resolution. The protein remains bound to the first *res* site whilst seeking the second by one-dimensional tracking along the DNA separating the two sites. The tracking mechanism may be envisaged as a loop of DNA being fed through the protein molecule; as the protein progresses along the DNA the displaced loop increases in size. When the second site is reached the sites are in correct alignment for recombination to occur. Such a model would explain the inefficiencies of inter-molecular recombination and of recombination between inverted *res* sites.

Res sites are represented by arrowheads.

binds tightly to one *res* site, then by some mechanism, sliding, hopping or both, moves in one dimension along the DNA molecule until it reaches the second *res* site. During this process the resolvase unit remains bound at the first site so that DNA would be fed through the unit, forming a loop. This loop would increase in size as the protein progresses along the intervening DNA to the second *res* site (figure 6.14B). The interaction of two *res* sites may induce a conformational change in the protein resulting in concerted cleavage at *res* sites followed by ligation to new partner strands. The ligation step may be dependent on the presence of Mg^{++} (Reed and Grindley, 1981).

The proposed tracking model adequately explains the data obtained from *in vitro* experiments. Tracking accounts for the observed preference for two *res* sites in *cis*, and may also provide an explanation for the lack of recombination between inverted *res* sites. One can imagine that if resolvase binds to one site then tracks to the second, and if the second *res* site is inverted, the alignment of sequences would be wrong, thus recombination could not take place. If instead resolvase used the random interaction mechanism, then one would predict the frequency of recombination between inverted *res* sites to be equivalent to that observed between directly repeated *res* sites. In comparison, Int works as efficiently on inverted *att* sites as it does on directly repeated sites. This probably reflects a major difference in the reaction mechanisms used by these two site-specific recombination proteins.

The tracking model suggests polarity in the mechanism used to search for the second *res* site. The tracking model also predicts that only one resolvase unit is required, bound at one *res* site .

Preliminary genetic evidence suggests that a shortened *res* site containing only the 96bp *Taq*I fragment, but not the neighbouring 66 and 192bp fragments (see figures 4.1 and 6.8), can recombine with a wild-type *res* site, but two directly repeated shortened sites do not recombine at detectable fraquency. The 96bp fragment contains the crossover point, but may lack sufficient sequence for resolvase binding. This mutant *res* site will be footprinted to analyze any differences in the resolvase binding pattern. It is possible that the shortened site cannot bind resolvase,or that the change in binding does not allow tracking to occur. The observed asymmetry is consistent with the tracking model. By lengthening or shortening the arms between the functional and mutant sites (on the plasmid which resolves), it may be possible to determine the directionality of tracking by determining the efficiency of these substrates to resolve *in vitro*.

By using chemical agents which crosslink DNA to protein it may be possible to catch some DNA molecules in the process of tracking. If, as hypothesized, the protein feeds through a loop of DNA between *res* sites until the two sites are brought into close proximity then one would expect the loop to increase in size as resolvase progresses along the intervening DNA. If these intermediates could be "frozen" by crosslinking agents then spread onto grids for electron microscopy, without disrupting the intermediates, it may be possible to visualize loops.

It has been observed that superhelicity is an absolute requirement for resolution *in vitro*. There are several stages in the reaction which may require an energy source, provided by negative superhelicity of the substrate: location and binding to *res* sites, tracking,and the break/rejoin reaction. The data obtained

from footprinting experiments suggests that resolvase binds specifically to linear DNA fragments; thus superhelicity is not required for location and binding to *res*. The cleavage/ligation reaction is predicted to be a conservative event. This occurs by cleavage of the DNA backbone at *res* followed by conservation of the energy contained within the phosphodiester bond by covalent attachment to the protein. The cleaved strands are then ligated to new partners using energy retained in the DNA-protein bond. This reaction may require an external energy source (provided by superhelicity) for the initial event to reach a certain activation level. During the recombination event limited strand rotation is required to align cleaved strands with their new partners; this step may be dependent on negative superhelicity. However, if the proposed binding model is correct (see figure 6.13B), then the actual cleavage site would be tightly associated with resolvase, maybe "inside" the resolvase "unit". In this position the affect of superhelicity may be negligible. The possibility remains that superhelicity may be required for the tracking mechanism. It has been proposed that the sliding model is dependent on ionic interactions between protein and DNA; presumably this would not be affected by superhelicity. The one dimensional tracking mechanism used to search for the initial target site may be different to that required to locate the second site. The tracking model for resolution demands that resolvase remains bound to one *res* site whilst feeding through DNA searching for a second *res* site. The process of feeding DNA through the protein would be different to that involved in sliding along a DNA molecule. Present data suggests that this is likely to be a stage at which superhelicity is required.

Results presented in the previous chapter suggested that resolvase may inhibit resolution when present at high concentration (>0.3µg/0.5µg DNA). This may be due to non-specific binding between the *res* sites inhibiting the tracking mechanism, or from resolvase "units" binding to both *res* sites on a molecule, when only one is required for resolution (by the tracking model). An alternative explanation is that if resolvase has DNA wrapped around the outside of each "unit" in a righthanded coil, then superhelical turns may be removed from the molecule on binding. When only one or two "units" are bound this would have little effect on the gross superhelicity of the molecule, but at ten times this concentration the reduction in superhelicity may be instrumental in inhibiting the reaction.

Reed and Grindley have determined the γδ resolvase binding sites on both γδ and Tn*3* *res* regions (Grindley *et al.*,1982). The source of Tn*3* *res* for these experiments was the 295bp *Sau*3AI/*Pvu*II fragment, labelled at the 5' terminus using T4 polynucleotide kinase; as described in this chapter. The footprints obtained, using either γδ or Tn*3* resolvase,have been compared. There appears to be only one noticeable difference, at positions 3178-3180 there are strongly enhanced DNase I cleavage sites (within site II). Based on this data Reed and Grindley have divided site II into two asymmetric regions. They have compared the DNA sequences of the three sites on Tn*3* and γδ to formulate a concensus sequence for binding by γδ resolvase (figure 6.15). The three sites share strong homology with the concensus sequence for the highly conserved bases, eg. TGT at the beginning of each site. Each site is of different length and therefore the axis of symmetry within each site differs. If there is a concensus sequence for resolvase recognition, why does

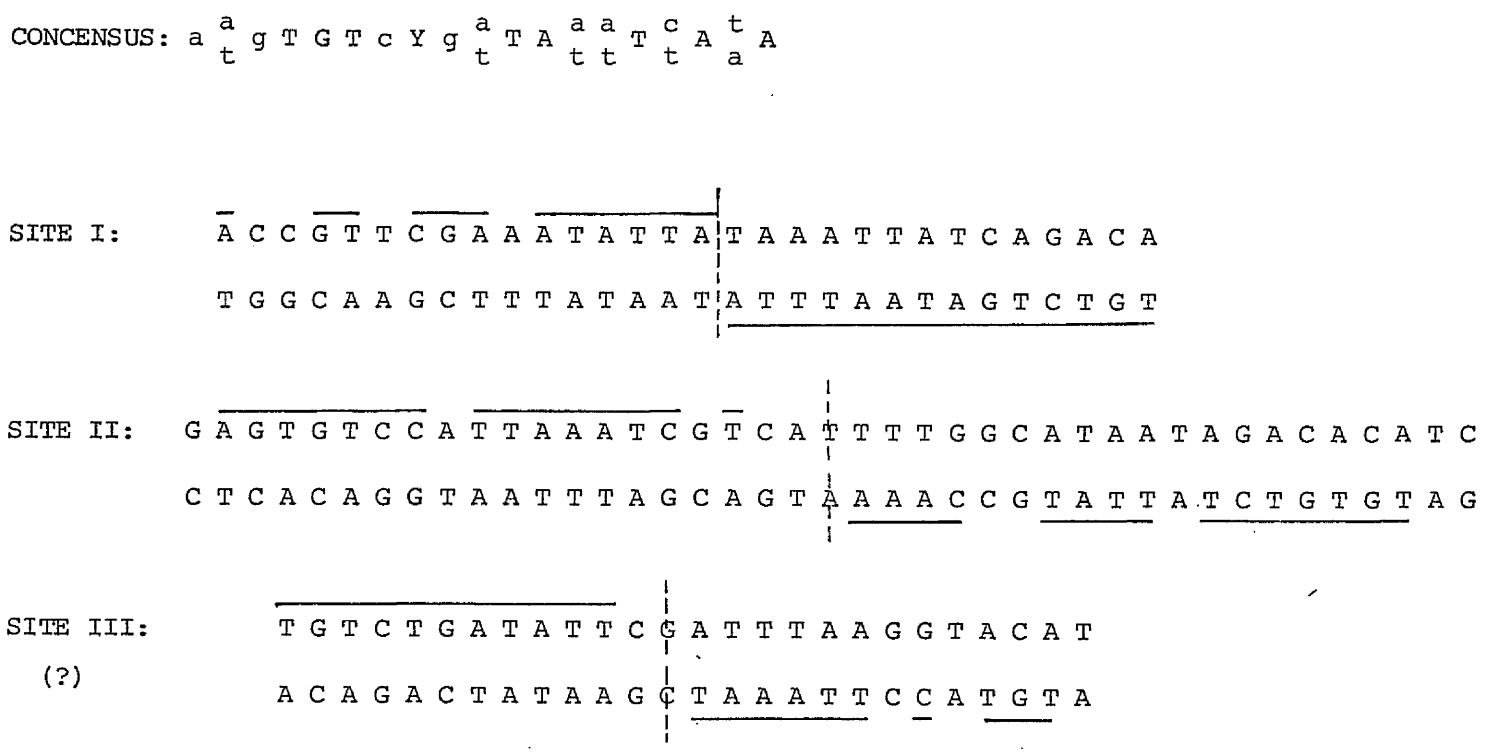


FIGURE 6.15 A comparison of the DNA sequences of the three resolvase binding sites

Site II, identified by Tn*3* resolvase binding, has arbitrarily been divided for purposes of comparison to the concensus sequence (Grindley *et al.*, 1982). Homologies to the concensus sequence are indicated by solid bars above or beneath bases; the axis of symmetry for each site is also indicated.

cleavage only occur at site I? It has been suggested that the high AT rich region in the centre of site I, in particular the Pribnow box sequence, identifies the cleavage site . There are , however, AT rich regions in the centre of each site. Resolvase would have to be very discriminating to distinguish between these sites, yet all the available evidence suggests that it must do so as there is no detectable cleavage at site II (and III),(Grindley *et al.*, 1982; Reed, pers.comm.). In fact the subtle differences which distinguish these sites may be a crucial factor in aligning the DNA into a topological state for recombination across directly repeated *res* sites.

It is of interest that the concensus sequence identified for resolvase binding within the *res* region is comparable with the binding sites for several other prokaryotic regulatory proteins. Gicquel-Sanzey and Cossart (1982) have compared the amino acid sequences of 13 regulatory proteins, including resolvase, and found several conserved regions. They have also compared the sites of action of these proteins to formulate a concensus sequence for recognition sites:

$$TGTGT\ N_{6-10}\ ACACA$$

The three resolvase binding sites conform to this concensus sequence though N is usually greater than 10bp; in the case of site II it is 27bp. Most regulatory proteins are multimeric, usually dimers or tetramers; one would therefore expect to find symmetries in their binding sites. Evidence from gel fitration of resolvase suggests that the protein exists as a dimer in 1M salt solution . These observations suggest that resolvase has evolved from a common ancestral regulatory protein, but must have diverged considerably

to be able to mediate site-specific recombination.

The topology of resolvase/*res* binding is still a mystery. Footprinting data suggests that there are three DNA sites which could bind resolvase dimers, but the DNase I protection patterns within these sites differ. The appearance of DNase I enhanced sensitivity sites within sites II and III has led me to speculate that these sites may be wrapped around the outside of a multimeric resolvase "unit". It is possible that dimers bind to the three sites, as shown in figure 6.13A; these could then fold on each other such that site I is completely protected from nuclease cleavage. At present two independent groups are trying to analyze the structure of the resolvase/DNA interaction by X-ray crystallography; hopefully their results will shed some light on the mystery of this DNA/protein complex.

CHAPTER 7

CONCLUDING REMARKS

The results presented in this thesis describe a detailed investigation into the Tn*3*-encoded site-specific recombination system. This recombination event is an integral part of the transposition process for Tn*3*, and closely related elements, ensuring efficient propagation through populations. Tn*3* and the related element γδ have been shown to encode interchangeable resolvase proteins, which mediate the site-specific recombination event; however, their transposase proteins do not appear to complement each other.

The mechanism of resolvase-mediated site-specific recombination has been investigated by constructing an *in vitro* system to study the resolution reaction. Only purified resolvase and a supercoiled substrate containing directly repeated copies of *res* are required for efficient resolution. The reaction proceeds by a reciprocal break/join event with no concomitant DNA synthesis or requirement for high energy cofactors. The reaction products are always catenated circles, suggesting that catenation is an intrinsic property of resolvase and is probably a direct consequence of the reaction mechanism.

Resolvase binds avidly to any DNA which is present in an *in vitro* resolution reaction. DNA with or without *res* sites efficiently binds resolvase, though preliminary results suggest that binding is tighter if the competitor DNA contains *res* sites. Binding to nonspecific sites may be dependent on the salt concentration in the reaction mixture (Berg *et al.*, 1981). The specific binding sites for resolvase within the Tn*3* *res* region were identified by DNase I footprinting. Three specific binding sites were revealed using this technique; the sequences contained within these

sites conform to the concensus sequence proposed by Grindley *et al.* (1982) for γδ resolvase binding sites. This concensus sequence also shares homology to a common recognition sequence formulated for prokaryotic regulatory proteins. There is suggestive evidence that the DNA within the *res* region may be wrapped around resolvase, analagous to the model proposed for the DNA-DNA gyrase interaction (Liu and Wang,1978; Kirkegaard and Wang,1981; Morrison and Cozzarelli,1981). A model invoking wrapping would introduce asymmetry into the interaction which may be significant in terms of the mechanism used for the recombination reaction.

Preliminary results suggest that recombination proceeds with greatest efficiency when the distance between *res* sites is reduced. This data, in conjunction with the observation that directly repeated *res* sites in the *cis* configuration are the preferred substrate for resolution, has led to the proposal of a tracking model for the recombination mechanism. This model predicts that as the distance between *res* sites is lengthened, the efficiency of recombination exponentially decreases (see Berg *et al.*, 1981; Winter *et al.*, 1981; Winter and von Hippel,1981). Experiments to test this hypothesis are currently being performed by other members of our group at Glasgow.

Throughout the results chapters I have tried to indicate the similarities and differences between resolvase and other site-specific recombination proteins. The lambda Int system has been subjected to detailed genetic and biochemical analysis and has, historically, provided the basic model for site-specific recombination. Although the Int and resolvase systems share the property of requiring only short sequence homologies for reciprocal break/rejoin recombination

events, many differences have become apparent through the course of these experiments. The Int system requires the presence of two host polypeptides encoded by the *himA* and *hip* genes, which together form IHF, or host integration factor (Nash and Robertson,1981). This plays a direct role in the recombination event, indicated by it s requirement *in vitro* and ability to bind specifically to *att* sites.(Craig and Nash, pers. comm.). The *himA* gene also plays an indirect role *in vivo* by controlling expression of several lambda genes required for establishment of lysogeny (Miller,1981). In contrast, resolvase-mediated recombination is independent of host factors.

Purified Int alone induces single-stranded concerted cleavage under certain conditions, classifying the protein as a type I topoisomerase (Kikuchi and Nash,1979b). This raises the possibility that IHF may be primarily involved in optimizing alignment of *att* sites while Int is specifically involved single-stranded breakage-reunion of strands for the recombination event. As the observed topoisomerase activity for Int is nonspecific IHF may be required to ensure that cleavage only occurs within *att* sites for recombination. Resolvase has been shown to mediate a specific double-stranded cleavage at *res*, but only when two *res* sites are directly repeated within a molecule (Reed and Grindley,1981). The break/join event appears to be mechanistically related to DNA gyrase-mediated cleavage (Sherratt *et al.*,1981b). Preliminary data suggests that resolvase has topoisomerase activity, but this has not been studied in detail.

Int and resolvase have different preferred substrates for recombination directly related to their functions *in vivo*. Int

recombines with equal efficiency intermolecularly and intramolecularly. The substrates for intermolecular recombination, the reaction required for integration of the circularized phage into the host chromosome to establish lysogeny, are *attP* and *attB* . The hybrid sites thus formed, *attR* and *attL*, are the substrates for intramolecular recombination to excise the prophage. For the excision reaction the product of the phage-encoded *xis* gene, Xis, is required in addition to Int. The slightly different requirements for these reactions *in vivo* ensures that the phage is not continually excised and integrated. Although lambda has no requirement during it s life cycle for recombination between inverted *att* sites, this reaction proceeds efficiently *in vitro* . There is no reason why this should not occur if the sites are aligned by a "random" mechanism.

The role of resolvase in propagation of transposable elements is to mediate the efficient reduction of transpositional cointegrates. Elements which proceed through an obligatory cointegrate stage during transposition have acquired site-specific recombination systems to resolve cointegrates, thus enabling efficient spread of the element through populations. This system is only required to work on directly repeated *res* sites in *cis*, which are generated during transposition, resulting in reduction of cointegrates. There is no advantage in joining molecules at *res* sites or recombination between inverted *res* sites as these processes do not aid propagation of the element. A tracking mechanism for resolution, as described in chapter 6, would ensure that only resolution occurs. As resolvase binds very tightly to *res* sites it is automatically in position to resolve transpositional cointegrates (or plasmid dimers) if the element should transpose again, as well

as controlling expression of both *tnp*A and *tnp*R genes. A similar mechanism to resolvase-mediated recombination may operate in bacteriophage P1.

The bacteriophage P1 exists as a plasmid during the lysogenic stage of it s life cycle. The copy number of the P1 plasmid is stringently controlled so that only one copy is normally present per cell. P1 encodes a site-specific recombination protein, *cre* which acts at a site, *lox*, present on the plasmid. Following replication many molecules dimerise due to homologous recombination mediated by the host-encoded general genetic recombination system. *Cre*-mediated recombination efficiently resolves dimers to monomers thus ensuring segregation of plasmids into daughter cells (Austin *et al.*, 1981). Without this system plasmids are rapidly lost from populations due to inefficient partitioning of daughter plasmids. Like resolvase-mediated recombination, the *cre* system acts preferentially on molecules containing directly repeated sites; this suggests that these two systems may have similar reaction mechanisms. Transposable elements, such as Tn*3*, could have a stabilizing effect on unstable or stringently controlled plasmids on which they reside by ensuring reduction of dimers to monomers. It is tempting to speculate that these two systems may have evolved from a common ancester.

The systems described above operate preferentially on directly repeated sites; there are also systems in prokaryotes which act only on inverted sites. The Hin protein encoded by *Salmonella typhimurium* mediates inversion of a 970bp segment of DNA resulting in alternate expression of the two flagellin proteins, H1 and H2 (Silverman *et al.*, 1981). This system is currently being investigated

to reveal the molecular mechanism underlying the inversion event (Simon, pers. comm.). Inversion of the "H" segment, and the analagous G loop in phage Mu, occurs at low frequency, only requires the Hin, or Gin, protein, and genetic evidence suggests works only on inverted repeats. The only apparent function of this system is to mediate inversion thus controlling expression of adjacent genes; in the case of Mu these genes are contained within the invertible segment. There would be no advantage in constant "flipping" allowing expression of both flagellin/host range proteins during each generation. The inefficiency of this recombination event could be due to low concentrations and/or activity of the protein or an inefficient mechanism used to align the inverted repeats. The Hin protein shares significant homology with resolvase, particularly over the amino terminal half of the protein; however, these proteins do not have interchangeable functions (Dyson, pers. comm.). These proteins have probably evolved from a common ancestor, but appear to have diverged to accomodate different substrate preferences and recombination efficiencies required for each particular system.

The Hin and Gin systems provide examples of recombinational gene switching in prokaryotes; similar systems are likely to be encoded by eukaryotes. The 2μ circle plasmid present in some yeast strains encodes a site-specific inversion system, though the functional significance of this is unclear (Broach *et al.*, 1982). In recent years the importance of site-specific recombination in generating diversity has been highlighted by the discovery of antibody gene splicing. During maturation from germline to somatic cells, the antibody genes are subject to a number of recombination events which bring together dispersed sequences in a variety of combinations (Leder 1982). Recombination occurs across short homologous

sequences flanking the component genes,or parts of genes. The proteins which mediate such events have yet to be identified. As observed in prokaryotic systems, the proteins involved are likely to have preferred substrate configurations and specific molecular mechanisms which optimize the efficiency of gene splicing.

Site-specific recombination can therefore be seen as an important mechanism,in a number of different systems, for generating rearrangements. Each system has a preferred substrate and a particular mechanism for optimizing the required recombination event. The Int and resolvase systems, which have been well characterized *in vitro*, demonstrate that these events do not always occur by like mechanisms, suggesting that they have diverged considerably from a common ancestor or have evolved independently. The ubiquitous occurrence of site-specific recombination systems in both prokaryotes and eukaryotes highlights their importance as a mechanism for generating diversity and controlling gene expression.

References

ADHYA,S.L. and SHAPIRO,J.A. (1969) The galactose operon of *E.coli* K-12. I. Structural and pleiotropic mutations of the operon. *Genetics* 62 231-247

ALBERTINI,A.M., HOFER,M., CALOS,M. and MILLER,J.H. (1982) On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. *Cell* 29 319-328

ALBERTS,B. and STERNGLANZ,R. (1977) Recent excitement in the DNA replication problem. *Nature* 269 655-661

ARTHUR,A.K. (1981) Molecular genetics of the transposition process. *Ph.D thesis Sussex University*

ARTHUR,A.K. and SHERRATT,D.J. (1979) Dissection of the transposition process: a transposon-encoded site-specific recombination system. *Mol. Gen. Genet.* 175 267-274

AUSTIN,S., ZIESE,M. and STERNBERG,N. (1981) A novel role for site-specific recombination in maintenance of bacterial replicons. *Cell* 25 729-736

BERG,O.G., WINTER,R.B. and VON HIPPEL,P.H. (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* 20 6929-6948

BERNARD,H.U. and HELINSKI,D.R. (1979) Use of the lambda phage promoter to promote gene expression in hybrid plasmid cloning vehicles. *Meths. Enzymol.* 68 482-492

BIRGE,E.A. and LOW,K.B. (1974) Detection of transcribable recombination products following conjugation in Rec^+, $RecB^-$ and $RecC^-$ strains of *E.coli* K-12. *J. Mol. Biol.* 83 447-457

BIRNBOIM,H.C. and DOLY,J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* 7 1513-1523

BOLIVAR,F., RODRIGUEZ,R.L., GREENE,P.J., BETLACH,M.C., HEYNEKER,H.L., BOYER,H.W., CROSA,J.H. and FALKOW,S. (1977) Construction and characterization of new cloning vehicles II: a multi-purpose cloning system. *Gene* 2 95-113

BROACH,J.R., GUARASCIO,V.R. and JAYARAM,M. (1982) Recombination within the yeast plasmid 2μ circle is site-specific. *Cell* 29 227-234

BROKER,T.E., CHOW,L.T. and SOLL,L. (1977) The *E.coli* gamma-delta recombination sequence is flanked by inverted repetitions. In *DNA insertion elements, plasmids and episomes.* Ed. Bukhari,A.I., Shapiro, J.A., and Adhya,S.L. pp 575-580. Cold Spring Harbor press.

BROWN,P.O. and COZZARELLI,N.R. (1979) A sign inversion mechanism for enzymatic supercoiling of DNA. *Science* 206 1081-1083

BROWN,P.O. and COZZARELLI,N.R. (1981) Catenation and knotting of duplex DNA by type I topoisomerases: a mechanistic parallel with type II topoisomerases. *Proc. Natl. Acad. Sci. USA* 78 843-847

BUKHARI,A.I. and AMBROSIO,L. (1978) The invertible segment of bacteriophage Mu determines the adsorption properties of Mu particles. *Nature* 271 575-577

CALOS,M.P. and MILLER,J.H. (1980) Transposable elements. *Cell* 20 579-595

CAMERON,J., LOH,E., and DAVIS,R. (1979) Evidence of transposition of dispersed repetitive DNA families in yeast. *Cell* 16 739-751

CAMPBELL,A. (1962) Episomes. *Adv. Genet.* 11 101-145

CASADABAN,M.J., CHOU,J., and COHEN,S.N. (1982) Overproduction of the Tn*3* transposition protein and it's role in DNA transposition. *Cell* 28 345-354

CASSUTO,E., WEST,S.C., MURSALIM,J., CONLON,S. and HOWARD-FLANDERS,P. (1980) Initiation of genetic recombination: homologous pairing between duplex DNA molecules promoted by the RecA protein. *Proc. Natl. Acad. Sci. USA* 77 3962-3966

CHANDLER,M., CLERGET,M. and CARO,L. (1982) The transposition frequency of IS*1* flanked transposons is a function of their size. *J. Mol. Biol.* 154 229-243

CHANG,A.C.Y. and COHEN,S.N. (1978) Construction and characterization of amplifiable multi-copy DNA cloning vehicles derived from P15A cryptic miniplasmid. *J. Bact.* 134 1141-1156

CHIANG,S.J. and CLOWES,R.C. (1980) Intramolecular transposition and inversion in plasmid R6K. *J. Bact.* 142 668-682

CHOU,J., CASADABAN,M.J., LEMAUX,P.J. and COHEN,S.N. (1979a) Identification and characterization of a self-regulated repressor of translocation of the Tn*3* element. *Proc. Natl. Acad. Sci. USA* 76 4020-4024

CHOU,J., LEMAUX,P.J., CASADABAN,M.J. and COHEN,S.N. (1979b) Transposition protein of Tn*3*: identification and characterization of an essential repressor-controlled gene product. *Nature* 282 801-806

CLARK,A.J. (1973) Recombination deficient mutants of *E.coli* and other bacteria. *Ann. Rev. Genet.* 7 67-86

CLARK,A.J. (1974) Progress towards a metabolic interpretation of genetic recombination of *E.coli* and bacteriophage lambda. *Genetics* 78 259-271

COELHO,A., LEACH,D., MAYNARD-SMITH,S. and SYMONDS,N. (1981) Transposition studies using a ColE1 derivative carrying bacteriophage Mu. *Cold Spring Harbor Symp. Quant. Biol.* 45 323-328

COHEN,S.N., CASADABAN,M.J., CHOU,J. and TU,C.P.D. (1979) Studies of the specificity and control of transposition of the Tn*3* element. *Cold Spring Harbor Symp. Quant. Biol.* 43 1247-1255

COLLINS,J. (1981) Instability of palindromic DNA in *E.coli*. *Cold Spring Harbor Symp. Quant. Biol.* 45 409-415

COX,M. and LEHMAN,I.R. (1981) Directionality and polarity in RecA protein promoted branch migration. *Proc. Natl. Acad. Sci. USA* 78

6018-6022

COZZARELLI,N. (1980) DNA gyrase and the superhelicity of DNA. *Science* 207 953-960

CUNNINGHAM,R.P., DASGUPTA,C., SHIBATA,T., and RADDING,C.M. (1980) Homologous pairing in genetic recombination: RecA protein makes joint molecules of gapped circular DNA and closed circular DNA. *Cell* 20 223-235

CUNNINGHAM,R.P., SHIBATA,T., DASGUPTA,C. and RADDING,C.M. (1979) Single strands induce RecA protein to unwind duplex DNA for homologous pairing. *Nature* 281 191-195

CUNNINGHAM,R.P., WU,A.M., SHIBATA,T., DASGUPTA,C., and RADDING,C.M. (1981) Homologous pairing and topological linkage of DNA molecules by combined action of *E.coli* RecA protein and topoisomerase I. *Cell* 24 213-223

DAGART,M. and EHRLICH,S.D. (1979) Prolonged incubation in calcium chloride improves the competence of *E.coli* cells. *Gene* 6 23-28

DASGUPTA,C., and RADDING,C.M. (1982) Polar branch migration promoted by RecA protein: effect of mismatched basepairs. *Proc. Natl. Acad. Sci. USA* 79 762-766

DASGUPTA,C., SHIBATA,T., CUNNINGHAM,R.P. and RADDING,C.M. (1980) The topology of homologous pairing promoted by RecA protein. *Cell* 22 437-446

DATTA,N. and HEDGES,R.W. (1972) Trimethoprim resistance conferred by W plasmids in *Enterobacteraceae*. *J. Gen. Microbiol.* 72 349-355

DAVIDSON,N., DEONIER,R.C., HU,S., and OHTSUBO,E. (1974) Electron microscope studies of sequence relations among plasmids of *E.coli* X. DNA sequence organization of F and of F primes and the sequences involved in Hfr formation. In *Microbiology 1974*, ed. Schlessinger,D., pp 56-65. Washington DC: Am. Soc. microbiol. press

DELEYS,R.J. and JACKSON,D.A. (1976) Dye titrations of covalently

closed supercoiled DNA analyzed by agarose gel electrophoresis. *Biochem. Biophys. Res. Commun.* 69 446-454

DOHERTY,M.J., MORRISON,P.T. and KOLODNER,R. (1982) Genetic recombination of bacterial plasmid DNA: physical and genetic analysis of the products of plasmid recombination in *E.coli*. Submitted

DOOLITTLE,W.F. and SAPIENZA,C. (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284 601-603

DOUGAN,G., SAUL,M., TWIGG,A., GILL,R. and SHERRATT,D.J. (1979) Polypeptides expressed in *E.coli* K-12 minicells by transposition elements Tn*1* and Tn*3*. *J. Bact.* 138 48-54

EISENBERG,S., GRIFFITH,J. and KORNBERG,A. (1977) ϕX174 *cistron A* protein is a multifunctional enzyme in DNA replication. *Proc. Natl. Acad. Sci. USA* 74 3198-3202

ESPEJO,R.T. and LEBOWITZ,J. (1976) A simple electrophoretic method for determining the superhelix density of closed circular DNAs and for observations of their superhelix density heterogeneity. *Analyt. Biochem.* 72 95-103

FARABAUGH,P., and FINK,G. (1980) Insertion of the eukaryotic transposable element Ty1 creates a 5bp duplication. *Nature* 286 352-356

FARABAUGH,P., SCHMEISSNER,U., HOFER,M. and MILLER,J.H. (1978) Genetic studies of the *lac* repressor VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of *E.coli*. *J. Mol. Biol.* 126 847-863

FELSENFIELD,G. (1978) Chromatin. *Nature* 271 115-122

FINK,G., FARABAUGH,P., ROEDER,G. and CHALEFF,D. (1981) Transposable elements (Ty) in yeast. *Cold Spring Harbor Symp. Quant. Biol.* 45 575-580

FINNEGAN,D.J. (1981) Transposable elements and proviruses. *Nature* 292 800-802

FISHEL,R., JAMES,A. and KOLODNER,R. (1981) RecA-independent general

genetic recombination of plasmids. *Nature* 294 184-186

FLAVELL,A. and ISH-HOROWICZ,D. (1981) Extrachromosomal circular copies of the eukaryotic transposable element *copia* in cultured *Drosophila* cells. *Nature* 292 591-595

FOSTER,T.J., LUNDBLAD,V., HANLEY-WAY,S., HALLING,S.M. and KLECKNER,N. (1981) Three Tn*10*-associated excision events: relationship to transposition and role of direct and inverted repeats. *Cell* 23 215-227

FRANKLIN,N.C. (1971) The N operon of lambda: extent and regulation as observed in fusions to the tryptophan operon of *E.coli*. In *The bacteriophage lambda*, ed. Hershey,A.D., pp 621-637. Cold Spring Harbor Lab., Cold Spring Harbor, New York.

GALOS,D.J. and CHANDLER,M. (1981) On the molecular mechanisms of transposition. *Proc. Natl. Acad. Sci. USA* 78 4858-4862

GALOS,D.J. and SCHMITZ,A. (1978) DNase footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucl. Acids Res.* 5 3157-3170

GELLERT,M. (1981) DNA topoisomerases. *Ann. Rev. Biochem.* 50 879-910

GILL,R., HEFFRON,F., DOUGAN,G. and FALKOW,S. (1978) Analysis of the sequences transposed by complementation of two classes of transposition-deficient mutants of Tn*3*. *J. Bact.* 136 742-756

GILL,R., HEFFRON,F. and FALKOW,S. (1979) Identification of the protein encoded by the transposable element Tn*3* required for it's transposition *Nature* 282 797-801

GILLEN,J.R., WILLIS,D.K. and CLARK,A.J. (1981) Genetic analysis of the RecE pathway of genetic recombination of *E.coli* K-12. *J. Bact.* 145 521-532

GRINDLEY,N.D.F. (1978) IS*1* insertion generates duplication of a 9bp sequence at it's target site. *Cell* 13 419-426

GRINDLEY,N.D.F., LAUTH,M.R., WELLS,R.G., WITYK,R.J., SALVO,J.T. and REED,R.R. (1982) Transposon-mediated site-specific recombination:

identification of three binding sites for resolvase at the *res* sites of gamma-delta and Tn*3*. Submitted

GRINDLEY,N.D.F. and SHERRATT,D.J. (1979) Sequence analysis at IS*1* insertion sites; models for transposition. *Cold Spring Harbor Symp. Quant. Biol.* 43 1257-1261

GUYER,M. (1978) The gamma-delta sequence of F is an insertion sequence. *J. Mol. Biol.* 126 347-365

HALLING,S.M. and KLECKNER,N. (1982) A symmetrical six-base-pair target site sequence determines Tn*10* insertion specificity. *Cell* 28 155-163

HARSHEY,R.M. and BUKHARI,A.I. (1981) A mechanism of DNA transposition. *Proc. Natl. Acad. Sci. USA* 78 1090-1094

HEDGES,R. and JACOB,A. (1974) Transposition of ampicillin resistance from RP4 to other replicons. *Mol. Gen. Genet.* 132 31-40

HEFFRON,F., BEDINGER,P., CHAMPOUX,J. and FALKOW,S. (1977) Deletions affecting the transposition of an antibiotic resistance gene. *Proc. Natl. Acad. Sci. USA* 74 702-706

HEFFRON,F., MCCARTHY,B. and OHTSUBO,E. (1979) DNA sequence analysis of the transposon Tn*3*: three genes and three sites involved in transposition of Tn*3*. *Cell* 18 1153-1163

HEFFRON,F., RUBENS,C. and FALKOW,S. (1975) Translocation of a plasmid DNA sequence which mediates ampicillin resistance: molecular nature and specificity of insertion. *Proc. Natl. Acad. Sci. USA* 72 3623-3627

HOLLIDAY,R. (1964) A mechanism for gene conversion in fungi. *Genet. Res.* 5 282-304

HOLLOMAN,W.K. and RADDING,C.M. (1976) Recombination promoted by superhelical DNA and the *recA* gene of *E.coli*. *Proc. Natl. Acad. Sci USA* 73 3910-3914

HSU,P.L., ROSS,W. and LANDY,A. (1980) The lambda phage *att* site: functional limits and interaction with Int protein. *Nature* 285 85-91

IKEDA,H., MORIYA,K. and MATSUMOTO,T. (1981) *In vitro* study of illegitimate recombination: involvement of DNA gyrase. *Cold Spring Harbor Symp. Quant. Biol.* 45 399-407

JAMES,A., MORRISON,P.T. and KOLODNER,R. (1982) Genetic recombination of plasmid DNA: analysis of the effect of recombination deficient mutants on plasmid recombination. *J. Mol. Biol.* (in press)

JOHNSRUD,L., CALOS,M.P. and MILLER,J.H. (1978) The transposon Tn*9* generates a 9bp repeated sequence during integration. *Cell* 15 1209-1219

JOYCE,C.M. and GRINDLEY,N.D.F. (1981) Analysis of the structure and function of the kanamycin- resistance transposon Tn*903*. *Cold Spring Harbor Symp. Quant. Biol.* 45 125-133

KAHN,R., CUNNINGHAM,R.P., DASGUPTA,C., and RADDING,C.M. (1981) Polarity of heteroduplex formation promoted by *E.coli* RecA protein. *Proc. Natl. Acad. Sci. USA* 78 4786-4790

KAMP,D., CHOW,L.T., BROKER,T.R., KWOH,D., ZIPSER D. and KAHMANN,R. (1979) Site-specific recombination in phage Mu. *Cold Spring Harbor Symp. Quant. Biol.* 43 1159-1167

KAMP,D and KAHMANN,R. (1981) The relationship of two invertible segments in bacteriophage Mu and *Salmonella typhimurium* DNA. *Mol. Gen. Genet.* 184 564-566

KAMP,D., KAHMANN,R., ZIPSER,D., BROKER,T.R. and CHOW,L.T. (1978) Inversion of the G DNA segment of bacteriophage Mu controls phage infectivity. *Nature* 271 577-580

KELLER,W. (1975) Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis. *Proc. Natl. Acad. Sci. USA* 72 4876-4880

KEMPER,B. and JANZ,E. (1976) Function of gene 49 of bacteriophage T4 1. Isolation and biochemical characterization of very fast sedimenting DNA. *J. Virol.* 18 992-999

KIKUCHI,Y. and NASH,H.A. (1978) The bacteriophage lambda *int* gene product. A filter binding assay for genetic recombination, purification of Int and specific binding to DNA. *J. Biol. Chem.* 253 7149-7157

KIKUCHI,Y. and NASH,H.A. (1979a) Integrative recombination of bacteriophage lambda: requirement for supertwisted DNA *in vivo* and characterization of Int. *Cold Spring Harbor Symp. Quant. Biol.* 43 1099-1109

KIKUCHI,Y. and NASH,H.A. (1979b) Nicking-closing activity associated with bacteriophage lambda *int* gene product. *Proc. Natl. Acad.Sci. USA* 76 3760-3764

KIRKEGAARD,K. and WANG,J.C. (1981) Mapping the topography of DNA wrapped around gyrase by nucleolytic and chemical probing of complexes of unique DNA sequences. *Cell* 23 721-729

KITTS,P.A. (1982) The transposition mechanism of Tn*3*. *Ph.D thesis Glasgow University*

KITTS,P.A., LAMOND,A. and SHERRATT,D.J. (1982a) Inter-replicon transposition of Tn*1/3* occurs in two sequentially separable steps. *Nature* 295 626-628

KITTS,P.A., SYMINGTON,L., BURKE,M., REED,R.R. and SHERRATT,D.J. (1982b) Transposon-specified site-specific recombination. *Proc. Natl. Acad. Sci. USA* 79 46-50

KLECKNER,N. (1979) DNA sequence analysis of Tn*10* insertions: origin and role of 9bp flanking repetition during Tn*10* translocation. *Cell* 16 711-720

KLECKNER,N. (1981) Transposable elements in prokaryotes. *Ann. Rev. Genet.* 15 341-404

KLECKNER,N. (1982) Transposon Tn*10*. Submitted

KOLODNER,R. (1980) Genetic recombination of bacterial plasmid DNA: electron microscope analysis of *in vitro* intramolecular recombination. *Proc. Natl. Acad. Sci. USA* 77 4847-4851

KORNBERG,A. (1980) *DNA replication*. San Francisco: W.M.Freeman press

KORNBERG,R.D. and KLUG,A. (1981) The nucleosome. *Sci. American* 244 48-60

KOSTRIKEN,R., MORITA,C. and HEFFRON,F. (1981) Transposon Tn*3* encodes a site-specific recombination system: identification of essential sequences, genes and actual site of recombination. *Proc. Natl. Acad. Sci. USA* 78 4041-4045

KRASNOW,M.A. and COZZARELLI,N.R. (1982) Catenation of DNA rings by topoisomerases. *J. Biol. Chem.* 257 2681-2693

KRESCHMER,P.J. and COHEN,S.N (1979) Effect of temperature on translocation frequency of the Tn*3* element. *J. Bact.* 139 515-519

KREUZER,K.N. and COZZARELLI,N.R. (1980) Formation and resolution of catenanes by DNA gyrase. *Cell* 20 245-254

LANDY,A. and ROSS,W. (1977) Viral integration and excision: structure of the lambda *att* sites. *Science* 197 1147-1160

LEDER,P. (1982) The genetics of antibody diversity. *Sci. American* 246 72-83

LITTLE,J.W., EDMISTON,S.H., PACELLI,L.Z. and MOUNT,D.W. (1980) Cleavage of the *E.coli lex*A protein by the RecA protease. *Proc. Natl. Acad. Sci. USA* 77 3225-3229

LITTLE,J.W. and MOUNT,D.W. (1982) The SOS-regulatory system of *E.coli*. *Cell* 29 11-22

LIU,L.F., LIU,C.C. and ALBERTS,B.M. (1980) Type II DNA topoisomerases: enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break. *Cell* 19 697-707

LIU,L.F. and WANG,J.C. (1978) DNA-DNA gyrase complex: the wrapping of the DNA duplex outside of the enzyme. *Cell* 15 979-984

LIVNEH,Z. and LEHMAN,I.R. (1982) Recombinational bypass of pyrimidine dimers promoted by the RecA protein of *E.coli*. *Proc. Natl. Acad. Sci. USA* 79 3171-3175

LUTTER,L.C. (1977) DNase I produces staggered cuts in the DNA of chromatin. *J. Mol. Biol.* 117 53-69

LUTTER,L.C. (1979) Precise location of DNase I cutting sites in the nucleosome core determined by high resolution gel electrophoresis. *Nucl. Acids Res.* 6 41-56

McCLINTOCK,B. (1952) Chromosome organization and genic expression. *Cold Spring Harbor Symp. Quant. Biol.* 16 13-47

McCLINTOCK,B. (1957) Controlling elements and the gene. *Cold Spring Harbor Symp. Quant. Biol.* 21 197-216

McCORMICK,M., WISHART,W., OHTSUBO,H., OHTSUBO,E. and HEFFRON,F. (1981) Plasmid cointegrates and their resolution mediated by Tn*3* mutants. *Gene* 15 103-118

McENTEE,K., WEINSTOCK,G.M. and LEHMAN,I.R. (1979) Initiation of general recombination catalyzed *in vitro* by the RecA protein of *E. coli*. *Proc. Natl. Acad. Sci. USA* 76 2615-2619

McGAVIN,S. (1971) Models of specifically paired like(homologous) nucleic acid structures. *J. Mol. Biol.* 55 293-298

MALAMY,M.H. (1966) Frameshift mutations in the lactose operon of *E.coli*. *Cold Spring Harbor Symp. Quant. Biol.* 31 189-201

MANIATIS,T., JEFFREY,A. and van de SANDE,H. (1975) Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. *Biochem.* 14 3787-3794

MAXAM,A.M. and GILBERT,W. (1977) A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* 74 560-564

MAXAM,A.M. and GILBERT,W. (1980) Sequencing end-labelled DNA with base-specific chemical cleavages. *Meths. Enzymol.* 65 499-560

MESELSON,M.S. and RADDING,C.M. (1975) A general model for genetic recombination. *Proc. Natl. Acad. Sci. USA* 72 358-361

MILLER,C. and COHEN,S.N. (1980) F plasmid provides a function that promotes RecA-independent site-specific fusions of pSC101 replicon.

Nature 285 577-579

MILLER,H.I. (1981) Multilevel regulation of bacteriophage lambda lysogeny by the *E.coli himA* gene. *Cell* 25 269-276

MILLER,J.H. (1972) *Experiments in molecular genetics*. Cold Spring Harbor Lab., Cold Spring Harbor, New York.

MIZUUCHI,K., KEMPER,B., HAYS,J. and WEISBERG,R.A. (1982) T4 endonuclease VII cleaves Holliday structures. *Cell* 29 357-365

MIZUUCHI,K. and MIZUUCHI,M. (1979) Integrative recombination of bacteriophage lambda: *in vitro* study of the intermolecular reaction. *Cold Spring Harbor Symp. Quant. Biol.* 43 1111-1114

MIZUUCHI,K. and NASH,H.A. (1976) Restriction assay for integrative recombination of bacteriophage lambda DNA *in vitro*: requirement for closed circular DNA substrate. *Proc. Natl. Acad. Sci. USA* 73 3524-3528

MIZUUCHI,K., WEISBERG,R., ENQUIST,L., MIZUUCHI,M., BURACYNSKA,B., FOELLER,C., HSU,P.L., ROSS,W. and LANDY,A. (1981) Structure and function of the phage lambda *att* site: size, Int-binding, and location of crossover point. *Cold Spring Harbor Symp. Quant. Biol.* 45 429-437

MORRISON,A., BROWN,P.O., KREUZER,K.N., OTTER,R., GERRARD,S.P. and COZZARELLI,N.R. (1980) Mechanisms of DNA topoisomerases. In *Mechanistic studies of DNA replication and genetic recombination*, ed. Alberts,B., pp. 785-807. Academic press.

MORRISON,A. and COZZARELLI,N.R. (1979) Site-specific cleavage of DNA by *E.coli* DNA gyrase. *Cell* 17 175-184

MORRISON,A. and COZZARELLI,N.R. (1981) Contacts between DNA gyrase and it's binding site on DNA: features of symmetry and asymmetry revealed by protection from nucleases. *Proc. Natl. Acad. Sci. USA* 78 1416-1420

NASH,H.A. (1975) Integrative recombination of bacteriophage lambda DNA *in vitro*. *Proc. Natl. Acad. Sci. USA* 72 1072-1076

NASH,H.A. (1981) Integration and excision of bacteriophage lambda: the mechanism of conservative site-specific recombination. *Ann. Rev. Genet.* 15 143-167

NASH,H.A., MIZUUCHI,K., ENQUIST,L.W. and WEISBERG,R.A. (1981) Strand exchange in lambda integrative recombination: genetics, biochemistry, and models. *Cold Spring Harbor Symp. Quant. Biol.* 45 417-428

NASH,H.A., MIZUUCHI,K., WEISBERG,R.A., KIKUCHI,Y. and GELLERT,M. (1977) Integrative recombination of bacteriophage lambda- the biochemical approach to DNA insertions. In *DNA insertion elements, plasmids and episomes*, ed. Bukhari,A.I., Shapiro,J.A. and Adhya,S.L., pp 363-373. Cold Spring Harbor Lab., Cold Spring Harbor, New York

NASH,H.A. and ROBERTSON,C.A. (1981) Purification and properties of the *E.coli* protein factor required for lambda integrative recombination. *J. Biol. Chem.* 256 9246-9253

NOLL,M. (1974) Internal structure of chromatin subunit. *Nucl. Acids Res.* 1 1573-1578

OHTSUBO,H., OHMORI,H. and OHTSUBO,E. (1979) Nucleotide sequence analysis of Tn*3*: implications for insertion and deletion. *Cold Spring Harbor Symp. Quant. Biol.* 43 1269-1277

ORGEL,L.E. and CRICK,F.H. (1980) Selfish DNA: the ultimate parasite. *Nature* 284 604-610

PATO,M.L. and REICH,C. (1982) Instability of transposase activity: evidence from bacteriophage Mu DNA replication. *Cell* 29 219-225

PEACOCK,A.C. and DINGMAN,C.W. (1967) Resolution of multiple ribonucleic acid species by polyacrylamide gel electrophoresis. *Biochem.* 6 1818-1827

POLLOCK,T.J. and ABREMSKI,K. (1979) DNA without supertwists can be an *in vitro* substrate for site-specific recombination of bacteriophage lambda. *J. mol. Biol.* 131 651-654

POLLOCK,T.J. and NASH,H.A. (1980) Catenation of the products of

integrative recombination: comparison of supertwisted and non-supertwisted substrates. In *Mechanistic studies of DNA replication and genetic recombination*, ed. Alberts,B., pp 953-961. Academic press

POTTER,H. and DRESSLER,D. (1979) DNA recombination: *In vivo* and *in vitro* studies. *Cold Spring Harbor Symp. Quant. Biol.* 43 969-985

PRIBNOW,D. (1975) Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proc. Natl. Acad. Sci. USA* 72 784-788

RADDING,C.M. (1978) Genetic recombination: strand transfer and mismatch repair. *Ann. Rev. Biochem.* 47 847-880

RADDING,C.M. (1981) Recombination activities of *E.coli* RecA protein. *Cell* 25 3-4

RADMAN,M., VILLANI,G., BOITEUX,S., KINSELLA.,A.R., GLICKMAN,B.W. and SPADARI,S. (1979) Mechanisms of mutation avoidance and mutation fixation. *Cold Spring Harbor Symp. Quant. Biol.* 43 937-946

RAVETCH,J.V., OHSUMI,M., MODEL,P., VOVIS,G.F., FISCHOFF,D. and ZINDER,N.D. (1979) Organization of a hybrid between phage f1 and plasmid pSC101. *Proc. Natl. Acad. Sci. USA* 76 2195-2198

REED,R.R. (1981a) Resolution of cointegrates between transposons γδ and Tn*3* defines the recombination site. *Proc. Natl. Acad. Sci USA* 78 3428-3432

REED,R.R. (1981b) Transposon mediated site-specific recombination: a defined *in vitro* system. *Cell* 25 713-719

REED,R.R. and GRINDLEY,N.D.F. (1981) Transposon mediated site-specific recombination: DNA cleavage and protein-DNA linkage at the recombination site. *Cell* 25 721-729

REED,R.R., YOUNG,R.A., STEITZ,J.A., GRINDLEY,N.D.F. and GUYER,M.S. (1979) Transposition of the *E.coli* insertion element γδ generates a 5bp repeat. *Proc. Natl. Acad. Sci. USA* 76 4882-4886

RHODES,D. and KLUG,A. (1980) Helical periodicity of DNA determined by enzyme digestion. *Nature* 286 573-578

ROBERTS,J.W., ROBERTS,C.W., CRAIG,N.L. and PHIZICKY,E.M. (1979) Activity of the *E.coli recA* gene product. *Cold Spring Harbor Symp. Quant. Biol.* 43 917-920

ROSENBERG,M. and COURT,D. (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. *Ann. Rev. Genet.* 13 319-353

ROSS,W., LANDY,A., KIKUCHI,Y. and NASH,H.A. (1979) Interaction of Int protein with specific sites on lambda *att* DNA. *Cell* 18 297-307

RUBIN,G.M., BROREIN,W.J., DUNSMUIR,P., FLAVELL,A.J., LEVIS,R., STROBEL,E., TOOLE,J.J. and YOUNG,E. (1981) *Copia*-like transposable elements in the *Drosophila* genome. *Cold Spring Harbor Symp. Quant. Biol.* 45 619-628

SANCAR,G.B., SANCAR,A., LITTLE,J.W. and RUPP,W.D. (1982) The *uvrB* gene of *E.coli* has both *lexA*-repressed and *lexA*-independent promoters. *Cell* 28 523-530

SAPIENZA,C. and DOOLITTLE,W.F. (1981) Genes are things you have whether you want them or not. *Cold Spring Harbor Symp. Quant. Biol.* 45 177-181

SANZEY,B.G. and COSSART,P. (1982) Homologies between different prokaryotic DNA-binding regulatory proteins and between their sites of action. *EMBO J.* 1 591-595

SHAPIRO,J.A. (1979) Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. *Proc. Natl. Acad. Sci. USA* 76 1933-1937

SHERRATT,D.J., ARTHUR,A.A. and BURKE,M.(1981a)Transposon-specified, site-specific recombination systems. *Cold Spring Harbor Symp. Quant. Biol.* 45 275-281

SHERRATT,D.J., ARTHUR,A.A. and DYSON,P. (1981b) Site-specific

recombination. *Nature* 294 608-610

SHIBATA,T., CUNNINGHAM,R.P., DASGUPTA,C. and RADDING,C.M. (1979a) Homologous pairing in genetic recombination: complexes of RecA protein and DNA. *Proc. Natl. Acad. Sci. USA* 76 5100-5104

SHIBATA,T., DASGUPTA,C., CUNNINGHAM,R.P. and RADDING,C.M. (1979b) Purified *E.coli* RecA protein catalyzes homologous pairing of superhelical DNA and single-stranded fragments. *Proc. Natl. Acad. Sci. USA* 76 1638-1642

SHIMOTOHNO,K. and TEMIN,H.M. (1981) Evolution of retroviruses from cellular movable genetic elements. *Cold Spring Harbor Symp. Quant. Biol.* 45 719-730

SHINE,J. and DALGARNO,L. (1974) The 3'-terminal sequence of *E.coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* 71 1342-1346

SHOEMAKER,C., GOFF,S., GILBOA,E., PASKIND,M., MITRA,S.W. and BALTIMORE,D. (1981) Structure of cloned retroviral circular DNAs: implications for virus integration. *Cold Spring Harbor Symp. Quant. Biol.* 45 711-717

SIGAL,N. and ALBERTS,B. (1972) Genetic recombination: the nature of a crossed strand-exchange between two homologous DNA molecules. *J. Mol. Biol.* 71 789-796

SILVERMAN,M., ZIEG,J., HILMEN,M. and SIMON,M. (1979) Phase variation in *Salmonella* : genetic analysis of a recombinational switch. *Proc. Natl. Acad. Sci. USA* 76 391-395

SIMON,M., ZIEG,J., SILVERMAN,M., MANDEL,G. and DOOLITTLE,R. (1980) Phase variation: evolution of a controlling element. *Science* 209 1370-1374

SOLLNER-WEBB,B., and FELSENFIELD,G. (1977) Pancreatic DNase cleavage sites in nuclei. *Cell* 10 537-547

SOLLNER-WEBB,B., MELCHOIR,Jr., W. and FELSENFIELD,G. (1978) DNase I,

DNase II and Staphylococcal nuclease cut at different, yet symmetrically located,sites in the nucleosome core. *Cell* 14 611-627

STARLINGER,P. and SAEDLER,H. (1972) Insertion mutations in microorganisms. *Biochimie* 54 177-185

STÜBER,D. and BUJARD,H. (1981) Organization of transcriptional signals in plasmids pBR322 and pACYC184. *Proc. Natl. Acad. Sci. USA* 78 167-171

SUGINO,A., HIGGINS,N.P. and COZZARELLI,N.R. (1980) DNA gyrase subunit stoichiometry and the covalent attachment of subunit A to DNA during DNA cleavage. *Nucl. Acids Res.* 8 3865-3874

SUTCLIFFE,J.G. (1978) pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. *Nucl. Acids Res.* 5 2721-2728

TACON,W., BHAMRA,B., SUNAR,B. and SHERRATT,D.J. (1981) Structure and function of plasmid ColK. *Plasmid* 6 358-359

TAYLOR,A. and SMITH,G.R. (1980a) Unwinding and rewinding of DNA by the RecBC enzyme. *Cell* 22 447-457

TAYLOR,A and SMITH,G.R. (1980b) Unwinding and rewinding of DNA by *exoV*.In *Mechanistic studies of DNA replication and genetic recombination*, ed. Alberts,B., pp 909-918. Academic Press

TELANDER-MUSKAVITCH,K.M. and LINN,S. (1980) Electron microscopy of *E.coli* RecBC enzyme reaction intermediates. In *Mechanistic studies of DNA replication and genetic recombination*, ed. Alberts,B., pp 901-908. Academic Press.

TONEGAWA,S., SAKANO,H., MAKI,R., TRAUNECKER,A., HEINRICH,G., ROEDER, W. and KUROSAWA,Y. (1981) Somatic reorganization of immunoglobulin genes during lymphocyte differentiation. *Cold Spring Harbor Symp. Quant. Biol.* 45 839-857

TOUSSAINT,A. and FAELEN,M. (1973) Connecting two unrelated DNA sequences with a Mu dimer. *Nature* 242 1-4

TSE,Y.C., KIRKEGAARD,K. and WANG,J.C. (1980) Covalent bonds between protein and DNA. *J. Biol. Chem.* 255 5560-5564

TU,C.P. and COHEN,S.N. (1980) Translocation specificity of the Tn*3* element: characterization of sites of multiple insertions. *Cell* 19 151-160

WANG,J.C. (1974) The degree of unwinding of the DNA helix by ethidium: 1. Titration of twisted PM2 DNA molecules in alkaline CsCl density gradients. *J. Mol. Biol.* 89 783-801

WANG,J.C., GUMPORT,R.I., JAVAHERIAN,K., KIRKEGAARD,K., KLEVAN,L., KOTEWICZ,M.L. and TSE,Y.C. (1980) DNA topoisomerases. In *Mechanistic studies of DNA replication and genetic recombination*, ed. Alberts,B., pp 769-784. Academic press

WARREN,G.J., TWIGG,A.J. and SHERRATT,D.J. (1978) ColE1 plasmid mobility and relaxation complex. *Nature* 274 259-261

WEINSTOCK,G.M., McENTEE,K. and LEHMAN,I.R. (1979) ATP-dependent renaturation of DNA catalyzed by the RecA protein of *E.coli*. *Proc Natl. Acad. Sci. USA* 76 126-130

WEST,S.C., CASSUTO,E., and HOWARD-FLANDERS,P. (1981a) Homologous pairing can occur before DNA strand separation in general genetic recombination. *Nature* 290 29-33

WEST,S.C., CASSUTO,E. and HOWARD-FLANDERS,P. (1981b) Heteroduplex formation by RecA protein: polarity of strand exchanges. *Proc. Natl. Acad. Sci. USA* 78 6149-6153

WINTER,R.B., BERG,O.G. and von HIPPEL,P.H. (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The *E.coli lac* repressor-operator interaction: kinetic measurements and conclusions. *Biochem.* 20 6961-6977

WINTER,R.B. and von HIPPEL,P.H. (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 2. The *E.coli lac* repressor-operator interaction: equilibrium measurements. *Biochem.* 20 6948-6960