The Xer Site-Specific Recombination System

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

by

Jennifer Ann Roberts

Institute of Genetics
University of Glasgow
Church Street
Glasgow
G11 5JS

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

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

This thesis is dedicated with love and thanks
to my parents.

Contents

Acknowledgements

Many people are due thanks for making this finished article a reality. Firstly, I must thank my supervisor, Dave Sherratt for his boundless enthusiasm about all things Xer, and constructive criticism of this thesis.

To Mary, my endless gratitude for keeping the sixth floor sane (almost), for making everything run like clockwork, for her support, and for being a friend. Thanks also to the Ladies and to the Secretaries for making everything so much more simple.

Very special thanks to Richard, Dave, Garry and Marshall who read various Chapters, to Marshall for lending me his computer, to Karen who carried out the statistical analysis for me one lunch-time, and to those people who listened to my ideas and told me their own, especially Dave, Richard, Marshall, Gerhard, Stephen, Martin and Karen. To the same people, love and thanks for being much more than colleagues.

Thanks to the "PTea Club" for coming when they get the nod, for paying up, for plenty of laughs, and Serious Scientific Discussion, and for letting me resign when I did (and special thanks to Michael for taking over) - they made each 11 o'clock and 4 o'clock enjoyable. Also included in these thanks are Martin McBride and Simon "Rudeboy" McCartin. Thanks to the other inhabitants of lab 66 over the last three years -namely Stephen, Sally, Gerhard and Nick for cheerfully tolerating my ranting - and to all of the other sixth and seventh floor members, who made the sentence endurable.

Thanks to the other members of the Genetics department who made mass gatherings worthwhile. I would especially like to thank Diane, Leslie, Jo and Audrey for actually meaningful chats over lunch and elsewhere.

Love and thanks to friends and family for their support, for listening, and for making time spent away from the lab happy. Special thanks to Marc for being Marc, and for rescuing me from eight-legged menaces.

My deepest love, thanks, and gratitude go to Dave, for always being there for me, for his unflinching support, for making me laugh, for plenty of hugs, for Troll earrings and for giving me something to look forward to.

Finally, I am grateful to the MRC for the studentship that (nearly) supported me during the course of this work.

Abbreviations

Chemicals

APS	ammonium persulphate
ATP	adenosine triphosphate
BSA	bovine serum albumin
CTAB	cetyltrimethylammonium bromide
DNA	2′ deoxyribonucleic acid
dNTP	2′ deoxyribonucleotide
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid (disodium salt)
IPTG	isopropyl-β-*D*-thiogalactoside
ONP	*o*-nitrophenol
ONPG	*o*-nitrophenyl-β-*D*-galactopyranoside
SDS	sodium dodecyl sulphate
TEMED	NNN′N′ tetramethylethylenediamine
Tris	tris-(hydroxymethyl)-aminoethane
X-Gal	5-bromo-4-chloro-3-indolyl-β-*D*-galactoside

Genotypes and phenotypes

xerC	the gene encoding XerC; a strain mutant in *xerC*
xerC$^+$	a strain wild-type for *xerC*
X^r	resistant to X
X^s	sensitive to X

Measurements

bp	base pair
kb	kilobase pair (10^3 bp)
kD	kilodalton (10^3 dalton)
A_x	absorbance at x nm
~x bp	approximately x bp

Miscellaneous

pfu	plaque forming units
UV	ultraviolet light
ORF	open reading frame

Summary

Three major themes were studied during the course of this work.

The role of the XerC binding site in resolution selectivity

Site-specific recombination catalysed by the *Escherichia coli* proteins XerC and XerD is required for stable inheritance of natural multi-copy plasmids at cell division, and for resolution of chromosome dimers to allow segregation of circular chromosomes. Recombination occurs in both cases at ~30 bp core recombination sites of similar sequence, which can be divided into a left arm (where XerC binds), a right arm (where XerD binds), and a central region of 6-8 bp between the two arms. Sites that are involved in plasmid resolution, such as *cer* from plasmid ColE1, also require 220 bp of accessory DNA, and two proteins ArgR (the arginine repressor) and PepA (aminopeptidase A) to permit recombination to occur.

To resolve plasmid dimers, thereby ensuring their heritable stability, site-specific recombination must occur preferentially between sites that are present in direct repeat within a DNA molecule; the molecular mechanisms underlying this "resolution selectivity" were investigated by studying recombination substrates with different left arm sequences. A series of plasmids was constructed varying in sequence at four positions in the left arm, and their abilities to recombine intermolecularly (i.e. without resolution selectivity) was assayed. One base within the left arm was found to be especially important in determining whether or not a site could support intermolecular recombination, but this effect was modified by the other three variant bases in the left arm. Six (of a total of sixteen) sites were analysed further. It was established that XerD bound equivalently to all of the six sites in the absence of XerC, but that the DNA-XerC/D complexes formed *in vitro* had different electrophoretic mobilities in gel retardation assays; this was proposed to be as a result of differences in protein-induced DNA bending. The relative affinities of XerC/D for the sites also varied depending on left arm sequence. The relative mobilities of the DNA-XerC/D complexes were correlated with both the relative affinities of the sites for XerC/D, and with their abilities or inabilities to support intermolecular recombination. It was therefore suggested that the protein-DNA conformation, and the affinity of the recombinases for a recombination site, could be affected by the sequence of the left arm, and that this affects whether or not a site is effective in resolution of plasmid dimers.

The role of Fis on Xer recombination

The role of the DNA-binding protein Fis on Xer recombination was investigated. It was established that recombination at an XerC/D-substrate, *dif*, was reduced in a *fis* mutant strain. Fis plays a variety of cellular roles, both in other recombination systems, and as a transcriptional and replication effector, therefore experiments were performed to establish whether the effect of Fis during Xer recombination is exerted directly at the recombination site, or indirectly. The results established that Fis acts indirectly. A *fis* derivative of an *xerC*:: mini Mu*d* *lacZ* strain was constructed, and expression of *lacZ* in the two strains was compared using β-galatosidase assays. The assays indicated that expression of *xerC* may be significantly reduced in a *fis* mutant, implying that Fis might stimulate expression of *xerC*.

A screen for additional *xer* mutants

A Tn*5* mutagensis of the *E. coli* chromosome was undertaken, with two major aims.

We are currently unable to detect products of Xer recombination *in vitro*, although it is possible to recreate a single strand-exchange either to produce or to resolve Holliday junctions. This has led to concern that there may be additional Xer proteins awaiting discovery, or that the existing reaction conditions are inappropriate. By performing a mutagenesis it was hoped that it would be possible to screen for novel *xer* mutants not falling into any of the four known complementation groups (*argR*, *pepA*, *xerC*, *xerD*). A reporter plasmid was constructed for screening of putative mutants, consisting of two *psi* sites (from plasmid pSC101) in direct repeat. This plasmid was utilised because recombination at *psi* was found to be ArgR-independent, but dependent on PepA, XerC and XerD; however, deletion analysis of *psi* indicated that another protein might be required to bind to this site to perform a similar role to that of ArgR during recombination at *cer*. The mutagenesis was therefore also undertaken with the second major objective of identifying the protein binding to this site. A mutant strain, JR29, was identified that failed to recombine at *psi* and at *cer*, but recombined at *dif*. JR29 was not complemented by XerC or by XerD, but was partially complemented by PepA. *In vivo* and *in vitro* assays also indicated that this strain was a *pepA* mutant; however, P1 transduction of the Tn*5* kanamycin resistance allele into another strain failed to produce a *pepA* mutant, but created a strain, JR30, in which recombination appeared to be at a lower rate. Therefore JR29 appears to contain two mutations: one in *pepA*, the other in an unidentified locus.

Chapter One

Introduction

Site-Specific Recombination

During site-specific recombination, DNA molecules are cleaved at precise positions, and their ends exchanged to produce recombinants. As a result of recombination, the DNA between the recombination sites is rearranged leading to a variety of structural and biological consequences. Site-specific recombination is responsible for a range of biological events, including the integration and excision of lambdoid bacteriophage into and out of the bacterial chromosome (Landy, 1989), resolution of transposon-mediated co-integrate molecules (Sherratt, 1989), resolution of multimeric plasmids and circular chromosomes to monomers (Austin *et al.*, 1981; Summers & Sherratt, 1984; Blakely *et al.*, 1991), inversion of DNA sequences to "switch" gene expression (Glasgow *et al.*, 1989), and control of plasmid copy number (Futcher, 1986).

Site-specific recombination occurs between short DNA sequences, which contain the recognition motifs for the recombinase protein(s) that mediate strand-exchange. These are the core recombination sites (or crossover sites). The recombinase proteins recognise and bind to the core recombination sites through these motifs, and bring the two recombining sites together by means of protein-protein and protein-DNA interactions to form a nucleoprotein complex (a synaptic complex or synapse). Once the sites are appropriately aligned in a synaptic complex, the DNA in both sites is cleaved at precise positions, and the ends are exchanged and re-ligated; depending on the system involved, strand-exchange can occur *via* a pair of double-strand cleavages, or two pairs of single-strand cleavages. No high-energy co-factors are required for strand-exchange, as there is no hydrolysis of phosphodiester bonds: at strand cleavage, the bond energy is preserved by formation of covalent intermediates in which the DNA is linked to the recombinase (catalysis of site-specific recombination is reviewed in Stark *et al.*, 1992).

For recombination to occur, however, many systems require additional "accessory factors", consisting of DNA outwith the core recombination site, and accessory proteins. The recombinase itself can act as an accessory protein by binding to sequences outwith the core recombination site (for example, Tn3 *res*/resolvase, see below); alternatively, other proteins, often with distinct cellular roles, can be required. Such accessory proteins are often DNA-binding proteins, such as Fis, HU, IHF, or the arginine repressor ArgR (Landy, 1989; Haykinson & Johnson, 1993; Stirling *et al.*, 1988b); however, the aminopeptidase PepA is required for recombination at ColE1 *cer* (Stirling *et*

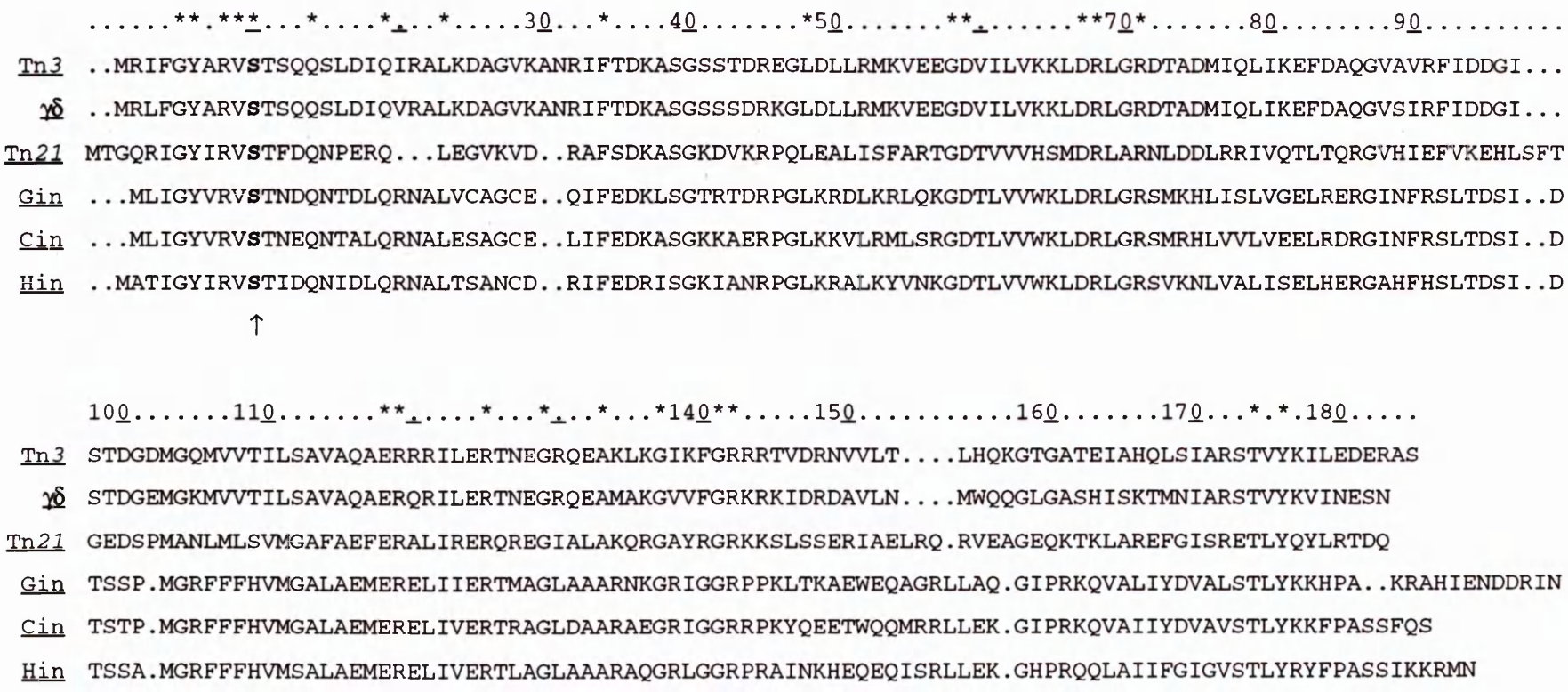


Figure 1.1:Alignment of predicted amino acid sequence of representative members of the resolvase and DNA invertase family of site-specific recombinases.

Position numbering is of Tn3 resolvase sequence. Every tenth position is underlined. Where possible, numbers are given. Identical residues are denoted by an asterisk. The active site serine (serine 10 in Tn3 resolvase) is denoted by an arrow (modified from Blake, 1993).

al., 1989), but has not been demonstrated to bind to DNA *in vitro* (G. Szatmari, unpublished observations).

The two classes of site-specific recombination system

Two families of site-specific recombination enzyme can be distinguished by homologies between the proteins, as well as recombination site organisation, and strand-exchange mechanism. The main characteristics of both families of enzymes are summarised below.

(a) Tn*3* resolvase family

The enzymes in this class are highly conserved in sequence, and are typically small proteins, of around 20 kD (although one 57 kD protein with homology to the resolvase family has been characterised; Sato *et al.*, 1990). Examples of proteins from this class are resolvases from the transposable elements Tn*3*, γδ and Tn*21*, and the DNA invertases Hin, Gin and Cin (reviewed in Hatfull & Grindley, 1988; Stark *et al.*, 1989b; Sherratt, 1989; Johnson, 1991; see Figure 1.1). These proteins are divisible into two functional domains: the DNA-binding domain is located in the carboxyl terminal, while the catalytic site is in the amino terminal of the protein (Abdel-Meguid *et al.*, 1984; Newman & Grindley, 1984; Klippel *et al.*, 1988a, b).

All of these recombinases catalyse strand-exchange within core recombination sites of ~30 bp, with imperfect dyad symmetry, but other DNA sequences are additionally required to permit recombination. For example, Tn*3* resolvase binds to three subsites within its recombination site, *res*, (Grindley *et al.*, 1982; see Figure 1.2). Subsite I is the core recombination site, where strand-exchange occurs, while the other two subsites are accessory DNA sites. It has recently been established that each subsite of Tn*3* *res* binds two monomers of resolvase, so that each recombination site requires a total of six monomers of resolvase (Blake, 1993). In this case, resolvase acts as an accessory protein as well as the recombinase. For efficient DNA inversion, the invertases Hin, Gin and Cin require a recombinational enhancer *in cis*, the recognition site of the DNA-binding protein, Fis. It is believed that Fis bound to the enhancer interacts directly with invertase bound at the core recombination site, stimulating site-specific inversion (Finkel & Johnson, 1992; Figure 1.2).

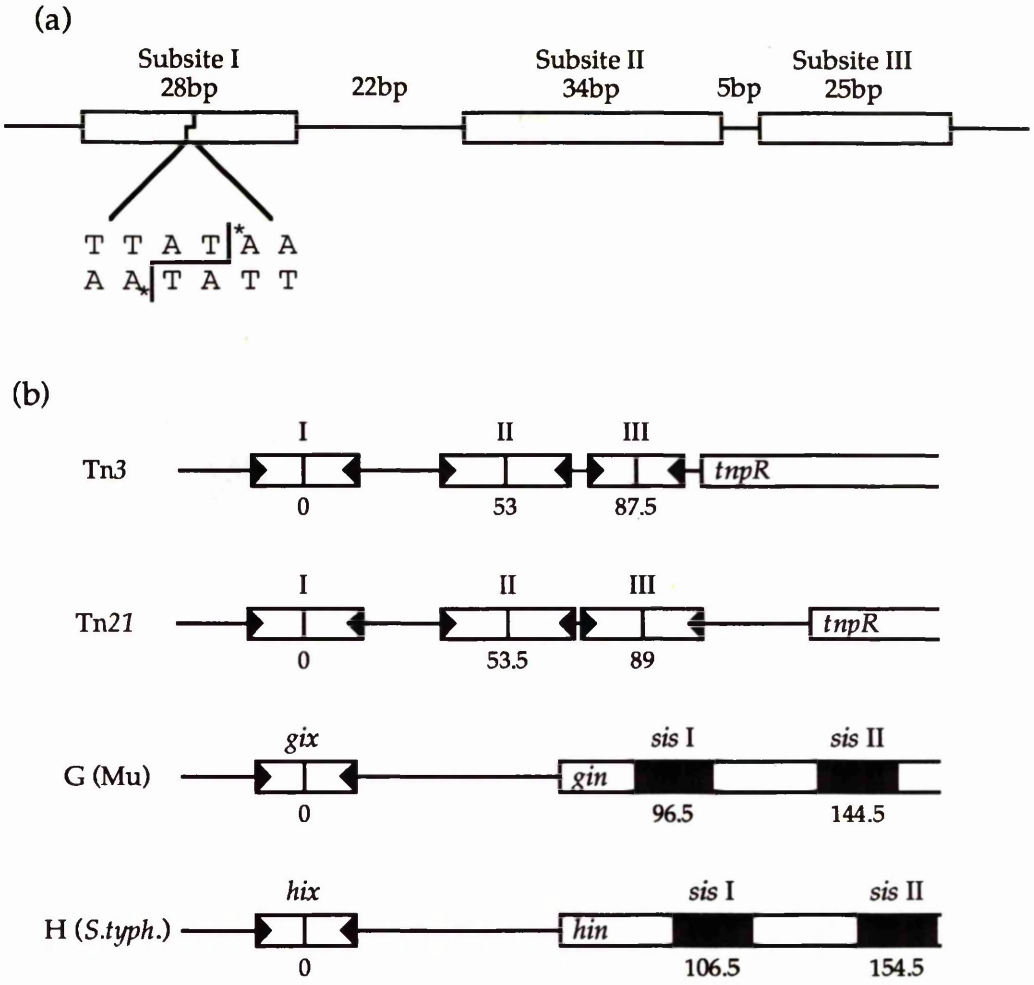


Figure 1.2: Organisation of recombination sites of members of the resolvase/invertase family.

(a) Structure of Tn3 *res*. Subsites I, II and III are indicated by boxes and the position of DNA cleavage by resolvase is also shown. An asterisk is used to denote positions at which resolvase becomes covalently linked to the DNA. All sizes/distances are in base pairs.

(b) Subsites of *res* are arranged as shown, where subsite I is the core recombination site. For the inversion sites both the core recombination sites (*gix* or *hix*: unfilled boxes) and enhancer sites (*sis*: filled boxes) are shown. Phosphodiester co-ordinates are defined from the centre of the core recombination sites. The number below each site indicates the co-ordinate of the site centre, and the (sub)site name is shown above. Arrowheads denote partial inverted repeat sequences within a site. The ends of the recombinase genes (i.e. *tnpR, gin, hin*) are also shown where possible.

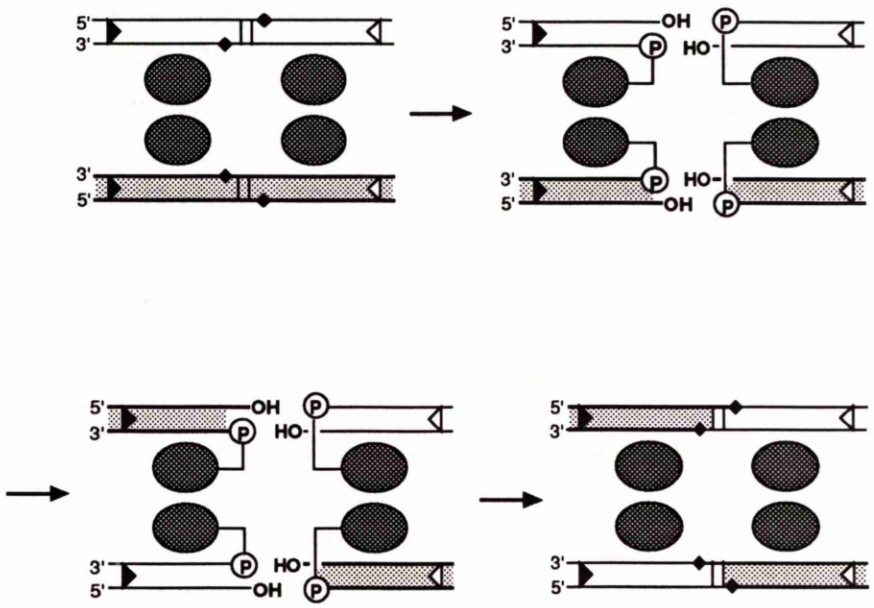


Figure 1.3: Events at the core recombination sites during catalysis of recombination by members of the resolvase family.

Recombinase subunits are represented by shaded ovals; the ends of the core recombination sites by inverted black and white arrowheads; the two base pairs in the overlap region by vertical lines; and the phosphates that are attacked by the enzyme by black diamonds. The DNA strands are shown as thick and thin lines, with shading to differentiate the two sites. (Adapted from Stark *et al.*, 1992).

Two recombining sites are brought together by protein-DNA (e.g. resolvase-*res*) and protein-protein (e.g. resolvase-resolvase) interactions , forming a synaptic complex. These interactions are also thought to ensure that the sites are correctly configured to allow recombination to proceed, by producing a substrate synapse of a particular topology (Stark *et al.*, 1989a, b; Bednarz *et al.*, 1990). Strand-exchange is accomplished by a pair of double-strand cleavages, leaving a 2 bp stagger, followed by a simple right-handed rotation of the ends by 180°, before ligation of the ends to new partners, forming recombinant products (see Figure 1.3). It is believed that the nucleophile responsible for strand cleavage is a serine residue near the amino terminal of the enzymes of this class, which becomes covalently attached to the 5′ end of the DNA at an intermediate step in recombination (Klippel *et al.*, 1988a; Hatfull & Grindley, 1986; see alignment in Figure 1.1).

(b) λ-integrase family

Members of this family of recombinases include λ-integrase (λ-Int), the *Escherichia coli* recombinases XerC and XerD, the bacteriophage P1 recombinase Cre and the *Saccharomyces cerevisiae* 2 μm plasmid FLP recombinase (Landy, 1989, Blakely *et al.*, 1993; Hoess & Abremski, 1984; Andrews *et al.*, 1985). This family of enzymes is more diverse than the resolvase family, with little conservation of recombinase sequence. The most highly conserved region is within two domains in the carboxyl-terminus, containing the active site. While the proteins of this class vary in size and sequence, the four amino acids believed to form the active site are highly conserved; an arginine in domain I, and an additional arginine, a histidine and a tyrosine in domain II (Argos *et al.*, 1986; Abremski & Hoess, 1992; see Figure 1.4). The conserved tyrosine residue is the nucleophile involved in strand cleavage, and forms a 3′ phosphotyrosyl bond with the DNA as an intermediate step during recombination (Gronostajski & Sadowski, 1985b; Pargellis *et al.*, 1988). Experiments with FLP mutants have indicated that the other three amino acids in the "invariant tetrad" contribute to the active site by activating the phosphodiester bond prior to cleavage (Lee & Jayaram, 1993). It has been shown that FLP bound at a recombination site is responsible for activation of the DNA at that site, but that it is a monomer bound at another site which is responsible for cleavage; in other words, activation is *in cis*, but cleavage is *in trans* (Chen *et al.*, 1992).

The proteins of this class mediate recombination at core sites of ~30 bp, which consist of two recombinase-binding arms with partial dyad symmetry,

Domain I

λ-Int	TGQ**R**VGDLCEMKWSD
FLP	NCG**R**FSDIKNVDPKS
Cre	TLL**R**IAEIARIRVKD
XerC	AGL**R**LSELVGLDIKH
XerD	TGL**R**VSELVGLTMSD

Domain II

λ-Int	**H**EL**R**SLSA-RLYEKYQ-ISDKFAQLLGHKS-DTMA-SQ**Y**RD
FLP	**H**IG**R**HLMTSFLSMKGLTELTNVVGNWSDKRASAVARTT**Y**TH
Cre	**H**SA**R**VGAARDMARAG-VSIPEIMQAGGWTN-VNIV-MN**Y**IR
XerC	**H**KL**R**HSFATHMLESS-GDLRGVQELLGHAN-LSTT-QI**Y**TH
XerD	**H**VL**R**HAFATHLLNHG-ADLRVVQMLLGHSD-LSTT-QI**Y**TH

Figure 1.4: Alignment of primary sequences of representatives of the λ-integrase family of site-specific recombinases.

The two most conserved domains of some of the λ-integrase class proteins are shown. The four amino acids that make up the active site (the invariant tetrad) are shown in bold-face font; the conserved tyrosine is the nucleophile that catalyses strand cleavage. The positions of the conserved domains vary within the protein sequences, although domain I is always located nearer the N-terminal than domain II. XerC and XerD are 37% identical throughout their primary sequence, and 70% identical in domains I and II. This alignment was modified from McCulloch, 1992.

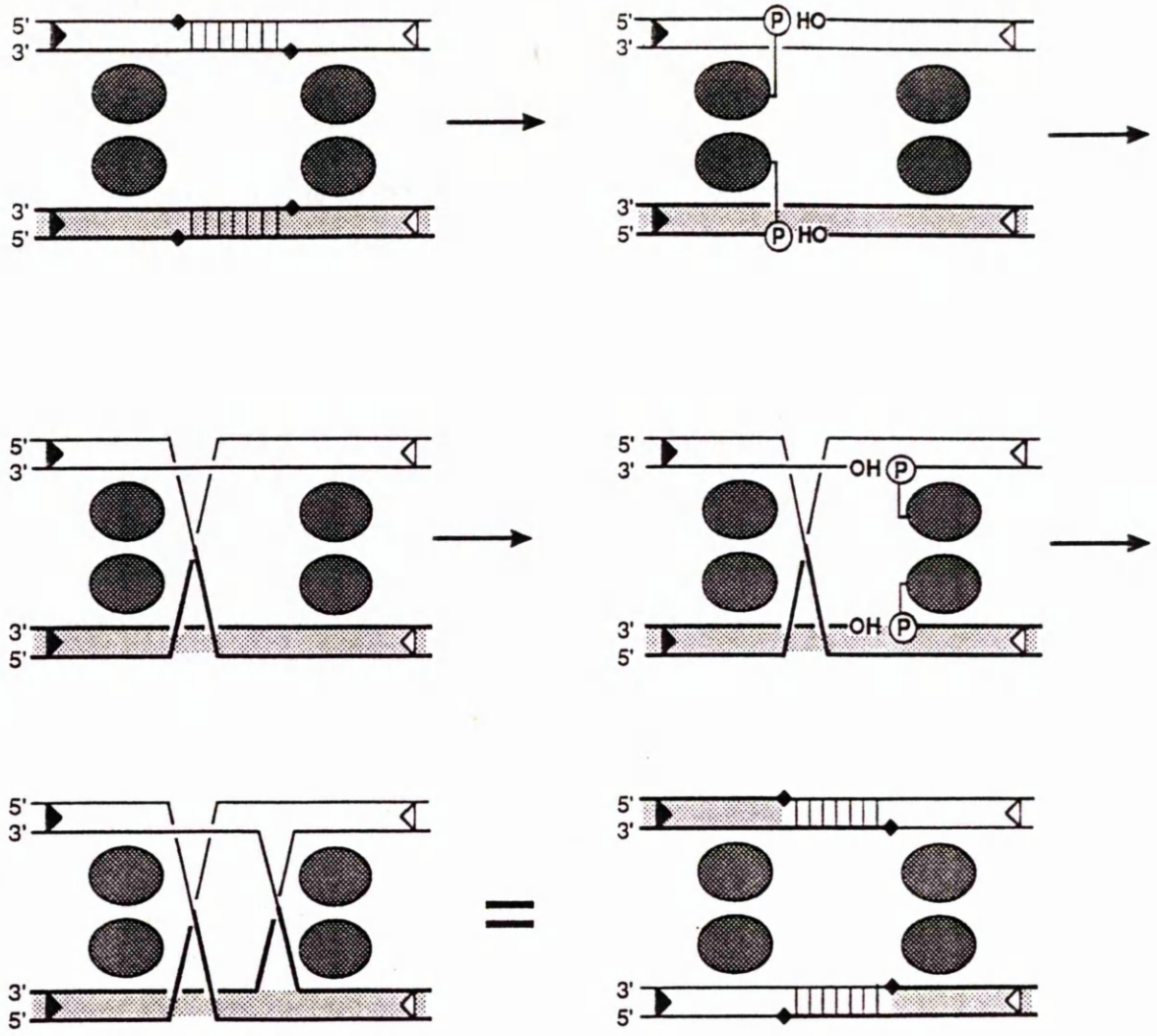


Figure 1.5: Events at the core recombination site during recombination mediated by λ-integrase family recombinases.
Recombinase subunits are represented by shaded ovals; the ends of the core recombination sites by inverted arrows; the 6-8 base-pairs between the positions of strand cleavage (the central region or overlap region) as vertical lines; and the phosphodiester bonds that are attacked by the recombinases as black diamonds. A single DNA strand (thick or thin line) from each core recombination site (the sites are differentiated by shading) is cleaved and covalently-linked protein-DNA intermediates are formed. The strands are exchanged and a Holliday intermediate is created. Recombinant product is formed by branch migration of the Holliday junction across the overlap region, and a second pair of strand-exchanges. It is not intended that this Figure imply that two four-way junctions are present simultaneously. In this Figure, the recombinase subunits are shown to cleave the DNA strand to which they are bound: this is cleavage *in cis*. However, experimental evidence with FLP indicates that cleavage might occur *in trans*, i.e. a recombinases subunit cleaves the phosphodiester bond diagonally opposite to its position of binding. This Figure was adapted from Stark *et al.*, 1992.

separated by a region of 6-8 bp. Recombination proceeds *via* two pairs of single-strand cleavages, exchanges and reunions (see Figure 1.5), through a Holliday junction intermediate, in which one pair of strands has been exchanged. These intermediates have been isolated *in vitro* and *in vivo* (Holliday, 1964; Hoess & Abremski, 1987; Nunes-Duby *et al.*, 1987; Jayaram *et al.*, 1988; Meyer-Leon *et al.*, 1988, 1990; McCulloch, 1992; McCulloch *et al.*, 1994a). The two pairs of strand-exchanges occur at either end of the 6-8 bp central region, resulting in a recombinant molecule in which one strand has been derived from each parent molecule in this region; for this reason, this region has been termed the "overlap" region for the systems in which the positions of strand cleavage have been determined (Landy, 1989).

Some enzymes of this class can mediate recombination at sites consisting of only core recombination sites, while others require accessory DNA sequences and proteins. For example, the minimal FRT site required for recombination by FLP *in vivo* and *in vitro* consists of two 13 bp FLP binding sites flanking an overlap region of 8 bp (Gronostajski & Sadowski, 1985a; Andrews *et al.*, 1985; Qian *et al.*, 1990); in contrast, λ-integrase requires an array of protein-binding sites and accessory proteins before integrative or excisive recombination will proceed (Landy, 1989; see Figure 1.6). The *E. coli* proteins XerC and XerD can perform recombination in the absence of accessory DNA and proteins, but sometimes accessory factors are required (Summers, 1989; Blakely *et al.*, 1991; Blakely *et al.*, 1993); this depends on the sequence of the core recombination site, as discussed in more detail in Chapter 3.

Selectivity in site-specific recombination

The initial organisation of two recombination sites determines the rearrangement event that occurs (Sadowski, 1986; Gellert & Nash, 1987; Craig, 1988). All wild-type sites have an intrinsic asymmetry in sequence so that during recombination the right half of one recombination site is joined to the left half of the other and *vice versa*. This ensures that recombination between sites in direct repeat in the substrate only results in resolution, while sites in inverted repeat recombine to invert the intervening DNA (Stark *et al.*, 1989b). Since two recombination sites can be arranged in three possible configurations, three types of rearrangement are possible, as illustrated in Figure 1.7. However, many site-specific recombination systems use

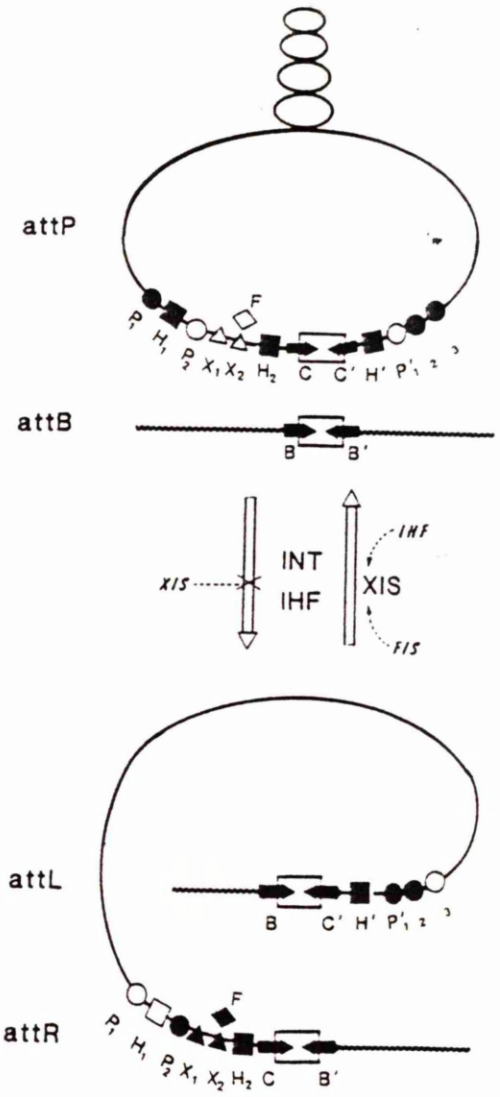


Figure 1.6: Bacteriophage λ integration and excision.
Bacteriophage λ requires accesory DNA sequences and proteins for both integrative and excisive pathways. Integration of λ into the bacterial chromosome requires recombination between the phage site *attP* and the chromosomal site *attB*. *attB* consists of two core-type Int binding sites flanking an overlap region of 7 bp; this is a λ core recombination site. *attP* has an array of additional protein binding sites. Binding sites for Int are shown as P, C or B; P are arm-type Int binding sites (Int has two binding domains: Moitoso de Vargas *et al.*, 1988), C and B are core-type binding sites. IHF binds to H sites, Xis binds to X sites, and Fis binds to F sites. Integration is stimulated by binding of Int and IHF, and inhibited by Xis. Integration resulting from recombination between *attB* and *attP* generates a pair of different sites, *attR* and *attL*. Recombination between *attL* and *attR* requires Int, IHF, Xis and Fis and results in phage excision, and recreates *attP* and *attB*. Under normal conditions *attP* recombines only with *attB*, and *attR* only with *attL*, therefore integration and excision show directionality. This diagram was adapted from McCulloch, 1992.

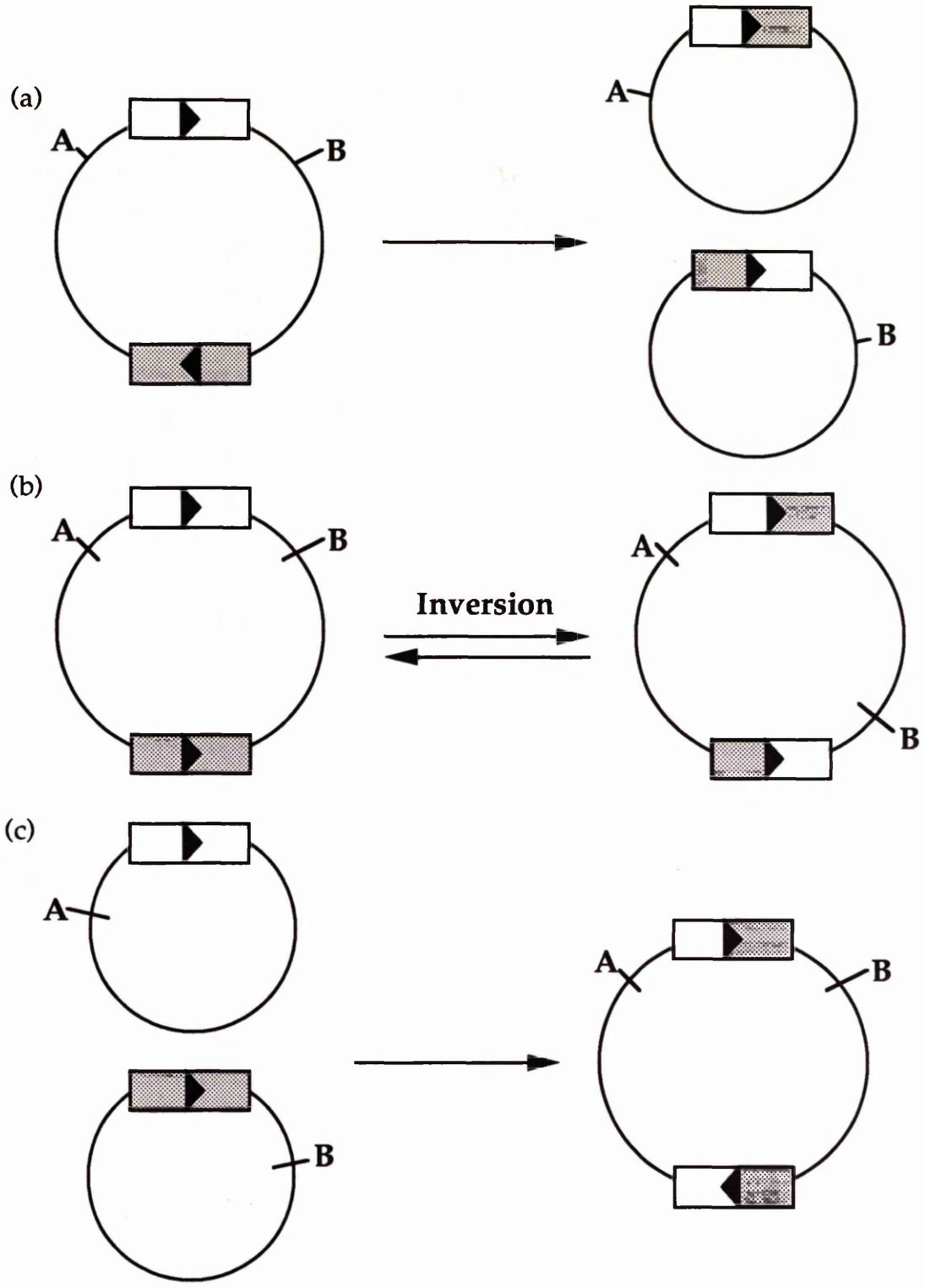


Figure 1.7: DNA rearrangements as a result of site-specific recombination. Three substrate site organisations are possible. In (a), two sites are in direct repeat in a circular molecule, and recombination between them results in the formation of a two circles each with one site; this reaction is called resolution, excision or deletion. In (b), two sites within a molecule are in inverted repeat, and recombination between tham results in inversion of the intervening DNA. In (c), two circular molecules each with a single recombination site recombine together by an intermolecular reaction to produce a single DNA circle with two sites in direct repeat; this reaction is termed fusion or integration, although in this thesis it is simply termed intermolecular recombination, and produces plasmid multimers.

mechanisms to ensure that one particular substrate configuration of sites is preferred over others, so that recombination will result in a controlled DNA rearrangement. This process has been termed selectivity (Bednarz *et al.*, 1990). Some examples of the mechanisms used to ensure reaction selectivity are described below, and are illustrated in Figure 1.8.

(a) Resolution catalysed by resolvases: To favour resolution reactions over DNA inversions or intermolecular fusions, resolvases preferentially recombine sites that are arranged in direct repeat in the substrate molecule (such as in a transposon-mediated co-integrate). It is believed that the accessory subsites of *res*, bound with additional units of resolvase, are responsible for ensuring that this configuration of sites undergoes recombination in preference to sites within different molecules or sites in inverted repeat within a molecule (Stark *et al.*, 1989a, b; Bednarz *et al.*, 1990). Models have been proposed for how this might be achieved, taking into account the topology changes produced as a result of resolution reactions catalysed by resolvase *in vitro* (Wassermann *et al.*, 1985; Boocock *et al.*, 1987; Stark *et al.*, 1989a, b; see Figure 1.8). In a negatively supercoiled circular DNA molecule, two *res* sites arranged in direct repeat are believed to be brought together as a result of interactions between resolvase units bound at the three subsites of *res*. The productive synapse has been proposed to have a specific topology, formed as a result of interactions between the three subsites of *res*, in which three interdomainal negative supercoils are trapped. Recombination occurs as a result of two double-strand cleavages, followed by a simple right-handed 180° rotation of the ends, which are then re-ligated; this is energetically favourable, since negative supercoiling is released as a consequence of the rotation. This specific topology is not achieved if the sites are in inverted repeat. Resolvases therefore show resolution selectivity, as sites in direct repeat are their favoured recombination substrates.

(b) Inversion reactions catalysed by DNA invertases: In contrast to the reactions catalysed by resolvases, the related DNA-invertases show a specificity for inversion reactions, by preferentially recombining sites in inverted repeat within a negatively supercoiled molecule. The invertases catalyse recombination at core recombination sites, with no requirement for additional invertase binding sites. However, to permit efficient recombination, an enhancer, the site of binding by Fis, is physically included in the synaptic substrate, the invertasome (Kahmann *et al.*, 1985; Heichman & Johnson, 1990). For recombination catalysed by Hin, it has been proposed that substrate DNA supercoiling favours an alignment of inverted sites into a

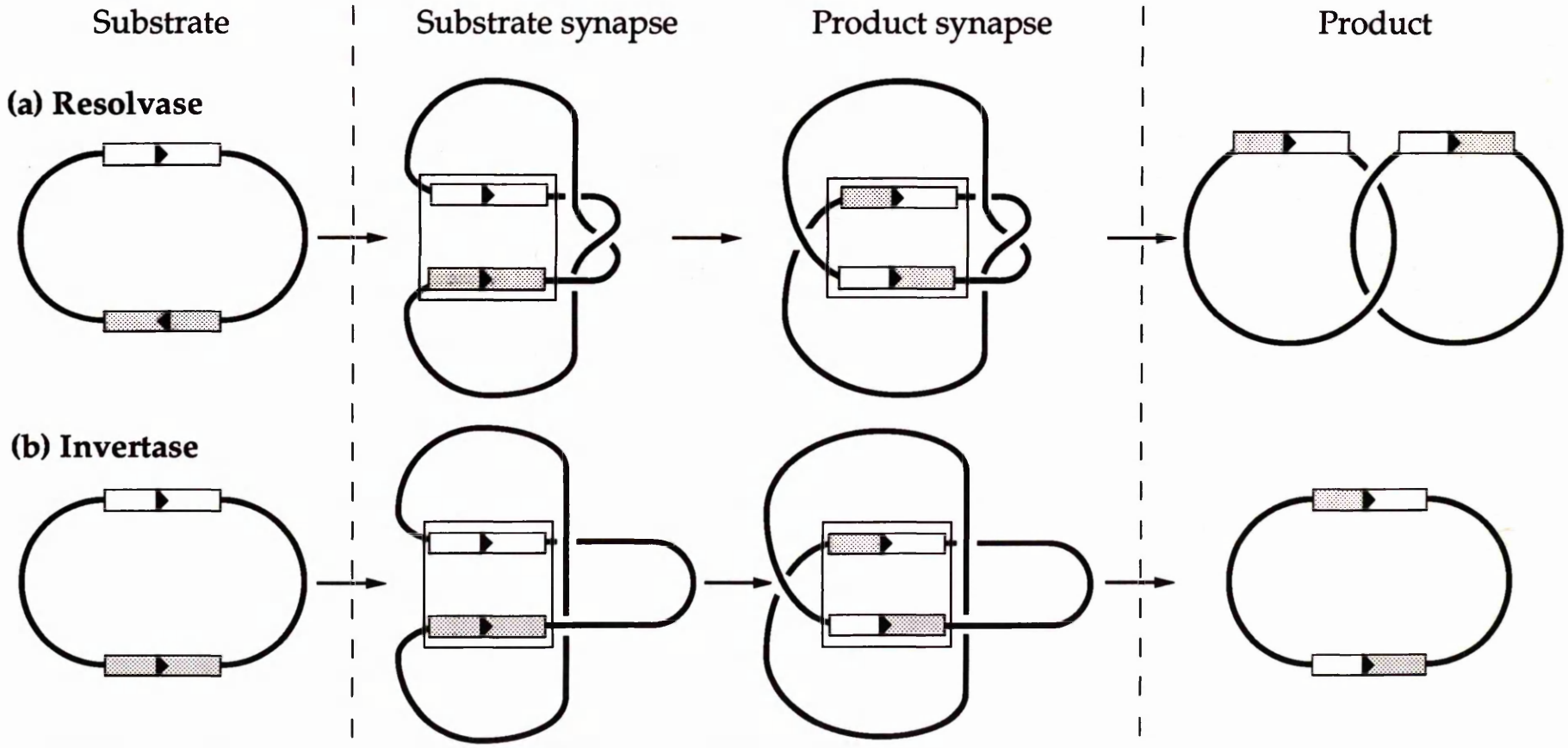


Figure 1.8: Recombination selectivity in the resolvase and invertase systems.
Topology of strand-exchange during recombination by (a) Tn3 resolvase; (b) Gin/Hin.
Boxes containing arrows represent core recombination sites. Recombinases and accessory sites/proteins are not shown. For *res*, interwrapping of the DNA strands is believed to be co-ordinated by resolvase bound at subsites II and III. The position (and function) of the Fis/*sis* enhancer in the invertasome is proposed to be analogous to resolvase bound at subsites II and III of *res*, i.e. to trap a specific number of (-) interdomainal supercoils between the core recombination sites (3 (-) supercoils for resolution, and 2 (-) supercoils in the case of inversion). (Adapted from Stark *et al.*, 1992).

synaptic complex in which two negative supercoils are trapped. Following strand cleavage and a 180° rotation of the DNA ends, inversion products are formed. Two determinants are believed to be required to ensure efficient inversion (Moskowitz *et al.*, 1991). The first is negative supercoiling of the substrate, which favours the formation of an invertasome structure in which a simple 180° rotation of the DNA ends will result in the formation of inversion products. The second determinant is asymmetry at the position of staggered cleavage within the core recombination site: this allows the orientation of sites to be tested during the strand-exchange process.

Deletion reactions can be catalysed at a low rate (< 10% of rate of inversion) by Hin *in vivo* and *in vitro* if the sites are in direct repeat and are well spaced (Scott & Simon, 1982; Moskowitz *et al.*, 1991). The extra spacing is proposed to allow additional DNA loops to form so that an interwrapped invertasome can be produced that would allow recombination to proceed.

(c) Integration and excision catalysed by λ-Int: While λ-Int catalyses both integration and excision reactions, each pathway shows selectivity (reviewed in Landy, 1989). λ integrates into the bacterial chromosome as a result of site-specific recombination between the phage locus *attP*, and the bacterial site *attB*. These sequences differ: *attB* consists of a core recombination site of ~30 bp, to which Int binds, while *attP* comprises a core recombination site and an array of sites, where the accessory proteins Int and IHF bind. Recombination between these two sites produces a pair of different sites, *attR* and *attL*, both of which have some of the accessory protein binding sites of *attP*. Recombination between *attR* and *attL* results in excision of the phage from the chromosome, and requires the accessory proteins IHF, Xis and Fis in addition to Int. Therefore the excisive and integrative pathways in λ recombination are not mere reversals of one another: they are distinct in terms of the proteins and DNA sequences involved. Each pathway utilises different substrates, and can thereby show selectivity, for excision in one case and integration in the other (see Figure 1.6).

Not all site-specific recombination systems show selectivity for one pathway as opposed to another. For example, FLP and Cre both require simple recombination sites, without accessory DNA or proteins, and can catalyse intermolecular and intramolecular (both resolution and inversion) events depending on the initial site organisation (Sternberg *et al.*, 1983; Abremski *et al.*, 1983; Gronostajski *et al.*, 1985c; Abremski *et al.*, 1986; see Figure 1.8). The Xer site-specific recombination system will catalyse recombination events that

show selectivity at some sites, but not at others (Summers, 1989; Blakely *et al*, 1993; Sherratt *et al.*, 1993); this is discussed further in Chapter 3.

The Xer site-specific recombination system

Xer recombination at *cer*

The Xer site-specific recombination system was originally discovered due to its involvement in the stable inheritance of the natural *E. coli* plasmid ColE1 (Summers & Sherratt, 1984). ColE1 plasmid units are believed to randomly partition to daughter cells at host cell division; therefore the probability of each daughter cell inheriting at least one copy of the plasmid is highest when the plasmid is in its monomeric form, when the number of segregating units is at its highest (Durkacz & Sherratt, 1973; Summers & Sherratt, 1984; Summers *et al.*, 1993). Plasmid multimers formed by homologous recombination between plasmid monomers result in instability at host cell division as the number of segregating units is reduced; this instability can be counteracted by site-specific recombination to resolve multimers back to monomers.

Site-specific recombination in ColE1 occurs at a ~250 bp site, *cer* (*cer* sequence is shown in Figure 1.9; Summers & Sherratt, 1984, 1988). During site-specific recombination at *cer*, the strand-exchange events occur within a ~30 bp core recombination site (Summers *et al.*, 1985; Sherratt *et al.*, 1993). The remaining ~220 bp of *cer* sequence play an accessory role during recombination. Several features of the accessory DNA sequence are important for this role: firstly, it contains a binding site for the arginine repressor, ArgR, at around 130 bp upstream of the recombination core (Stirling *et al.*, 1988b), and secondly, it contains periodically spaced A-tracts that are believed to affect the three-dimensional organisation of *cer*, perhaps influencing the ability of the site to form an appropriate synaptic complex to allow recombination to proceed (Summers & Sherratt, 1988). A third feature of the accessory DNA that is important for plasmid stability is a functional promoter, located within *cer*, from which a transcript, Rcd, is expressed. Transcription of Rcd occurs in cells containing multimeric plasmids, and this coincides with inhibition of cell division. It is therefore proposed that the formation of this transcript stops cells dividing, allowing time for the conversion of multimers to monomers by site-specific recombination within *cer* (Patient & Summers, 1993). While this latter function does not appear to affect recombination, it does contribute to plasmid stability.

GAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGTAATGTGGC

-35 -10

CTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGCG

CGATCGCGGCAGTTTTTCGGGGTGGTTTGTTGCCATTTTTACCTGTCTGC

Core recombination site

TGCCGTGATCGCGCTGAACCGCGTTTTAGCGGTGCGTACAATTAAGGGAT

TATGGTAAATCCAACTTACTGTCTGCCCCTCGTAGCATCGA

Figure 1.9: The *cer* site from plasmid ColE1.
The 241 bp functional *cer* site is shown, as defined by Summers and Sherratt, 1988. The core recombination site is indicated, and underlined. The Arg-box is shown in italics, centred at 123 bp upstream of the centre of the core recombination site. Overlapping the Arg-box is the -35 region of the promoter from which the *cer* Rcd transcript is expressed.

Four host-encoded proteins are required for recombination at *cer*

Four chromosome-encoded proteins are involved in recombination at *cer*. Two of these proteins, XerC and XerD are proteins of the λ-integrase class of recombinases (Colloms *et al.*, 1990; Blakely *et al.*, 1993; see Figure 1.4). XerC and XerD are encoded in separate transcriptional units on the *E. coli* chromosome. *xer*C is within an operon containing four open reading frames at the 85-minute region of the chromosome (Colloms, 1990; Colloms *et al.*, 1990). Two of the ORFs have unknown function; the other characterised gene is *dapF*, the most upstream gene in the operon, which encodes diaminopimelate reductase (Richaud *et al.*, 1987; Richaud & Printz, 1988). *xerD* (originally named *xprB*) is within the same multicistronic unit as *recJ*, at around 74 minutes on the genetic map (Lovett & Clark, 1985; Lovett & Kolodner, 1991; Blakely *et al.*, 1993). XerC and XerD are both apparently required for recombination, as *xer*C mutant strains are not complemented by XerD and *vice versa*. Both proteins are apparently required to be catalytically active for recombination: site-directed mutants in the active sites of either protein prevent recombination *in vivo*, although very low activity can be detected with XerD mutants in the presence of wild-type XerC (Blakely *et al.*, 1993).

The other two proteins are accessory factors for recombination, with other known roles in the cell. ArgR, the arginine repressor, is a regulator of arginine biosynthesis; it binds, in the presence of its co-repressor *L*-arginine, to DNA sequences upstream of arginine biosynthetic genes (Arg-boxes) to prevent transcription (Lim *et al.*, 1987). During recombination at *cer*, ArgR binds to the Arg-box region of *cer*, causing a DNA bend (Stirling *et al.*, 1988b; A. Merican, personal communication). The second accessory protein is PepA, the *E. coli* aminopeptidase A (Miller & Schwartz, 1978; Stirling *et al.*, 1989). This was identified by homology to the bovine lens leucine aminopeptidase, however, it has been established that PepA is not acting as a peptidase during recombination at *cer*, since site-directed mutants lacking all detectable peptidase activity are wild-type for recombination at *cer in vivo* (McCulloch, 1992; McCulloch *et al.*, 1994b). The role of PepA during recombination has not yet been established. It is believed that the accessory proteins ArgR and PepA, in conjunction with the accessory DNA sequences, are involved in ensuring that recombination occurs preferentially between directly repeated *cer* sites,

thereby favouring resolution of multimeric plasmids back to monomers (Summers, 1989; Sherratt *et al.*, 1993).

Xer recombination at *dif*

While obligatory for recombination at *cer in vivo*, accessory factors are not essential for recombination at all *cer*-like sites. A related site from the terminus region of the *E. coli* chromosome, *dif*, recombines efficiently as a ~30 bp recombination core, when cloned into a plasmid (Blakely *et al.*, 1991; Kuempel *et al.*, 1991; see Figure 1.10). Recombination at *dif* requires XerC and XerD, but is independent of ArgR and PepA. Plasmids containing *dif* sites undergo intermolecular XerC/D-mediated recombination *in vivo* to produce plasmid multimers, as well as intramolecular recombination to resolve multimers to monomers, and will mediate inversion between sites in inverted repeat, although at lower efficiency (Blakely *et al.*, 1991, 1993). *dif* does not, therefore, make an efficient plasmid resolution site due to its lack of resolution selectivity.

Mutant strains in which *dif* has been deleted, or mutants of *xerC* or *xerD*, form filamentous cells with aberrant nucleoids, as a consequence of their inability to segregate their chromosomes to daughter cells at cell division (Blakely *et al.*, 1991; Kuempel *et al.*, 1991; G. Blakely, N. Leslie, unpublished results). It is believed that the role of *dif* in *E. coli* is to resolve chromosome dimers formed as a result of homologous recombination between sister chromosomes during replication (Blakely *et al.*, 1991; see Figure 1.10), and models have been proposed for how this could be achieved, taking into account the lack of resolution selectivity displayed during recombination between *dif* sites (Sherratt *et al.*, 1993; Sherratt, 1993; see Figure 1.10). Recent results have also indicated that the location of *dif* in the bacterial chromosome terminus region is important to its function, since strains deleted for *dif* cannot be complemented by insertion of a *dif* site into ectopic positions in the chromosome (N. Leslie, personal communication).

Other DNA substrates for Xer recombination

In addition to *cer* and *dif*, several other DNA substrates for Xer-mediated recombination have been identified or constructed. Many of these originate from plasmids related to ColE1, although a site that undergoes XerC/D-mediated recombination in the plasmid pSC101 has also been discovered (see Chapter 5). As well as natural plasmid and chromosomal sites, several artificial sites have also been created, which have been informative in terms of the role of core recombination site sequence during

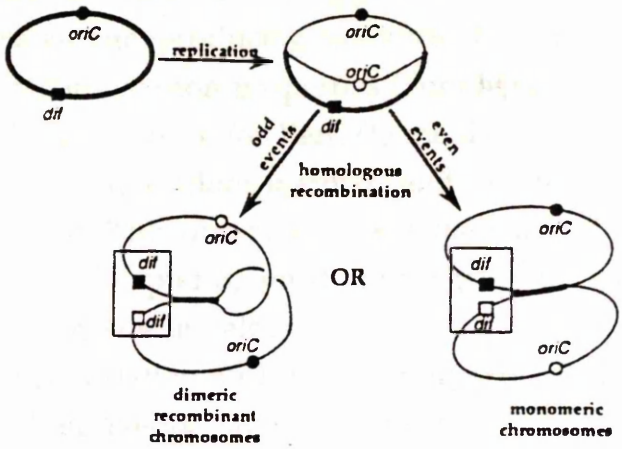
Figure 1.10a: *dif* core recombination site sequence.

TTGGTGCGCATAA TGTATA TTATGTTAAAT

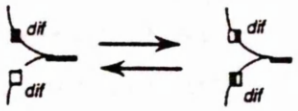
Unlike *cer*, *dif* recombines efficiently as a core recombination site. No accessory DNA or proteins are required. Recombination at *dif* does not show resolution selectivity.

Figure 1.10b: Two models proposed to account for how recombination at *dif* in the chromosome terminus region could ensure that chromosome monomers are partitioned to daughter cells at cell division (facing page) .
In the first model, recombination occurs repeatedly once *dif* has been replicated. The chromosome would be expected to be in monomeric form 50% of the time, and in dimeric form for the remainder of the time. It is supposed that when chromosomes are in monomeric form they will be separated by the motive force that separates daughter nucleoids. In the second model, a single strand-exchange occurs at *dif* to produce Holliday junctions. The Holliday junctions are then resolved either by the Xer system, or by cellular Holliday junction-resolving enzymes. In this case, if Holliday junctions were formed between chromosome monomers, a reversal of the Xer first strand-exchange would be required, while if the junction was formed between *dif* sites in a chromosome dimer, then a second Xer strand-exchange would be required. Model 2 would account for the readiness with which we are able to observe Xer-mediated Holliday junctions *in vitro* and *in vivo*.This Figure was adapted from Sherratt *et al.*, 1993.

Model 1



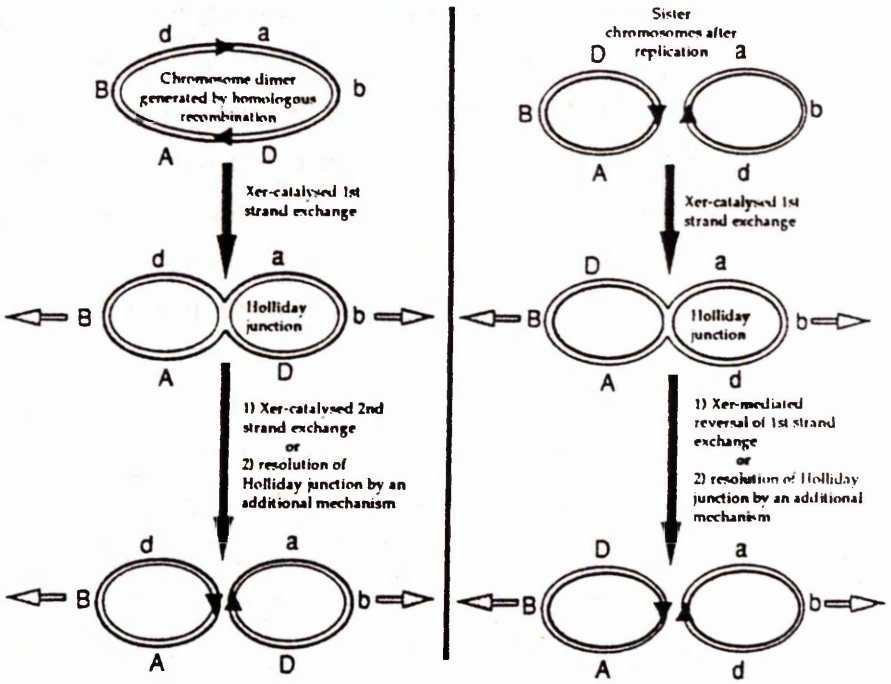
Repeated Xer-mediated site-specific recombination, after *dif* replication would lead to the chromosomes being in the monomeric configuration 50% of the time, when they can be partitioned to daughter cells.



unreplicated double stranded DNA

replicated double stranded DNA

Model 2



The choice of resolution mechanism must be imposed by external factors, eg., the putative partitioning mechanism for separating sister chromosomes.

Top strand

Bottom strand

Indicates the chromosomes under motive force during partition

recombination. The type I and type II hybrids were created by recombining *cer* with a related site, producing two sites of slightly different sequence with different recombination properties (Summers, 1989). The type II hybrid was found to be a substrate for XerC/D-mediated recombination in an ArgR/PepA-independent manner, and to undergo intermolecular recombination. This observation was the basis of the experiments that were carried out in Chapter 3, which investigated the role of the left arm of the core recombination site in selectivity.

Gel retardation and footprinting experiments have shown that XerC and XerD bind co-operatively to the core recombination site, with XerC binding the left arm, and XerD binding to the right; binding produces bends in the DNA (Blakely *et al.*, 1993; G. Blakely, unpublished results). Informative derivatives of *dif* have also been constructed. A derivative with a symmetrical central region can undergo normal recombination with a similar site, indicating that asymmetry in the central region is not required to ensure that the right arm of one site is joined to the left arm of the other. This contrasts with other members of this class of site-specific recombination system, where such asymmetry is essential (Hoess *et al.*, 1986; Senecoff & Cox, 1986). Other *dif* derivatives with two right arms or two left arms will not recombine, further suggesting that both recombinases are required to allow recombination to occur. Taken together, these results suggest that the recombinases are not functionally interchangeable, and that the asymmetry imposed as a result of having one XerC-binding and one XerD-binding site might be responsible for ensuring that the correct left-right joins are made during recombination (Blakely *et al.*, 1993).

Xer recombination *in vitro*

Several site-specific recombination systems have been reconstituted *in vitro*, using purified DNA substrates and proteins in simple buffers (Vetter *et al.*, 1983; Landy, 1989; Hoess & Abremski, 1985; Hatfull & Grindley, 1988; Stark *et al.*, 1989a). Following the relatively recent discovery of XerD, it has been possible to perform Xer site-specific recombination with partial success *in vitro*. To date, however, only one strand-exchange has been recreated. Two types of Xer *in vitro* system have been studied. In one, a supercoiled plasmid containing two *cer* sites in direct repeat is incubated in a simple buffer with purified ArgR, PepA, XerC and XerD, before separating the products by agarose gel electrophoresis; resolution products have not yet been isolated by this method, but Holliday junctions can be easily detected (Holliday, 1964: R.

McCulloch, K. Grant, unpublished results; also see Chapter 4). In the second type of reaction, when synthetic Holliday junctions are incubated with XerC and XerD, strand-exchange occurs, although this recreates parental, rather than recombinant, molecules (L. Arciszweska, unpublished results). The inability to recreate an entire resolution reaction *in vitro* has led to some concern that a component of the Xer site-specific recombination system may currently be missing, or that the existing conditions are inappropriate. To address possible reasons for this, in Chapter 4 the effect of Fis on the Xer site-specific recombination system was tested, while in Chapter 5 a mutagenesis of the *E. coli* chromosome was undertaken with one possible aim of uncovering further Xer mutants.

Aims

This project is divisible into three major topics.

The first subject of investigation was the role of the left arm of the core recombination site of the type II hybrid in allowing XerC and XerD to catalyse intermolecular recombination. This was undertaken by constructing a series of plasmid recombination substrates differing at single positions within the left arm, and assaying their ability to support intermolecular recombination *in vivo*. Several of these sites were then selected for further analysis to attempt to address the molecular basis of Xer-mediated intermolecular recombination, to infer possible mechanisms behind resolution selectivity.

The aim of the experiments described in Chapter 4 was to investigate a potential role for Fis in Xer recombination. It was first established that recombination at *dif* was reduced in a *fis* strain *in vivo*, implying that Fis might affect XerC/D-mediated recombination. Possible reasons for this effect were then investigated, to determine whether Fis might be acting directly at the recombination site, or in an indirect manner.

In Chapter 5, a chromosomal mutagenesis was performed, with two major aims. As mentioned above, an Xer recombination site has been isolated from plasmid pSC101. Recombination at this site was found to be independent of ArgR, but dependent on PepA. Preliminary analysis of the site implied that an additional protein might be required, therefore it was hoped that a mutagenesis would help identify this protein. In addition, it was considered feasible that an additional Xer protein is required for

recombination at *cer* and *dif*, since current *in vitro* recombination reaction conditions do not form products.

Chapter Two

Materials & Methods

Bacterial strains

Derivatives of *Escherichia coli* K12 used in this study are tabulated below:

Strain	Genotype	Source/Reference
AB1157	*thr-1, leuB6, hisG4, thi-1, ara-14 Δ(gpt-proA)62, argE3, galK2, supE44, xyl-5, mtl-1, tsx-33, lacY1, rpsL31*	Bachmann, 1972
DS941	AB1157, but *recF143, lacZ*ΔM15, *lacI*^{q}	D. Sherratt
DS945	AB1157, but *recBC, sbcA*	D. Sherratt
DS953	DS941, but *sup*^{o}	D. Sherratt
DS956	DS941, but *xerA9* (*argR::fol*)	D. Sherratt
DS957	DS941, but *xerB1* (*pepA*::Tn*5*)	C. Stirling
DS981	DS941, but *xerC2* (*xerC*::Km^{r})	S. Colloms
DS984	DS941, but *xerCY17* (*xerC*::Mu *d*II PR13)	S. Colloms
DS9008	DS941, but *xerD2* (*xerD*::Tn*10*-9)	D. Sherratt
DS9009	DS984, but *xerD2*(*xerD*::Tn*10*-9)	D. Sherratt
CSH50 (DS887)	*ara*, Δ(*lac-pro*), *thi*	R. Kahmann
CSH50 *fis*::Km		R. Kahmann
JR10	DS941, but *fis*::Km^{r}	Chapter 4
JR20	DS984, but *fis*::Km^{r}	Chapter 4
JR29	Tn*5*-derived mutant of DS953	Chapter 5
JR30	Tn*5*-derived mutant of DS941	Chapter 5

Bacteriophage

Tn*5* mutagenesis was carried out using the suicide λ phage, NK467, which was provided by N. Kleckner. The genotype of NK467 is Δb221, *rex*::Tn*5*, cI857, O*am*, P*am*. P1 transduction was performed using the generalised transducing phage P1*kc* (M. Masters).

Plasmids

Unless otherwise stated, all plasmids are Ap^{r} pMB1 derivatives. In "reporter" plasmids, two *cer/dif/ psi* sites are present in direct repeat. Recombination between the two sites forms two resolution products, only one with an origin of replication. The replicative resolution product can be detected genetically by loss of the marker between the two sites, and by electrophoretic mobility on gels. The plasmids used throughout this work are tabulated below:

Plasmid	Description	Reference/ Source
pUC18	High copy number cloning vector	Yanisch-Perron *et al*, 1985
pCB104	λ–*dv* based vector, Cm^r, Ap^s	A.C. Boyd
pBEND2	Circular permutation analysis vector	Kim *et al.*, 1989
pKS492	~280 bp *cer* site in pUC18	D.K. Summers
pKS455	2-*cer* reporter with deletable Cm^r	Summers & Sherratt, 1988
pKS456	Resolution product of pKS455	Summers & Sherratt, 1988
pCS210	pACYC184-based *cer* reporter; deletable *lacZ* gene; Tc^r, Ap^s	Stirling *et al.*, 1988a
pCS211	Resolution product of pCS210	Stirling *et al.*, 1988a
pSDC115	2-*cer* reporter	S. Colloms, 1990
pMIN33	*dif* site in pUC18	Blakely *et al.*, 1991
pGB300	*dif*-8 site in pUC18	G. Blakely
pSDC124	2-*dif* reporter with deletable Km^r	Blakely *et al.*, 1991
pAL21*sis*	Contains a Fis binding site,*sis*	Bednarz, 1989
pSDC105	*xerC*-expression plasmid; p15a origin	Colloms *et al.*, 1990
pRM130	*xerD*-overexpressing plasmid	Blakely *et al.*, 1993
pCS112	*pepA*-overexpressing plasmid	Stirling *et al.*, 1988b
pMAY5	*xerC* & *xerD* expression under P_{tac} control; p15a origin	G. May
pJEN1	Minimal *cer* site in pUC18	Chapter 3
pJEN2	Minimal type I hybrid in pUC18	Chapter 3
pJEN3	Minimal type II hybrid in pUC18	Chapter 3
pJEN4	Minimal *parB* site in pUC18	Chapter 3
pJEN71-714	14 derivatives of pJEN3 & pJEN4 with variant left arm sequences	Chapter 3
pJEN33	Type II minimal site in pBEND2	Chapter 3
pJEN43	*parB* minimal site in pBEND2	Chapter 3
pJEN723	Variant 2 site in pBEND2	Chapter 3
pJEN7113	Variant 11 site in pBEND2	Chapter 3
pJEN7143	Variant 14 site in pBEND2	Chapter 3
pSDC133	363 bp psi^+ *Dde*I fragment in pUC18	S. Colloms, Chapter 5
p752	2-*psi* reporter with deletable Km^r	M. Burke
p751	Resolution product of p752	M. Burke
p741	Deletion derivative of pSDC133	M. Burke, Chapter 5
p733	Deletion derivitive of pSDC133	M. Burke, Chapter 5
pJEN81	Gm^r derivative of p752	Chapter 5
pJEN96	2-*psi* reporter in λ-*dv* replicon	Chapter 5

Chemicals and reagents

Reagent	Supplier
General chemicals, biochemicals, dyes and organic solvents	BDH, May & Baker, Sigma, BRL, Fluka
Media	Difco, Oxoid
Agarose	BRL
Deoxyribonucleoside triphosphates	Boehringer Mannheim
Antibiotics	Sigma
Restriction & DNA modification enzymes	BRL, Promega, Boehringer Mannheim, New England Biolabs, United States Biochemicals
Other proteins	As stated in the text
Radiochemicals	NEN-DuPont
Phenol (buffered against 0.5 M Tris. HCl, pH 8)	Fisons
30:0.8 acrylamide solutions (Protogel)	National Diagnostics
40:1 acrylamide solutions	Biorad
Reagents for oligonucleotide synthesis	Cruachem Ltd., Applied Biosystems
DNA sequencing reagents	United States Biochemical, Promega

Media

Bacterial cultures were generally propagated in complex liquid and solid media. When liquid culture was used, this was either 2 x YT or LB broth, as stated in the text. For growth of colonies on plates, LB agar was used. The composition of these culture media is described in Sambrook *et al.*, 1989. Cultures propagated on both solid and liquid media were routinely incubated at 37°C, with liquid cultures undergoing vigorous shaking. Bacterial strains were stored on LB agar slopes at room temperature, or in 50% LB broth, 20% glycerol and 1% peptone at -70°C.

Where minimal media were used, this is stated in the text. Minimal media consisted 75 ml 2% agar in distilled water (dH_2O) supplemented with 25 ml Davis & Mingioli salts (28 g K_2HPO_4, 8 g KH_2PO_4, 1g sodium citrate, 0.4 g $MgSO_4H_2O$ per litre dH_2O) and 0.2% glucose, 20 $\mu g.\ ml^{-1}$ thiamine and amino acids as necessary. For val-leu plates, leucine was omitted from the supplementing amino acids, and the dipeptide amide *L*-val-leu-NH_2 was added to a final concentration of 0.5 mM.

Sterilisation

Growth media were sterilised by autoclaving at 120°C for 15 minutes, supplements and buffer solutions at 108°C for 10 minutes, and $CaCl_2$ at 114°C for 10 minutes.

Antibiotics

Where antibiotic selection was required, this is stated in the text. The concentrations and storage conditions of the antibiotics used is described below:

Antibiotic	Selective concentration
Ampicillin (Ap)	100 $\mu g.ml^{-1}$
Kanamycin (Km)	50 $\mu g.ml^{-1}$
Tetracycline (Tc)	10 $\mu g.ml^{-1}$
Chloramphenicol (Cm)	25 $\mu g.ml^{-1}$
Gentamycin (Gm)	5 $\mu g.ml^{-1}$

With the exception of chloramphenicol and tetracycline, all antibiotic solutions were prepared in sterile distilled, deionised water, and stored at 4°C. Chloramphenicol and tetracycline were dissolved in ethanol and stored at -20°C. Where antibiotic concentrations are written in the text, the are in the form Ap_{100}, where the number refers to the final concentration in $\mu g.ml^{-1}$.

Other indicators

X-Gal (5-bromo-4-chloro-3-indolyl-β-*D*-galactoside) solutions were made to a stock concentration of 40 mg. ml^{-1} in dimethylformamide, and was added to agar to a final concentration of 20 µg. ml^{-1}. Solutions of X-Gal were stored at -20°C, in the dark.

Stock IPTG (isopropyl-β-*D*-thiogalactoside) solutions were at 25 mg. ml^{-1} in sterile dH_2O, and used at a final concentration of 25 µg. ml^{-1}.

ONPG (*o*-nitrophenyl-β-*D*-galactopyranoside) stock solutions were made to a concentration of 20 mM in β-galactosidase assay buffer. Standard curves of ONP concentrations were performed by making dilutions of a stock solution of 0.5 mM ONP(*o*-nitrophenol) in β-galactosidase assay buffer, and measuring absorbance at 420 nm.

Bradford's Reagent was used to assay the amount of protein in cell extracts spectrophotometrically (Bradford, 1976).

Buffer solutions

EDTA:	Solutions of EDTA were always at pH 8.0.
50 x E buffer:	1210 g Tris, 410 g, sodium acetate, 93 g $Na_2EDTA.2H_2O$, 355 ml glacial acetic acid in 5 litres dH_2O, (pH 8.2).
10 x TBE:	108 g Tris, 55 g boric acid, 40 ml 0.5 M EDTA per litre dH_2O, giving a pH of 8.3.
TE:	10 mM Tris. HCl (pH 8), 1 mM EDTA.
GTE:	50 mM glucose, 25 mM Tris. HCl (pH 8), 10 mM EDTA.
Alkaline SDS:	0.2 M NaOH, 1 % SDS.
"5M" potassium acetate:	300 ml 5 M potassium acetate, 57.5 ml glacial acetic acid, 142.5 ml H_2O.
0.5 M Tris. glycine (pH 9.4):	60.55 g Tris, 7.43 g glycine per litre dH_2O.
STE:	20 mM Tris. HCl (pH 7.5), 100 mM NaCl, 10 mM EDTA.
Gel loading buffers (4x):	25% sucrose, 0.05% bromophenol blue, 0.05% xylene cyanol in dH_2O.
Single colony buffer:	2.5% Ficoll, 1.25% SDS, 0.01% bromophenol blue, 0.01% xylene cyanol in 1 x E buffer. Stored at room temperature, adding RNase A to a final concentration of 1 $\mu g.\ ml^{-1}$ just before use.
5 x KGB:	500 mM potassium glutamate, 125 mM Tris. acetate (pH 7.5), 50 mM magnesium acetate, 250 $\mu g.\ ml^{-1}$ B.S.A. (Sigma, Fraction V), 2.5 mM β-mercaptoethanol. Stock solutions of 5 x KGB were stored at -20°C; aliquots in use were stored at 4°C. The appropriate final concentration of KGB was used for each enzyme, as tabulated in Sambrook *et al.*, 1989.
Restriction enzyme buffers:	if KGB was unsuitable, the buffer recommended by the supplier of the enzyme was used.
Ligase buffer:	supplied by BRL.
Mung bean nuclease buffer:	supplied by Promega.
β-gal assay buffer:	100 mM NaH_2PO_4 (pH 7.25), 5 mM $MgSO_4$.
2 x peptidase assay buffer:	40 mM Tris. HCl (pH 8.2), 200 mM KCl, 2 mM $MnCl_2$, 0.2 mM EDTA.

1x Xer buffer:	50 mM Tris.HCl (pH 8), 50 mM NaCl, 1 mM EDTA, 1 mM *L*-arginine, 10% glycerol, 500 ng. ml^{-1} poly dI.dC.
GBA1:	5 mM Tris. glycine (pH 9.4), 1 mM EDTA, 100 mM NaCl, 10% glycerol.
GBA2:	5 mM Tris. glycine (pH 9.4), 1 mM EDTA, 10% glycerol.
GBA3:	5 mM Tris. glycine (pH 9.4), 1 mM EDTA, 25 mM NaCl, 10% glycerol.
λ–dil:	10 mM Tris. HCl (pH 7.4), 5 mM MgSO4, 0.01% gelatine.
Phage buffer:	7 g Na_2HPO_4, 3 g KH_2PO_4, 5 g NaCl, 0.25 g $MgSO_4$, 15 mg $CaCl_2.2H_2O$ and 1 ml 1% gelatine per litre dH_2O.
Stop solution:	95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05 % xylene cyanol.
Fis dilution buffer:	50 mM Tris. HCl (pH 8), 1 mM NaCl, 1 mM EDTA.

Oligonucleotides

Oligonucleotides were synthesised on an Applied Biosystems 391 PCR-Mate oligonucleotide synthesiser. The sequences of the oligonucleotides used throughout the course of this study are described in Chapter 3.

Oligonucleotides were de-protected as follows. The glass bead support containing the oligonucleotide was removed from its column, and incubated with 1 ml 30% aqueous ammonia for 1-2 hours at room temperature. The solution was then vortexed, centrifuged, and the supernatant removed. To this, a further 1 ml 30% aqueous ammonia was added, and the oligonucleotide solution heated at 55°C overnight. The de-protected oligonucleotide could then be stored at -20°C.

The oligonucleotides were purified before cloning. One quarter of each ammoniac solution was precipitated with ethanol, and the pellet resuspended in 20 µl TE. An equal volume of formamide was added, the solution vortexed, then heated at 55°C for five minutes, before loading onto a pre-run denaturing polyacrylamide gel. Electrophoresis was carried out at 15 W for approximately 1 hour. The gel was then stained with Stains-All or ethidium bromide , and the oligonucleotide excised. The DNA was recovered from the gel-chip, and precipitated with ethanol. The pellet was then resuspended in 50 µl TE.

The complementary strands of a double-stranded oligonucleotide were annealed by mixing equal quantities of each oligonucleotide (in TE) by vortexing, then heating at 100°C in a beaker, before allowing to cool slowly to room temperature.

Oligonucleotides were not phosphorylated prior to cloning, therefore the vector could not be de-phosphorylated. Oligonucleotides were typically cloned as *Xba*I-*Sal*I fragments into pUC18. For example, 0.5-1 µg pUC18 was digested with *Xba*I and *Sal*I, and the digested DNA purified an agarose gel. The DNA was then precipitated with ethanol, and the pellet resuspended in 5 µl TE. 10 µl ligation reactions were set up by adding 2 µl 5 x ligase buffer, 2 µl annealed oligonucleotide, and 1 µl (1 U) ligase.

P1 transduction

(a) Preparation of P1 lysates

The donor strain was grown to late log phase in LB broth, and cells from a 100 µl aliquots were harvested and resuspended in 100 µl LB broth, 100 µl 50 mM $CaCl_2$ and 100 µl 100 mM $MgCl_2$. To each aliquot of cells, 100 µl P1 phage in phage buffer was added (several dilutions of a ten-fold dilution series of phage were used), and the cells and phage incubated together at 37°C for 30 minutes. To each cell/phage aliquot, 3 ml molten 45°C soft agar (0.6% agar in dH_2O) was added, and this mixture was overlaid on a fresh, un-dried LB plate. The plates were then incubated right side up until plaques appeared on the plates, but were not quite confluent. The phage were harvested from the plate by adding 2.5 ml phage buffer, and allowing this to stand for 15 minutes at room temperature, before lifting off both soft agar and phage buffer and placing them in a centrifuge tube. Two drops of chloroform were added, followed by vortexing for 30 seconds to lyse the remaining cells. After 30 minutes at room temperature, the vortexing was repeated, and the debris was removed from the phage by centrifugation at 17500 *g* for ten minutes. The supernatant, containing approximately 10^9 pfu. ml^{-1}, was stored at 4°C.

(b) P1 transduction

The recipient strain was grown to mid-log phase in LB broth, and 100 µl aliquots were resuspended in 100 µl LB broth, 100 µl 50 mM $CaCl_2$ and 100 µl 100 mM $MgCl_2$. 0.1-10 µl P1 phage was added, and the cells and phage incubated together for 20 minutes at 37°C. The phage infection was then stopped by the addition of 200 µl 1M sodium citrate, and 500 µl of LB was added to the cells

before incubating for 60 minutes at 37°C. The cells were then plated onto media containing the appropriate antibiotic to select for the allele being transduced. After overnight growth at 37°C, transductants appeared on the selective plate. These were streaked to single colonies at least twice, to remove any contaminating phage.

Mutagenesis of the *E. coli* chromosome with Tn5

(a) Preparation of NK467 lysates

Phage plaques were prepared and stocks titred by growing a 5 ml culture of DS941 in LB broth, then harvesting the cells, and re-suspending them in 2.5 ml 100 mM Mg_2SO_4. 100 µl aliquots of cell suspension were added to 2.5 ml aliquots of molten, 45°C soft agar, and this mixture was poured onto a fresh, un-dried LB agar plate, and allowed to set, to produce a lawn of cells. Onto this, 10 µl spots of serially diluted NK467 were dropped, and the plate incubated right side up at 37°C overnight. (The phage stock could then be titred by assuming that if single plaques are present in the 10^{-6} spot, the titre was 10^{10}pfu. ml^{-1}).

To make a new lysate stock, a 5 ml overnight culture of DS941 was harvested and resuspended in 2.5 ml 10 mM Mg_2SO_4, then between 1 and 5 plaques (from the plates prepared as described above) were added to 100 µl aliquots of cell suspension. 2.5 ml of LB broth, and 2.5 ml molten, 45°C soft agar were added, and the mixture poured over fresh, un-dried LB agar plates. The plates were incubated right side up in a wet box at 37°C until confluent lysis occurred (after approximately 5 - 7 hours). The top agar was then removed to a 15 ml Corex tube. The plate was washed with 2 ml λ-dil, and this was combined with the soft agar. 2 drops of chloroform were added, and the mixture vortexed for 30 seconds, then centrifuged at 4500 *g* for 10 minutes. The lysate supernatants were removed, and stored in the dark at 4°C. The new lysates were titred as described above, and found to have a titre of 10^{13} pfu. ml^{-1}.

(b) Mutagenesis with NK467

A 20 ml overnight culture of DS953 was grown in 2 x YT, at 37°C. To 4 x 2.5 ml aliquots of cells, 100 µl NK467 was added, and the cultures were incubated for a further 2 hours at 30°C, with shaking, to allow uptake and transposition of the Tn5, and expression of the Km^r gene. Half of the cells were then plated onto LB Km_{50} plates, to estimate the number of mutants in each culture. Each plate contained approximately 2500 colonies, indicating that 2000 Km^r colonies were obtained per ml culture of DS953. The remaining half of the

cultures was used to inoculate cultures to prepare competent cells for transformation by the appropriate reporter plasmid. The selection of mutants is described in Chapter 5.

Transformation of *E. coli* with plasmid DNA

Bacteria were made competent for transformation by growing 20 ml cultures to A_{600} = 0.4-0.6, then harvesting the cells by centrifugation, and re-suspending them in 6 ml ice-cold 50 mM $CaCl_2$, and incubating on ice for 10 minutes. The cells were then harvested and resuspended in 4 ml ice-cold 50 mM $CaCl_2$, and incubated on ice for a further 30 minutes. Plasmid DNA was added to 100 µl aliquots of competent cells, the cells and DNA mixed gently together, then incubation on ice continued for 30 minutes. The cells were then subjected to a 2 minute 42°C heat-shock. If the selective antibiotic was ampicillin, the transformants were then plated directly onto selective plates (or inoculated into cultures containing ampicillin). If another antibiotic was being used, after heat-shock 1 ml 2 x YT or LB broth was added to the cells, and they were incubated at 37°C for 60 minutes, before plating onto plates containing the appropriate antibiotic.

Rapid screening of plasmid content by single colony gels

The plasmid content of a small number of cells (typically a large colony on a plate, or a small patch prepared from a single colony) could be rapidly analysed by lysis of the cells in single colony buffer. Cells were recovered from a plate using a wooden toothpick, then transferred into 100 µl single colony buffer in a small Eppendorf tube by rotation of the stick between thumb and forefingers. Lysis was allowed to continue at room temperature for about 5 minutes, before centrifugation at 14000 rpm in a microcentrifuge for 15 minutes. The supernatant was then loaded into the wells of an agarose gel, taking care to avoid the viscous pellet. The plasmid DNA was then analysed by agarose gel electrophoresis.

This technique was used for quickly screening the size of potential clones, and also for rapidly analysing a population of transformants direct from a transformation plate (or from liquid culture) by recovering a pool of colonies into the same aliquot of buffer.

Preparation of plasmid DNA

(a) Large scale plasmid DNA preparations

Two variations of the alkaline lysis plasmid preparation were used. In the first method, used to make high quality plasmid DNA, the sealed-up method (Birnboim & Doly, 1979, as modified in this laboratory) was used. 100 ml overnight cultures were grown in LB broth or 2 x YT, the cells harvested by centrifugation at 12000 *g*, 4°C for 5 minutes, then resuspended in 4 ml cold GTE, and incubated on ice for 5 minutes. 8 ml freshly prepared alkaline SDS was added, and the cell suspension mixed by gentle inversion followed by incubation on ice for 4 minutes. 6 ml cold "5 M" potassium acetate was added, and the suspension mixed gently, before spinning down the white precipitate at 35000 *g*, 4°C, 15 minutes, with no brake. The nucleic acid was then recovered from the supernatant by adding 12 ml isopropanol, mixing well, and allowing precipitation to occur at room temperature for 10-15 minutes, before centrifuging at 35000 *g* at 20°C, for 20 minutes. The pellet was then rinsed carefully in 70 % ethanol, dried, and resuspended in 2.5 ml TE by vortexing, or shaking on an orbital shaker for 30 minutes. Insoluble debris was removed by centrifugation. To prepare a CsCl/ethidium bromide gradient, 2.1 ml of DNA solution was added to 4.32 ml CsCl solution (5 g CsCl in 3 ml TE, pH 8) and 0.27 ml 15 $mg.ml^{-1}$ ethidium bromide. This mixture was transferred to a Beckmann Quick-Seal (16 x 76 mm) tube, the tube filled with liquid paraffin, then heat-sealed. Tubes were then centrifuged at 200000 *g* for 16 hours at 20°C in a Beckmann Ti70 fixed-angle rotor. The plasmid bands were visualised using a 365 nm UV transilluminator and were harvested from the tubes by piercing with a syringe needle, and recovered to a test-tube. The ethidium bromide was removed from the plasmid DNA by first diluting the plasmid solution approximately two-fold (to 3 ml) then extracting with butanol; 2 ml butanol was added to the DNA, the solution vortexed, and the upper layer removed (this was repeated until the upper layer was no longer pink). The DNA was then precipitated with ethanol, as described below.

The alternative method to quickly prepare plasmid DNA (of slightly inferior quality) was to use QIAGEN tips, using the manufacturer's protocol (DIAGEN GmbH, 1990).

(b) Small scale plasmid DNA preparations

Mini-prep DNA was generally made by alkaline lysis preparations, using the same solutions as above. 2.5 ml overnight cultures were centrifuged, and the

cells resuspended in 100 µl GTE then incubated at room temperature for 5 minutes. 200 µl alkaline SDS was added, the suspension mixed by gentle inversion, then incubated on ice for 5 minutes. The cell debris was precipitated by the addition of "5 M" potassium acetate, incubation on ice for 5 minutes, followed by centrifugation at 14000 rpm in an microcentrifuge for 5 minutes. The supernatant was recovered and extracted with phenol:chloroform as described below (if the plasmid DNA was to be used for sequencing, at least two phenol:chloroform extractions were performed, followed by a chloroform extraction). The DNA was then precipitated by adding 750 µl ice-cold ethanol, incubating on ice for 10 minutes, then centrifugation at 4°C, 14000 rpm for 15 minutes. The pellet was washed with 70 % ethanol, then dried and resuspended in 50 µl TE containing 10 $\mu g.ml^{-1}$ RNase A.

Phenol:chloroform extractions

Contaminating proteins, including restriction enzymes before subsequent steps such as ligation, were removed from DNA solutions by extraction with phenol:chloroform. Phenol:chloroform solutions contained an equal mixture of Tris-buffered phenol with a 24:1 mixture of chloroform:isoamyl alcohol. An equal volume of phenol:chloroform was added to the DNA solution, followed by thorough mixing by careful vortexing. The organic and aqueous layers were separated by centrifugation at 14000 rpm for 5 minutes in a microcentrifuge, and the aqueous phase recovered. (This was sometimes followed by a second phenol:chloroform extraction step, or by a chloroform step, when the DNA purity was very important, e.g. when it was for sequencing.) The DNA was then concentrated by precipitation by ethanol, re-suspending the pellet in a smaller volume.

Precipitation of DNA fragments

Concentration of DNA solutions was generally carried out by ethanol precipitation. 1/10 volume ammonium acetate and 2 volumes ice-cold ethanol were added to the DNA solution. The solution was mixed thoroughly but gently, and allowed to precipitate on ice for 15 minutes (or where small fragments or small quantities of DNA were used, at -70°C for 30 minutes), before recovering the DNA pellet by centrifugation for at least 15 minutes, 35000 *g*, at 4°C. The pellet was then rinsed with 70% ethanol, dried, then resuspended in a suitable volume of buffer, e.g. TE.

Calculating DNA concentrations

DNA concentrations were calculated by measuring A_{260}, as described in Sambrook *et al.*, 1989.

Preparation of plasmid monomers and dimers for *in vivo* recombination assays

Where monomeric plasmids were required, DNA fragments of the correct mobility to be supercoiled monomer was excised from an agarose gel and purified from the gel-slice. The DNA solution was then used to transform DS981 or DS9008, and the transformants were used to prepare mini-prep DNA containing largely plasmid monomers.

If dimers or higher multimeric forms were required, DS945 was transformed with the plasmid, and transformants were used to make mini-prep DNA containing multimeric plasmids formed by the RecF-mediated recombination pathway. This mini-prep was then transformed into DS981 or DS9008 and single colonies were patched out. The multimeric state of plasmids obtained from several patches was determined by single colony gels. Those colonies containing plasmid of the correct oligomeric form were used to inoculate cultures to prepare mini-prep DNA that could be utilised for subsequent *in vivo* recombination assays.

Restriction digestion

Restriction digestions were routinely performed using the potassium glutamate buffer, KGB, a buffer suitable for use with most restriction endonucleases and DNA modification enzymes (McClelland *et al.*, 1988). This was used at the optimum concentration for the restriction enzyme, as tabulated in Sambrook *et al.*, 1989. Where satisfactory conditions for digestion could not be achieved with KGB, the supplier's recommended buffer was used for these enzymes.

Restriction digests were generally performed in volumes of 10 or 20 µl, at 37°C for 60 - 120 minutes, unless otherwise directed by the supplier of the enzyme. 10 U of each enzyme was routinely used per µg of DNA being digested. Restriction enzymes were removed from DNA prior to subsequent manipulations by phenol:chloroform extraction or by heating to 70°C.

Ligation of DNA fragments

DNA fragments were ligated using ligase buffer and enzyme supplied by BRL (typically in a volume of 10 µl containing 1 U ligase). Ligations were

generally performed at room temperature for at least 60 minutes, or overnight at 16°C.

Removal of 5′ and 3′ extensions from double-stranded DNA fragments with mung bean nuclease

Mung bean nuclease was used to remove the 5′ overhang from an *Xba*I fragment when constructing pJEN42 and pJEN62 (see Chapter 3). Mung bean nuclease and mung bean nuclease buffer were supplied by Promega, and were used in accordance with their instructions.

Filling in recessed 3′ ends and end-labelling with DNA polymerase I Klenow fragment

Klenow enzyme was used to fill-in recessed 3′ termini as follows. After restriction digestion in a volume of 10 µl, all four dNTPs were added to a final concentration of 10 µM, and the volume was increased to 18 µl with dH_20, after first adjusting the final KGB concentration so that it would be at 1 x in a 20 µl reaction volume. 2 µl (2U) Klenow enzyme was added, and the reaction was allowed to proceed for 15 minutes at room temperature. The Klenow enzyme was then removed by phenol:chloroform extraction, and the DNA recovered by ethanol precipitation.

End-labelling with Klenow enzyme was performed similarly to above, with several exceptions. Firstly, only three of the dNTPs were added, the fourth supplied as its α^{32}P-radioisotope (5 µCi per pmol DNA ends). After 15 minutes incubation with Klenow as above, the reaction was "chased" by adding all four unlabelled dNTPs to a final concentration of 50 µM to inhibit the exonuclease reaction of Klenow. The reaction volume was then made up to 70 µl with STE, and the labelled fragment purified from unincorporated dNTPs using NucTrap Push Columns in accordance with the manufacturer's instructions (Stratagene, 1990).

β-galactosidase assays

Overnight cultures of DS941, DS984 and JR20 were grown in 2.5 ml 2 x YT, then 1 ml aliquots were harvested, washed in 1 ml β-gal assay buffer , then resuspended in 1 ml β-gal assay buffer. Protein was released from the cells by adding 10 µl 5 $mg.ml^{-1}$ CTAB, vortexing for 20 seconds, then incubating at 37°C for 10 minutes, before placing on ice. To 2.25 ml assay buffer, 0.5 ml 20 mM ONPG and 0.25 ml cell suspension was added, at time zero. The reactions were incubated at 37°C for fifteen minutes, then the reaction stopped by the addition

of 2 ml 1 M Na_2CO_3. The reactions were cleared by centrifugation, before spectrophotometric determination of β-galactosidase activities against a "no protein" blank. β-galactosidase activities were calculated as moles ONP released, by measuring A_{420}, and using a standard curve for ONP (not shown). To calculate specific activities, the protein content of each cell extract was determined by Bradford's assays, and the β-galactosidase activities were determined as a function of time and protein concentration. This gave specific activities as nmoles ONP released per μg protein per minute.

Statistical analysis was carried out on my behalf by Karen Grant using the StatWorks Data Student t-test package.

Peptidase assays *in vitro*

Peptidase assays were performed *in vitro* using the substrate leucine-*p*-nitroanalide, which is cleaved by peptidases to *p*-nitroanaline; this can be detected spectrophotometrically. 100 ml cultures of DS941, DS957, JR29 and JR30 were grown overnight in LB broth at 37°C. The cells were then harvested and resuspended in 1 ml 2x peptidase assay buffer (40 mM Tris.HCl (pH 8.2), 200 mM KCl, 2 mM $MnCl_2$, 0.2 mM EDTA), and lysed by adding 50 μl 25 mg. ml^{-1} CTAB, vortexing, and incubating at 37°C for 10 minutes. Extracts were then heated at 70°C for ten minutes, then cleared by centrifugation, and the peptidase activity of the supernatant assayed. Reactions were carried out in 1 ml volumes, containing a final concentration of 1 mM leucine-*p*-nitroanalide in 1 x assay buffer, and 20 μl of extract. Reactions were continued for 30 minutes at 37°C, then stopped by adding EDTA at a final concentration of 10 mM. The release of *p*-nitroanaline was determined by reading A_{400} against a "no protein" blank, then specific activities determined as described in Chapter 5.

Bradford assays

10 μl cell suspension was added to 990 μl Bradford's reagent (Bradford, 1976), and A_{595} measured. Bradford's reagent was kindly provided by K. Grant. The protein concentration could then be calculated using a standard curve for bovine serum albumin (kindly provided by A. Merican).

Preparation of crude protein extracts containing XerC or XerD

Crude extracts were prepared using the following protocol which was supplied by Gerhard May. 20 ml overnight cultures of DS9009 pSDC105 or DS9009 pRM130 were grown in 2 x YT or LB broth, at 37°C, with shaking. The cultures were then diluted 1/100 and allowed to grow for a further 2-3 hours

Note: Potential uses of polydI.dC

PolydI.dC is used as a "carrier" DNA during gel retardation assays. This synthetic DNA polymer is used as a sink for non-sequence specific DNA-binding activity, and prevents smearing-back of complexes. At high concentrations of polydI.dC, it is possible to compete for sequence-specific DNA-binding activity. This property could be exploited to address the question of XerC/D affinities for different sites (Chapter 3), by competing for XerC/D binding to labelled fragments using increasing concentrations of polydI.dC. Those sites with higher relative affinities for XerC/D would continue to produce DNA-XerC/D complexes at higher concentrations of polydI.dC than those with lower relative affinities.

before inducing expression of *xerC* or *xerD* with IPTG overnight. The cell pellet was washed in 10 mM Tris.HCl (pH 7.5), 100 mM NaCl, then resuspended in 500 µl buffer A10 (50 mM Tris.HCl (pH 7.5), 10 mM NaCl, 1 mM EDTA, 100 µM DTT). The bacterial cell walls were then degraded by adding lysozyme to a final concentration of 0.4 $mg.ml^{-1}$, and incubating at 37°C for 1-2 hours. The resultant spheroplasts were then lysed by freezing in liquid nitrogen and subsequent thawing at room temperature, three times. The cell debris was removed by centrifugation for 20 minutes at 14000 rpm in a microcentrifuge. The supernatant, containing XerC or XerD as 1-5% of total protein, was recovered, mixed with glycerol (50% v/v), then stored at -20°C. Generally 1 µl of each extract was used for each binding reaction.

Gel retardation experiments (band shift assays)

Binding reactions were prepared containing DNA, and the appropriate proteins in a suitable binding buffer. Three variations of the same gel binding buffer (GBA) were used, containing differing amount of salt: GBA1 , GBA2 and GBA3 . The choice of buffer depended on the salt concentration of the protein(s) being used. In addition, in some experiments the carrier DNA polydI.dC was used at a final concentration of 500 $ng.ml^{-1}$; where this is so, it is stated in the text (in later experiments it was generally omitted, as it was found not to improve binding). Binding reactions were typically carried out in 10 µl or 20 µl volumes. The proteins used for gel retardation were either crude protein extracts containing XerC or XerD (prepared as described above), partially purified XerC or XerD (gifts from Gerhard May and Karen Grant), or purified Fis (a gift from Regine Kahmann). Binding reactions were mixed very gently using a pipette tip. Incubation times and temperatures for each experiment are described in the Results, after which the reactions were generally quenched on ice before electrophoresis. Reactions were always electrophoresed on pre-run, non-denaturing polyacrylamide gels buffered with Tris. glycine (pH 9.4), 0.1 mM EDTA as described below, at 200 V in the cold-room. Generally the DNA fragments used for gel retardation experiments were end-labelled (each track contained approximately 10 cps), with detection of complexes autoradiography; unlabelled fragments were also occasionally used, and these were detected by ethidium bromide staining.

DNA Footprinting

(a) Footprinting with DNase I

For footprinting reactions, the amount of radiolabelled DNA required was much greater than for standard gel retardation (to ensure that all of the fragments can be detected on a gel), therefore at least 100 cps was used for each reaction. In addition, it was essential to separate the fragment containing the XerC/XerD binding site from the plasmid backbone: this was achieved by excising the appropriate fragment from a polyacrylamide gel, and recovering the fragment from the gel chip. Binding reactions for DNase I footprinting were prepared similarly to ordinary binding reactions, using GBA3 and 1 µl of each partially purified protein as appropriate (see Chapter 3), and incubated at 37°C for 55 minutes. At 55 minutes, 1 µl 8 $ng.ml^{-1}$ DNase I (in 10 mM $MgCl_2$, 25 mM $CaCl_2$) was added, and the incubation was continued for a further 14 minutes. The reactions were stopped by the addition of 3.5 µl Stop solution. Half of each sample was then electrophoresed on a sequencing gel.

(b) Footprinting with 1,10-phenanthroline-copper *in situ*

Binding reactions were performed similar to normal, but the amount of DNA used was scaled-up fifty-fold (500 cps per reaction). The volume of each binding reaction used was 20 µl, using GBA3 buffer, and 2 µl of each partially purified protein as appropriate. Binding reactions were incubated at 37°C for 60 minutes, then the complexes separated by electrophoresis on a non-denaturing polyacrylamide gel, in the cold-room. Complexes were detected by autoradiography of the wet gel, then excised from the gel and placed in Eppendorf tubes. The gel slices were chopped coarsely, then immersed in 100 µl 50 mM Tris. HCl (pH 8) before the addition of 10 µl OP-Cu mix (2 mM 1, 10-phenanthroline and 0.45 mM $CuSO_4$) and 10 µl 58 mM mercaptoproprionic acid. The mixture was vortexed gently, and incubated at room temperature for 15 minutes, then stopped by the addition of 20 µl 2, 9-dimethyl-1, 10-phenanthroline (Sigman *et al.*, 1991). The DNA fragments were eluted from the gel by the addition of 500 µl STE, and incubation overnight at 60°C; the supernatants were recovered, and the DNA precipitated by the addition of 2 volumes ice-cold ethanol. The pellets were washed in 70% ethanol, dried and resuspended in Stop solution, before electrophoresis on denaturing polyacrylamide gels.

Sequencing double-stranded plasmid DNA

Double-stranded plasmid sequencing was carried out by the dideoxy method (Sanger *et al.*, 1977)using the Sequenase Version 2.0 Sequencing Kit (United States Biochemical), but with a variation in the method of denaturation and annealing to allow sequencing of a double-stranded template. Alkaline denaturation of plasmid DNA was used: to 2-3 µg plasmid DNA, 3 µl 2 M NaOH and dH_2O to 30 µl were added, followed by incubation at 37°C for 10-15 minutes. The plasmid was annealed with the primer (typically -40 or rev primers for pUC18, supplied in the Sequenase kit, or by Promega) by adding 10 pmoles primer, then vortexing briefly. 120 µl ice-cold ethanol, then 12.75 µl 7.5 M ammonium acetate (pH 7.5) were added, and the solution mixed then incubated at -70°C for 5 minutes. The pellet was recovered by centrifugation at 14000 rpm, 30 minutes in a microcentrifuge, rinsed in 70% ethanol and dried. The annealed template and primer was resuspended in 10 µl 1x Sequenase reaction buffer, and labelling and termination reactions performed as in the Sequenase Version 2.0 protocol. The reactions were separated on a sequencing gel, as described below.

Recombination at *cer in vitro*

Xer-mediated recombination *in vitro* was carried out using the *cer* reporter plasmid pSDC115, in a simple buffer. The reactions described in Chapter 4 were performed with approximately 500 ng pSDC115 in a total reaction volume of 20 µl in 1 x Xer buffer with combinations of the four Xer proteins (ArgR, PepA, XerC, XerD) with or without Fis. Purified ArgR and PepA were provided by M. Burke, and crude protein extracts containing XerC and XerD were prepared from DS941 or JR10 containing pSDC105 or pRM130. Recombination reactions were incubated at room temperature for 90 minutes, then inactivated by heating at 75°C for 10 minutes. The reactions were then divided in half: one half was retained to electrophorese undigested, the other half was digested with *Eco*RI, which has a single recognition site in pSDC115. Digestions were performed by adding 3 µl 5 x KGB and 1 µl (10 U) *Eco*RI to the 10 µl half-reaction, and incubating at 37°C for 60 minutes. Loading dyes were then added to all reactions, which were electrophoresed on an agarose gel, and the DNA visualised by ethidium bromide staining.

Gel electrophoresis of DNA

Agarose gels:

Unless otherwise stated in the text, 1% agarose gels were routinely used. 1 g agarose per 100 ml 1 x E buffer was dissolved at medium setting in a microwave oven, then cooled to ~55°C, before pouring into horizontal gel formers. Electrophoresis was carried out in 1 x E buffer, in horizontal gel tanks. The DNA was stained with ethidium bromide before being visualised on a UV transilluminator. If the gel was only to be photographed, a short-wave (254 nm) transilluminator was used; if DNA fragments were to be recovered from the gel, a long-wave (365 nm) transilluminator was used to visualise the DNA while causing the minimum of UV-induced damage to the DNA.

Polyacrylamide gels:

Several types of polyacrylamide gels were used to separate small fragments of DNA. In the first type, small restriction fragments could be separated (for example to analyse the size of inserts in putative clones, or to isolate radioactively labelled fragments). Using the second type of gel, protein-DNA complexes could be electrophoresed for gel retardation assays and for circular permutation analysis. On the third type of gel, single-stranded DNA molecules could be electrophoresed, either to provide DNA sequence of a footprint, or for purification of oligonucleotides.

(a) Separation of small double-stranded DNA fragments

Gel composition: For these gels the proportion of acrylamide:bisacrylamide was always 30:08; the final percentage of acrylamide used depended on the size of the DNA fragments being separated, but tended to be in the range 5-12%. These gels were buffered with 1 x TBE, and electrophoresed in the same buffer. For example, a gel solution to prepare a 6% gel of this type would contain 6 ml 30:0.8 acrylamide, 3 ml TBE and 21 ml dH_2O. To polymerise this solution, 250 µl 10% ammonium persulphate and 15 µl TEMED were added, the solution mixed by swirling, and the gel cast between clean, glass plates immediately. After polymerisation was complete (approximately 45 minutes), gels were pre-run in 1 x TBE, samples were generally mixed with an appropriate volume of loading buffer, then loaded into the wells. Electrophoresis was usually carried out under a constant voltage of 200 V. After the dyes had migrated a suitable distance, the plates were prised apart, and the DNA detected either by staining with ethidium bromide, or by autoradiography depending on whether or not the fragments

were labelled. If labelled fragments were to be excised from the gel, autoradiography was carried out using the wet gel, which was wrapped in cling-film, and left attached to one glass plate. However, wherever suitable, gels were transferred to a double layer of Whatmann 3 MM paper and dried on a Biorad vacuum drier at 80°C, before exposing to X-ray film to make an autoradiograph.

(b) Non-denaturing gels for gel retardation assays

Great care was always taken to ensure no SDS was present on the plates or in the gel tank by retaining a set of equipment specifically for gel retardation purposes. Gel composition: 6% polyacrylamide gels were generally used, which had the following constituents for 30 ml gel solution: 6 ml 30:0.8 acrylamide, 3 ml 10 x Tris. glycine (pH 9.4), 21 ml dH_2O, 15 µl 200 mM EDTA, 250 µl 10% ammonium persulphate, 20 µl TEMED. Occasionally, 8% gels were used which contained 8 ml 30:0.8 acrylamide, 3 ml 10 x Tris. glycine (pH 9.4), 19 ml dH_2O, 15 µl 200 mM EDTA, 250 µl 10% ammonium persulphate, 20 µl TEMED per 30 ml gel solution. The running buffer used during electrophoresis was 50 mM Tris. glycine (pH 9.4), 0.1 mM EDTA. Gels were always pre-run in the cold-room at 200 V for at least 30 minutes, until the current reached 5 mA. Binding reactions were loaded onto non-denaturing gels without the addition of loading buffers to the samples; dyes were loaded into separate tracks to monitor electrophoresis. During electrophoresis, the buffer was continuously re-circulated from the bottom reservoir to the top using a peristaltic pump. After electrophoresis was complete, the gel was processed as above.

(c) Denaturing gel electrophoresis

Denaturing gels containing 8 M urea were used to separate single-stranded DNA fragments, such as oligonucleotides during their purification, sequencing reaction products and the products of footprinting reactions. Sequencing gels and footprinting gels were generally 8 M urea, 8% polyacrylamide gels made as follows: 50 g urea, 20 ml 40:1 acrylamide: bisacrylamide, 10 ml 10 x TBE, 30 ml dH_2O in 100 ml gel solution. To dissolve the urea, gel solutions were heated at 37°C for 10 - 15 minutes, then stirred on a magnetic stirrer. To polymerise, 1 ml 10% ammonium persulphate and 20 µl TEMED were added. The plates (supplied in the IBI Baserunner kit) were thoroughly cleaned with ethanol and one plate was treated with dimethyldichlorosilane solution prior to use. For sequencing oligonucleotides cloned in the pUC18 polylinker, short plates (43 x 21.5 cm) were used with straight spacers. For footprinting, or to read longer sequences, long plates (58 x 21.5 cm) were used with straight or wedge spacers. After gels were

polymerised, they were pre-run at 45 W (short plates) or 55 W (long plates) for at least 30 minutes in 1 x TBE buffer, until the gel temperature reached 55°C. The wells were flushed out with TBE, and the samples heated to 80°C for 2 minutes prior to loading onto the gel. Electrophoresis was carried out at 45 W or 55 W in an IBI Baserunner apparatus. The gel was processed by prising the plates apart, then soaking the gel in 10% methanol, 10% acetic acid for 20 - 30 minutes, before transferring it to three layers of Whatmann 3 MM paper and drying at 80°C on a vacuum drier. The dried gel was then used to make an autoradiograph.

For the purification of oligonucleotides, 8 M urea gels were also used, but the concentration of acrylamide was generally 12%. Such a gel had the following composition: 15 g urea, 9 ml 40:1 acrylamide, 3 ml 10 x TBE, 6 ml dH_2O per 30 ml gel solution. 200 µl 10% ammonium persulphate and 15 µl TEMED were added immediately prior to casting. The gel was electrophoresed at 15 W for at least 15 minutes before loading samples. Oligonucleotides were mixed in equal volumes with formamide, and heated to 55°C for 5 minutes before loading. The gel was run at 15 W for about an hour, then the gel stained with Stains-All and the oligonucleotide recovered as the band with slowest electrophoretic mobility.

Staining of gels

Ethidium bromide staining was carried out by soaking gels in electrophoresis running buffer or tap water containing ethidium bromide at a final concentration of 0.6 µg. ml^{-1}. Staining was continued for at least 30 minutes, then the gel de-stained in tap water. The DNA was visualised on a UV transilluminator, at 254 nm if the gel was to be photographed only, or at 365 nm if the DNA fragments were to be recovered, to prevent photo-chemical damage.

Stains-All (1-ethyl-2-[3-(1-ethylnaphtho[1,2-*d*]thiazolin-2-ylidene)-2-methylpropenyl]naphtho[1,2-*d*]thiazolium bromide; supplied by Aldrich) can be used to stain DNA, RNA or protein on polyacrylamide gels. This was generally used to stain gels for the purification of oligonucleotides, as there is no requirement for UV irradiation. The staining solution was made by mixing 70 ml dH_2O, with 20 ml isopropanol and 10 ml 0.1% (w/v) solution of Stains-All (in formamide). The gel was soaked in this staining solution until sufficient staining was observed (5 - 20 minutes). The gel was then de-stained in water. Single-stranded oligonucleotides give cyan or purple bands depending on their size: small fragments are cyan, while larger ones are purple.

Recovery of DNA fragments from gels

Agarose gels:

The appropriate band was excised from the gel using a scalpel. The DNA was recovered from an agarose gel chip by one of two methods: either using commercially available Spin-X columns (manufactured by Costar), or by spinning through siliconised glass wool (Heery *et al.*, 1990).

Polyacrylamide gels:

Gel slices containing the fragment of interest were recovered using a scalpel, then crushed in an Eppendorf tube using a pipette tip. 500 µl STE was then added, and the tube vortexed vigorously, then heated to 65°C for 2 - 3 hours, vortexing intermittently. The gel pieces were then centrifuged to the bottom of the tube and the supernatant recovered to another Eppendorf. The DNA was precipitated by the addition of 2 volumes ice-cold ethanol, incubation at -70°C for 10 minutes, followed by centrifugation at 14000 rpm for 30 minutes. The pellet was washed in 70% ethanol, dried, then resuspended in a suitable buffer.

Autoradiography

Dried or un-dried polyacrylamide gels were used to produce autoradiographs. Un-dried gels containing ^{32}P labelled DNA fragments were exposed to X-ray film while attached to one glass plate. The gel was wrapped with cling-film before exposing to the film. Autoradiographs produced in this way could be used to locate the fragments within the gel, which could then be excised and purified from the gel-slice. Dried gels were exposed directly to the film without being wrapped with cling-film. Fuji-RX film was generally used for all autoradiography.

Photography

Gels stained by ethidium bromide were visualised on a 254 nm UV transilluminator, and photographed using Polaroid type 667 land film or using a Pentax 35 mm SRL loaded with Ilford HP5 film. Both cameras were fitted with a Kodak Wratten filter, No. 23a. 35 mm negatives were developed using Ilford Microphen developer and Amfix. Prints were made with Ilford photographic paper (Grade 3), Ilfospeed developer and Amfix fixer. Where possible, contact prints were prepared from autoradiographs onto Ilford paper (Grade 3), but if

necessary, autoradiographs were photographed using PanF film and the negatives used to make prints of an appropriate size.

Thesis preparation

This thesis was word processed using the Apple Macintosh Microsoft Word package, in Palatino font. Figures were drawn using Macdraw Pro. Plasmid maps were prepared using the Macplasmap package, and modified using Macdraw Pro. Graphs were plotted using Microsoft Excel. The thesis was printed using an Apple Macintosh Laserprinter.

Chapter Three

The Role of the XerC Binding Site in Resolution Selectivity

Selectivity in site-specific recombination

Site-specific recombination results in rearrangement of the DNA between the two recombining sites; the rearrangement that occurs depends on the initial orientation of the sites in the recombination substrate. Two recombination sites can be organised in three possible ways, with recombination between them producing three different outcomes. If the sites are present in different DNA molecules, intermolecular recombination between the sites results in fusion of the two substrate molecules to produce a larger DNA molecule with two recombination sites in direct repeat; if two sites are in direct repeat within a single molecule, intramolecular recombination between the two results in deletion of the intervening DNA, to produce two smaller molecules each with a single recombination site; two sites arranged in inverted repeat will recombine to invert the intervening DNA segment (see Chapter 1, Figure 1.7).

The structural outcomes of recombination between sites in different orientations will often have biological consequences, for example phage λ integration into the bacterial chromosome is a result of intermolecular recombination, while excision of the phage is a consequence of intramolecular recombination between sites in direct repeat. While some site-specific recombination systems can perform recombination between sites in any arrangement, many systems use mechanisms to ensure that one particular organisation of sites is favoured over the other two possible arrangements; in other words, they show selectivity for a specific relative orientation, and this results in a directionality in the reaction pathway. For example, during resolution of cointegrate structures formed during Tn3 transposition, it is essential that the two *res* sites are in direct repeat within a supercoiled DNA molecule. Before recombination will proceed, this arrangement of sites is ensured by resolvase by formation of a substrate synaptic complex with a specific topology, believed to arise as a result of inter-wrapping of the two sites such that three interdomainal supercoils are trapped; such a structure is obtained due to the interactions between resolvase subunits bound at subsites II and III of *res* (Stark *et al.*, 1989a, b). In some systems, other proteins are required to ensure directionality of recombination. For example, in the integration/excision system of λ, IHF stimulates integration, while Xis and Fis stimulate excision *in vitro* (reviewed in Landy, 1989). In addition, the sites involved in integrative and excisive recombination differ, therefore both site sequence and differences in protein components contribute to selectivity during recombination catalysed by λ-Int.

Resolution selectivity in Xer-mediated site-specific recombination

To convert ColE1 multimers to monomers, Xer-mediated recombination occurs preferentially between *cer* sites that are arranged in direct repeat within a single DNA molecule; this process is termed resolution selectivity. If recombination were to occur in the absence of resolution selectivity, the effect would be to allow multimerisation of plasmids, and might lead to less stable inheritance at host cell division (Summers & Sherratt, 1984; Summers, 1989).

Recombination at *cer* and related sites occurs within ~30 bp core recombination sites of similar sequence, mediated by the recombinases XerC and XerD (Summers *et al.*, 1985; Colloms *et al.*, 1990; Blakely *et al.*, 1993). Gel retardation and footprinting experiments have shown that XerC binds to the left arm of the chromosomal site *dif*, while XerD binds its right arm; binding is highly co-operative (Blakely *et al.*, 1993). Because of the conservation of arm sequences it is expected that for *cer* and other related sites, XerD also binds to the right arm and XerC to the left. An alignment of core recombination sites from a number of *cer*-like sites is presented in Table 3.1. The region between the left and right arms is termed the central region, and may correspond to the "overlap" region of other systems from the λ-integrase class (Landy, 1989).

Table 3.1: Alignment of some known Xer core recombination sites (modified from Sherratt, 1993)

left arm XerC binding	**central region**	**right arm** XerD binding	
GCGGTGCGTACAA	TTAAGGGA	TTATGGTAAAT	ColE1 (*cer*)
GCGGTGCGTACAA	TAAGGGA	TTATGGTAAAT	ColN
GCGGTGCGTACAA	CGGATG	TTATGGTAAAT	ColA
GCGGTGCGCGCAA	GATCCA	TTATGTTAAAC	pSC101
GCGGTACCGATAA	GGGATG	TTATGGTAAAT	CloDF13 (*parB*)
GCGGTGCGTACAA	TTGGGATG	TTATGGTAAAT	type I hybrid
GCGGTGCGTACAA	GGGATG	TTATGGTAAAT	type II hybrid
TTGGTGCGCATAA	TGTATA	TTATGTTAAAT	*dif*
TTGGTGCGCATAA	TTTGTATA	TTATGTTAAAT	*dif*-8

(Recombination core site sequences were derived from: Summers *et al.*, 1985; Morlon *et al.*, 1988; Hakkaart *et al.*, 1984; Kolot, 1990; Summers, 1989; Kuempel, 1991; Blakely *et al.*, 1991; G. Blakely, personal communication).

A pair of *cer* core recombination sites are insufficient for recombination: "accessory factors" are also required. These consist of the ~220 bp of DNA upstream of the core recombination site, and two additional proteins, ArgR, the arginine repressor (along with its co-repressor *L*-arginine), and PepA, aminopeptidase A (Summers & Sherratt, 1988; Stirling *et al.*, 1988a, 1988b, 1989). The accessory factors are believed to play an important role in the resolution selectivity process during Xer-mediated recombination, and are currently under investigation in this laboratory.

The generation and characterisation of the type I and type II hybrids allowed the first demonstration that accessory factors are not always an obligatory feature of Xer recombination (Summers, 1989). These two sites were produced by recombining *cer* with the related plasmid resolution site from plasmid CloDF13 (*parB*; Hakkaart *et al.*, 1984). The resulting hybrid sites both comprised the accessory DNA and left arm of *cer*, with the right arm of *parB* (which is identical to the *cer* right arm; see Table 3.1). However, the recombination events had produced sites with central regions of different lengths: the type I hybrid has an 8 bp central region, like *cer*, while that of the type II hybrid is 6 bp. The sequences of the core recombination sites of the type I and type II hybrids are shown in Table 3.1. Recombination at the type I hybrid *in vivo* requires accessory factors, and shows resolution selectivity; in other words the type I hybrid resembles *cer*. In contrast, the type II hybrid was found capable of supporting recombination in an *argR* or *pepA* host, and a deleted recombination site of 45 bp could still recombine. The deleted site, or a full-length site in an *argR* or *pepA* mutant, does not show resolution selectivity: Xer-mediated intermolecular recombination leads to the formation of plasmid multimers. This evidence implies that some *cer*-like sites can recombine without accessory factors, but where such recombination occurs, resolution selectivity is abolished. A requirement for accessory factors might therefore be associated with resolution selectivity.

What determines whether or not XerC/D-mediated recombination can occur at a site in the absence of accessory factors? This must be dependent on the core recombination site sequence, since the only differences in sequence between the type I and type II hybrids reside within this region. In fact, the 2 bp deletion within the central region of the type II hybrid compared with the type I hybrid must be responsible for the ability of this site to support intermolecular recombination. When cloned into a plasmid, the core recombination site of the chromosomal site, *dif*, promotes intermolecular recombination; this site also has a central region of 6 bp (Blakely *et al.*, 1991; see Table 3.1). All of the known *cer*

family sites that can recombine without showing resolution selectivity have central regions of 6 bp (e.g. type II hybrid, *dif*), while many of the sites which show resolution selectivity have central regions of 8 bp (e.g. *cer*, type I hybrid). This has led to the proposal that sites with 8 bp central regions cannot recombine without accessory factors, perhaps as a result of the increased distance between the between XerC and XerD binding sites, and the predicted 72° relative rotation between XerC and XerD bound to the DNA, in those sites compared with those with 6 bp central regions; this might be expected to weaken interactions between the recombinases (Blakely *et al.*, 1993). However, it is not true that all sites with central regions of 6 bp will recombine without resolution selectivity; *parB* has a 6 bp central region, yet is an effective plasmid resolution site (Hakkaart *et al.*, 1984). Resolution selectivity cannot, therefore, be attributed simply to length of the central region; other features of the core recombination site must contribute to the ability of sites to recombine in the absence of accessory factors, and their consequent lack of directionality. The right arm sequences of all of the *cer*-like sites so far characterised are more highly conserved than those of the left arm (see Table 3.1). Where sites have a central region of 6 bp in length, the left arm sequence must therefore be important in determining whether or not resolution selectivity is shown.

Aim

The left arms of the type II hybrid and *parB* core recombination sites were studied, to attempt to attribute loss of resolution selectivity in the type II hybrid to a particular left arm sequence. Experiments were aimed at gaining insights into the molecular mechanisms underlying the resolution selectivity process, by studying how left arm sequence might affect interactions between XerC and XerD bound at the core recombination site, and recombination *in vivo*.

Results

The type II hybrid recombines as a "minimal site" *in vivo*

To test whether the core recombination sites of *cer*, the type I and type II hybrids, and *parB* could support Xer-mediated intermolecular recombination *in vivo*, a series of "minimal site" plasmids was constructed. These plasmids were designed to be analogous to the minimal *dif* site plasmid pMIN33, to test whether these four sites were able to support recombination in a similar manner to *dif* (Blakely *et al.*, 1991). The smallest type II hybrid site previously shown to function as a substrate for intermolecular recombination was a 45 bp *Mlu*I-*Taq*I fragment (Summers, 1989). By construction of oligonucleotides, the sites could be further reduced in size, to determine whether all of the necessary sequence for recombination was within a similar 32 bp region of the type II hybrid or *parB* (or 34 bp for *cer* and the type I hybrid) as in *dif*. In addition, it would be possible to test similar sites with *cer*, *parB* or type I hybrid core recombination sites sequences for recombination proficiency.

Oligonucleotides consisting of the top and bottom strands of the core recombination sites of *cer*, the type I and type II hybrids and *parB* were synthesised; the sequences of the oligonucleotides are shown below.

```
   XbaI
5'CTAGAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCG3'        cer top strand
    3'TCGCCACGCATGTTAATTCCCTAATACCATTTAGGCAGCT5'    cer bottom strand
                                          SalI

   XbaI
5'CTAGAGCGGTGCGTACAATTGGGATGTTATGGTAAATCCG3'        type I top strand
    3'TCGCCACGCATGTTAACCCTACAATACCATTTAGGCAGCT5'    type I bottom strand
                                          SalI

   XbaI
5'CTAGAGCGGTGCGTACAAGGGATGTTATGGTAAATATG3'          type II top strand
    3'TCGCCACGCATGTTCCCTACAATACCATTTATACAGCT5'      type II bottom strand
                                        SalI

   XbaI
5'CTAGAGCGGTACCGATAAGGGATGTTATGGTAAATATG3'          parB top strand
    3'TCGCCATGGCTATTCCCTACAATACCATTTATACAGCT5'      parB bottom strand
                                        SalI
```

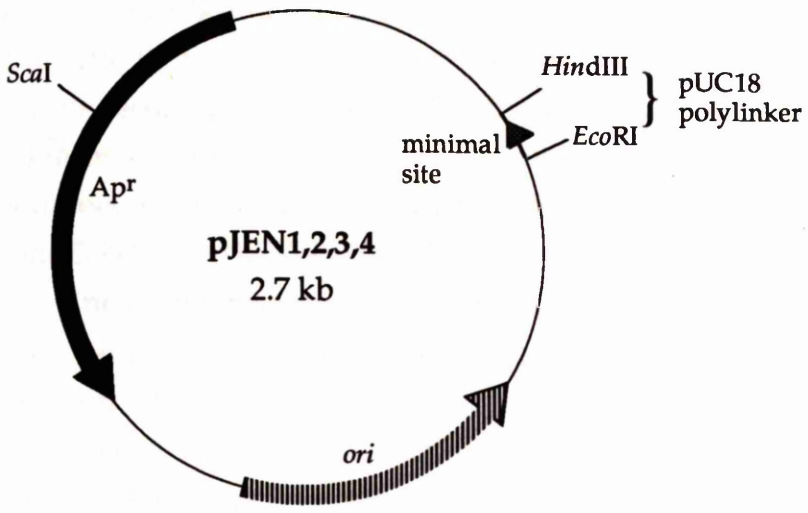


Figure 3.1: pJEN1, 2, 3, 4.
The "minimal site" plasmids pJEN1, 2, 3, and 4 were constructed by cloning oligonucleotides consisting of the core recombination sites of *cer*, the type I hybrid, the type II hybrid, and *parB* into pUC18 at the *Xba*I to *Sal*I sites in the polylinker. The orientation of the polylinker region is indicated by depicting the flanking *Eco*RI and *Hin*dIII sites. All of the polylinker restriction sites of pUC18 are retained in these plasmids. The sequences of the cloned oligonucleotides are given in the text.

The oligonucleotides were purified, annealed and cloned into the pUC18 polylinker as *Xba*I to *Sal*I fragments, as described in Chapter 2. The resulting plasmid constructs, pJEN1 (*cer*), pJEN2 (type I hybrid), pJEN3 (type II hybrid), pJEN4 (*parB*) are shown in Figure 3.1.

Intermolecular recombination with pJEN1-4 was assayed *in vivo*. Intermolecular recombination between sites on different plasmid molecules results in fusion of the plasmids, and can be assayed by the appearance of plasmid multimers. To assay Xer-mediated site-specific recombination (rather than homologous recombination), the *recF* strain DS941 and its *argR*, *pepA*, *xerC* and *xerD* derivatives were used. The *in vivo* assay was performed as follows: plasmid monomers were purified as described in Chapter 2, then transformed into competent DS941 (*xer*$^+$), DS956 (*argR*), DS957 (*pepA*), DS981 (*xerC*) and DS9008 (*xerD*). The transformants were propagated overnight in 2 x YT (or on LB agar plates) with ampicillin to select for the plasmids. A representation of the population multimerisation state of the plasmids in each strain was obtained by making mini-prep DNA of an overnight population sample followed by electrophoresis on agarose gels (as described in Chapter 2).

The results of such an experiment are shown in Figure 3.2. Multimeric pJEN3 and pMIN33 plasmid forms are detected on the gel in those tracks in which the DNA originated from DS941, DS956 (*argR*) or DS957 (*pepA*) (pJEN3, lanes 16-18; pMIN33, lanes 21-23). No such multimers were detected when DS981 (*xerC*) or DS9008 (*xerD*) were transformed with pJEN3 (lanes 19-20) or pMIN33 (lanes 24-25). This indicates that a type II hybrid minimal site plasmid is able to undergo intermolecular recombination mediated by XerC and XerD, but independent of ArgR and PepA. The type II hybrid core recombination site can therefore recombine without resolution selectivity, giving similar results to those obtained with *dif*. In contrast, the minimal site plasmids of *cer* (pJEN1), the type I hybrid (pJEN2) and *parB* (pJEN4) did not undergo intermolecular XerC/D-mediated recombination, as no multimeric forms were detected in transformants of DS941, DS956 or DS957.

These results are in agreement with the findings of David Summers (Summers, 1989). His experiments showed that the type II hybrid could support recombination in the absence of ArgR and PepA, and that deletions into the accessory DNA of the type II hybrid did not prevent recombination, but under these conditions, recombination now favoured the intermolecular reaction, leading to the formation of plasmid multimers. The data presented in Figure 3.2 confirm that a type II hybrid core recombination site functions as a substrate for

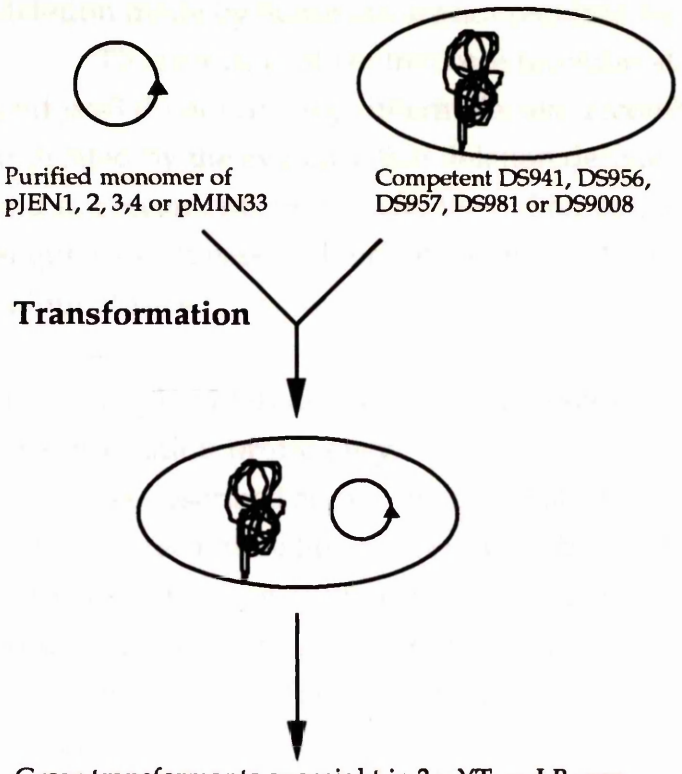
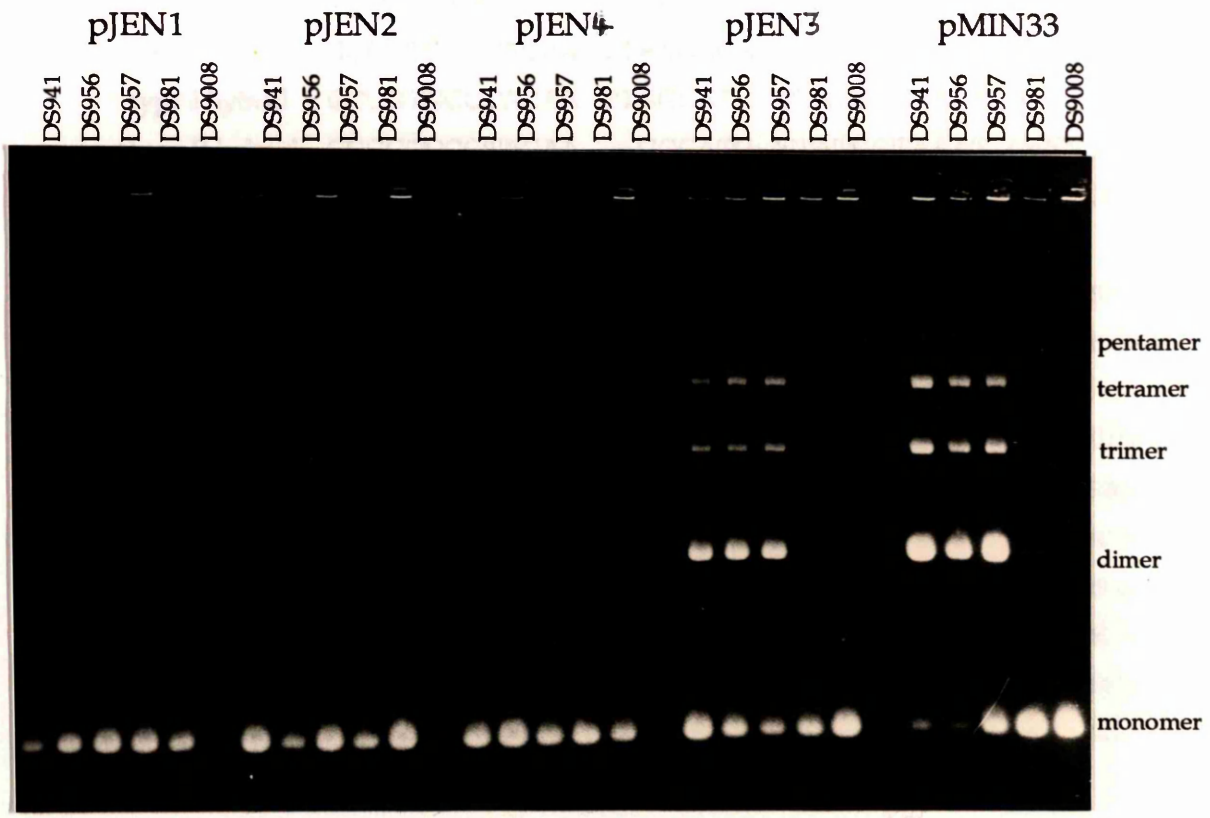


Figure 3.2: Assay for intermolecular Xer-mediated recombination between minimal sites *in vivo*. Purified plasmid monomers were transformed into competent DS941, DS956, DS957, DS981 and DS9008. After overnight propagation of transformants, the multimerisation state of the minimal site plasmids was analysed. The agarose gel below shows the results of such an experiment.The plasmid multimerisation state is annotated along the right edge of the gel.



XerC/D-dependent recombination: the remaining sequences left in the smallest deletion made by Summers are not required for recombination.

The results confirm that core recombination sites of *cer*, the type I hybrid and *parB* do not undergo intermolecular recombination. This was previously indicated by the evidence that deletion derivatives of *cer* and the type I hybrid did not recombine (Summers, 1989). This suggests that the intact accessory DNA sequences are essential for recombination between sequences that are inactive as minimal sites.

Type II hybrid left arm sequence is responsible for its minimal site recombination proficiency

As described above, the minimal site plasmids pJEN1, 2, 3 and 4 contain only core recombination site sequence from *cer*, the type I and type II hybrids and *parB*, and therefore differences in recombination proficiency must be attributed to sequence differences in these regions. The sequences of the core recombination sites are aligned in Table 3.2 to highlight the differences in sequences between the sites:

Table 3.2: Alignment of core recombination site sequences from *cer*, *parB* and the type I and type II hybrids

	left arm XerC binding	central region	right arm XerD binding	
cer	GCGGTGCGTACAA	**TTAAGGGA**	TTATGGTAAAT	(pJEN1)
type I hybrid	GCGGTGCGTACAA	**TT**GGGATG	TTATGGTAAAT	(pJEN2)
type II hybrid	GCGGTGCGTACAA	GGGATG	TTATGGTAAAT	(pJEN3)
parB	GCGGT**A**C**CG**A**T**AA	GGGATG	TTATGGTAAAT	(pJEN4)

All four of these sites have identical XerD binding sites. The differences between the type II hybrid and the three other sites are shown in bold-face font. *cer* and the type II hybrid differ in the length and sequence of their central regions indicating that the central region contributes to the ability or inability of a minimal site to mediate intermolecular recombination. The type I hybrid has a 2 bp insertion in its central region compared with the type II hybrid. This must be responsible for its inability to recombine as a minimal site as the sequences are otherwise identical. In fact, a similar insertion in the *dif* central region results in a minimal site that does not undergo intermolecular recombination (G. Blakely, unpublished data). This indicates that changing the relative rotation of (and

distance between) the XerC and XerD binding sites by altering the length of the central region produces a *dif* variant which cannot undergo recombination as a minimal site.

However, the type II hybrid and *parB* both have identical central region sequences. The length of the central region is therefore not the only determinant of whether or not recombination will occur, since the results of the experiment in Figure 3.2 showed that *parB* clearly does not support Xer-mediated recombination as a minimal site. There are four single base substitutions in the left arm of the type II hybrid compared with *parB*. It must therefore be as a result of these changes that pJEN3 can undergo XerC/D-mediated intermolecular recombination, whereas pJEN4 cannot.

Construction of left arm variant sites

Since the type II hybrid can recombine as a minimal site, in contrast to *parB*, the changes in the left arm sequence that have allowed this to occur must be important in allowing minimal site recombination. A series of experiments was carried out to address the role of the left arm sequence of the type II hybrid in allowing minimal site recombination, and, hence, the role of the left arm in selectivity.

To determine which of the four substitutions in the left arm are responsible for changing a minimal site that does not recombine (*parB*) into one that does (type II hybrid), a series of plasmids was constructed in which each of the four changes was made alone or in every possible combination. To obtain these plasmids without the expensive process of synthesising two oligonucleotides for each of the fourteen possible variants, a pair of degenerate oligonucleotides of the following sequence was constructed, where the degeneracies are indicated by showing the sequence in brackets:

*Bam*HI *Xba*I

*GATCCTCTAGA***GCGGT(AG)C(CG)(GT)A(TC)AAGGGATGTTATGGTAAATAT***G*

*GAGATCT***CGCCA(TC)G(GC)(CA)T(AG)TTCCCTACAATACCATTTATA***CAGCT*

*Sal*I

These oligonucleotides were designed so that the resulting plasmids would be comparable to pJEN3 and pJEN4. The additional bases at the left end, to form a *Bam*HI site, were added to aid annealing of the oligonucleotides through the region of mismatch, and correspond to the polylinker sequences in pJEN3 and pJEN4, thereby preserving the flanking sequence.

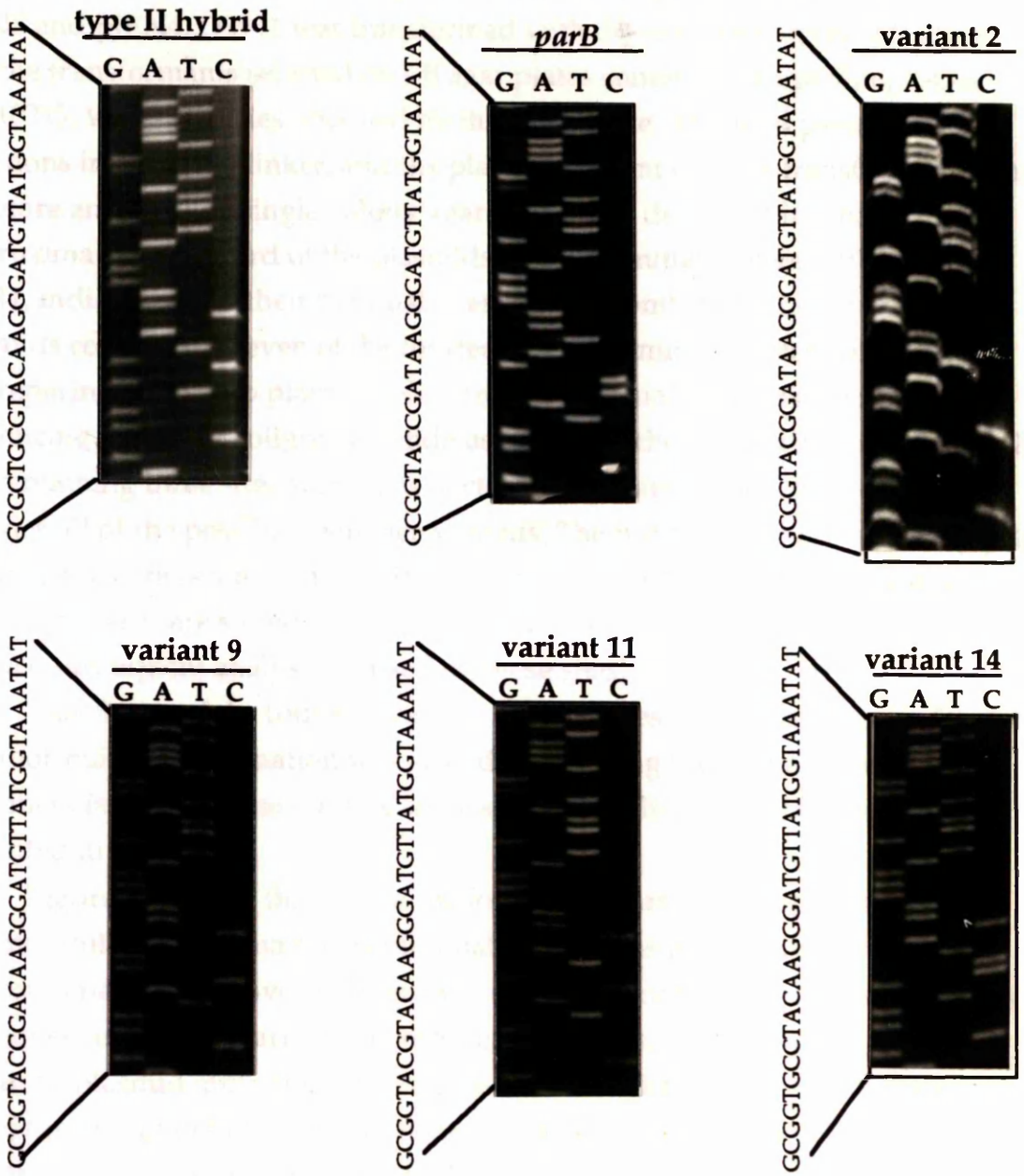


Figure 3.3: The sequences of the core recombination sites from pJEN3, pJEN4, pJEN72, pJEN79, pJEN711 and pJEN714.
The core recombination sites of the fourteen possible left arm variants were isolated by sequencing a pool of clones derived from degenerate oligonucleotides. This Figure shows the relevant portion of the sequencing gels on which the type II hybrid, *parB*, and variants 2, 9, 11 and 14 sequences were obtained. The sequences of these clones are shown since these sites were investigated further in this Chapter. The sequences of all of the clones are given in the text.

The oligonucleotides were purified and annealed together, then ligated into pUC18 as a *Bam*HI to *Sal*I fragment, thus creating equivalent plasmids to pJEN3 and pJEN4. DS941 was transformed with the resulting ligated plasmids, and the transformants selected on LB agar plates containing ampicillin, X-Gal and IPTG. White colonies obtained by this procedure contained plasmids with insertions in their polylinker, and the plasmid content of these transformants was therefore analysed on single-colony agarose gels (as described in Chapter 2). Approximately one-third of the plasmids screened initially were multimeric in DS941, indicating that their minimal sites could recombine (data not shown). Plasmids containing eleven of the fourteen possible minimal sites were isolated by preparing mini-prep plasmid DNA from individual transformants and sequencing the cloned oligonucleotide using the methods described in Chapter 2. The remaining three sites were constructed as separate oligonucleotides, to produce all of the possible sequence variants. The portions of the sequencing gels showing the core recombination site sequences of pJEN3, pJEN4 and four of the left arm variants are shown in Figure 3.3 (these four variants are presented because subsequent analysis focused on these sites).

The ability of the fourteen variant left arm sites to act as substrates for Xer intermolecular recombination was scored by assaying the formation of plasmid multimers *in vivo* on agarose gels (as described in Chapter 2). The results are presented in Figure 3.4.

Figure 3.4 shows that five of the fourteen variants can support intermolecular recombination as minimal sites. It was noticeable that the extent of recombination after overnight growth of transformants varied, depending on the sequence of the left arm of the recombination site. For example, pJEN72 produced plasmid multimers, but to a lesser extent than pJEN711. An attempt to assign the amount of plasmid multimerisation as a result of XerC/D-mediated recombination between minimal sites has been made in Table 3.3. In this Table, the four variable bases are highlighted in bold-face font. If the plasmid was almost exclusively in monomeric form irrespective of the Xer status of the cell, no Xer-mediated intermolecular recombination was said to occur, and this is denoted by (-). If extensive plasmid multimerisation had occurred in DS941, DS956 and DS957, but not in DS981 or DS9008, so that higher multimeric forms in addition to monomers and dimers were easily detected, this was assigned the value (+++). Where the extent of recombination was intermediate, this was assigned a value of (+) or (++) depending on the distribution of oligomeric forms observed in Figure 3.4. (For recombination at *parB* and the type II hybrid, refer to Figure 3.2)

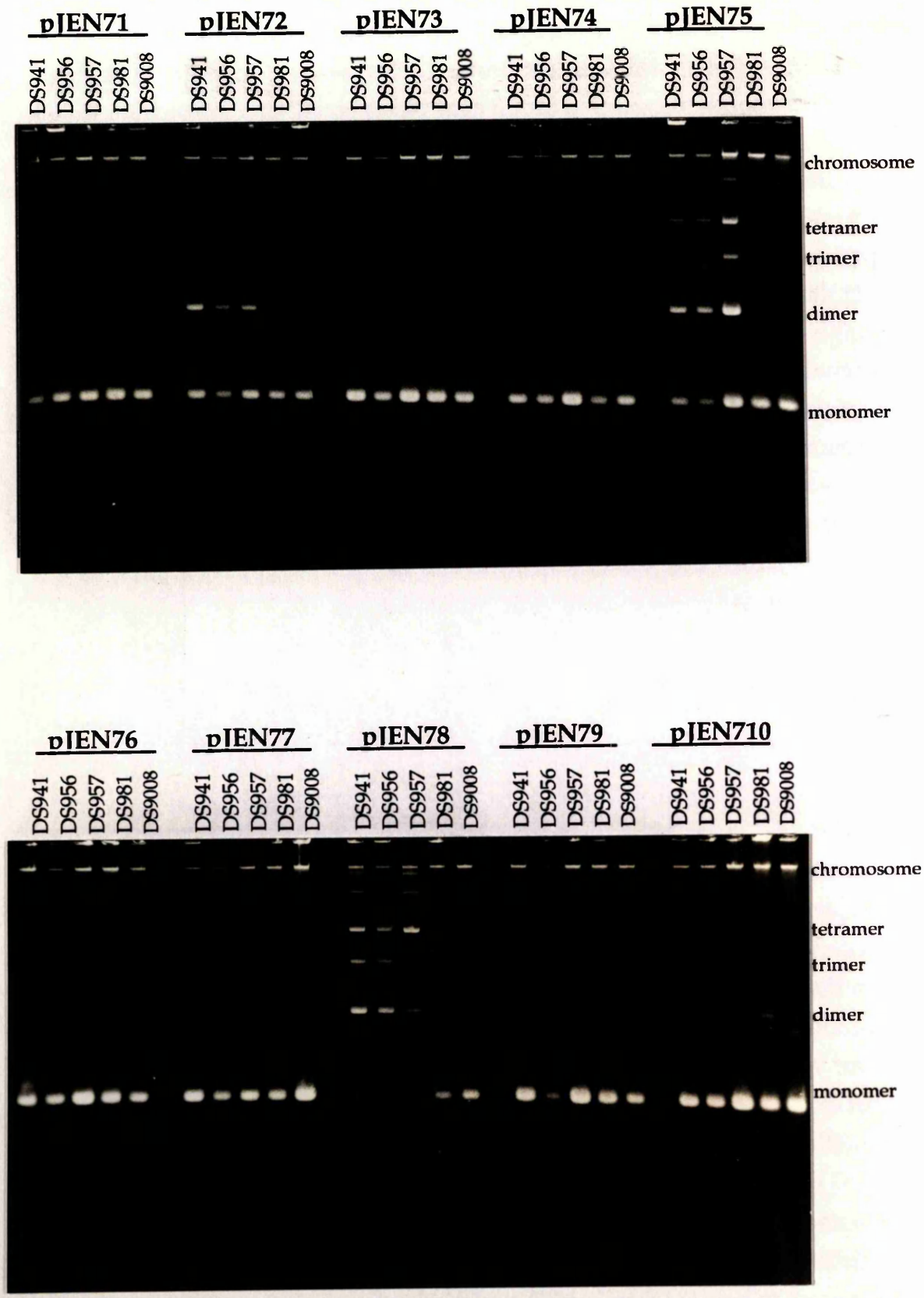


Figure 3.4: For description see facing page.

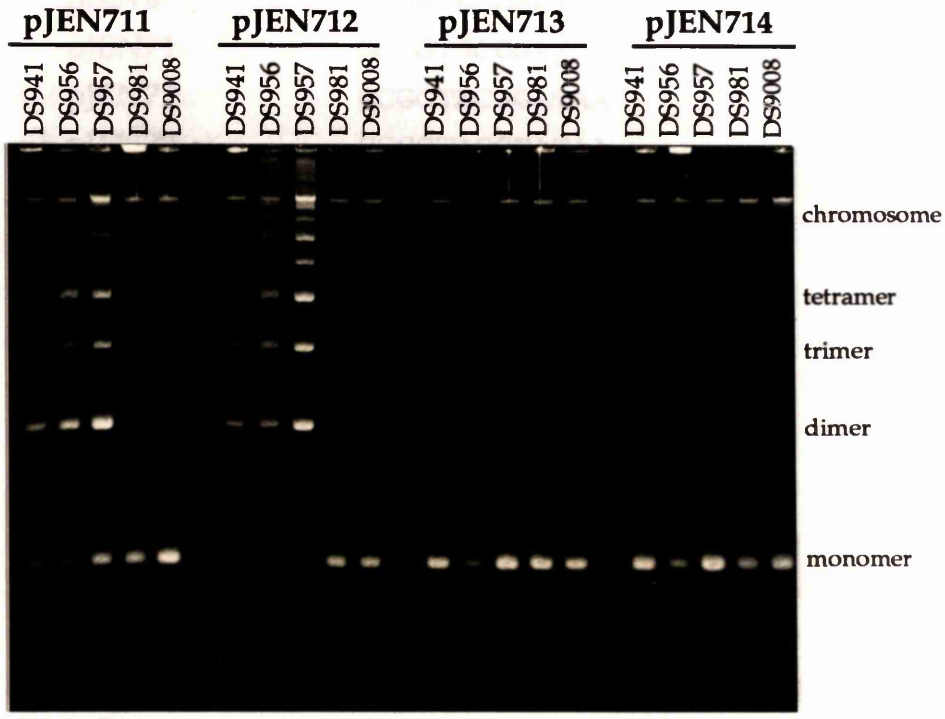


Figure 3.4: Intermolecular recombination at variant sites *in vivo*. *In vivo* intermolecular recombination was assayed as in Figure 3.2. The plasmid content of a population of cells scraped from the transformation plate was analysed by lysis in single colony final sample buffer, then electrophoresis on a 1% agarose gel. The major plasmid oligomerisation states are annotated alongside the photograph, although some higher forms can also be seen. An attempt to assign extent of multimerisation is given in the text.

Table 3.3: Intermolecular recombination between variant left arm sites *in vivo*

	Plasmid name	Left arm sequence	Intermolecular recombination
		1 23 4	
(*parB*)	pJEN4	GCGGT**A**C**CG**A**T**AA	-
	pJEN71	GCGGT**G**C**CG**A**T**AA	-
	pJEN72	GCGGT**A**C**GG**A**T**AA	+
	pJEN73	GCGGT**A**C**CT**A**T**AA	-
	pJEN74	GCGGT**A**C**CG**A**C**AA	-
	pJEN75	GCGGT**G**C**GG**A**T**AA	++
	pJEN76	GCGGT**G**C**CT**A**T**AA	-
	pJEN77	GCGGT**G**C**CG**A**C**AA	-
	pJEN78	GCGGT**A**C**GT**A**T**AA	+++
	pJEN79	GCGGT**A**C**GG**A**C**AA	-
	pJEN710	GCGGT**A**C**CT**A**C**AA	-
	pJEN711	GCGGT**A**C**GT**A**C**AA	+++
	pJEN712	GCGGT**G**C**GT**A**T**AA	+++
	pJEN713	GCGGT**G**C**GG**A**C**AA	-
	pJEN714	GCGGT**G**C**CT**A**C**AA	-
(type II)	pJEN3	GCGGT**G**C**GT**A**C**AA	+++

The sequences of the variant sites indicate which bases in the left arm determine whether or not minimal site recombination occurs. pJEN72 has one base of an otherwise *parB* left arm substituted with a single type II hybrid base at the position marked "2". This allows this sequence to function as a minimal site. Therefore a single base change is sufficient to alter a site that would not otherwise recombine without accessory factors into one that does. All of the sequences that recombine have this C to G change at position (2). The context of this base is important; the extent of recombination achieved with pJEN72 is increased by making other changes such as the A to G change at position (1) in pJEN75, and the G to T change at position (3) in pJEN78. However, pJEN79 also has a G at position (2), but will not recombine as a minimal site. This must be due to the additional T to C change at position (4) that is present in this plasmid. Therefore, while this base is necessary for minimal site function, it is not always sufficient. In order for a site with the T to C change at position (4) to recombine, it must also have the third substitution of G to T at position (3) as in plasmid pJEN711. The G to T change at position (3) also appears very important: all of the

minimal sites that recombine well (pJEN78, 711, 712 and pJEN3) have a T at this position as well as a G at position (2).

pJEN714 had the reciprocal G to C change at position (2) to the C to G change in pJEN72. This substitution prevents an otherwise type II hybrid sequence recombining. This confirms that the base at position (2) is pivotal in allowing or disallowing a minimal site to function as a substrate for intermolecular recombination.

In summary, the results of this experiment confirm that the sequence of the left arm of *cer*-like sites strongly influences the ability of minimal sites to recombine. It appears to be most critical that the base at position (2) is a G, although the other three bases modulate whether or not recombination occurs, and the proficiency of recombination, as gauged by the proportion of plasmid multimers formed in an *in vivo* assay.

Binding of XerC and XerD to variant left arm sites

The simplest and most obvious explanation for the observed differences in recombination proficiency at minimal sites is that different sequences interact differently with the recombinases. Changes in left arm sequence are likely to affect interactions with the one or both of the recombinases, and this in turn might affect recombination proficiency. Gel retardation and footprinting experiments have shown that XerC binds to the left arm of *dif* and XerD to the right *in vitro* (Blakely *et al.*, 1993), so it seems more likely that the interaction with XerC would be directly affected by altering the sequence of the left arm. However, since XerC and XerD bind co-operatively to *dif* and *cer in vitro*, it is possible that left arm mutations affect the ability of both recombinases to bind.

The gel retardation or band shift assay (Fried & Crothers, 1981) is a widely used approach for investigation of protein-DNA interactions (for reviews see Carey, 1991; Lane *et al.*, 1992). The popularity of the method is a result of its technical simplicity, and the requirement for small quantities of protein and DNA. The technique is based on the observation that when proteins bind to linear DNA, the migration of that DNA fragment is retarded during electrophoresis on non-denaturing gels. The extent of retardation depends on the mass and charge of the protein, and on any changes in the conformation of the DNA caused by protein binding. The effect of protein mass on DNA mobility was first established for binding of the Lac repressor to operator-containing DNA fragments: the relative mobilities of the complexes obtained by binding correlated with those expected if successive additions of the repressor tetramer were binding (Fried & Crothers, 1981). The effect of protein mass can be observed

since the DNA fragments used (typically 60-500 bp; 40 to 330 kDa) are of similar mass to typical DNA-binding proteins (10-100 kDa; XerC & XerD ~34 kDa). That protein charge can affect complex mobility was demonstrated by studies with the Trp repressor (Carey, 1988). This relatively acidic protein (pI ~6) effectively retards operator DNA only at low pH; at pH above its isoelectric point, the net negative charge of the protein results in acceleration of the complex, thereby compensating for any effects of mass or DNA conformation effects. Thus under some conditions, no retardation of operator fragments by the Trp repressor is detected. The effect of proteins on DNA conformation, and the subsequent effect on complex mobility is discussed later.

Sites from the following plasmids were selected for gel retardation analysis as their sequences were considered the most informative about the roles of particular bases in minimal site function: pJEN3 (type II hybrid), pJEN4 (*parB*), pJEN72 (variant 2), pJEN79 (variant 9), pJEN711 (variant 11) and pJEN714 (variant 14). 1 μg of each of these plasmids was digested with *Bam*HI and *Pst*I and end-labelled at the *Bam*HI end with 10 μCi of ^{32}P-αdGTP, and the radiolabelled fragments purified using NucTrap Push columns as described in Chapter 2. Labelling gave DNA fragments of approximately the same specific radioactivities. Binding was carried out in GBA1 buffer, with polydI.dC carrier DNA as described in Chapter 2, and incubation was for 10 minutes at 37°C. Crude protein extracts containing XerC (in the absence of XerD) or XerD (in the absence of XerC) were prepared from DS9009 pSDC105 (*xerC, xerD* strain containing an XerC-overexpressing plasmid) or DS9009 pRM130 (*xerC, xerD* strain containing an XerD-overexpressing plasmid) respectively in order to assay binding of either protein individually. In each binding reaction, the same amount of the same extracts were used; however, the concentration of active recombinase in each extract was not known. Gel conditions were as described in Chapter 2.

Figure 3.5 shows the results of such a binding experiment. Several observations can be made from these gels. Firstly, it was impossible to detect DNA-XerC complexes with any of these sites, indicating either that XerC binds weakly to all of these sequences, or that the amount of active XerC in the extracts was low. In no such experiment was the XerC complex visible, unfortunately eliminating the possibility of comparing binding of XerC alone to any of these fragments. However, this is consistent with experiments which showed that binding by XerC is reproducibly weaker than that of XerD to *dif*, the XerC-DNA complex being seldom visible. In addition, both proteins bind less well to *cer* (both separately and together), and XerC-DNA complexes are usually not observed at *cer* (Blakely *et al.*, 1993; G. May, personal communication).

Figure 3.5a

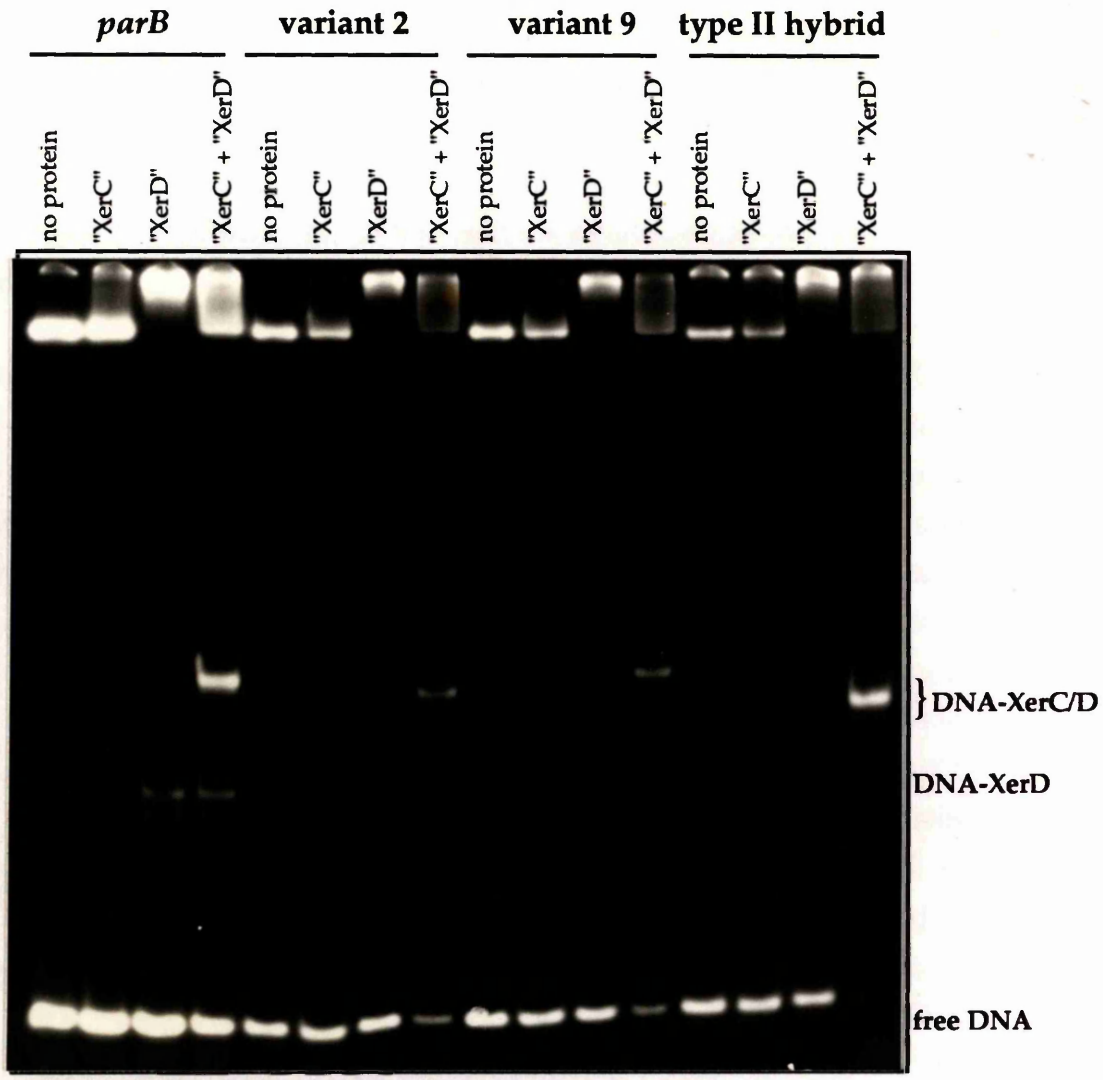
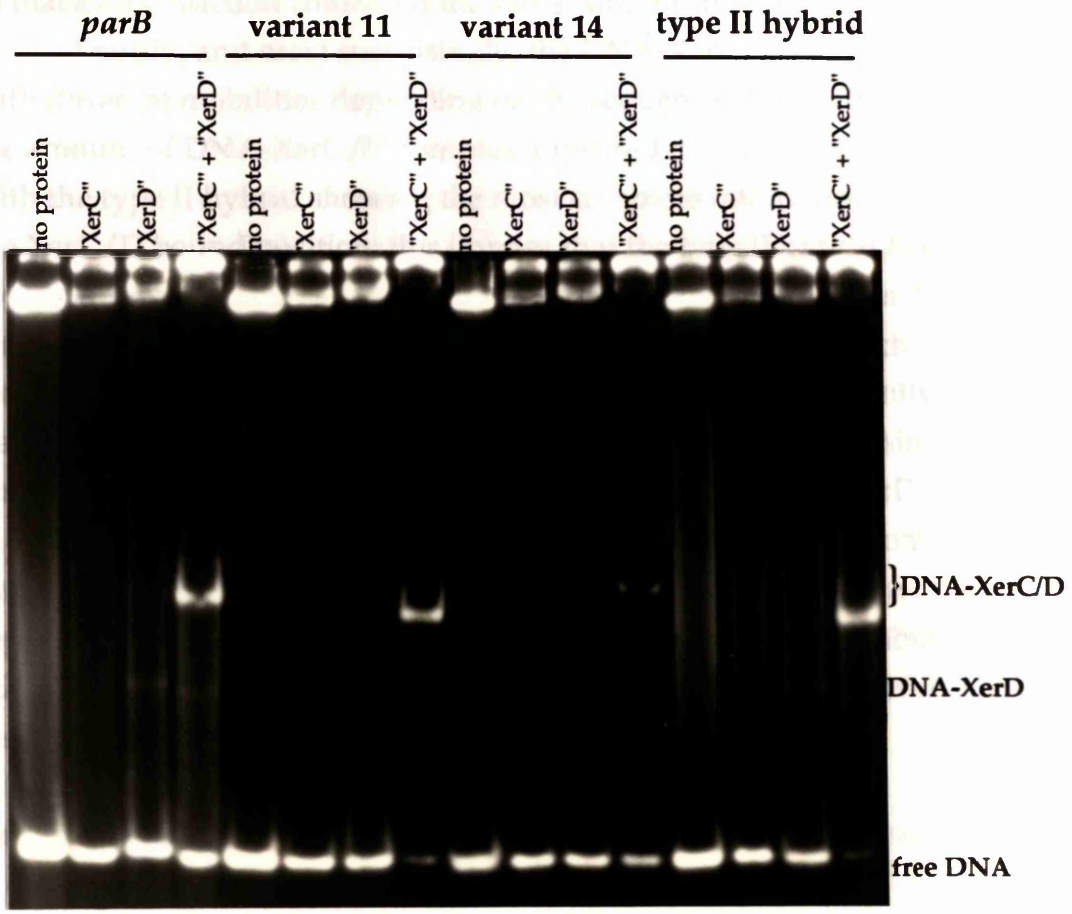


Figure 3.5: Gel retardation assays of left arm variant sites bound with XerC and XerD.

Gel retardation assays were performed as described in the text. The protein extracts added to the reactions loaded into each track are shown at the top of the photographs, along with the identity of the DNA site. The protein names are in inverted commas to indicate that crude extracts were used: "XerC" refers to the extract from DS9009 pSDC105, "XerD" means DS9009 pRM130. The positions of the XerD and XerC/D complexes are indicated alongside the gels. Figures 3.5a and b (facing page) show the binding of XerC and XerD individually and in combination to the six sites. The same source of protein extract was used for each reaction mixture, and the concentration of DNA in each was equal.

Figure 3.5b


parB
variant 11
variant 14
type II hybrid
no protein
"XerC"
"XerD"
"XerC" + "XerD"
no protein
"XerC"
"XerD"
"XerC" + "XerD"
no protein
"XerC"
"XerD"
"XerC" + "XerD"
no protein
"XerC"
"XerD"
"XerC" + "XerD"
DNA-XerC/D
DNA-XerD
free DNA

The second observation that can be made from Figure 3.5 is that all of the sites appear to have bound XerD comparably, both in terms of the amount of complex formed, and in the mobility of the complex: this is not unexpected given that they all share identical XerD binding sites. This comparison can be made given that the DNA concentrations for the six sites tested were equal, and because equal amounts of the same XerD-containing extract were used for each so that every reaction contained the same concentration of XerD.

Thirdly, and most surprisingly, the DNA-XerC/D complexes migrated with different mobilities depending on the sequence of the left arm. In addition, the amount of DNA-XerC/D complex appeared to depend on left arm sequence, with the type II hybrid showing the most complete retardation of the DNA into the XerC/D bound position; this implies that the type II hybrid has the highest overall affinity for XerC and XerD compared with the other sites. Variant 11 was retarded to a similar extent, implying that those sequences which produce a less retarded DNA-XerC/D complex may have a higher overall affinity for XerC and XerD. Again, these observations require that each reaction contained, within the limits of error, equivalent concentrations of DNA, XerC and XerD.

The relative mobilities (R_f values) of the DNA-XerC/D complexes of these fragments compared with the unbound DNA were calculated on at least two separate occasions, and the means are shown in Table 3.4. The ability, or otherwise, of each of the DNA substrates to support intermolecular recombination as minimal sites is also shown.

Table 3.4: Relative mobilities of DNA-XerC/D complexes with variant sites

DNA site	R_f value	Minimal site function
type II hybrid	0.60	+++
parB	0.56	-
variant 2	0.58	+
variant 9	0.56	-
variant 11	0.60	+++
variant 14	0.55	-

The ability of minimal sites to support intermolecular Xer recombination may therefore be correlated with the formation of a less retarded DNA-XerC/D complex as detected by a gel retardation assay. It is noteworthy that the variant 2-XerC/D complex migrated at a position intermediate to those of the type II

hybrid and *parB*; the extent of plasmid multimerisation as a result of Xer-mediated recombination at this site *in vivo* was also intermediate (see Figure 3.3).

What is the basis of the differences in DNA-XerC/D complex mobility? In the above discussion it has been assumed that all of the sites have bound the same proteins (XerC and XerD) under identical experimental conditions, and that differences in the rates of electrophoretic migration cannot, therefore, be attributed to differences in the mass or charge of the proteins bound to the DNA fragments. While this cannot be absolutely certain because crude extracts were used, it seems likely that all of the sites should bind XerC and XerD, rather than other unidentified proteins, given their site sequence homologies. Differences in relative mobility might therefore be as a result of changes in complex conformation, perhaps as a result of sequence-directed changes in protein-induced DNA bending as a result of binding by XerC and XerD.

XerC and XerD induce different DNA bends depending on left arm sequence

Many DNA-binding proteins produce bends in DNA upon binding; for example, bends induced by the CAP, Fis, GalR, Xis and IHF proteins have been calculated (Thompson & Landy, 1988; Kim *et al.*, 1989; Zinkel & Crothers, 1991). The crystal structures of the restriction endonuclease *Eco*RV bound to cognate and non-cognate DNA fragments have recently been solved, revealing that the enzyme induces a kink of ~50° at the position of cleavage with a cognate site, in contrast to the non-cognate site, with implications for the mechanism of strand-cleavage (Winkler *et al.*, 1993). However, crystal structures are available for few protein-DNA complexes, and bending induced by protein binding is therefore generally determined indirectly. The methods used most frequently rely on gel electrophoretic techniques. Two methods are generally applicable to the study of DNA bends: circular permutation analysis (Wu & Crothers, 1984) and phasing analysis (Zinkel & Crothers, 1987; Salvo & Grindley, 1987). Circular permutation analysis was adopted to determine whether XerC and XerD produce different bends upon binding sites with different left arm sequences.

The rationale for using circular permutation analysis is as follows. If a DNA fragment is bent, the number, magnitude and position of the bend(s) within the DNA molecule determine the rate of migration of that fragment through polyacrylamide gels (Wu & Crothers, 1984; Koo *et al.*, 1986). This tenet is derived from the mode by which DNA migrates through a polyacrylamide gel. DNA is believed to move by a snake-like "reptation" through the pores of the gel as a vertical rod (de Gennes, 1971; Deutsch, 1988; Lumpkin *et al.*, 1985; Lumpkin & Zimm, 1982; Levene & Zimm, 1989). Bends within the DNA decrease the end-

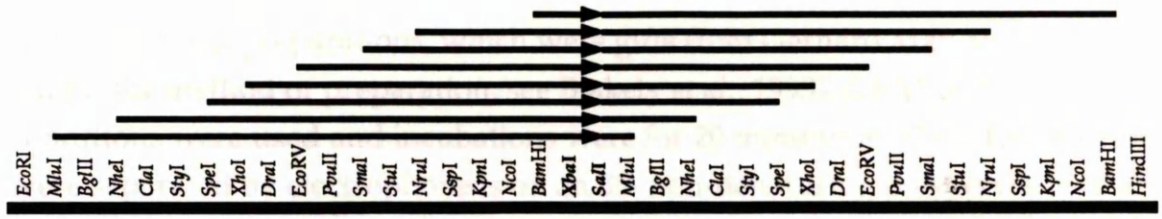
to-end distance of the molecule and retard electrophoretic mobility; the closer the bend is to the centre of the DNA fragment the slower the migration, as the cross-sectional area of the fragment is at its largest. This point is represented diagrammatically in Figure 3.6. If there is more than one bend in a DNA fragment, the relative directions and magnitudes of the bends will affect the rate of migration of the fragment through the gel: two bends in the same plane and direction will decrease the end-to-end distance of the molecule and maximally retard migration, bends in opposing directions will have a minimal effect on migration (reviewed in Lane *et al.*, 1992).

Bends within a DNA fragment may be caused in two main ways: the sequence of the DNA itself can result in intrinsic bending, such as the cumulative effect of periodically spaced A residues (Hagerman, 1985; Wu & Crothers, 1984), or proteins binding to a DNA fragment can induce a bend, or stabilise an intrinsic bend, as first described for the CAP protein of *E. coli* with the *lac*, *mal*, and *gal* operators (Kolb *et al.*, 1983; Wu & Crothers, 1984). The sequence of the DNA at the binding site can cause the same protein to induce bends of different magnitudes. For example CAP has been shown to produce different bends with binding sites of different sequence (Crothers *et al.*, 1991).

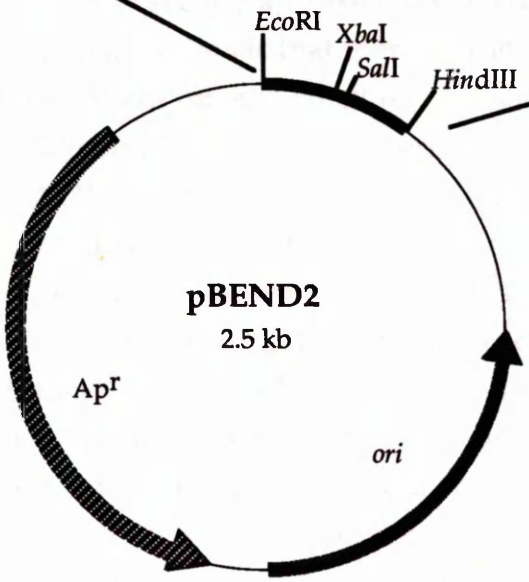
In the following experiments, circular permutation was used to analyse the differences in mobilities between DNA-XerC/D complexes of sites with different left arm sequences, using the vector pBEND2 (Kim *et al.*, 1989). The polylinker cloning site of this plasmid consists of two unique restriction sites, *Xba*I and *Sal*I, flanked by an additional seventeen sites in tandem repeat. Restriction digestion with several of these additional enzymes produces a series of DNA fragments of identical size in which the cloned sequence, containing the potential bending site, is moved progressively throughout the fragment. A map of pBEND2, and a description of the plasmids constructed for circular permutation analysis, is given in Figure 3.6.

The *Xba*I to *Sal*I fragments from pJEN3, pJEN4, pJEN72, pJEN711 and pJEN714 containing the core recombination sites were sub-cloned into pBEND2 at the unique *Xba*I and *Sal*I sites. The resulting clones were named pJEN33, pJEN43, pJEN723, pJEN7113 and pJEN7143 respectively. No such similar clone was obtained for variant 9.

Two similar approaches for bend analysis were undertaken. Under the first conditions, four restriction enzymes were selected and, after digestion of pJEN33, pJEN43, pJEN723 and pJEN7143, the resulting fragments were end-labelled. The restriction enzymes used in this case were *Mlu*I, *Xho*I, *Xma*I and *Bam*HI. The DNA fragments were used in binding assays with partially purified



pBEND2 polylinker



Circular permutation analysis was carried out using the vector pBEND2. The order of the restriction sites within the polylinker is shown above: the sites are in tandem repeat, such that digestion with a series of these enzymes produces a set of fragments of similar length, within which the putative bending site is moved progressively from one end to the other. This point is represented diagramatically at the top of the Figure, where the cloned fragment is shown as a filled triangle, and its position within the fragment is indicated for each of the restriction enzymes used in these experiments. pJEN33 consists of the type II hybrid minimal site cloned into pBEND2, *XbaI* to *SalI*. *parB* is cloned in pJEN43, variant 2 in pJEN723, variant 11 in pJEN7113, and variant 14 in pJEN7143.

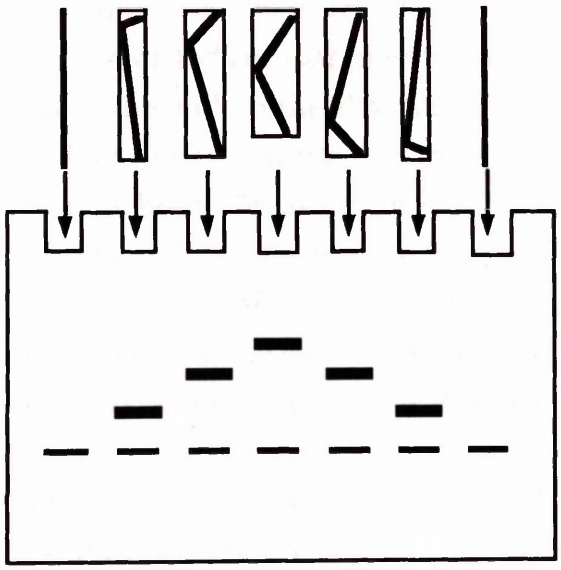


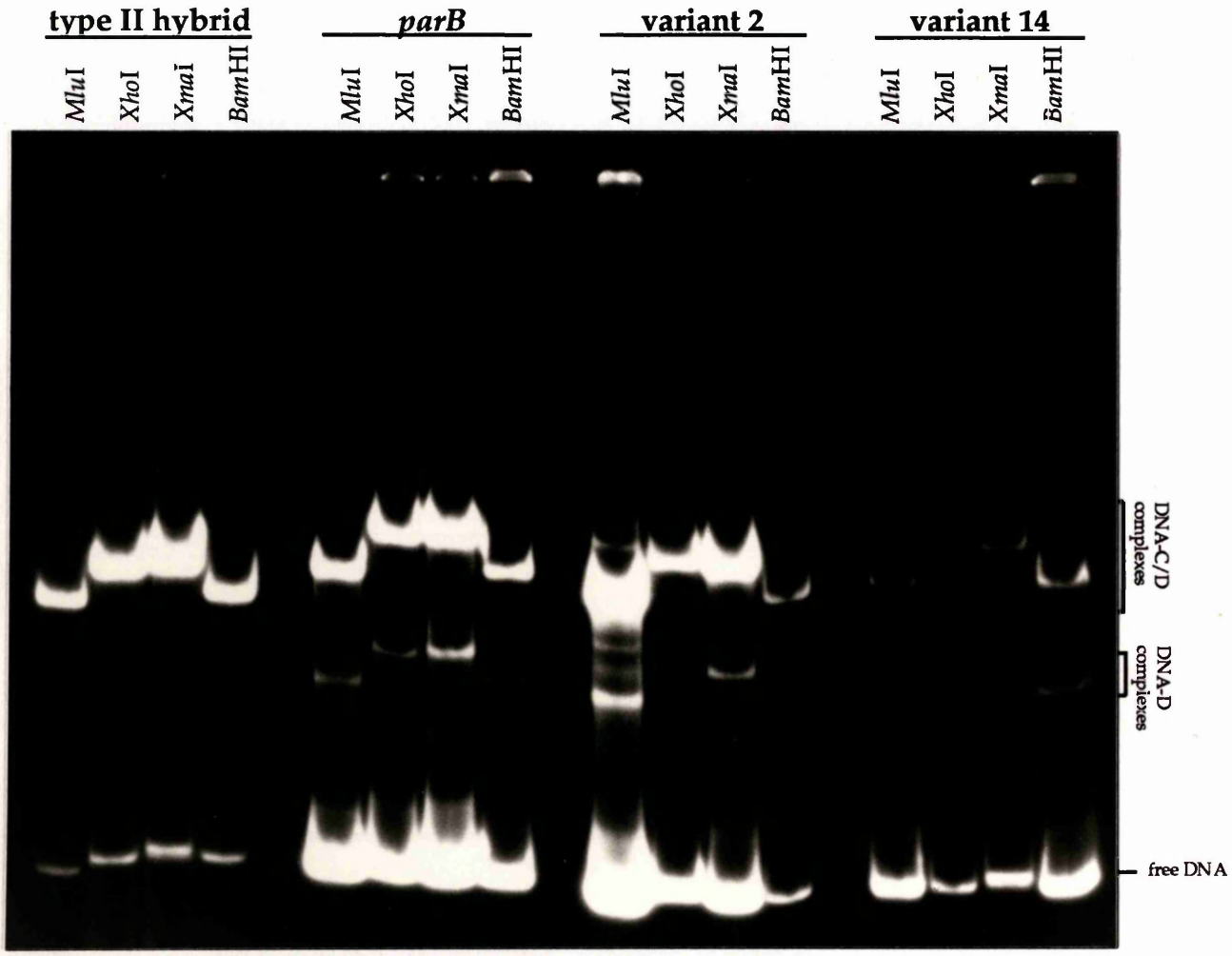
Figure 3.6: Circular permutation analysis of distorted DNA fragments.
If a DNA fragment is bent, its rate of migration through a polyacrylamide gel will depend on the position of the bend. The most retarded fragments are those with the bend nearest the centre, since in these fragments the end-to-end distance of the molecule is least, and the cross-sectional area greatest, therefore reptation is most hindered. The cartoon on the left illustrates this point: the position of bending in each fragment is shown at the top, and the expected position of the complex on a polyacrylamide gel is shown as a thick black line. The position of an unretarded fragment of the same size is indicated as a thin line.

XerC and XerD preparations, which were gifts from Gerhard May and Karen Grant (for method of preparation, see Blakely *et al.*, 1993). GBA3 buffer conditions were used and incubations were for 20 minutes at 37°C. The reactions were separated by electrophoresis on an 8% non-denaturing polyacrylamide gel, as described in Chapter 2. The results of such an experiment are shown in Figure 3.7.

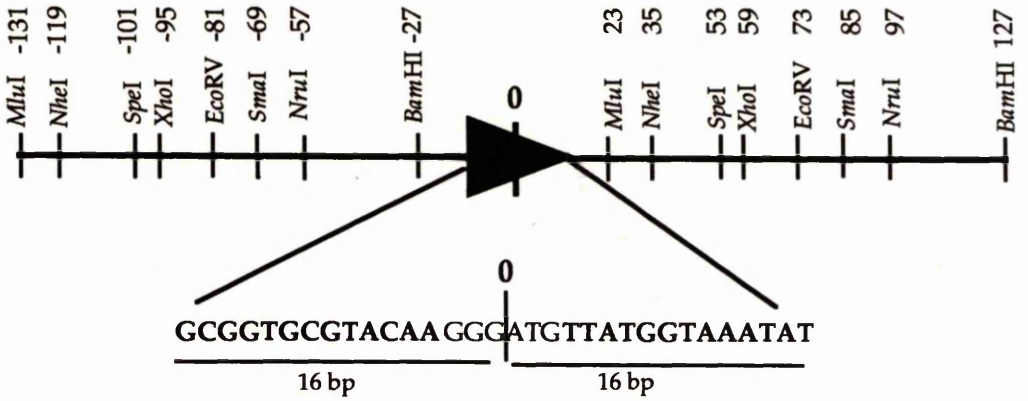
The second approach involved using non-radioactive fragments. The advantage of this is that more restriction enzymes can be used, since there is no requirement for end-labelling. A disadvantage is reduced sensitivity in detection of the complexes. In this experiment, 2 µg of each of pJEN33, pJEN43, pJEN723, pJEN7113 and pJEN7143 were digested with *Mlu*I, *Nhe*I, *Spe*I, *Eco*RV, *Sma*I, *Nru*I and *Bam*HI in 1x KGB buffer in a volume of 10 µl, then 10 µl of 2x GBA2 was added. (The composition of GBA2 is described in Chapter 2; this was used so that no salt was added to the binding reactions in addition to the 50 mM potassium glutamate from the KGB buffer, and the 50 mM NaCl from the proteins). Partially purified XerC and XerD were used (from Gerhard May). The proteins were mixed at 2 volumes XerC to 1 volume XerD, and 4 µl of this mixture was added to each binding reaction. Binding reactions were incubated for 30 minutes at 37°C, then electrophoresed on 8% non-denaturing polyacrylamide gels. The DNA was then visualised by ethidium bromide staining. Figure 3.8a shows the results of this experiment.

Figures 3.7 and 3.8 show that binding of XerC and XerD to these DNA fragments did indeed produce bends: as the core recombination site was moved towards the centre of the DNA fragment the DNA-XerC/D complex became more retarded compared with those fragments with the recombination site nearer the end. In the absence of protein, the fragments migrated in a relatively straight line indicating that there was no intrinsic DNA distortion (data not shown), however, as can be seen in Figure 3.7a, the unbound DNA did not produce an entirely straight line on this gel in the presence of protein. This is probably an artefact of electrophoresis on this particular gel. To correct for these perturbations, the distance migrated from the well to the XerD or XerC/D complex was divided by the distance migrated by the unbound DNA, giving R_f values. The distance migrated was measured from the well to the front of the band in each case. To present the data in the form of a graph, the "intercentre distance" for each restriction site was calculated and plotted on the X-axis. To calculate this value, the centre of the core recombination site within each of the cloned fragments was assigned (as position zero), and the mid-point of the restriction fragment was determined. Using these two values, the intercentre

Figure 3.7: Circular permutation analysis of XerD and XerC/D-induced bends at the type II hybrid, *parB*, variant 2 and variant 14.
Figure 3.7a (below) shows the results of a gel retardation experiment with circularly permuted fragments containing the XerC/D binding sites from the type II hybrid, *parB*, variant 2 and variant 14. The restriction fragments of pJEN33, pJEN43, pJEN723 and pJEN7143 are detailed at the top of the photograph.



On the facing page, the relative mobilities of the XerD- and XerC/D-induced bends are plotted as a function of the intercentre distance for each of the restriction fragments. An explanation of how the intercentre distance was derived is given at the top of the page. The graphs are shown in Figure 3.7b and 3.7c. The y-axis values are plotted so that the curve of the graph mimics the curve on the gel in Figure 3.7a.



In the cartoon above, the positions of cutting by the restriction enzymes used in both Figure 3.7 and 3.8 are shown. The positions are given positive or negative values depending on whether they fall to the right or left of "position zero". Position zero (shown here by **0**) is in the centre of the central region: its position within the sequence of the type II hybrid is shown. The "intercentre distance" is the distance between position zero and the midpoint of each restriction fragment. The intercentre distances for all of the restriction enzymes used in both experiments are listed below: *Bam*HI= +50, *Nru*I= +20, *Sma*I= +8, *Xma*I= +6, *Eco*RV= -4, *Xho*I= -18, *Spe*I= -24,*Nhe*I= -42, *Mlu*I= -54

Figure 3.7b: Circular permutation analysis of XerC/D-induced bends

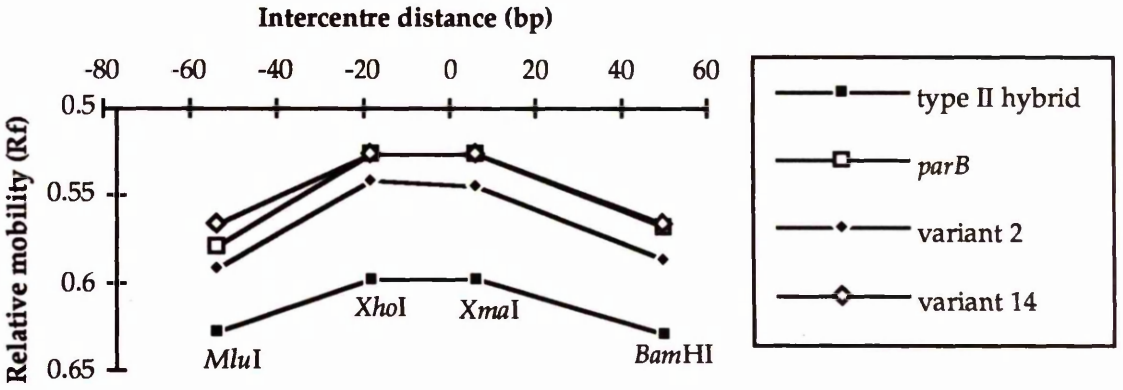


Figure 3.7c: Circular permutation of XerD-induced bends

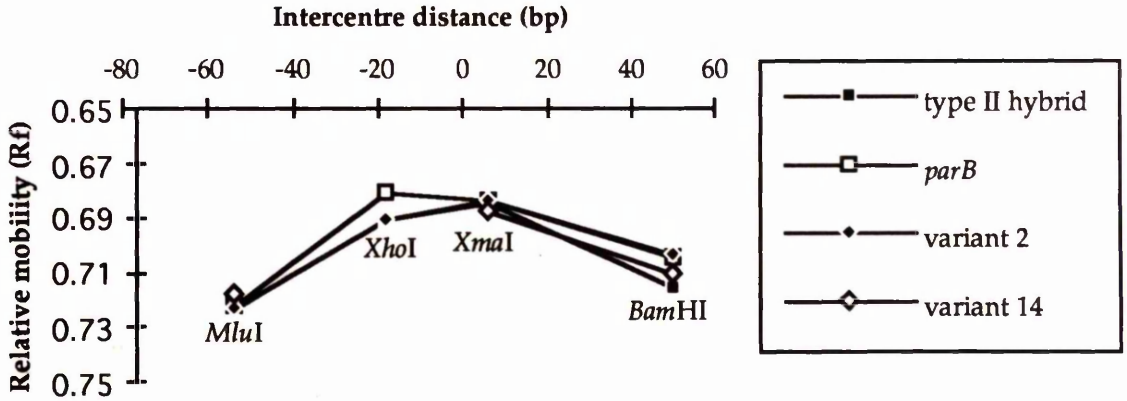
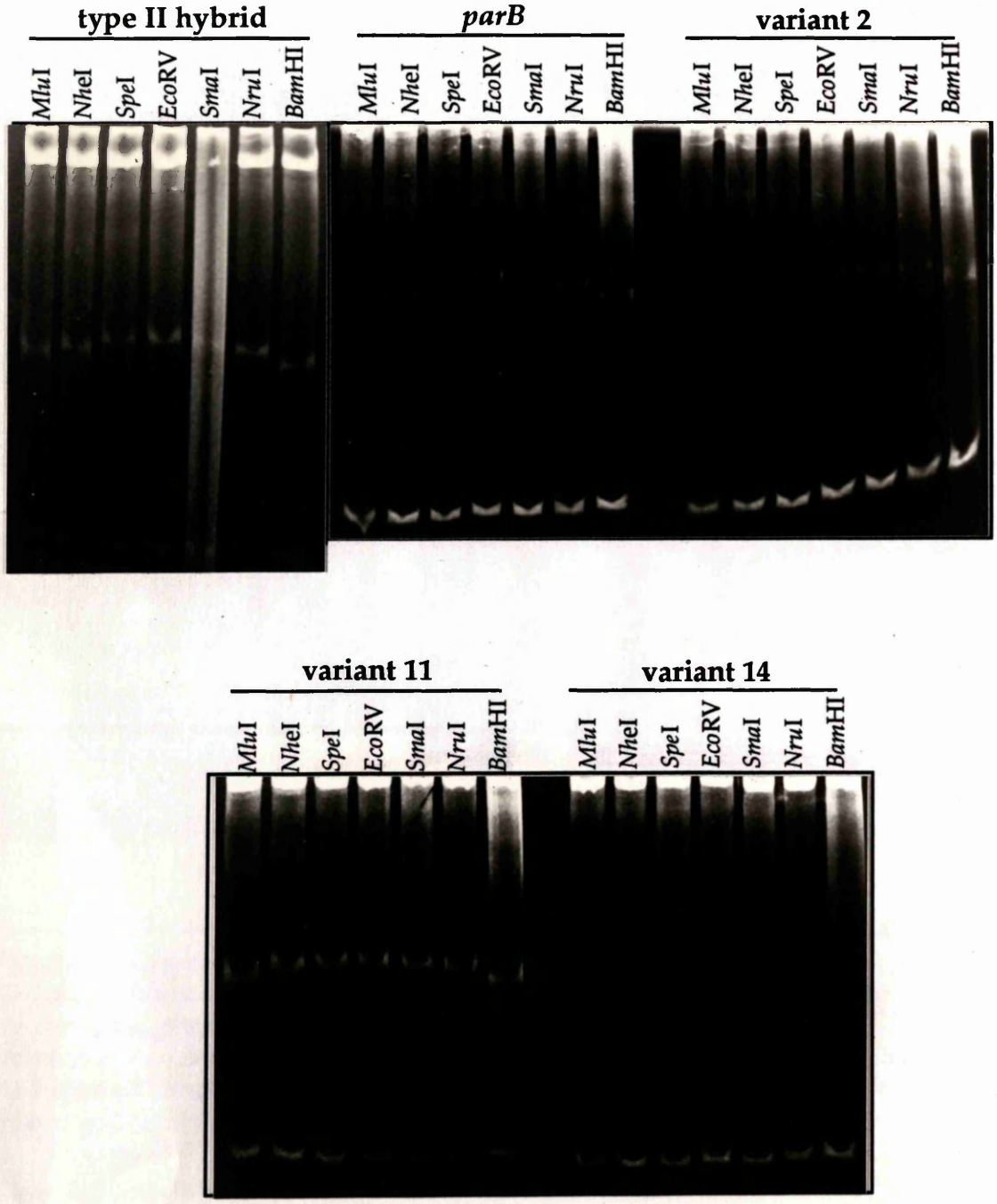


Figure 3.8: Circular permutation analysis of DNA-XerC/D fragments of the type II hybrid, *parB*, variant 2, variant 11 and variant 14.
Equal concentrations of pJEN33, 43, 723, 711 and 714 were digested with *Mlu*I, *Nhe*I, *Spe*I, *Eco*RV, *Sma*I, *Nru*I and *Bam*HI and the resulting fragments were used in binding reactions with XerC and XerD. The reactions were electrophoresed on 8% non-deneturing polyacrylamide gels, and the complexes were visualised using ethidium bromide staining. Photographs of the gels are shown below in Figure 3.8a. The relative mobilities of the DNA-XerC/D complexes were measured and are plotted in Figure 3.8b (facing page).

Figure 3.8a



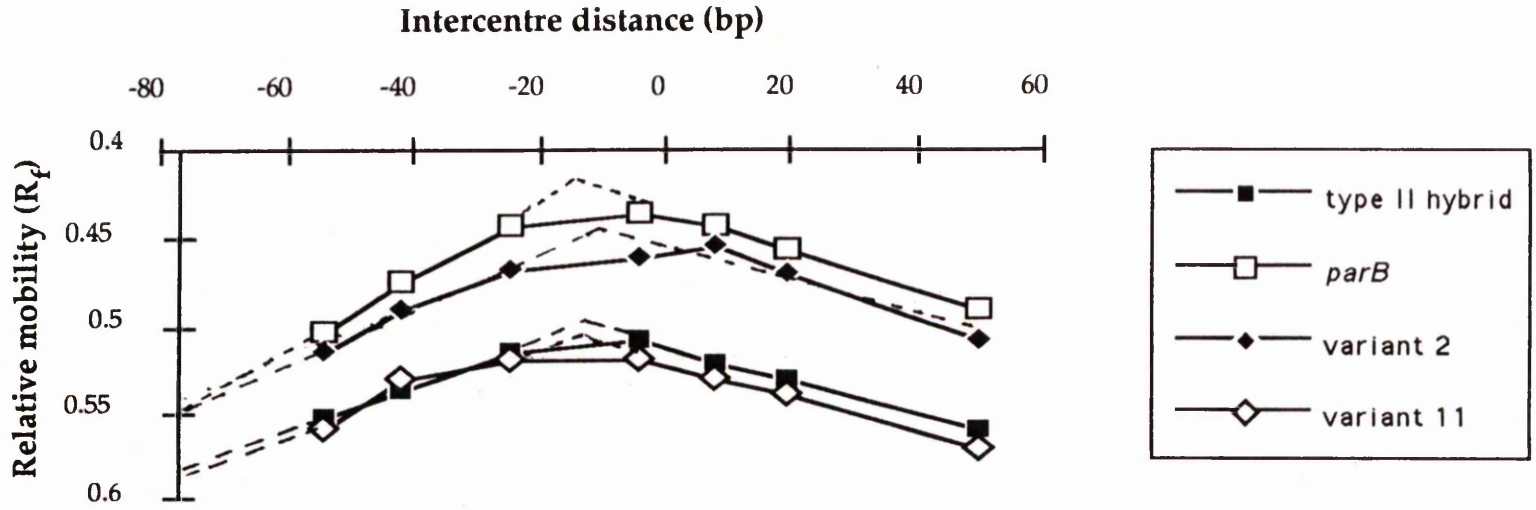


Figure 3.8b: Plot of relative mobilities of DNA-XerC/D complexes
The R_f values for the XerC/D complexes produced with circularly permuted fragments containing the type II hybrid, *parB*, variant 2 and variant 11 were plotted as a function of intercentre distance for each of the restriction fragments used. The solid lines join the actual values as measured, while the dashed lines are extrapolations to show μmax (the point of maximum retardation) and μmin (the point of minimum retardation), to allow estimation of overall bend angles. In addition, μmax allows the overall centre of bending to be estimated: it appears to be in the left arm for the type II hybrid, *parB* and variant 11, but is apparently shifted towards the central region for variant 2.

distance was derived by calculating the distance between these two points; assignment of a negative value was made if the restriction fragment mid-point was left of position zero, a positive value was given to those falling to the right.

The pattern of retardation of DNA-XerC/D complexes on the gel in Figure 3.7 appeared to mirror the pattern for DNA-XerD, indicating that XerD might be largely responsible for determining the position of bending, or the magnitude of bending; the simplest interpretation is that the more retarded DNA-XerC/D complex may be due to the added mass or charge of XerC alone. However, different XerC/D-induced bends are produced with the different sites: the type II hybrid is least retarded by XerC and XerD at all four positions, indicating that (since the same amount of protein is expected to have bound) the other sites may be bent to a greater extent, or at a different position (see Figure 3.7b). This is not true for the XerD-induced bend: when these R_f values were plotted on a graph, all four sites appeared similar (Figure 3.7c). These observations imply that XerC contributes to the overall bend, and that it bends sites with different left arm sequences to different extents or at a different position, so that the XerC- and XerD-induced bends may be more additive for those sequences that are more retarded (*parB*, variant 14) or less additive for the type II hybrid. Alternatively, the two proteins together might produce a particular DNA conformation that depends on the sequence of the site, but which is not necessarily a result of two independent bends.

The DNA-XerC/D complexes could be visualised by ethidium bromide staining, however, no DNA-XerD complexes were seen. This could be a consequence of using excess XerC relative to XerD, since initial experiments in which XerC and XerD were mixed in equal volumes had visible DNA-XerD complexes (data not shown). Again, as in Figure 3.7, the difference in magnitude of the overall bends induced at the different sites was clear. The electrophoretic mobility of all of the type II hybrid and variant 11 fragments is greater than those containing *parB*, implying that the overall bend at these sequences may be smaller. Again, the bend at variant 2 appeared intermediate. Unfortunately, the complexes with variant 14 were too weak to be detected by ethidium bromide staining, reflecting the low apparent affinity of this site for XerC/D (see Figure 3.5). The R_f values of the XerC/D complexes were calculated from this gel and are presented as a function of intercentre distance for each restriction fragment in a graph in Figure 3.8b.

The use of more restriction sites meant that a better estimate of the position of the bends induced by XerC and XerD could be derived from these data. By analysing the graph, it appeared that the position of the overall XerC/D-

induced bend with variant 2 was shifted towards the centre of the fragment relative to the other sites. For *parB*, it appears that the overall centre of bending is at around position -12. For the type II hybrid and variant 11, the best estimate for the centre of bending is at approximately position -10, which is in the centre of the left arm, while for variant 2, it is more likely to be at around position -5, at the right end of the left arm. However, in Figure 3.7b, this was not the case; this observation might be due to the gel in Figure 3.8a running unevenly, leading to a skew in the results. Even in the absence of experimental deviations, circular permutation analysis cannot accurately be used to determine the position of bend centre to the base pair, as errors on the order of 1% in the measured complex mobilities will give rise to several base pair deviations in the apparent centre of the distortion (Kerppola & Curran, 1993). It should also be remembered that the bends plotted on the graphs are likely to reflect the overall DNA deformation induced by XerC and XerD, and it is impossible to determine from these data the individual contributions of either of the proteins in the DNA-XerC/D complex to DNA conformation.

It must be noted that the R_f values for the XerC/D-induced bends with the labelled and unlabelled fragments are different. This might simply be due to using buffers with different salt composition for the two experiments, as it is known that the relative mobilities of bent DNA fragments are influenced by electrophoresis conditions, including ionic strength (Kerppola & Curran, 1993). Unfortunately, this means that the data on the two gels cannot be compared quantitatively. Nevertheless, it is clear that the qualitative results are the same. On both gels, the type II hybrid was less retarded at all positions than *parB*, and variant 2 migrated intermediate to them. In Figure 3.7, there are additional data for variant 14, showing that complexes with this fragment migrated similarly to *parB*, indicating that the same single base change (G to C at position (2), as described above) that prevents this site recombining intermolecularly as a minimal site is also responsible for altering the electrophoretic mobility of an XerC/D-complex with this site. The data for variant 11 are presented in Figure 3.8, revealing that the variant 11-XerC/D complexes migrated similarly to those of the type II hybrid, consistent with the ability of variant 11 to recombine as a minimal site.

It is theoretically possible to calculate bend angles from such data, using mathematical equations derived from plots of relative mobility versus bending angle for DNA fragments constructed as standards with various degrees of intrinsic curvature (Thompson & Landy, 1988). The following equation was applied to the data derived from the plot in Figure 3.8b:

Estimation of error during calculation of bend angles:

Bend angles were calculated from extrapolation and interpolation of the data in Figure 3.8b. The graphs were in turn derived by determining R_f values, by measuring the distance migrated by the DNA-XerC/D complex divided by the distance migrated by the unbound DNA fragment. Error could arise during the calculation of bend angles depending on the position within each band on the gels that was taken to represent the distance migrated. Errors were calculated for the XerC/D-induced bend angles with *parB* and the type II hybrid supposing that the measurements were made to the front or back of each band in Figure 3.8a (i.e. error was calculated taking into account the spread of each band on a gel). After so-doing, it was found that the potential error involved in calculation of the bend angle for the type II hybrid was 5° (between 56° and 61°), while that for *parB* was found to be 2° (between 79° and 80°). Similar errors would be expected for the other sites (not tested).

$$R_L = \cos(\alpha/2)$$

where R_L is the R_f value for the position of maximum retardation (μ_{max}) divided by the R_f value at the position of minimum retardation (μ_{min}), and α is the bend angle. To calculate R_L, μ_{max} was derived by extrapolation of the curves to estimate the position at which the most retardation would be expected (extrapolation is shown by broken lines in the plot), μ_{min} was derived by extrapolating the curves to meet the y-axis at position -77, which is the position at which minimal retardation would occur, i.e. if a bend at the centre of the core recombination site could be exactly at the end of the DNA fragment. The following R_L values and angles were obtained from Figure 3.8b:

Table 3.5: Calculation of bend angles from the data in Figure 3.8b

Site	R_L value	XerC/D-induced bend angle
type II hybrid	0.86	61°
parB	0.76	81°
variant 2	0.82	69°
variant 11	0.86	60°

These calculations are liable to be over-simplifications of these data, because in the DNA-XerC/D complexes the distortion observed is likely to be due to two bends, one induced by XerD (as seen in Figure 3.7c), and one induced by XerC. Furthermore, the data presented here could not distinguish between alterations in the relative directions of two bends as opposed to changes in the angle of bending. For these reasons, it was considered inappropriate to attribute a change in DNA-XerC/D complex mobility simply to a change in bend angle.

To conclude: from these data, the differences in the relative mobilities of XerC/D complexes with DNA fragments containing sites with left arm sequence variations can probably be attributed to differences in XerC/D-induced bends at these sites. The data from the labelled fragments indicated that the XerD-induced bend may be similar irrespective of left arm sequence, consistent with the fact that all of the sites share right arms of the same sequence, and with the previous observation that the mobility of the XerD complex was the same for all of the sites tested. Both sets of data suggested that the formation of DNA-XerC/D complexes with altered mobilities is due to differential bending of the DNA depending on the left arm sequence. The individual contributions of XerC and

XerD to the DNA distortion could not be determined, as it has been impossible to detect DNA-XerC complexes with these sites to date. It is therefore impossible to attribute differences in complex mobility either to differences in bend angle, or phasing of bends. However, it is clear that the mobilities of the complexes are correlated with the ability of sequences to act as minimal sites, and the formation of different DNA-protein conformations may therefore be a determinant of whether or not a site can recombine in the absence of accessory factors.

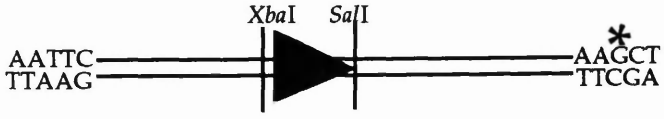
Footprinting of variant sites

Footprinting techniques were adopted to attempt to detect whether the proteins bind similarly to all of the sites. Two protocols were attempted, with the objectives of determining whether the same bases are protected by XerC and XerD, and to possibly detect helical distortions that might indicate reasons for the differences in DNA-XerC/D complex mobility, thereby extending the results of the circular permutation analysis.

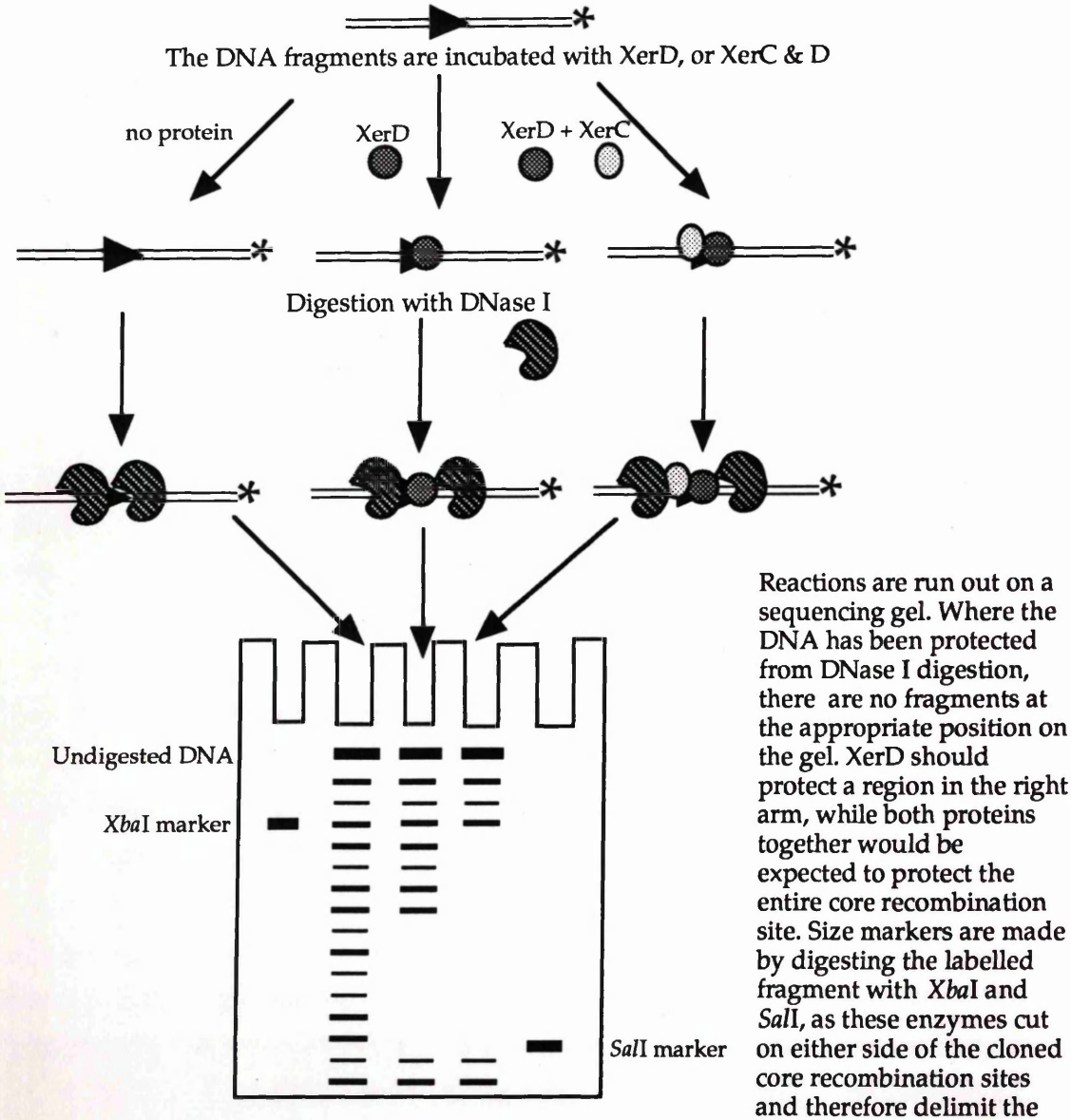
To attempt to assign the binding sites of XerC and XerD in the variant left arm derivatives, DNase I footprinting was used (Galas & Schmitz, 1978). The principle of DNase I footprinting is simple: if a fragment of DNA contains a binding site for a protein, pre-binding of the DNA and protein will protect this region of the DNA from subsequent attack by the endonuclease DNase I. To carry out such an experiment, one end of a DNA fragment containing the potential binding site is labelled at a single position, and the fragment is used in a binding reaction with the protein(s) under test. After a period of binding, DNase I is added at a suitable concentration to produce an average of one nick per DNA molecule during the time of digestion (determined empirically for each DNA concentration). Following digestion, the nicked DNA is electrophoresed on a sequencing gel so that the positions of cleavage can be determined relative to the labelled end. For each sample, a 'no protein' control is carried out so that the cleavage patterns in the presence and absence of protein can be compared. This should indicate which phosphodiester bonds are protected from DNase I digestion by the protein. This technique is quite crude, as DNase I, being a protein, is a large reagent and therefore cannot cleave at the bases immediately adjacent to the protein binding site, due to steric hindrance by the bound protein. However, it was hoped that DNase I footprinting would enable the positions of binding of XerC and XerD to be compared for all of these sites.

The protocol followed is described in Chapter 2, and the results are shown in Figure 3.9. All of the sites showed similar patterns of XerD protection of their top strands, consistent with the results reported above which suggest that XerD

Figure 3.9: DNase I footprinting of the top strands of six sites with variant left arms.
The top strands of the type II hybrid, *parB*, variant 2, variant 9, variant 11 and variant 14 were labelled at the *Hin*dIII end of an *Eco*RI to *Hin*dIII fragment of pJEN3, pJEN4, pJEN72, pJEN79, pJEN711 and pJEN714 respectively, as described in Chapter 2. The labelled G is asterisked below:



A schematic representation of DNase I footprinting with these sites is shown below:



Reactions are run out on a sequencing gel. Where the DNA has been protected from DNase I digestion, there are no fragments at the appropriate position on the gel. XerD should protect a region in the right arm, while both proteins together would be expected to protect the entire core recombination site. Size markers are made by digesting the labelled fragment with *Xba*I and *Sal*I, as these enzymes cut on either side of the cloned core recombination sites and therefore delimit the XerC and XerD binding sites.

The results of DNase I footprinting with XerD and XerC & D are shown on the facing page . The protein additions are shown at the top of the photograph. The sequence of the top strand of the type II hybrid is shown alongside for reference.

type II hybrid
parB
variant2
variant 9
variant 11
variant 14
dif
no protein
XerD
XerC + XerD
Markers
GCGGTGCGTACAAGGGATGTTATGGTAAAT

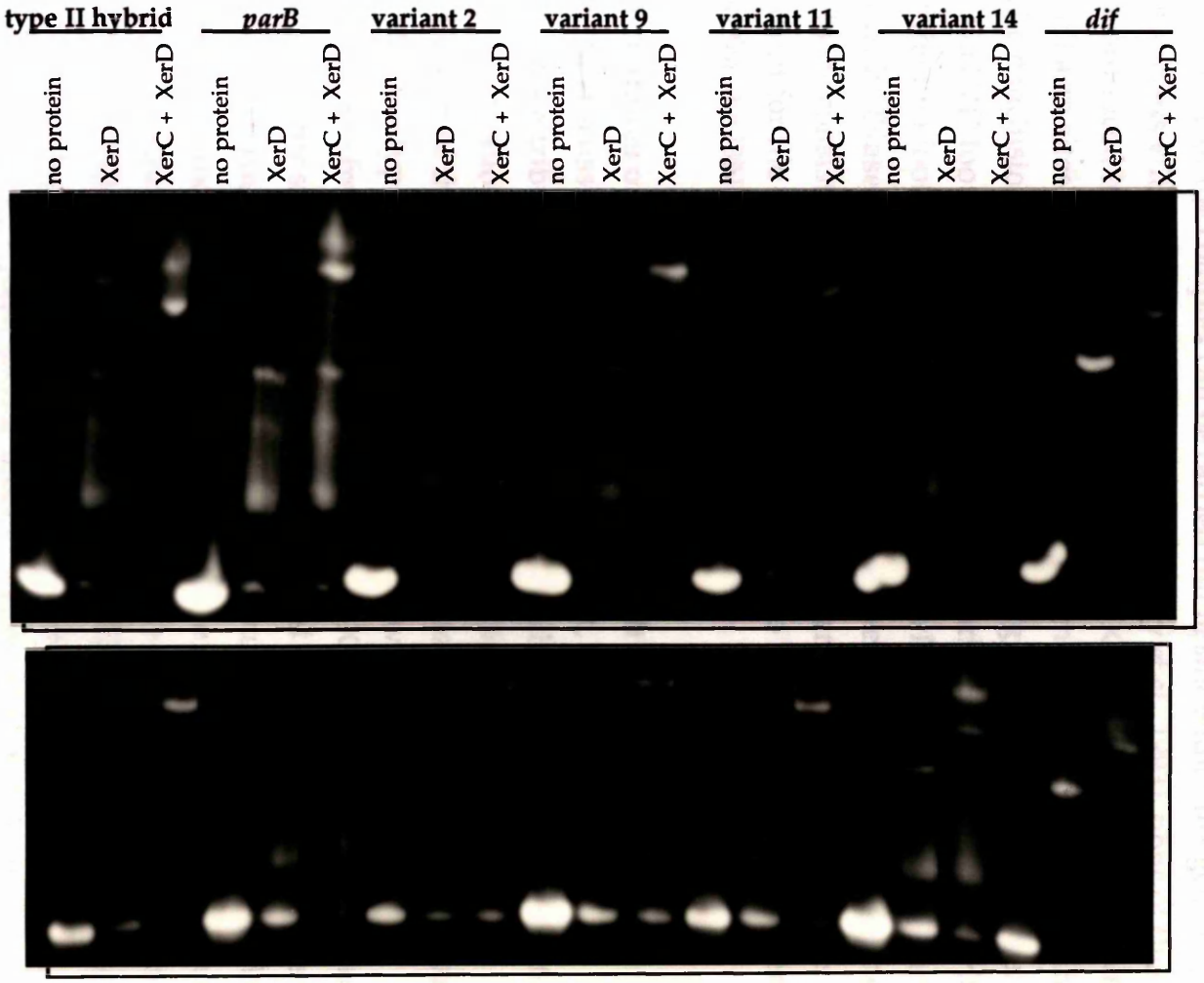
binds similarly to all of the sites. The same bases were protected from DNase I and there were no regions of enhanced cleavage outside of the XerD binding region in the different variant sites. There was one slight enhancement to DNase I cleavage of the unbound variants compared with *dif*, at the single position which differs between these sites and *dif* in the right arm.

No additional protection of the variant sites' DNA occurred when XerC was included in the binding reaction along with XerD, since the same protections were observed as with XerD alone. To ensure that the addition of the DNase I and its required salts did not abolish XerC binding, samples of the binding reactions from before and after the addition of DNase I were electrophoresed on a non-denaturing gel as for standard gel retardation assays (Figure 3.10). XerC/D complexes were detected for both samples, indicating that the failure to produce an XerC/D footprint was not due to the fragments failing to bind XerC, nor to DNase I dislodging XerC from the DNA. XerC may interact weakly or transiently with these sites so that it might be leaving the DNA unprotected for sufficiently long periods to allow DNase I to nick. Alternatively, the XerC/D complexes that were detected with these variant sites might be due to XerC interacting with bound XerD, but not with the DNA itself.

An XerC/D footprint was obtained for *dif*, implying that the affinity of XerC for all of the variant left arm sites (including those which recombine intermolecularly as minimal sites) is lower than it is for *dif*. This must be due to additional sequence changes in the left arm of *dif*, possibly in conjunction with its altered right arm, and consequent better XerD binding (G. May, unpublished data). XerD affinity must contribute to the binding of XerC, since XerC and XerD bind co-operatively to *dif* and *cer in vitro* (Blakely *et al.*, 1993), and because only in the presence of XerD does XerC bind to these DNA sites (see Figure 3.5). Since it is possible to detect an XerC/D footprint for *dif*, this suggests that under the correct conditions, a similar footprint might be obtained for the other sites. It is likely that the main problem is a low specific activity of XerC in the preparations used (although these were the best available at the time of the experiments) combined with a reduced affinity for these sites. All further attempts, such as increasing the concentration of DNase I while reducing the incubation time, also failed to produce footprints in which the left arms were protected (data not shown).

Attempts were also made to footprint DNA-XerC/D complexes from binding gels *in situ*, using the reagent 1,10-phenanthroline-copper. This small reagent attacks DNA in the minor groove at the C1′ position of deoxyribose (Sigman *et al.*, 1991), and can be used to detect protections of the DNA by a

Figure 3.10: Non-denaturing gels to detect the presence of DNA-XerD and DNA-XerC/D complexes before and after the addition of DNase I to the incubation mixture.
1 µl of each reaction was loaded onto a binding gel both before and after the addition of DNase I. Complexes can be detected on both gels, indicating that XerD and XerC are both binding and it should therefore be possible to obtain an XerC/D footprint. Notice, however, that the bands are smeary: this implies that the complexes may be dissociating on the gel.



The top gel was loaded after 35 minutes of binding, 10 minutes before the addition of DNase I.

The bottom gel was loaded after 50 minutes of binding, 5 minutes after the addition of DNase I.

protein as long as that protein either binds through the minor groove to sterically hinder the nuclease activity at the position of binding, or distorts the minor groove upon binding. This reagent also detects protein-induced changes in DNA structure, such as underwinding of the helix. The major advantages of this technique are that a protein-DNA complex can be isolated from a gel prior to footprinting so that the electrophoretic species under examination is known, and that footprinting of the complex can be carried out *in situ* within the gel. Despite previous success with *dif* using this protocol (Blakely *et al.*, 1993), none of the attempts produced informative data with regard to XerC or XerD binding to the variant left arm sites. Consequently, no such footprinting gels are presented here.

The conclusions derived from the footprinting experiments are sparse. No XerC or XerC/D footprints were obtained for the variant left arm mutants with either of the two footprinting reagents used. The XerD footprints for all of the variants with DNase I appear identical, and are all very similar to that obtained for *dif*. This is consistent with the fact that the right arm sequences are all the same (except for the single change in *dif*) and that no bending differences were observed for the DNA-XerD complexes.

Relative affinities of XerC and XerD for variant left arm sites

In the discussion of the previous experiments, the failure to produce footprints was attributed to the affinities of XerC and XerD for the variant sites being too low to allow sufficiently stable complexes to form. Figure 3.5 gave some indication of the relative overall affinities of the variant sites for XerC and XerD, in that the type II hybrid and variant 11 sites were substantially more retarded from the free DNA position into the DNA-XerC/D position compared with the other four sites tested. It was also suggested that the affinities of the sites for XerD were all roughly equivalent, since the amount of DNA-XerD complex detected was approximately equal irrespective of the sequence of the DNA site. An obvious conclusion from this is that changing the sequence of the left arm may affect the affinity of a site for XerC, but not XerD; this is in agreement with the observed binding of the *dif* left arm by XerC (Blakely *et al.*, 1993). The affinities of these sites for XerC should ideally be compared to indicate whether a basis for minimal site recombination might be in the affinity of XerC for particular left arm sequences.

To establish the affinity of a DNA binding protein for a DNA site in a simple bimolecular reaction, the dissociation constant K_d is calculated. K_d is the value assigned to the concentration of protein required to bind half of the DNA at equilibrium. This value can be determined using gel retardation. By

performing binding reactions containing different concentrations of protein, and a constant DNA concentration, then the fraction of free DNA is plotted against log protein concentration; this plot allows determination of which concentration of protein will bind half of the DNA (K_d). The fraction of free DNA is generally plotted, because as a binding reaction is loaded onto a gel under current, the free DNA enters the gel first, so that the free DNA band on the gel will report the fraction of free DNA in the binding mixture; while the fraction of bound DNA can also be used, the complexes might dissociate during electrophoresis therefore giving an underestimate of the amount of complexes DNA at the time of loading.

To calculate absolute affinities, it is essential that the concentration of active protein is known, and that this concentration is in excess to that of the DNA in the binding reactions. This presents two immediate problems for calculation of the absolute affinities of XerC (or XerD) for the variant sites. The first is that the concentration of active recombinase in the protein preparations is not known, and the second is that none of the sites could be 50% retarded into DNA-XerC (or DNA-XerD) complexes even with the most concentrated extracts available at the time of the experiments; in fact, DNA-XerC complexes were not detected at all (see Figure 3.5). These problems made calculation of K_d values for XerC (and XerD) binding to these sites impossible. In theory, it might be possible to use the overall affinities of XerC and XerD together for variant sites to determine the affinity of XerC for them. It could be imagined DNA-XerC/D complexes in a reaction mixture containing DNA and both proteins could be produced by three possible pathways:

$$\text{DNA} + \text{XerC} \underset{}{\overset{K_d1}{\rightleftharpoons}} \text{DNA-XerC} \overset{\text{XerD},\ K_d2}{\rightleftharpoons} \text{DNA-XerC/D} \qquad (1)$$

$$\text{DNA} + \text{XerD} \underset{}{\overset{K_d3}{\rightleftharpoons}} \text{DNA-XerD} \overset{\text{XerC},\ K_d4}{\rightleftharpoons} \text{DNA-XerC/D} \qquad (2)$$

$$\text{DNA} + \text{XerC} + \text{XerD} \overset{K_d5}{\rightleftharpoons} \text{DNA-XerC/D} \qquad (3)$$

The overall K_d values for all three routes should be equivalent. It is therefore possible, in theory, to determine K_d1 (the affinity of XerC for variant sites) as long as K_d2, K_d3, and K_d4 can be measured. In practice, however, it was impossible to measure K_d3 (because the concentration of XerD to bind half of the

DNA could not be determined) and to separate K_d2 and K_d4, therefore this could not be achieved.

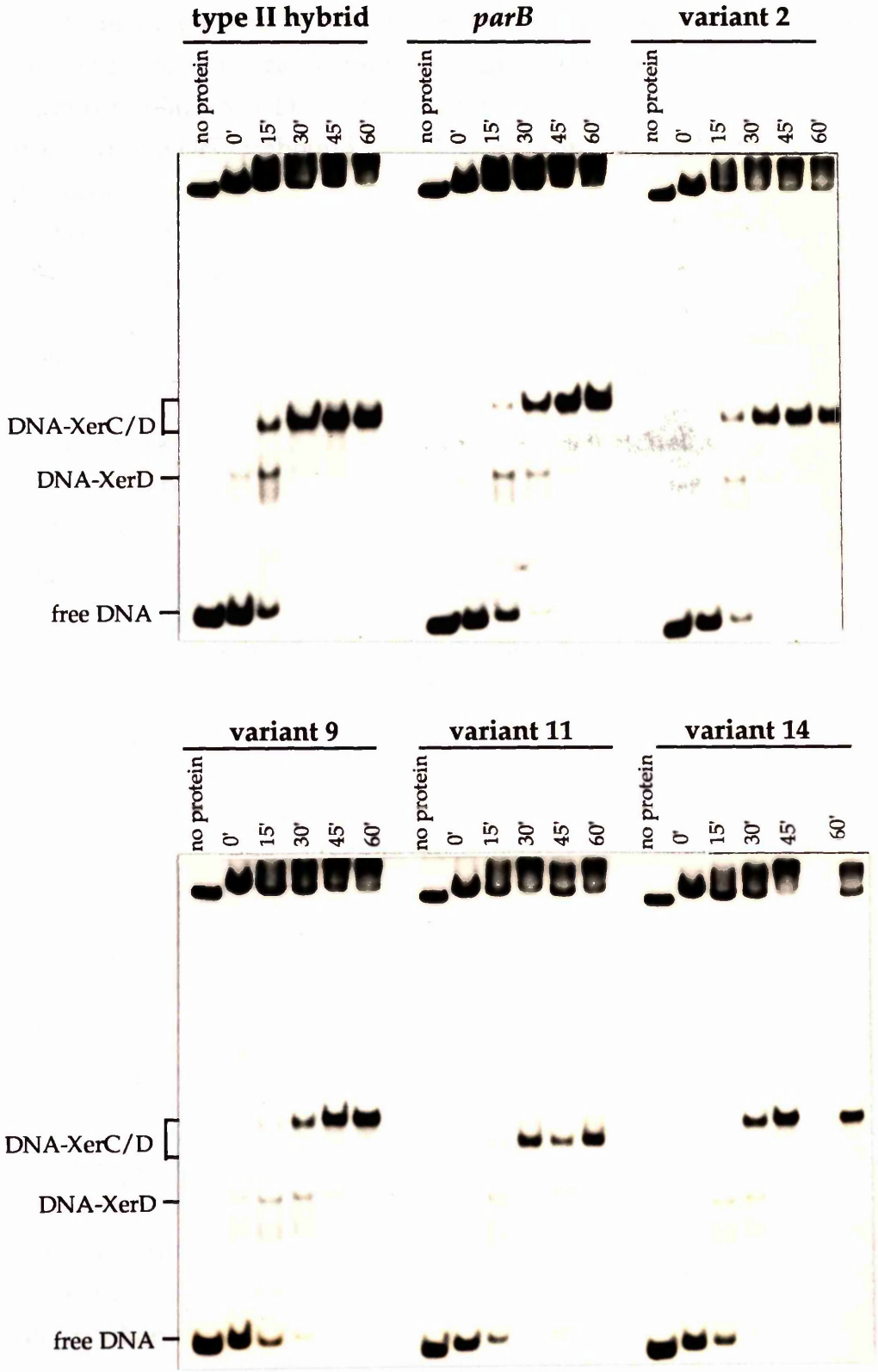
It is possible to detect DNA-XerC/D complexes for all of the sites, therefore it might be possible to calculate overall relative affinities for XerC and XerD for variant sites. This might be informative in terms of determining whether the ability of certain minimal sites to recombine can be correlated with affinity for both recombinases. Calculation of overall K_d values for these sites is marred by the apparent low specific activity in the protein preparations, which means that even undiluted extracts do not bind half of the DNA of all sites into XerC/D complexes under the conditions used in Figure 3.5 (especially notice variant 14). This means that it is impossible to prepare a dilution series of protein to produce binding curves which would allow calculation of overall K_d values for XerC/D binding.

To address the question of binding affinities for these variant left arm sites, time-courses of binding were used to attempt to determine whether reactions were at equilibrium after 10 minutes incubation (as used in Figure 3.5). Binding reactions were prepared and incubated for different periods before separating the complexes on non-denaturing gels. In addition, time-courses were carried out at two different temperatures, 37°C and 0°C. The same six sites as above were used: type II hybrid, *parB*, variant 2, variant 9, variant 11 and variant 14. Equivalent concentrations (2 µg) of pJEN3, 4, 72, 79, 711 and 714 were digested with *Bam*HI and *Pst*I, the DNA was end-labelled using the protocol described in Chapter 2, to give radiolabelled fragments of the same specific activities.

In the first time-course experiment, the binding buffer was GBA2 (+ polydI.dC), and partially purified XerC and XerD were used (from Gerhard May). At time zero, 1 µl of each protein was added to the "60 minute" reaction for each site, and the reaction was incubated at 37°C. This was repeated at fifteen minute intervals, to give 45, 30, and 15 minute time-points for each of the sites. The zero minute time-points were prepared by adding the proteins to the DNA, mixing, then loading onto the gel immediately, under current. The samples from the other time-points were then loaded direct from the 37°C waterbath in ascending order of incubation times, and electrophoresis was continued as normal. The results of this experiment are shown in Figure 3.11.

XerD was able to bind to the sites almost immediately, as there are DNA-XerD complexes visible at the zero minute time-point. The time taken to mix and load the samples onto the gel was less than thirty seconds, so the complexes must have been able to form in this time or during the clearance time taken for the

Figure 3.11: Time-course of XerC & XerD binding at 37°C.
In each binding reaction, the concentrations of DNA and protein were constant. Incubation times were varied between 0 and 60 minutes at 37°C. The gel was kept running during loading of samples. The protocol for time-courses is described in the text. Gel conditions are as standard for gel retardation assays, and are described in Chapter 2.



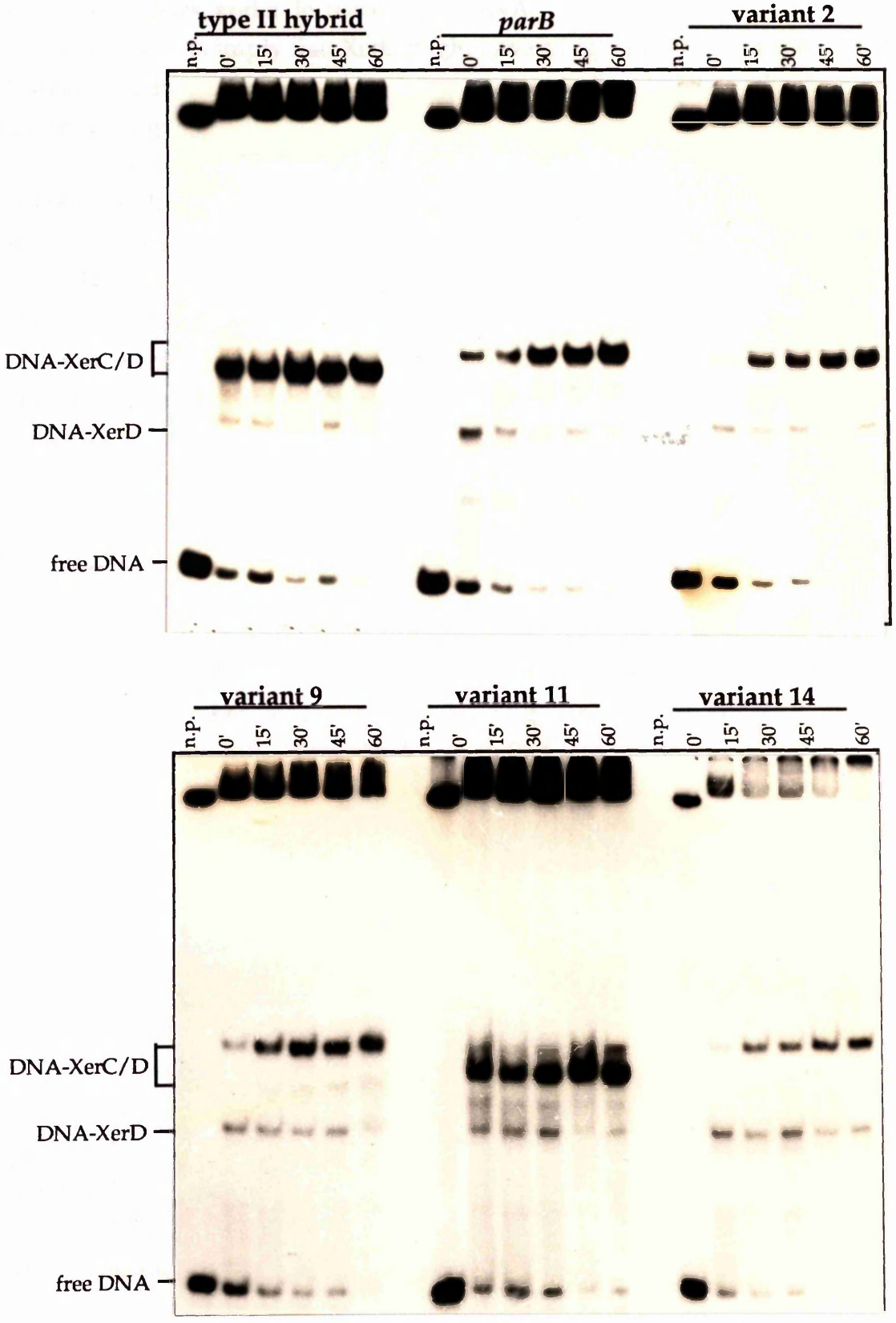
DNA to enter the gel. No DNA-XerC/D complexes were detected by this time. By the 15 minute time-point, DNA-XerC/D complexes had begun to form with all of the sites. The amount of DNA-XerD complex did not decrease as the DNA-XerC/D complex formed. This suggests either that the DNA-XerD complex is not an intermediate in the formation of the DNA-XerC/D complex, or that an equilibrium had been reached between unbound DNA, DNA-XerD and DNA-XerC/D. At this time-point, there were no obvious differences between any of the six sites in terms of the ratio of unbound to bound DNA. By thirty minutes, there was substantially less DNA in the unbound position for all of the sites, and there appeared to be less unbound type II hybrid compared with the other sites. Binding continued to improve with all of the sites until 60 minutes, by which point most of the DNA was at the XerC/D-bound position. These data suggest that the formation of DNA-XerC/D complexes with these DNA sites may take a long time to reach equilibrium (approximately an hour); this further suggests that the reactions in Figure 3.5 were not at equilibrium at the time of loading.

In a second "time-course" experiment, the reaction times were kept constant, but the temperatures of incubation were varied. In this experiment, the same protocol as above was followed, with the following exceptions. All of the time-points were begun simultaneously; the "60 minute" time-point was incubated at 37°C for 60 minutes, the "45 minute" reaction had 45 minutes at 37°C, but was then continued on ice for the remaining 15 minutes. This was carried out for all of the reactions, so that the zero time-point had no 37°C incubation, but was kept on ice for 60 minutes. The incubation conditions are summarised below:

Time-point	**Incubation at 37°C (mins)**	**Incubation at 0°C (mins)**
"0 minutes"	0	60
"15 minutes"	15	45
"30 minutes"	30	30
"45 minutes"	45	15
"60 minutes"	60	0

The results of this experiment are shown in Figure 3.12. Incubation temperature appeared to make little difference to the formation of DNA-XerD and DNA-XerC/D complexes with the type II hybrid and with variant 11. Both sites were substantially retarded even when no incubation at 37°C had occurred. However, incubation at 37°C did increase complex formation, as detected by the

Figure 3.12: Time-course of XerC & XerD binding at 37°C and 0°C. All of the binding reactions on this gel had a total of 60 minutes. However, the temperature was varied so that the reactions had some incubation time at 37°C, and some time on ice. The incubation time at 37°C is indicated at the top of the gel (the abbreviation "n.p." means "no protein"). The protocol for this experiment is described in the text.



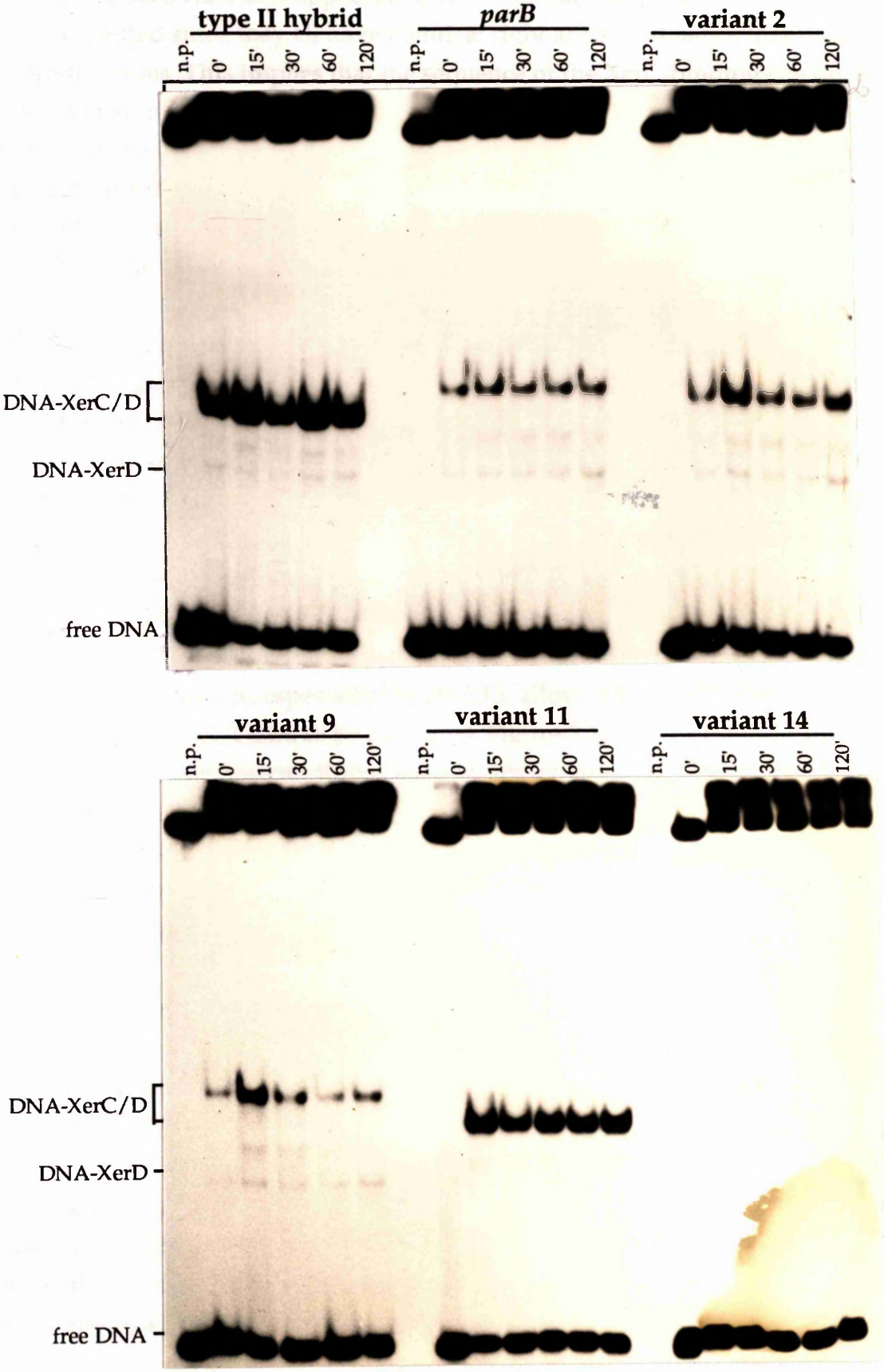
decrease in the amount of free DNA as the time at 37°C was increased. While the other sites could bind XerD reasonably well irrespective of the temperature (as the amount of DNA-XerD complexes appeared relatively constant), the formation of DNA-XerC/D complexes was increased by incubating the reactions at 37°C. This was apparent both from an increase in the formation of DNA-XerC/D complexes, and a decrease in free DNA.

These results imply that XerC might have a higher affinity for the type II hybrid and variant 11 than for the other sites tested, since lowering the incubation temperature had little effect on the formation of DNA-XerC/D complexes. To test the proficiency of the sites to bind at 0°C, time-courses were performed without incubation at 37°C. In this experiment, the buffer conditions were altered since in the gels shown in Figures 3.11 and 3.12, there is some smearing of the bands, and there are also additional bands to the normal DNA-XerD and DNA-XerC/D complexes, complicating interpretation of the data. To attempt to overcome these problems, GBA3 (+ polydI.dC) conditions were used, so the salt concentration was increased by 20 mM (to a final concentration of 80 mM). The time-course was carried out using the same general method as for the 37°C time-course in Figure 3.11, except for this slight change in buffer conditions. The times of incubation were 0, 15, 30, 60, and 120 minutes, at 0°C. The results of this experiment are shown in Figure 3.13.

Under these conditions, an increased binding of the type II hybrid and variant 11 by XerC and XerD was most apparent. For both sites, a ~50% shift of the DNA into the DNA-XerC/D complex was achieved very early in the time-course, then no further improvement in binding was detected. For *parB*, variant 9 and variant 14, very little of the DNA was shifted into the bound position at any of the time-points; even after 120 minutes incubation on ice, no more than ~10% of the DNA was bound by XerC and XerD. Slightly better binding was achieved with variant 2 than with *parB*, however much less of this site was retarded than either variant 11 or the type II hybrid.

The complexes formed under these conditions are much more defined than those in Figure 3.11. This suggests that either the slightly increased salt concentration or incubation at 0°C decreased complex dissociation on the gel. It was also noticeable that under these conditions, very little DNA-XerD complex was detected with any of the sites. This implies that the conditions used in this experiment might favour the formation of DNA-XerC/D complex rather than DNA-XerD. However, these conditions could clearly not be used to maximise the amount of DNA-XerC/D complex for *in situ* footprinting, since the total DNA-XerC/D complex formation at all sites was reduced at 0°C. (After this

Figure 3.13: Time-courses of XerC & XerD binding at 0°C.
In this experiment, the concentrations of protein and DNA were kept constant and the time of incubation at 0°C was varied. There was no incubation at 37°C. The abbreviation "n.p." means no protein addition was made.



experiment, however, all of the binding experiments which followed used these buffer conditions.)

These results give an indication of the differences in affinity of XerC and XerD for these six sites. XerD appeared to bind comparably to all of the sites, as might be expected since they all have identical right arms, and hence identical XerD-binding sites. This implies that the sequence of the XerC-binding site does not affect the ability of XerD to bind. Given sufficient time, all of the sites are retarded into the DNA-XerC/D position when incubated at 37°C. However, in the experiments in which some or all of the incubation time was on ice, the greatest differences in binding at different sites were detected. These experiments were perhaps the most informative in terms of indicating the affinities of XerC and XerD for the variant sites: they showed that variant 11 and the type II hybrid were bound substantially better than any of the other sites. During binding on ice, the length of incubation seemed relatively unimportant; the extent of binding changed little between the 0 and 120 minute time-points for any of the sites. This suggests that an equilibrium is reached quickly at lower temperature. It must be remembered, however, that the three experiments could not be compared with each other since the salt concentration was altered; but within each experiment the six sites can be compared since care was taken to ensure that the DNA and protein concentrations were always the same (within the limits of experimental error).

These experiments, especially Figure 3.13, allow relative affinities of XerC and XerD for these DNA sites to be proposed. Figure 3.13 is the best indicator of affinity, because under the conditions used in this experiment, binding reached an equilibrium that did not alter in two hours, whereas under the conditions in Figure 3.11 and 3.12, it was not clear whether equilibrium had been reached for all of the sites. The apparent relative affinities of these sites for XerC and XerD suggested by Figure 3.13 are:

type II hybrid = variant 11>variant 2> *parB*> variant 9> variant 14.

Since all of the sites bind XerD similarly well, these might be extrapolated to represent relative affinities for XerC, despite the inability to DNA-XerC complexes.

It was not possible to calculate absolute affinities of XerC and/or XerD for the variant sites during the period of study. This was because at the time of the experiments, the specific activities of the available XerC and XerD preparations were too low to allow K_d values to be determined. Work is on-going in this

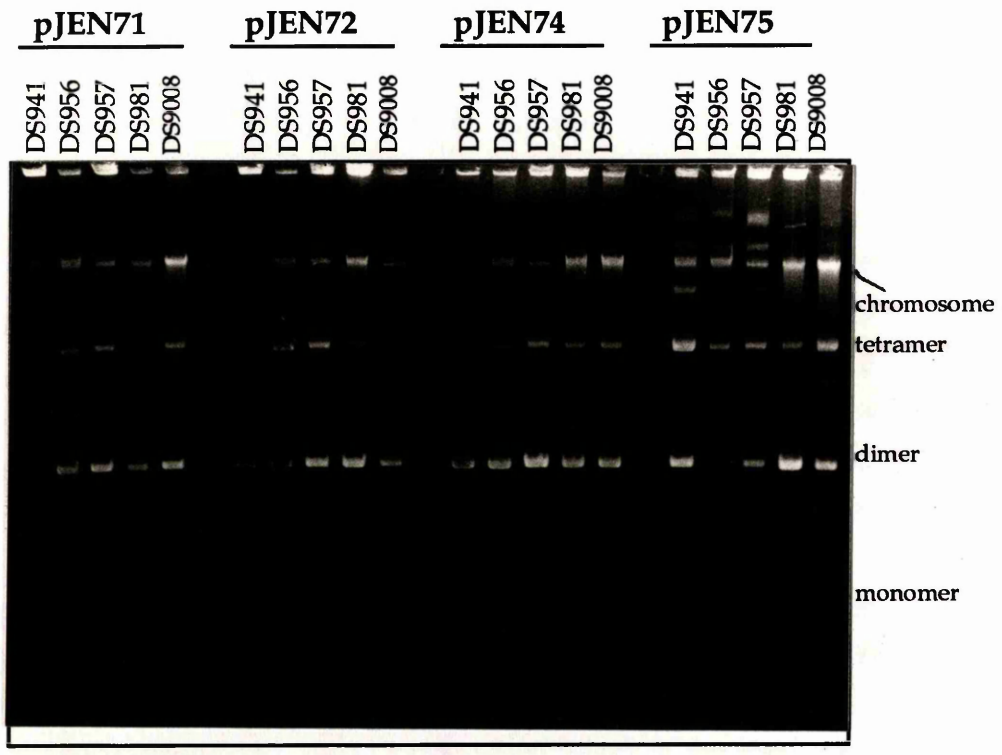
laboratory to prepare extracts of high specific activity, with fair success. Once highly active proteins are available, K_d values should be determined for both XerC and XerD alone, and together. This would indicate whether or not the relative affinities of XerC/D for the variant sites reflect alterations in XerC affinity depending on the site sequence, or if the co-operative interaction between XerC and XerD bound at the right and left arms is affected by these mutations.

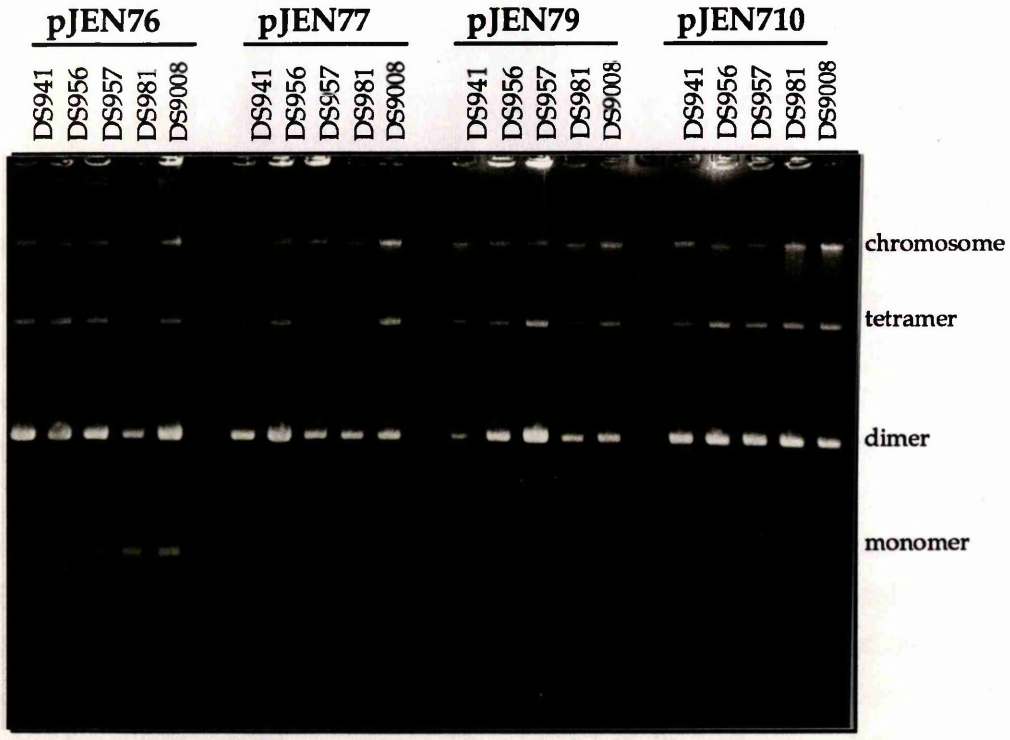
None of the minimal sites recombines exclusively intramolecularly

The experiments described above investigate possible reasons for the ability of the type II hybrid to recombine as a minimal site. Recombination proficiency was measured by the formation of plasmid multimers from monomers *in vivo* as a result of Xer-mediated intermolecular recombination (Figure 3.4). Plasmid resolution sites such as *cer* and *parB* are efficient substrates for the opposite reaction, intramolecular recombination or resolution, and deletions into the accessory DNA of sites like *cer* abolish recombination altogether. The experiments described in this Chapter were, therefore, based on the premise that if minimal sites could recombine at all, there would be no resolution selectivity and recombination could be assayed by the formation of plasmid multimers by intermolecular events. However, it is possible that this assumption is false and that those variant sites which do not mediate the formation of plasmid multimers are in fact able to act as resolution systems.

To test this possibility, *in vivo* recombination assays were carried out in which DS941, DS956, DS957, DS981 and DS9008 were transformed with dimers of pJEN71, 72, 74-77, 79-714 to assay for site-specific resolution at these minimal sites. The results of such an experiment are shown in Figure 3.14.

The results indicated that none of the minimal sites is able to exclusively undergo intramolecular recombination. The sites that do promote some Xer-mediated intramolecular recombination to produce plasmid monomers are most likely to be the ones that also recombine intermolecularly producing multimers (pJEN75, 711 and 712). Monomers were also detected for some of the plasmids that do not recombine intermolecularly (pJEN76, 714), but this is likely to be as a result of homologous recombination, as monomers could also be detected in the tracks derived from the *xerC* or *xerD* mutants. This is probably because homologous recombination activity is not altogether abolished in a *recF* strain, merely reduced, and there is sufficient homology within plasmid dimers to allow some resolution to occur. This idea is supported by the observation that a little intermolecular recombination occurred in DS981 and DS9008 to produce some plasmid multimers independent of Xer recombination. With more time, this


pJEN71
pJEN72
pJEN74
pJEN75
DS941
DS956
DS957
DS981
DS9008
chromosome
tetramer
dimer
monomer


pJEN76
pJEN77
pJEN79
pJEN710
DS941
DS956
DS957
DS981
DS9008
chromosome
tetramer
dimer
monomer

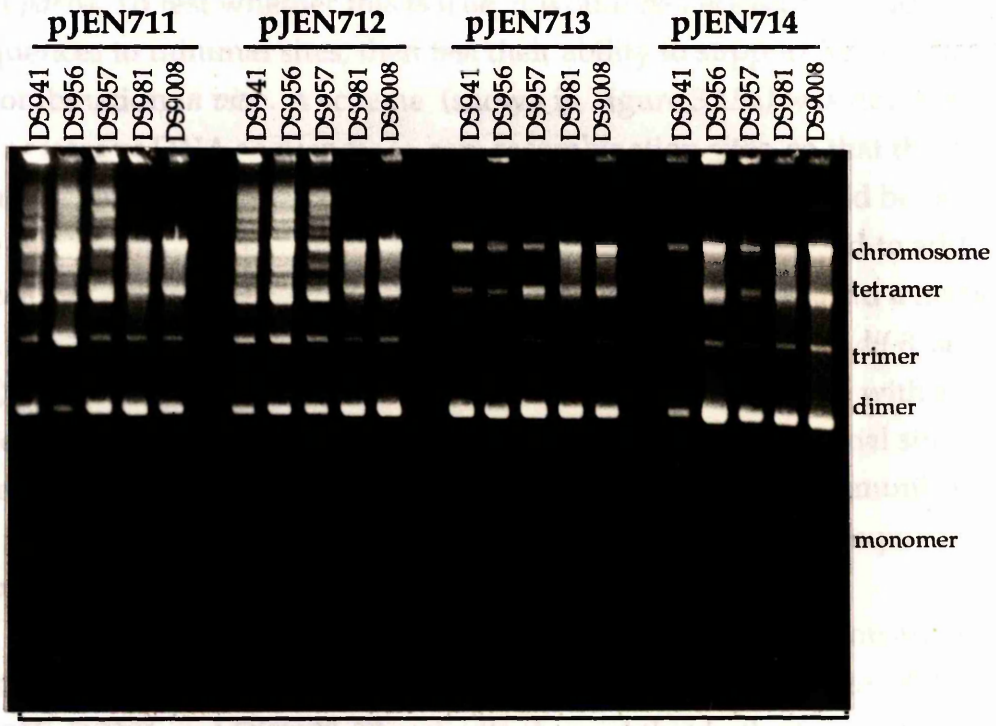


Figure 3.14: Intramolecular recombination at variant left arm sites. Dimers of pJEN71, 72, 74, 75, 76, 77, 79 -714 were purified and transformed into DS941, DS956, DS957, DS981 and DS9008 to assay intramolecular recombination *in vivo*. Populations of transformants with each plasmid were lysed in single colony final sample buffer and electrophoresed on 1% gels. The major oligomeric forms are indicated along the right hand edge of the gels, however, for pJEN72, 75, 711 and 712 higher forms are also visible.

problem could have been overcome either by constructing reporter plasmids for all of the variant sites rather than using dimers (to limit the homology), or by repeating the experiments using a *recA* strain, and its *xer* derivatives.

It is likely that those minimal sites that do not recombine fail to do so because of their lack of accessory DNA sequences (as do *cer*, the type I hybrid and *parB*). To test whether this is true, it would be necessary to add accessory sequences to minimal sites, then test their ability to support Xer-mediated recombination *in vivo*. A scheme (shown in Figure 3.15a) was developed to add *cer* accessory DNA sequences to core recombination sites, so that the exact spacing between the core recombination site and Arg-box could be retained, and the accessory DNA sequence preserved. This strategy was used to add the accessory DNA of *cer* to the core recombination sites of *parB* and a derivative of *dif*, to create pJEN42 and pJEN62 respectively (for sequence of *dif*-8, see Table 3.1). This *dif* derivative consists of a *dif* core recombination site with a 2 bp insertion in the central region; a plasmid containing this minimal site, pGB300, does not multimerise in Xer^+ strains (G. Blakely, personal communication). The 2 bp insertion effectively converts the functional *dif* site into the equivalent of the type I hybrid minimal site in pJEN2.

Figure 3.15b shows the results of an *in vivo* recombination assay in which trimers of pJEN42 and dimers of pJEN62 were transformed into DS941, DS956, DS957, DS981 and DS9008. The results showed that both of these sites will resolve plasmid multimers in DS941. These two sites are therefore competent for recombination, but require accessory factors. It is noteworthy that pJEN62$_{dimers}$ did not resolve to monomers as well as pJEN42$_{trimers}$. Sequencing of the two recombination sites showed that there is an additional base inserted between the core recombination site and the accessory DNA in pJEN62 compared with pJEN42. This extra base did not derive from this cloning strategy, however, but during the oligonucleotide synthesis, as it is also present in pGB300 (G. Blakely, personal communication). It is possible that the additional base is responsible for reduced recombination at this site, perhaps by altering the phase of the Arg-box and the core recombination site.

This strategy could be adopted to test all of the variant sites for their ability to recombine. It is likely that all of the variant sites would undergo XerC/D-dependent recombination in the presence of ArgR and PepA, if given accessory sequences, since all of the sites tested could bind XerC and XerD, if with varying efficiencies. However, this awaits confirmation.

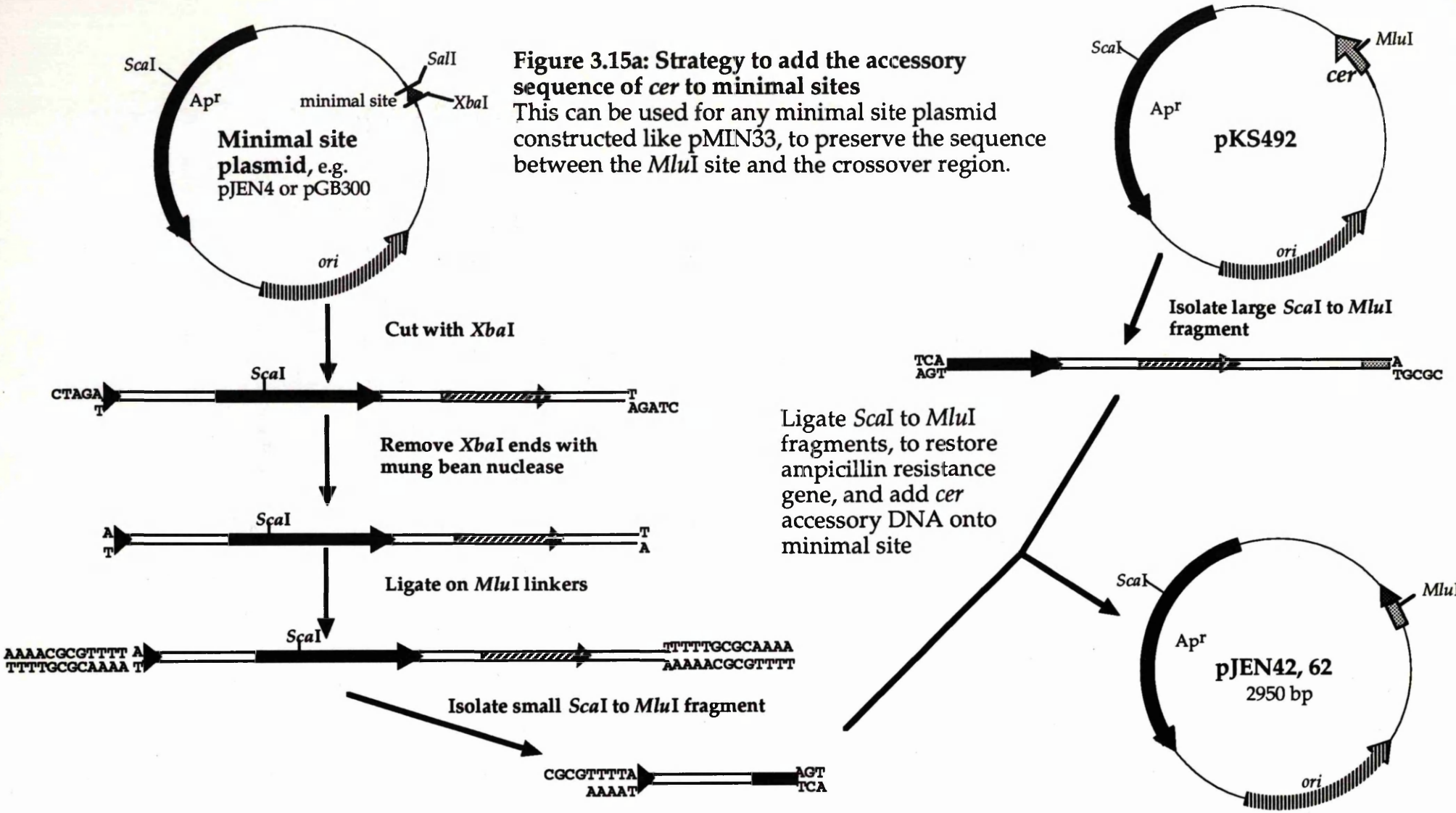


Figure 3.15a: Strategy to add the accessory sequence of *cer* to minimal sites
This can be used for any minimal site plasmid constructed like pMIN33, to preserve the sequence between the *Mlu*I site and the crossover region.

Figure 3.15b

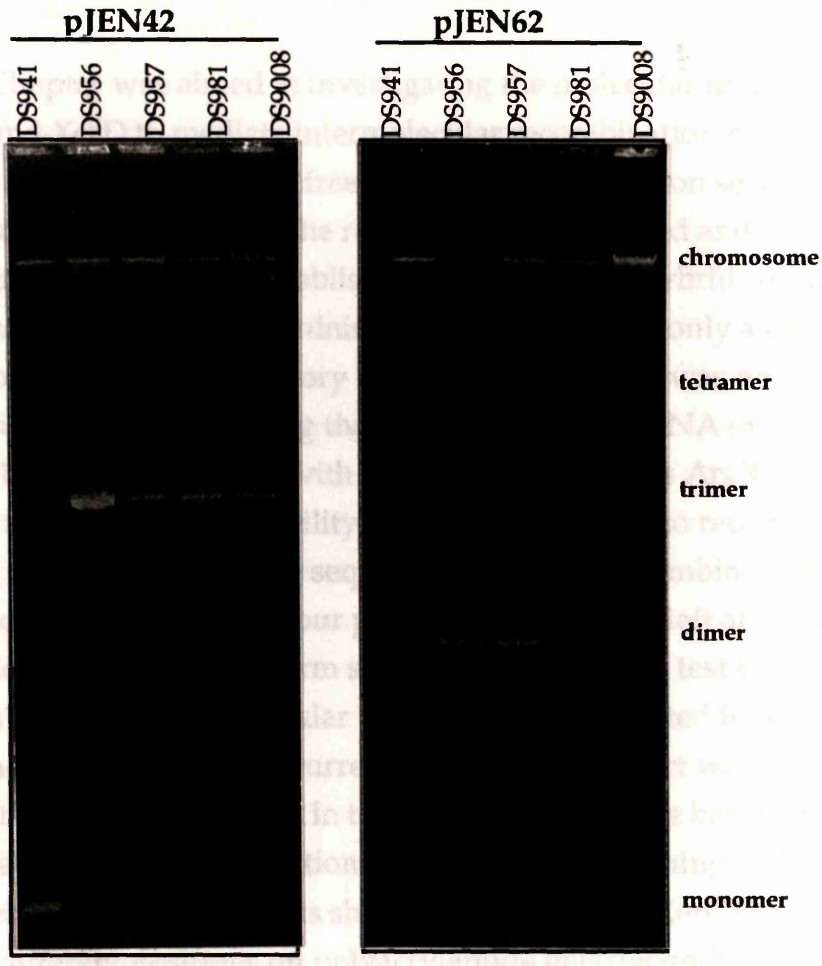


Figure 3.15: *In vivo* intramolecular recombination of "full-length" sites. For those minimal sites which do not recombine to produce plasmid multimers, a scheme was devised to test recombination proficiency in the presence of accessory factors. This strategy is shown in Figure 3.15a, and can be used to add *cer* accessory DNA sequence onto any of the minimal sites described in this Chapter. pJEN42 and pJEN62 were constructed in this way to make plasmids consisting of *parB* or *dif*-8 core recombination sites and *cer* accessory DNA. Both of these plasmids resolve in an ArgR-, PepA-, XerC-, XerD-dependent manner, as shown in the *in vivo* recombination assay in Figure 3.15b above. In this experiment, DS941, DS956, DS957, DS981 and DS9008 were transformed with trimers of pJEN42 and dimers of pJEN62.

Discussion

This Chapter was aimed at investigating the molecular mechanisms that allow XerC and XerD to mediate intermolecular recombination between type II hybrid sites, and the concomitant freedom from the resolution selectivity process that constrains recombination at the related *cer*, type I hybrid and *parB* sites.

The initial experiments established that the type II hybrid could recombine intermolecularly as a minimal site, consisting of only a core recombination site with no accessory DNA sequences. This was not true of the other three sites tested, confirming that it is the accessory DNA of *cer*, *parB* and the type I hybrid, in conjunction with the accessory proteins ArgR and PepA, that allows them to recombine. The ability of the type II hybrid to recombine as a minimal site was attributed to the sequence of the core recombination site. The type II hybrid and *parB* differ at four positions within their left arm sequences. A set of plasmids with variant left arm sequences was used to test the roles of each of these substitutions. One particular base pair was implicated in determining whether or not recombination occurred. However, this effect was not entirely clear-cut, as the sequence context, in terms of the other three bases tested, could modulate the extent of recombination, in some cases abolishing it altogether.

Gel retardation experiments showed that the DNA-XerC/D complexes migrated to different positions on polyacrylamide gels depending on the sequence of the left arm of the core recombination site. Those minimal sites which formed more retarded DNA-XerC/D complexes in gel retardation assays failed to support intermolecular recombination *in vivo*. The extent of retardation of the DNA-XerC/D complex was found to be due to distortion of the DNA, most likely as a result of protein-induced DNA bending. Circular permutation analysis was used to show that overall bending of the DNA in XerC/D complexes occurred to different extents depending on the sequence of the left arms of the sites. Two lines of evidence indicate that this is probably due to changes in XerC-induced bending: firstly, the DNA sequence changes are within the left arm, which is implicated in XerC binding, and secondly, the relative mobilities of the DNA-XerD complexes are all similar. The changes in bending on binding of XerC are not necessarily due to bending induced by XerC itself, however, but could be a result of interactions between XerC and XerD causing a conformational change in the overall structure of the complex. It was impossible to analyse the XerC-induced bends because DNA-XerC complexes could not be detected. If both proteins induce separate bends in the DNA upon binding, the

changes in the complex mobilities could be due either to an increase in overall bend magnitude, or to changes in the relative position or directions of XerC- and XerD-induced bends.

Time-course experiments at 0°C strongly suggested that the type II hybrid and variant 11 have greater relative affinities for XerC and XerD than any of the other four sites tested. No data was available for the affinity of XerC for any of these sites in the absence of XerD due to the inability to detect DNA-XerC complexes; this is unfortunate since the simplest interpretation of the differences in behaviour between the sites is an alteration in the interactions between XerC and the variant sites as a result of the changes in their left arms.

There are precedents for sequence differences within or adjacent to protein-binding sites influencing both the affinity of a protein for the site, and the structure of the resultant protein-DNA complex. Much of this information has been derived from the structure of the bacteriophage 434 repressor bound to its DNA site (recently reviewed by Koudelka, 1993). Non-contacted bases within a binding site can indirectly affect binding affinity by modulating interactions between the functional groups on the protein and DNA. These effects can result from sequence-dependent alterations in the deformability of a DNA site: where a protein must bend the DNA to make the correct interactions, the affinity of the protein for a particular site will depend on the ability of that DNA sequence to bend appropriately. In general, the more flexible the DNA sequence, or the more flexible the protein, the higher the affinity is expected to be.

The data reported in this Chapter allow a similar interpretation for binding of XerC and XerD to different sequences. It appeared that XerD binds all of the sites with equal affinity and that the bending induced by XerD was very similar for the sites tested. This indicates that the sequence of the left arm does not influence XerD-induced bends. Different results were obtained from the XerC/D complexes. These results suggest that XerC might bend the sites differently, or that the DNA-XerC/D complex differs in structure depending on the left arm sequence. The extent of XerC/D-induced bending correlated well with the estimated relative affinities: those sites that were most retarded by XerC/D binding had the lowest relative affinities for the proteins. These data are consistent with a model in which XerC and XerD must bend the DNA in order to interact with each other while still making contact with the appropriate functional groups within the DNA. If a large bend must be induced so that all of the correct contacts can be made, more energy would be required to do so, and therefore the affinity of the proteins for the DNA would be expected to be low.

This could explain why those sequences which appear to bend more also have a lower relative affinity for XerC and XerD (*parB*, variant 9, variant 14).

The differences in relative affinities and bending reported here could be accounted for if the sequence changes in the left arm alter the disposition of functional groups in the XerC-binding site. The base substitutions which play the greatest role in determining whether or not a minimal site will recombine are a C to G change at position (2) (variant 2) and the reciprocal G to C mutation (variant 14). While there will still be a G/C base-pair at position (2), this substitution could influence the ability of the DNA to be distorted into a particular conformation, perhaps by determining the favoured bending direction. The preferred bending direction for the variant 14 sequence might be unfavourable for a co-operative interaction between XerC and XerD, resulting in the observed reduced relative affinity of this site for the two proteins. In the case of variant 2, the single change in this site compared with *parB* might improve XerC and XerD interactions, thereby slightly increasing the binding affinity, allowing minimal site recombination. This site is perhaps the most informative of all, since it has intermediate affinity, bending and recombination compared with the type II hybrid and *parB*, indicating that this is not an "all or nothing" effect.

For all of the minimal sites that cannot recombine, accessory factors might promote recombination. The accessory factors of *cer* consist of two proteins, ArgR and PepA, and approximately 220 bp of DNA. It is known that ArgR binds to the accessory DNA at the Arg-box sequence, and has been shown to induce a bend (A. Merican, personal communication). It has also been proposed that the accessory DNA may play a role in organising the structure of the *cer*-like site, as it shows the characteristics of flexible DNA (Summers & Sherratt, 1988). The role of the accessory factors may be to improve interactions between XerC and XerD at those core recombination site sequences where their affinities are low. This could be achieved by the accessory factors interacting directly with the core recombination site and XerC and XerD to stimulate binding or recombination, in a similar way to that proposed for the enhancer and Fis in the DNA inversion reactions catalysed by Hin, Gin and Cin (Finkel & Johnson, 1992); alternatively, it could be an indirect effect, for example, the three-dimensional structure imposed by the accessory DNA sequence, in conjunction with ArgR and PepA, could favour the binding of XerC and XerD at the core recombination site. A recent report has demonstrated that a tract of bent DNA within a supercoiled molecule can direct the overall structure of that molecule (Kremer *et al.*, 1993). A similar role could be imagined for the accessory factors of *cer*: bends produced by binding of ArgR to the Arg-box could organise a higher-order nucleoprotein

complex that is competent for recombination. These are, however, currently purely speculative suggestions.

However the accessory factors exert their effect, it must be due to them that resolution selectivity is shown at certain sites. If a minimal site can recombine, there is no requirement for the sites to be arranged within a single molecule. Therefore interactions between the proteins and DNA at these sites must be sufficiently strong to allow productive synapses to form after collision between sites on different molecules. This might result simply from an increase in XerC or XerD binding affinity so that the recombination sites tend to be associated with the recombinases for longer, or more often. Conversely, if the affinity of the recombinases for a particular sequence is reduced, that site will be associated with XerC and XerD less often, and will therefore be less likely to form synapses with other sites. It would consequently be unlikely recombine.

The sites which show resolution selectivity require accessory factors to recombine: if a minimal site can recombine, recombination occurs intermolecularly as well as intramolecularly. The sites that cannot recombine without accessory factors are therefore superior dimer resolution systems. For sites that must fulfil this role, a balance must be struck so that the sequence is effectively bound by XerC and XerD under the correct conditions (in direct repeat within a single molecule), but does not have such a high affinity for XerC and XerD that intermolecular recombination can occur. This can be achieved in several ways. Affinity for XerC can be reduced or increased by altering left arm sequence, as described in this Chapter. Affinity for XerD is reduced in *cer* (and all of the other sites described here) compared with *dif*, by making a T to G substitution in the right arm (G. May, unpublished results). The XerC/D interaction can be weakened by increasing the length of the central region, as in the type I hybrid, *cer* and *dif*-8, and this will reduce the binding affinities (as proposed in Blakely *et al.*, 1993, and discussed in the Introduction to this Chapter). Any mutation that reduces the possibility of forming a stable DNA-XerC/D complex should necessitate the addition of accessory factors for recombination and will result in the formation of a site showing resolution selectivity. Conversely, for a site like *dif*, that is required to recombine intermolecularly, the DNA-XerC/D interaction must be optimised, and this is achieved by having a 6 bp central region, and strong XerC- and XerD- binding sites.

The data discussed in this Chapter begin to address mechanisms behind the resolution selectivity process adopted by plasmid resolution systems, typified by *cer*, so that the plasmid is maintained in a monomeric form. However, more

experiments are required to fully understand this process. Once more active XerC and XerD is available, binding affinities of these proteins for different sites will be calculable, and it will be possible to determine whether it is the affinity of XerC for different sites that is affected *per se*, or if it is the co-operative interaction between XerC and XerD that is influenced. It seems likely that the binding of the recombinases is an important level of control over Xer recombination. This point will be discussed again in Chapter 4.

Chapter Four

The Effect of Fis on Xer Recombination

Introduction

Fis is a small, basic DNA-binding protein, which falls into the class of "histone-like" proteins. The multifarious roles of this protein have recently been reviewed (Finkel & Johnson, 1992; Saenger *et al.*, 1993). Fis has been shown to function in various biological processes, including site-specific recombination systems of both the λ integrase and resolvase/DNA invertase classes, as well as in control of transcription from several operons (including its own), and DNA replication from *oriC*.

The Fis protein, its DNA binding and bending

Fis is a 12 kD protein, consisting of a homodimer of two 98 amino acid subunits. The crystal structure has been solved to 2 Å resolution (Kostrewa *et al.*, 1991), revealing that in each subunit there are four regions of α-helix, called the A, B, C and D helices, and a disordered region of 24 amino acids at the N-terminus. The DNA-binding domain is a helix-turn-helix motif comprising the C and D helices. Fis binds DNA with some degree of sequence specificity, at a 15 bp motif showing partial dyad symmetry. A degenerate consensus sequence has been derived from a number of known Fis binding sites (Hübner & Arber, 1989; Finkel & Johnson, 1992; Lazarus & Travers, 1993; presented later in this Chapter). It has been postulated that this degeneracy is tolerated because the recognition helix, helix D, contains six positively charged residues, thereby allowing the protein to make largely non-specific interactions with the DNA (Kostrewa *et al.*, 1991).

While Fis itself has been successfully crystallised, protein-DNA co-crystals are yet to be obtained. In the absence of such co-crystals, speculative models for DNA-binding have been proposed. The observed distance between the recognition helices in the 3D-protein structure is too short to allow the protein to interact with adjacent major grooves on straight B-form DNA. It has therefore been postulated that the Fis dimer must bind to bent DNA to make the appropriate contacts (Kostrewa *et al.*, 1991). This possibility is supported by DNA footprinting data which suggest that the minor groove width is altered by Fis binding, with the DNA bent towards the major groove (Bruist *et al.*, 1987). The bending angles induced by Fis binding have been calculated for several sites and have been shown to vary between 40-90° depending on the sequence of the binding site and the flanking DNA (Thompson & Landy, 1988; Gille *et al.*, 1991).

The role of Fis in site-specific recombination

Fis was named as the Factor for Inversion Stimulation due to its original discovery as a component of the DNA inversion systems catalysed by Hin and Gin (Johnson & Simon; 1985: Kahmann *et al.*, 1985). In addition to the DNA binding sites to which these recombinases bind to carry out strand-exchange, a recombinational enhancer is required *in cis*. It is here that Fis binds to stimulate recombination. During site-specific inversion, the recombination sites and the enhancer, together with their associated recombinase and Fis, form a nucleoprotein complex called the "invertasome" (Heichmann & Johnson, 1990). Mutational analysis has implicated amino acids 10-36 of Fis in invertasome assembly (Osuna *et al.*, 1991; Koch *et al.*, 1991). There is evidence to suggest that DNA strand-exchange during site-specific DNA inversion may be accomplished by rotation of the recombinase subunits relative to one another following double-strand cleavage (Kanaar *et al.*, 1990; Heichmann *et al.*, 1991). Such rotation would normally be inhibited due to the dimerisation interactions between the protein subunits. It has been proposed that the role of Fis is to weaken these interactions, thereby allowing rotation to occur (Finkel & Johnson, 1992). This hypothesis was supported by the characterisation of Fis-independent Gin and Cin mutants which, in the absence of Gin or Cin crystals, have been mapped to the putative dimerisation interface of the related γδ resolvase (Klippel *et al.*, 1988b; Haffter & Bickel, 1988; Hughes *et al.*, 1990), suggesting that mutations which weaken the protein-protein interactions can overcome the requirement for Fis.

In addition, Fis has been demonstrated to participate in integrative and excisive recombination of bacteriophage λ *in vitro* and *in vivo*. In the absence of Xis *in vivo*, Fis stimulates integration by binding to a single Fis binding site, F, which is in *attP* (or *attR* in the prophage), and influences the stability of lysogens in an unknown manner (Ball & Johnson, 1991b). However, in the presence of Xis the excisive pathway is favoured, and the role of Fis in excision seems to be in promoting the binding of Xis to one of its binding sites, X1, under conditions where Xis is limiting *in vivo* (Ball & Johnson, 1991a) and *in vitro* (Thompson & Landy, 1988). Mutations which affect the binding or bending of DNA by Fis abolish its role in λ recombination; however, mutations in the N-terminal region do not affect λ recombination, unlike site-specific inversion (Osuna *et al.*, 1991; Koch *et al.*, 1991).

Fis as a transcriptional activator and repressor

Fis has been reported to be involved in activation of transcription of stable ribosomal and transfer RNA genes (Nilsson *et al.*, 1990; Ross *et al.*, 1990; Lazarus

& Travers, 1993). Fis has been shown to bind to conserved sequences upstream of the *thrU/tufB*, *rrnB* and *tyrT* genes, and activation of expression of these genes by Fis *in vivo* and *in vitro* has been described (Nilsson *et al.*, 1990; Ross *et al.*, 1990). However, more recently, it has been observed that Fis only stimulates transcription of *tyrT* under conditions when the promoter-RNA polymerase interaction is impaired by promoter down-mutations or by certain physiological conditions. Under such unfavourable conditions, Fis activates transcription by facilitating binding of RNA polymerase to the promoter. It has been proposed that this may be achieved by the localisation of stable RNA promoters to an apex of negatively supercoiled DNA by means of a Fis-induced bend upstream of the start of transcription, thereby allowing these promoters to be distinguished from the rest of the chromosome (Lazarus & Travers, 1993).

Fis represses its own expression by binding to sites overlapping with the promoter of its own operon, thereby competing out RNA polymerase binding (Ninnemann *et al.*, 1992). However, this is not the only level of control of Fis expression, as this operon is also subject to stringent control resulting in a growth-phase dependent expression pattern (Ninnemann *et al.*, 1992).

Mutations that affect the ability of Fis to stimulate transcription map to a region in the N-terminus, partially overlapping the region involved in inversion stimulation (Finkel & Johnson, 1992). Although this domain is separate from the DNA binding domain, mutations abolishing binding also prevent transcriptional stimulation.

The role of Fis in stimulation of DNA replication from *oriC*

Fis has been found to be involved in DNA replication at *oriC* (Filutowitz *et al.*, 1992). *fis* mutants produce the wrinkled colonies typical of *E. coli* cells with altered cell division, and poorly maintain *oriC* minichromosomes. On microscopic examination, the cells were found to be filamentous and to possess aberrant nucleoids, particularly at high temperatures. It was also established that Fis acts synergistically with DNA gyrase. This suggests that, at high temperatures, when DNA topology is affected, Fis is most required for replication. Fis was shown to be required at, or before, the transcription step in chromosome replication. These data have lead to the proposal that Fis binding at *oriC* could be playing a topological role during chromosome replication, perhaps by favouring the formation of an initiation complex by bending the DNA either to facilitate strand separation, or to improve interactions between other proteins bound at the origin.

Aim

These experiments attempted to investigate whether Fis also affects Xer recombination and, if so, how. Its effect might be exerted in a number of ways: for example, during Xer site-specific recombination, it might bind at or near the core recombination site to facilitate interactions between XerC and XerD with each other, with the DNA, or with either or both of the accessory proteins ArgR and PepA. Alternatively, it might affect expression of any of the four Xer genes. Some of these possibilities are addressed, and a possible role for Fis in Xer recombination is proposed.

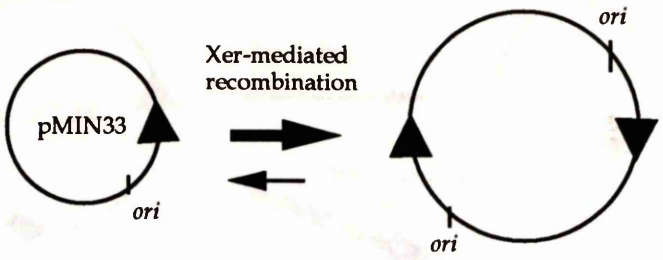
Results

The effect of Fis on Xer-mediated site-specific recombination at *dif* and *cer in vivo*

To analyse the effect of Fis on Xer-mediated site-specific recombination at *dif* and *cer*, recombination of plasmids containing one or two copies of these sites was assayed in *fis*$^+$ and *fis* cells. In order to assay Xer-mediated plasmid recombination, as opposed to homologous recombination, the *recF* strain DS941 was used, and a *fis* mutant of this strain was constructed by P1 transduction of the kanamycin resistance marker from CSH50*fis*::Km. The resultant DS941*fis*::Km strain was named JR10.

DS941 and JR10 were transformed with purified monomers of pMIN33, pSDC124 and pKS455, and transformants selected on LB agar containing either ampicillin alone to select for the plasmids in DS941, or ampicillin and kanamycin to select for the plasmids in JR10. These "reporter" plasmids are described in Figures 4.1, 4.2 and 4.3.

From each transformation plate, the plasmid content of each of five separate transformants was analysed by single colony gel electrophoresis, as described in Chapter 2. The results are shown in Figures 4.1, 4.2 and 4.3. Figure 4.1 shows that in the *fis* mutant JR10, substantially less Xer-mediated intermolecular recombination between pMIN33 plasmids occurred than in DS941. This indicates either that intermolecular recombination at *dif* has been reduced *per se*, or that in JR10 there is increased resolution selectivity such that the preferred oligomeric state for the plasmid is monomer. By studying Figure 4.2, in which pSDC124 has been used as a substrate for intermolecular and intramolecular recombination, the latter possibility can be eliminated. On this



pMIN33 contains a minimal *dif* site within pUC18. Xer-mediated recombination between plasmids can occur at *dif* to produce plasmid multimers. pMIN33 can therefore be used to assay intermolecular recombination at *dif in vivo*.

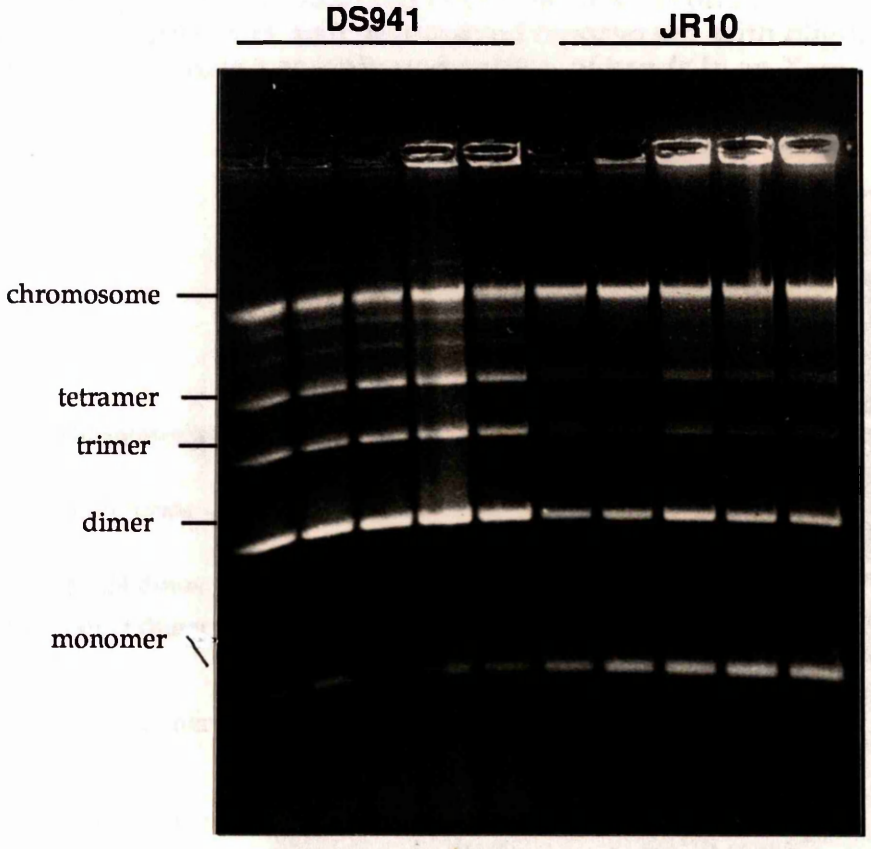
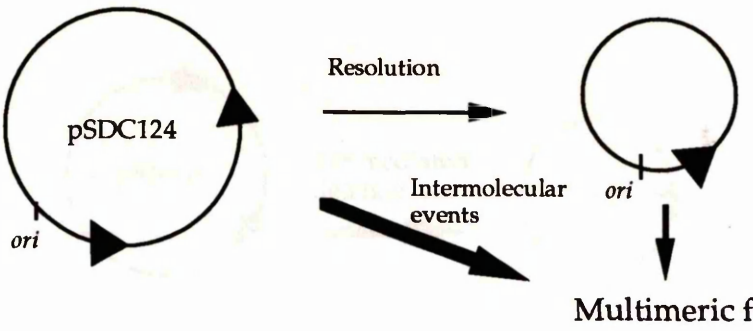


Figure 4.1: Xer-mediated intermolecular recombination between *dif* minimal sites.
DS941 and JR10 were transformed with monomers of pMIN33, and five individual transformants of each isolated and their plasmid content analysed by lysis in single colony final sample buffer then electrophoresis on a 1% agarose gel. The main oligomer sizes are indicated at the left of the photograph. The first five tracks are transformants of DS941, the second set are from JR10.



pSDC124 is a *dif* "reporter" plasmid containing two minimal *dif* sites in direct repeat that can be used to assay inter- and intra-molecular recombination events. Intermolecular recombination produces plasmid multimers on a gel. Intramolecular recombination produces a replicative resolution product of approximately the same size as pMIN33. This product can undergo intermolecular recombination with other resolution products, with unresolved reporters or with oligomers of pSDC124 to produce a complicated pattern of bands in an Xer+ strain.

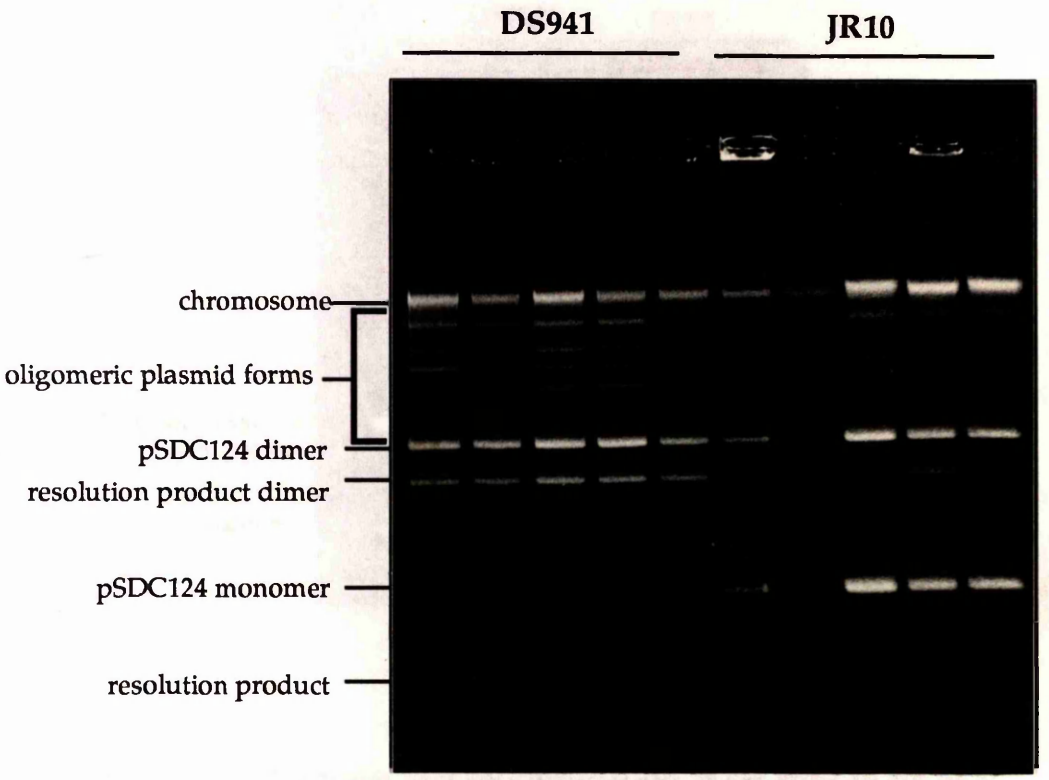
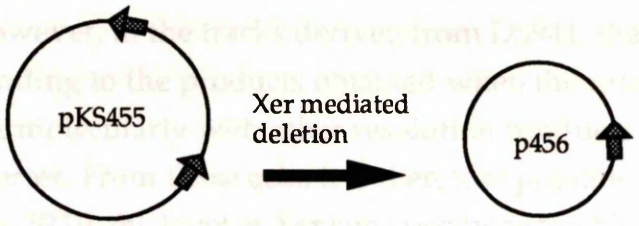


Figure 4.2: Intramolecular and intermolecular recombination of pSDC124 in DS941 and JR10.
Monomers of pSDC124 were transformed into DS941 and JR10. Five transformants were isolated from each strain, the cells lysed in single colony final sample buffer, and the plasmid content analysed by electrophoresis on a 1% agarose gel. The main plasmid forms are indicated along the left side of the photograph.



The pUC18-based *cer* reporter plasmid pKS455 has two *cer* sites in direct repeat flanking a chloramphenicol resistance marker. Xer-mediated recombination occurs almost exclusively in the direction shown to produce the resolution product, p456.

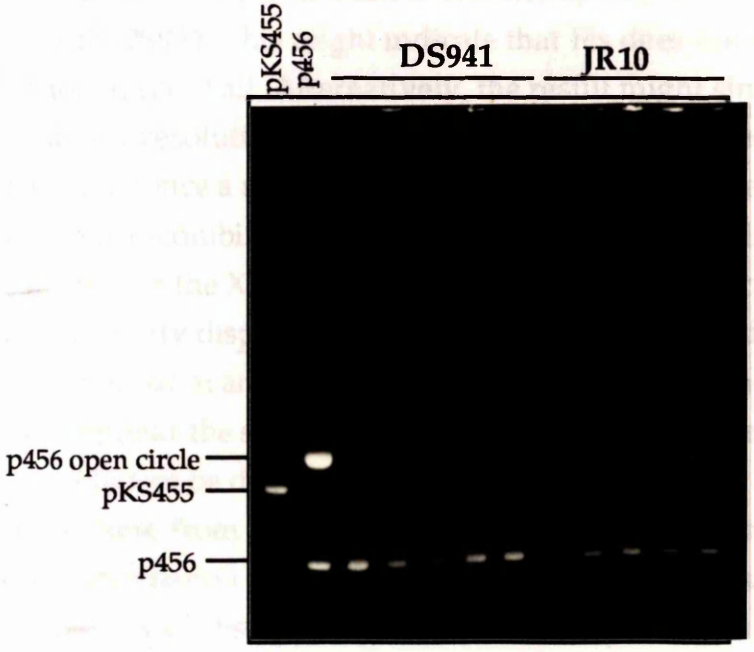


Figure 4.3: Intramolecular recombination at *cer* sites in pKS455. Monomers of pKS455 were transformed into DS941 and JR10. Five transformants from each strain were isolated and lysed in single colony final sample buffer before electrophoresis on this 1% agarose gel. pKS455 and p456 markers are indicated alongside the gel.

gel, similarly low amounts of monomeric resolution product were obtained in both strains. However, in the tracks derived from DS941, there is a ladder of bands corresponding to the products obtained when the resolution product recombines intermolecularly with other resolution products and with the unresolved reporter. From these gels, it is therefore possible to suggest that in the *fis* mutant strain, JR10, the level of Xer site-specific recombination at *dif* is reduced, and that this is true both for the inter- and intra-molecular reactions.

A similar experiment was performed using the 2-*cer* reporter plasmid pKS455. Recombination between *cer* sites shows strong resolution selectivity, as introduced in Chapter 3, so that the products of recombination *in vivo* are almost exclusively deletion products, and the ladder of plasmid multimers similar to that observed with *dif* substrates is not seen. The results of the experiment in Figure 4.3 show that when a *cer* reporter substrate was used, there was no discernible reduction in the amount of Xer site-specific recombination in JR10 compared with DS941. This might indicate that Fis does not affect Xer-mediated recombination at *cer* at all. Alternatively, the result might simply reflect the fact that the assay for resolution is relatively insensitive compared with that for multimerisation. Once a resolution event has occurred between *cer* sites, unless the homologous recombination system produces more multimers, there are no more substrates for the Xer system to recombine. In other words, due to the resolution selectivity displayed at *cer*, only one event is required. This means that resolution can occur at any point between transformation and analysis and the result would appear the same: while recombination could be at a lower level in JR10, this would not be detected using a *cer* reporter in an *in vivo* assay.

To conclude from these initial experiments, Fis was shown to affect Xer-mediated recombination at *dif in vivo*. No reduction in resolution of the *cer* reporter pKS455 was observed, but this might simply reflect the limitations of the *in vivo* assay used.

Does Fis bind to *dif* or *cer* to affect Xer-mediated recombination?

One possible explanation for how Fis could affect recombination is that it might bind to *dif* and/or *cer*, thereby influencing the ability of XerC or XerD to bind, or to improve interactions between these two proteins. The *dif* core recombination site fits the Hübner and Arber Fis binding consensus reasonably well (Hübner & Arber, 1989). The same is not true for the *cer* core recombination site despite the high level of sequence homology between the two sites.

TTGGTGCGCATAATGTATATTATGTTAAAT	*dif* core recombination site
GNNYRNNTNNYRNNC	Hübner & Arber
T A A	
GCGGTGCGTACAATTAAGGGATTATGGTAAAT	*cer* core recombination site

[Note: *dif* does not conform to more recently published Fis binding consenses (Finkel & Johnson, 1992; Lazarus & Travers, 1993), but at the time of these experiments the Hübner and Arber sequence was the best published consensus.]

To test whether Fis binds to *dif* or to *cer in vitro*, gel retardation assays were employed. The following DNA fragments were used for binding:
a) a 320 bp *Bam*HI-*Hin*dIII fragment from pKS492 containing *cer*
b) a 295 bp *dif*-containing *Nde*I-*Eco*RI fragment from pMIN33
c) a 180 bp *Bam*HI fragment from pAL21*sis* containing the control Fis binding site, *sis*, the enhancer of Gin-mediated recombination.
For each end-labelling reaction, approximately 0.5 µg plasmid was digested, and labelling was performed as described in Chapter 2. Also see Chapter 2 for descriptions and references for these plasmids.

The stock Fis protein used in the binding experiments was a gift from Regine Kahmann, at 9 $mg.ml^{-1}$. This was diluted in Fis dilution buffer (see Chapter 2) to give final concentrations in the binding reactions of 312.5 nM, 62.5 nM, 12.5 nM, 2.5 nM. The binding conditions used were the GBA1 conditions described in Chapter 2, with polydI.dC carrier DNA at a final concentration of 0.5 $\mu g.\ reaction^{-1}$. Reactions were incubated for 10 minutes at 37°C.

The results of such a gel retardation experiment are shown in Figure 4.4. This binding gel shows that Fis bound relatively well to *cer* (Fig 4.4a) compared with *dif* (Fig 4.4c), which was bound poorly. While Fis was not able to bind *cer* to the same extent as the positive control fragment, *sis* (Fig 4.4b), it did bind reasonably well. Notice that more than one complex was produced with each of these fragments. This could be as a result of using these relatively large DNA fragments which might contain several poor Fis binding sites to which Fis could bind through largely non-specific interactions, especially at high protein concentrations.

A Fis binding consensus sequence was identified in the *cer* accessory DNA, adjacent to the Arg box. The sequence of the *Hin*dIII-*Bam*HI *cer* fragment used in the experiments above is shown below. The putative Fis binding site in large, bold font, the *Pvu*I site shown underlined, the Arg box italicised and the core recombination site in bold.

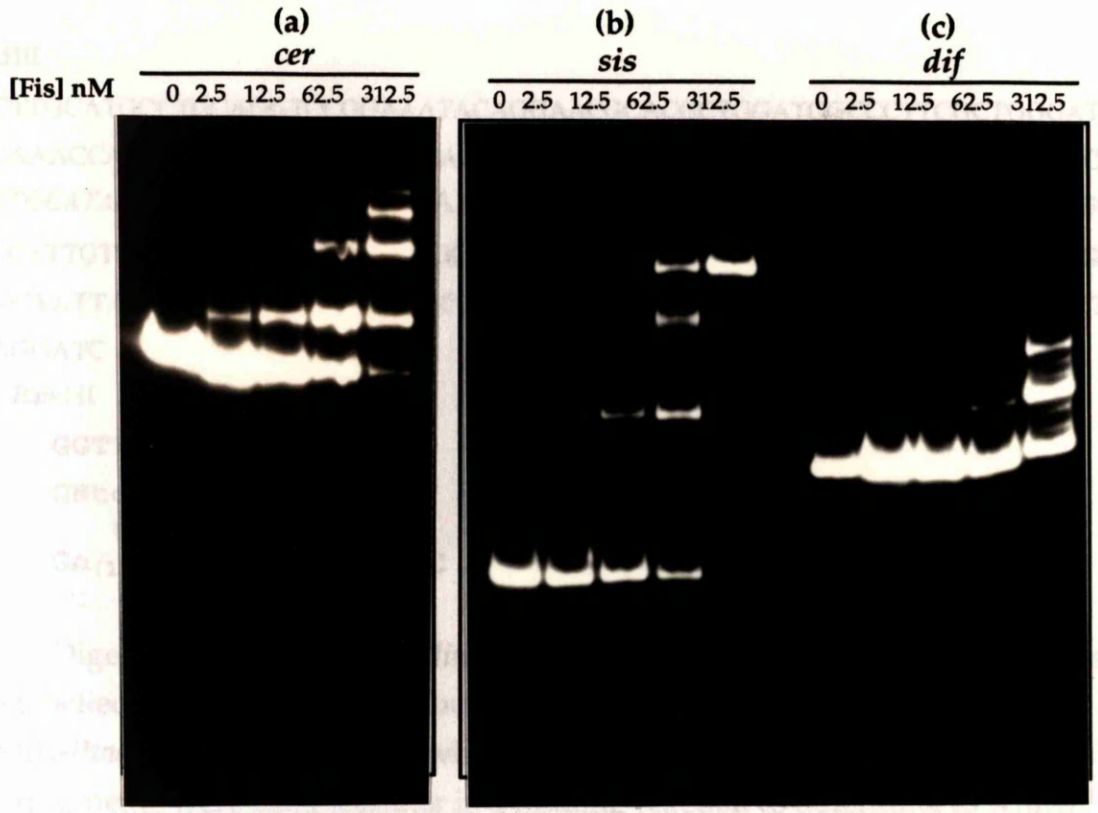


Figure 4.4a,b,c: Binding of Fis to *cer* and *dif*.
Binding of Fis to the *cer*, *dif* and *sis* fragments as described in the text. The concentrations of Fis in the binding reactions in nM is shown above each track.

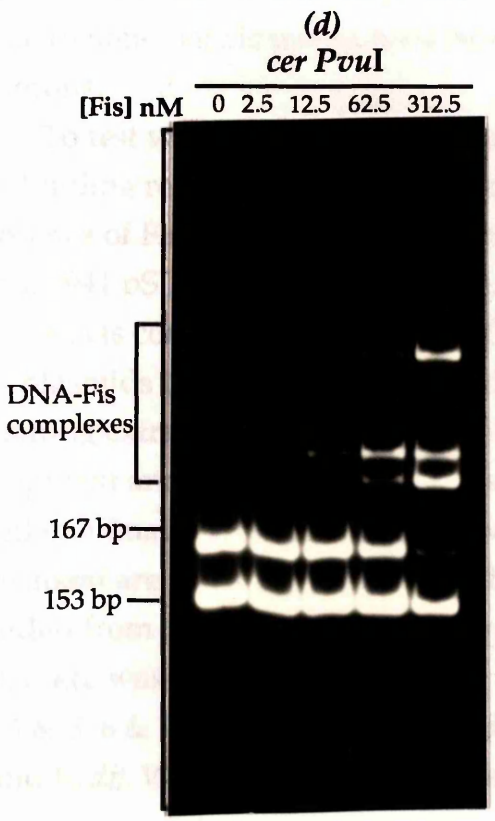


Figure 4.4d: Binding of Fis to *PvuI* digested *cer*.
The *cer* fragment used for binding in Figure 4.4a was digested with *PvuI*, and the digested DNA used in a Fis binding experiment. The concentrations of Fis used are as above, and are annotated at the top of the gel. The larger of the two fragments is the *Hin*dIII-*Pvu*I fragment, and contains the putative Fis binding site.

*Hin*dIII

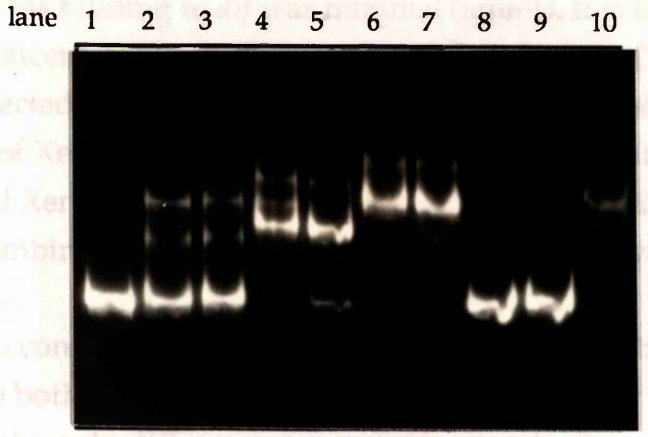
AGCTTGCATGCCTGCAGGTCCGGAAATACAGGAACGCACGCTGGATGGCCCTTCGCTGGGATG
GTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGTAATGTGGCCTGTACCCTCT
*GGTTGCATAGGTATTCAT*AC**GGTTAAAATTTATCA**GGCGCGATCGCGGCAGTTTTTCGGG
GTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATCGCGCTGAACCGCGTTTTA**GCGGTGC**
GTACAATTAAGGGATTATGGTAAATCCACTTACTGTCTGCCCTCGTAGCCATCGAGCACTCT
AGAGGATC

*Bam*HI

GGTTAAAAATTTATC	**Putative Fis binding site in *cer***
GNtcAaaTttTgaNC	**Fis consensus (Finkel & Johnson, 1992)**
tg tAa ca	
$Gn_{(1-3)}YRn_{(4-6)}YRn_{(1-3)}C$	**Fis consensus (Lazarus & Travers, 1993)**

Digestion of this labelled *Hin*dIII-*Bam*HI *cer* fragment with *Pvu*I produces two labelled fragment of similar, but electrophoretically distinct, sizes. The *Hin*dIII-*Pvu*I fragment is 167 bp, while the *Pvu*I-*Bam*HI fragment is 153 bp. These two fragments were used together in a binding reaction to determine to which Fis bound with greater relative affinity. The results of such a binding experiment are shown in Figure 4.4d: Fis had a greater affinity for the larger of the two fragments. This indicates that there may indeed be a Fis binding site in this fragment; if so, it is most likely to be the sequence that is highlighted in bold. Notice that Fis also bound to the other fragment, but to a lesser extent; this might be due to non-specific interactions between this fragment and Fis under these conditions.

To test whether Fis might stimulate the binding of XerC or XerD to *dif in vitro*, binding reactions were carried out with XerC and/or XerD in the presence or absence of Fis. Crude protein extracts were prepared as described in Chapter 2 from DS941 pSDC105, DS941 pRM130, JR10 pSDC105 and JR10 pRM130: these four extracts contained either XerC (pSDC105) or XerD (pRM130) over-expressed from plasmids in either a *fis*$^+$ or *fis* background. Binding of the XerC- and XerD-containing extracts prepared from the two strains was compared with the same *dif* fragment and binding conditions as described above. In addition, Fis was added to a final concentration of 20 nM in one reaction. The results of this experiment are shown in Figure 4.5. This experiment showed that if Fis is entirely excluded from the binding reactions by preparing XerC and XerD extracts from JR10, there was no effect on the ability of XerC or XerD to bind (compare lanes 2 & 3, 4 & 5, 6 & 7), indicating that Fis did not stimulate the ability of XerC or XerD to bind to *dif*. When Fis was added to supplement a similar reaction to that in



Protein addition by lane number:

1. no protein
2. DS941 pSDC105
3. JR10 pSDC105
4. DS941 pRM130
5. JR10 pRM130
6. DS941 pSDC105 + DS941 pRM130
7. JR10 pSDC105 + JR10 pRM130
8. Fis 10 nM
9. Fis 20 nM
10. JR10 pSDC105 + JR10 pRM130 + Fis (20 nM)

Figure 4.5: Does Fis stimulate binding of XerC or XerD to *dif in vitro*? Crude protein extracts made from either DS941 or JR10 containing the XerC over-expressing plasmid pSDC105 or the XerD over-expressing plasmid pRM130 were used in binding reactions with *dif*. For each reaction, 1 µl of each extract being used was added; the proteins that were added to each reaction are noted below the gel, numbered according to lane number. Fis was added at a final concentration of 20 nM to a reaction containing XerC and XerD extracts from JR10 to see if binding was stimulated (lane 10).The protein additions made to each reaction are shown at the top of the Figure.

lane 7, no difference in binding was observed (lane 10). One possible proviso to this result is that, at the concentration of Fis used to supplement this reaction, the extent of Fis binding to *dif* was minimal (lane 9). It is therefore possible that at higher concentrations of Fis, some stimulation of XerC/D binding might have been detected. This experiment also excludes the possibility that Fis inhibits binding of XerC or XerD to *dif*, as no increase in binding was observed when the XerC and XerD extracts were prepared from JR10. This is consistent with the *in vivo* recombination assays which showed that there was less recombination at *dif* in JR10.

To conclude, these *in vitro* binding experiments showed that while Fis bound to both *cer* and to *dif*, its relative affinity for *cer* was greater, despite the fact that the only difference in recombination *in vivo* was observed with a *dif* substrate. Fis was also shown not to stimulate the binding of XerC and XerD to *dif in vitro*. It therefore seems unlikely that the effect of Fis on Xer-mediated recombination is exerted through altering XerC/D interactions with the core recombination site.

The effect of Fis on Xer recombination *in vitro*

To test whether Fis stimulates Xer recombination between *cer* sites *in vitro*, an *in vitro* assay was employed (R. McCulloch, pers. comm.). In this assay, when ArgR, PepA, XerC and XerD are added to a supercoiled *cer* reporter plasmid in a simple buffer, Holliday junctions are formed. This is similar to the generation of Holliday junctions *in vivo* when XerC is over-expressed from the inducible *lac* promoter in the strain RM40 (McCulloch, 1992; McCulloch *et al.*, 1994a).

An experiment was performed to investigate the role of Fis on recombination at *cer in vitro*. For this experiment, crude protein extracts containing XerC or XerD were prepared from four strains- DS941 pSDC105, DS941 pRM130, JR10 pSDC105 and JR10 pRM130- in order to over-express XerC (pSDC105) or XerD (pRM130) from *fis*$^+$ and *fis* backgrounds, as above. In this way, it was hoped that Fis could be eliminated from the reactions when JR10 extracts were used; however, purified ArgR and PepA (kindly provided by Mary Burke) had been prepared from a *fis*$^+$ background, so the possibility did exist for some contamination by Fis. Reactions were carried out with combinations of the four Xer proteins with or without added purified Fis: the protocol for the *in vitro* recombination reaction is described in Chapter 2.

The products of the *in vitro* recombination reaction were separated by electrophoresis on a 1% agarose gel both before and after digestion with *Eco*RI. This gel, shown in Figure 4.6, implies (but does not prove) that Fis does not

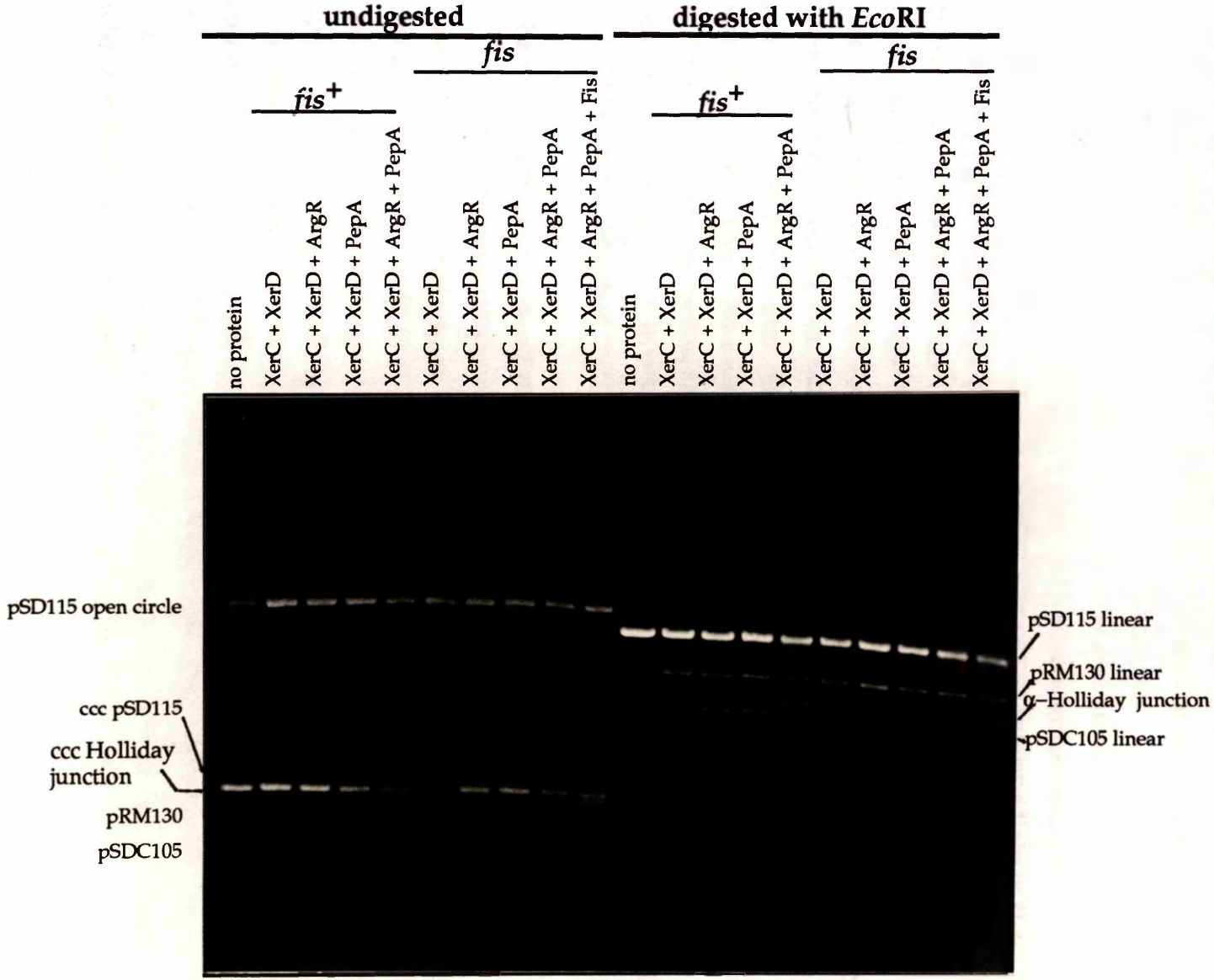


Figure 4.6: The effect of Fis on Xer recombination *in vitro*.
Recombination of the *cer* reporter plasmid pSD115 was assayed *in vitro* using crude protein extracts of XerC and XerD made from the over-expressing plasmids pSDC105 and pRM130 in DS941 and JR10. The proteins used in each reaction are indicated at the top of the gel: for simplicity, "XerC" means crude extract from DS941 pSDC105 or JR10 pSDC105 (the strain is referred to as *fis*$^+$ or *fis)* and "XerD" likewise means extracts from DS941 pRM130 or JR10 pRM130. The same extracts were used as in Figure 4.5. Purified ArgR and PepA were provided by Mary Burke. On the left half of the gel, the reactions were electrophoresed prior to restriction digestion: some supercoiled Holliday junction can be seen at the position indicated (supercoiling indicated by ccc).The right half shows the reactions after digestion with *Eco*RI, which cleaves the Holliday junction in one domain to generate an α structure (with one supercoiled domain and two "arms"): its mobility is faster than linear pSD115 due to the supercoiled region.

contribute to the formation of Holliday junctions during Xer site-specific recombination at *cer* sites *in vitro*. This is evident for two reasons. Firstly, Holliday junction α structures were visible whether or not the source of XerC and XerD extracts was a *fis* mutant strain (compare lanes 15 and 19), and secondly, the amount of Holliday junction did not increase if excess Fis was added (lane 20). Of course, the possibility remains that the reactions already contained Fis from either of the ArgR or PepA preparations, as discussed above.

There was one obvious difference between corresponding reactions that originated either from a *fis*$^+$ or *fis* background: there was a clear DNA band in lanes 16 & 17 that migrated slower than the linear pSD115, which was not visible in the corresponding lanes from the DS941 background (lanes 12 &13). This species was found to be artefactual, as later experiments failed to reproduce it (data not shown); it will therefore not be discussed further.

This experiment indicates that Fis plays no role in recombination at *cer in vitro*. It must be borne in mind that an entire resolution reaction was not reproduced here: Holliday junctions are believed to be an intermediate in the recombination reaction only, although it is not known why they appear under these conditions while we have to date failed to produce resolution products (this point is discussed further in McCulloch, 1992; McCulloch *et al.*, 1994a). The possibility remains that Fis is involved in the conversion of Holliday junctions to recombination product, but that the conditions described here do not permit the reaction to go to completion: perhaps another factor is required for the second strand-exchange step. This experiment does indicate that the addition of Fis to the existing *in vitro* recombination reaction conditions is insufficient to drive the reaction to form product. In other words, Fis is not the "missing factor" required to reproduce the complete recombination reaction *in vitro*.

The expression of *xerC* is affected by Fis

If Fis does not contribute to the recombination reaction *per se*, another possibility could be envisaged: it may affect the expression of one of the recombinases, XerC or XerD. Fis has been demonstrated to play a variety of cellular roles, including transcriptional activation of the *rrnB*, *thrU/tufB* and *tyrT* operons under certain conditions, and transcriptional repression of its own synthesis (reviewed in Finkel & Johnson, 1992). It is therefore possible that Fis could influence Xer recombination by activating or repressing the expression of one or both of the recombinases. While it might also affect *argR* or *pepA* expression, since the greatest effect on Xer recombination *in vivo* was observed with a *dif* substrate, *xerC* and/or *xerD* seemed the likelier candidates.

The results reported in Chapter 3 indicated that the sequence of the left arm of *cer*-like core recombination site strongly influences the resolution selectivity process, and that this might be due to changing the affinity of XerC for the DNA sequence. The right arms of *cer*-like sites including *dif* are all much more similar in sequence than the left arms, and, indeed, XerD binds reasonably well to all such sites. This may suggest that XerC is a more likely "control point" for Xer-recombination, and so the effect of Fis on expression of *xerC* was investigated.

As a tool to study the effect of Fis on the expression of *xerC*, the strain DS984 was used. DS984 is a derivative of DS941, with a Mu *d*II PR13 insertion in the *xerC* reading frame at nucleotide 1058. This mutation was originally constructed by Richaud *et al.* in 1987 but was transduced into DS941 by Sean Colloms (Colloms, 1990; Colloms *et al.*, 1990).This mini-Mu contains a copy of *lacZ* lacking transcriptional and translational starts, such that if the insertion is within the open reading frame of a gene, the level of expression of that gene can be assayed by the production of β-galactosidase. Consequently, DS984 is blue on plates containing X-Gal, while DS941 colonies are white (not shown).

To make use of this mini-Mu insertion to assay expression levels of *xerC*, a *fis* mutant of this strain was constructed by P1 transduction of the kanamycin resistance allele from JR10 into the Cm^r DS984. The resulting Km^rCm^r strain was named JR20. JR20 colonies are also blue on X-Gal plates indicating that the mini-Mu co-transduced with the Km^r allele. However, X-Gal plates cannot accurately assay the β–galactosidase activity. To compare the levels of expression of *xerC* in *fis*$^+$ (DS984) and *fis* (JR20) backgrounds, β-galactosidase assays were carried out on cell extracts from JR20 and DS984 using the substrate *o*-nitrophenyl-β-*D*-galactopyranoside (ONPG) which is cleaved to the yellow product *o*-nitrophenol (ONP) by β-galactosidase. The protocol for β-galactosidase assays which was used is described in Chapter 2. To calculate the specific activities of β-galactosidase in DS984 and JR20, and therefore gain a better picture of *xerC* expression in these two strains, the total protein in the cell extracts was also calculated using Bradford assays, as also described in Chapter 2. The results of the β-galactosidase assays are shown in Figure 4.7.

Figure 4.7 shows that the level of β-galactosidase expression in JR20 extracts was approximately 0.55 of those of DS984 suggesting that the expression of *xerC* might be reduced similarly in a *fis* mutant. The difference in β-galactosidase specific activities between JR20 and DS984 was found to have a 95% probability of being statistically significant using a Student's t-test (kindly performed by Karen Grant). The ratio of β-galactosidase produced from these

Strain	β-galactosidase activity nmol ONP released. ml cells^{-1}.min^{-1}	Total protein μg protein. ml cells^{-1}	β-galactosidase specific activity nmol ONP. μg protein^{-1}.min.$^{-1}$
DS941	5.33 x 10^3	140	38
	1.60 x 10^3	78	20
	1.33 x 10^3	114	11
DS984	4.20 x 10^4	200	210
	3.44 x 10^4	114	302
	3.33 x 10^4	158	210
JR20	3.00 x 10^4	250	120
	2.56 x 10^4	146	175
	1.86 x 10^4	176	106

Figure 4.7: The effect of Fis on the expression of XerC.
β-galactosidase assays were carried out for overnight cultures of DS941, DS984 and JR20. DS941 is a negative control. This experiment was carried out on three independent occasions, and the results of the three experiments are tabulated above. The values in the second column were obtained by assaying the cleavage of ONPG to ONP by β-galactosidase spectrophotometrically, and converting the A_{420} to nmoles ONP produced using a standard curve. In order to determine specific activity of β-galactosidase in the three strains, total protein was calculated using Bradford' s assays. These values are in the third column. The specific activities were obtained by dividing the β-galactosidase activities by total protein, and these are shown in column four. Notice that while the actual values vary slightly between experiments, the trend remains the same: in JR20, reproducibly less β–galactosidase expression occurs than in DS984.

If the β-galactosidase activities from JR20 are presented as a proportion of those from DS984, the following values are obtained: 0.57, 0.58 and 0.50 (mean = 0.55).

two strains was consistent in three separate assays (although the absolute specific activities varied slightly) implying this is a result that is a feature of these strains.

To summarise, use of the strains DS984 and JR20 allowed the level of expression of *xerC* to be analysed by the use of β-galactosidase assays. While these assays are a rather indirect indicator of *xerC* expression, the results indicate that in the *fis* mutant JR20 the level of expression of β-galactosidase was reproducibly and significantly lower than in the *fis* wild type strain DS984. This strongly suggests that *xerC* expression is reduced in a *fis* mutant strain, and therefore that Fis might stimulate expression of *xerC*. These experiments were only possible due to the availability of the mini-Mu insertion in *xerC*. It would be desirable to repeat the experiments described here with similar *xerD* mutant strains when they become available.

Discussion

The data described in this Chapter suggest that Fis is involved in Xer site-specific recombination, and have attempted to address the basis of this role.

In the *fis* mutant strain JR10, the level of inter-and intra-molecular Xer-mediated recombination at *dif* was found to be reduced in an *in vivo* assay. The same was not observed for *cer*. It is possible that Fis does not affect Xer recombination at *cer*, or this might be explainable by the insensitivity of the *in vivo* assay for resolution. An *in vitro* recombination assay with a *cer* reporter plasmid implied that the formation of Holliday junction intermediates between *cer* sites does not require Fis, and is not stimulated by the addition of excess purified Fis. This does not exclude the possibility that Fis aids in the resolution of Holliday junctions to product, as the second strand-exchange cannot be reproduced *in vitro* as yet. However, it does indicate that Fis is not the factor that is required to drive the *in vitro* reaction to completion. The effect of Fis on a *dif* substrate has not been tested by a recombination assay *in vitro* to date, but no great stimulation of recombination would be expected if it is true that the observed differences in the extent of recombination at *cer* and *dif in vivo* are due only to the limitations of the assay system. No effect of Fis on XerC or XerD binding to *dif* was observed *in vitro* in Figure 4.5. Taken together, these data suggest that Fis does not directly participate in Xer-mediated recombination.

Although the data are consistent with the theory that Fis does not interact with the components of the Xer recombination system to stimulate recombination, the protein has been shown to bind to both *dif* and *cer* fragments

in vitro. This binding might be irrelevant to recombination, and could even be largely artefactual, due either to the recognition by Fis of a degenerate DNA sequence, or to the *in vitro* binding conditions used. Fis bound neither *cer* nor *dif* as well as the control fragment, *sis*; however, *cer* did display better binding than *dif*, and a reasonable Fis binding site could be identified in *cer*. It must also be remembered that the *dif* fragment used in these binding assays had ~260 bp of pUC18 DNA in addition to the *dif* sequence, and the observed binding could have been within these sequences.

Significantly lower β-galactosidase specific activity was measured in JR20 compared with its *fis*$^+$ counterpart, DS984. This infers that in the *fis* mutant strain the expression of *xerC* is lower, since the expression of *lacZ* in these strains is driven by the signals normally directing the expression of the *dapF/xerC* operon. One possibility is that this could be due to Fis binding to the promoter region to stimulate transcription, perhaps by bending the DNA to facilitate interactions between RNA polymerase and other factors controlling expression. It is formally possible, of course, that the amount of β-galactosidase is reduced in JR20 as a result of other phenomena not specifically related to its transcription and translation, for example increased proteolysis, or an overall reduction in gene expression. While this possibility cannot be ruled out, no previous reports of such phenotypes in a *fis* mutant have been described, and total protein measured from JR20 was generally found to be, if anything, slightly higher than that from DS984. It is also possible that the effect could be exerted indirectly, for example if Fis affects the expression of another factor that stimulates *xerC* transcription.

The difference in β-galactosidase activity between these two strains, though statistically significant, is not large. This could be due to the fact that levels were assayed from overnight cultures of cells, which would be at stationary phase. At this point in the cell cycle, there are <100 molecules of Fis protein per cell (Finkel & Johnson, 1992). If growth were to be re-initiated, a rapid increase in the production of Fis would be expected, until levels peak at around 50000 molecules per cell. This suggests that under conditions of exponential growth a much larger difference in β-galactosidase activity between DS984 and JR20 might have been observed. This possibility could be simply addressed by carrying out β-galactosidase assays during different points of a growth curve. Alternatively, if DS984 were to be transformed with a plasmid over-expressing *fis*, the β-galactosidase levels would be expected to increase dramatically if it is true that Fis stimulates expression from the *dapF/xerC* promoter. Of course, *in vitro* transcription from the natural *xerC* promoter in the presence or absence of Fis would demonstrate whether Fis directly affects transcription. These

experiments have not yet been attempted.

Limited sequence information upstream of the *dapF* coding region is available (Richaud & Printz, 1988). Consequently, it is impossible to determine by eye whether there are any Fis binding consenses upstream of the start of transcription of the operon containing *xerC*. No further experiments were carried out during the course of this work to establish unequivocally whether Fis controls the expression of *xerC*, or indeed, whether it also affects the expression of *xerD*. The data obtained to date provide evidence that Fis may stimulate the expression of the XerC recombinase, and that this is the mechanism by which Xer-mediated recombination is reduced in a Fis mutant. It seems likely that the binding of XerC plays a pivotal role in Xer recombination, and fluctuations in its intracellular concentration will undoubtedly greatly affect the level of recombination at different sites depending on their affinities for XerC (for example, due to the sequence of their left arm as described in Chapter 3). If Fis does activate *xerC* expression, this might provide a means by which Xer-mediated recombination can be regulated depending on the growth phase, and could be a valuable mechanism for ensuring that most recombination at *dif* occurs during periods of exponential growth when it is most likely to be required.

Finally, it has been demonstrated that *fis* mutants are filamentous and that the nucleoids are abnormally distributed within the filament, producing densely staining masses (Filutowitz *et al.*, 1992). While this observation has been attributed to the impaired DNA replication experienced by these cells, it is noteworthy that a similar phenotype has also been reported for *xerC* mutants (Blakely *et al.*, 1991), and *dif* mutants (Kuempel *et al.*, 1991). It is therefore tempting to speculate that this phenotype of *fis* cells could be due in part to a reduction in the intracellular XerC concentration and the concomitant reduction in recombination at *dif*.

Chapter Five

A Chromosomal Mutagenesis and Screen for Further Xer Mutations

Introduction

A mutagenesis of the *E. coli* chromosome was performed to attempt to identify additional *xer* genes. This mutagenesis was considered worthwhile for two main reasons, which are discussed below.

The *psi* site from plasmid pSC101

Many multi-copy plasmids contain sites that ensure their stable inheritance at host cell division by acting as the DNA substrates for site-specific recombination mediated by XerC and XerD. This was originally established for ColE1 (Summers & Sherratt, 1984), but sites similar to *cer* have subsequently been identified in other related plasmids. An alignment of the core recombination sites of some such sites is presented in the Introduction to Chapter 3 (Table 3.1). In addition, many sequences that act as plasmid resolution sites also require ArgR and PepA, as well as several hundreds of base pairs of accessory sequences to permit recombination (Summers & Sherratt, 1988; Stirling *et al.*, 1988b., 1989; also see Chapter 3).

A plasmid resolution site has been identified in the natural oligo-copy plasmid pSC101 (Cohen & Chang, 1973, 1977); the site, called *psi* , is located between co-ordinates 6500 and 6863 on the pSC101 map (Bernardi & Bernardi, 1984; Kuempel *et al.*, 1991). To our knowledge, this is the first plasmid unrelated to ColE1 in which an Xer recombination site has been identified. The entire sequence required for resolution of plasmid dimers is located within a 363 bp *Dde*I fragment; this was cloned into the *Sma*I site of pUC18 to construct the plasmid pSDC133 (S. Colloms, personal communication). Resolution of dimers of pSDC133 was found to require XerC, XerD and PepA, but was independent of ArgR (S. Colloms, personal communication; M. Burke, personal communication). Recombination at *psi* shows resolution selectivity, therefore it seemed likely that another protein might be required to fulfil the role of ArgR, since, as discussed in Chapter 3, it is believed that sites that show resolution selectivity require accessory factors. However, since the role of PepA during recombination is not known, it is not clear whether or not it could act alone as an accessory factor.

By creating deletions into the accessory sequences, the minimum site required for Xer-mediated recombination was defined, showing that accessory DNA was essential for recombination at *psi* (M. Burke, personal communication). In Figure 5.1, the sequence of the 363 bp *psi* fragment is shown, with the core recombination site, and the deletion end-points

Figure 5.1: The *psi* site from plasmid pSC101
This Figure shows the sequence of the *psi* site cloned into pSDC133, p741 and p733 (Figure 5.1a), the putative protein binding motif in p741 (Figure 5.1b) and an alignment of *cer* with *psi* (Figure 5.1c, overleaf).

*Dde*I

[CTGAGACAACTTGTTACAGCTCAACAGTCACACATAGACAGCCTGAAACAGGCGATGCT

GCTTATCGAATCAAAGCTGCCGACAACACGGGAGCCAGTGACGCCTCCCGTGGGGAAAAA

ATCATG]GCAATTCTGGAAGAAATAGCGCTTTCAGCCGGCAAACCTGAAGCCGGATCTGCG

ATTCTGATAACAAACTAGCAACACCAGAAGAGCCCGTTTGCGGGCAGCAAAACCCGTACT

TTTGGACGTTCCGGCGGTTTTTTGTGGCGAGTGGTGTTCGG**GCGGTGCGCGCAAG**

ATCCATTATGTTAAACGGGCGAGTTTACATCTCAAAACCGCCCGCTTAACACCAT

CAGAAATCCTCAG

*Dde*I

Figure 5.1a: The *psi* site from pSC101
The 363 bp *Dde*I fragment that is cloned into pSDC133 is shown above (top strand only). The entire functional *psi* site is within this sequence; the putative core recombination site is shown in large, bold font. The region that was deleted in plasmid p741 is shown within large brackets, to indicate that all of the sequences required for recombination are outwith this bracketed region. The end-point of the deleted site found in p733 is shown as an underlined *Nae*I site. The putative protein binding site is shown in bold font, and below: this lies just within the sequences that are necessary for site-specific recombination, as delimited by the deletions in p741 and p733, and is centred at 155 bp upstream of the centre of the core recombination site.

Figure 5.1b: Putative protein binding site from *psi*
CTGGAAGAAAT | AGCGCTTTCAG
The sequence showing partial dyad symmetry within the deletion end-point is shown here, with a vertical line to indicate the centre of symmetry.

```
TTACAGCTCAACAGTCACACATAGACAGCCTGAAACAGGCGATGCTGCTT  psi
                          *** ** *       * *
                          AGCTTGCATGCCTGCAGGTCCGGA  cer

ATCGAATCAAAGCTGCCGACAACACGGGAGCCAGTGACGCCTCCCGTGGG  psi
*    *  ** **   *   *   *      * * ** *          *
AATACAGGAACGCACGCTGGATGGCCCTTCGCTGGGATGGTGAAACCATG  cer

GAAAAAATCATGGCAATTCTGGAAGAAATAGCGCTTTCAGCCGGCAAACC  psi
 ***         ** *     ***     *   * *   * *      *
AAAATGGCAGCTTCAGTGGATTAAGTGGGGGTAATGTGGCCTGTACCCTC  cer

TGAAGCCGGATCTGCGATTCTGATAACAAACTAGCAACACCAGAAGAGGC  psi
**    *  *  *     *  * * * ** ** **    *    * ***
TGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGCGCGATCGCGGC  cer

CCGTTTGCGGGCAGCAAAACCCGTACTTTTGGACGTTCCGGCGGTTTTTT  psi
   *** **** *           * ****    **          *  *
AGTTTTTCGGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATC  cer

GTGGCGAGTGGTGTTCGGGCGGTGCGCGCAA   GATCCATTATGTTAAAC  psi
* *  **   * ***   ******** ***   *   ****** ****
GCGCTGAACCGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAAT  cer
```

Figure 5.1c: Alignment of *cer* with *psi*

The *psi* sequence from pSC101 is aligned here with *cer*. Sequence matches are indicated by asterisks. The core recombination sites are highlighted in bold; a 2 bp gap has been inserted in the central region of *psi* to give the best sequence match across the right and left arms. The Arg-box of *cer* is shown underlined and is centred at 127 bp upstream of the core recombination site, while the putative protein binding site from *psi* is also underlined and is 159 bp upstream of the centre of the core recombination site.

delimited. The smallest site to act as a substrate for XerC/D-mediated recombination is in plasmid p741, and has 125 bp deleted from the right end of the site. A deleted site, in plasmid p733, that is ~25 bp smaller fails to recombine. Within this 25 bp region, is a 22 bp sequence showing partial dyad symmetry (M. Burke & S. Bell, personal communication). We therefore hypothesised that this sequence might be the recognition motif for a DNA-binding protein that might be replacing ArgR during Xer-mediated recombination at *psi*. This putative protein-binding site is indicated in Figure 5.1. An alignment of *cer* and *psi* is also given in this Figure 5.1 to compare the position of this putative protein binding site with the Arg-box of *cer*.

This Chapter describes a mutagenesis of the *E. coli* chromosome that had the objective of preventing recombination at *psi* by mutating the gene encoding the putative protein binding to this sequence. The identification of such an accessory factor for Xer-mediated recombination at *psi* might provide a valuable insight into the role of ArgR in recombination at *cer*, as well as being of interest in its own right. During Xer recombination at *cer*, ArgR is believed to function by binding to the Arg boxes in the two recombining *cer* sites, using protein-protein interactions to bring the sites together, thereby aiding the formation of a synaptic complex (Sherratt *et al.*, 1993; Sherratt, 1993). This theory was proposed because of the observation that ArgR is a hexameric protein in solution (Lim *et al.*, 1987), and is therefore expected to be able to bind up to three Arg-boxes, bringing them into proximity by protein-DNA and protein-protein interactions. If this is true, by extrapolation, a protein involved similarly in recombination at *psi* would be expected to be at least tetrameric, so that it could bring two sites together. This assumption is based on the following reasoning. The putative binding site in *psi* shows partial dyad symmetry, suggesting that a protein dimer might be required to bind each 22 bp motif within a single *psi* site; if such a protein functions to bring two *psi* sites together by protein-protein interactions, by this reasoning the native protein must be able to form at least tetrameric contacts. By identifying the protein involved, it was hoped that it would be possible to determine whether or not this protein can make higher-order interactions which might support or refute the hypothesis described above for the role of ArgR during recombination at *cer*.

Xer recombination *in vitro*

A mutagenesis of the *E. coli* chromosome was also considered worthwhile due to the current inability to reproduce the complete Xer recombination reaction *in vitro*. Several site-specific recombination systems have been reproduced *in vitro* using purified DNA and protein components in a simple buffer (Landy, 1989; Hoess & Abremski, 1985; Hatfull & Grindley, 1988; Stark *et al.*, 1989a); for example, adding purified resolvase to a supercoiled plasmid containing two *res* sites in direct repeat forms two interlinked product circles (Stark *et al.*, 1989a).

At present it is possible to reproduce a single strand-exchange between *cer* sites *in vitro*, by adding ArgR, PepA, XerC and XerD to a supercoiled plasmid containing two directly-repeated sites, in a simple buffer system (*in vitro* reaction conditions are described in Chapter 2). Such reactions generate Holliday junctions in sufficient quantity to be detected by ethidium bromide staining of agarose gels (an example of such a gel is shown in Chapter 4, Figure 4.6), but two product circles have not yet been recreated *in vitro*. It is possible that the Xer system only catalyses a single strand-exchange, and that the Holliday junctions formed are resolved by another cellular Holliday-junction resolving enzyme, such as RuvC or RecG (Connolly *et al.*, 1991; Iwasaki *et al.*, 1991; Lloyd, 1991). However, an equally likely possibility is that we are currently lacking another protein component of the system, or that the existing conditions are inappropriate. A second major aim of the chromosomal mutagenesis described in this Chapter was to attempt to isolate additional Xer mutants that might allow a complete Xer resolution reaction to be recreated *in vitro*.

Aim

The experiments described in this Chapter were carried out with the objectives of mutagenising the *E. coli* chromosome with Tn*5*, then screening for mutants failing to perform site-specific recombination at *psi* sites. Complementation assays would then be carried out *in vivo* to test whether putative mutants were affecting previously characterised *xer* genes. If any mutations appeared to be within novel *xer* genes, the ability of the putative mutant to carry out site-specific recombination at *cer* and *dif* would be tested. Finally, an attempt to characterise the mutant gene would be made, by identifying the DNA sequences flanking the Tn*5* insertion, after verifying that the mutation was as a result of disruption of a gene by Tn*5* .

Results

The advantages of using Tn*5* as a mutagen

The rate of mutation in bacteria can be increased by various methods, including irradiation with UV light, treatment with chemical mutagens or by insertional mutagenesis with transposons such as Tn*5*. Mutagenesis with the composite transposon Tn*5* has a number of advantages over other methods. Firstly, mutants obtained by insertion of Tn*5* into the chromosome can be selected by virtue of the kanamycin resistance conferred by the neomycin phosphotransferase II gene encoded by the transposon. Secondly, Tn*5* transposes with very little sequence specificity, such that mutation of the chromosome by insertion of Tn*5* can be considered random, occurring at many positions within many genes on the *E. coli* chromosome (Berg *et al.*, 1980; Shaw & Berg, 1979). A third advantage is that a Tn*5* insertion can be used to map the mutated gene due to its expected close physical proximity to the mutant allele; and the mutation can be transduced between strains to construct mutant strains in different genetic backgrounds by selection of kanamycin resistance. Finally, insertion of Tn*5* into a gene is likely to result in the total inactivation of that gene, and, if the gene is in a multi-cistronic unit, the downstream genes will also be inactivated by the insertion. These properties make Tn*5* a suitable mutagen to produce and select a large pool of mutations in the *E. coli* chromosome and to aid in their subsequent characterisation. (The properties of Tn*5* are reviewed in Berg & Berg, 1983).

NK467 as a mutagenic agent

Previous chromosomal mutageneses performed by researchers in this laboratory have successfully utilised the mutagenic λTn*5*, NK467 (Stirling, 1987; Stirling *et al.*, 1988; Colloms, 1990; Colloms *et al.*, 1990). This is a suicide λ phage which acts as a vector for Tn*5* delivery into the bacterial cell (de Bruijn & Lupski, 1984). This λ phage has a deletion removing *attP*, *xis* and *int* so that it does not integrate into the *E. coli* chromosome by site-specific recombination at *attB*. In addition, the *O* and *P* genes, which are essential for phage replication, carry amber mutations, so that the phage cannot replicate in a *sup*^{o} host; for this reason, the *E. coli* strain used for mutagenesis was the DS941 *sup*^{o} derivative, DS953, with which the majority of Km^{r} cells arising after mutagenesis with NK467 are as a result of transposition of Tn*5* into the chromosome.

Construction of vectors to select for chromosomal mutants failing to recombine *psi* sites

Plasmid p752 was constructed by Mary Burke as a reporter plasmid to detect site-specific recombination between directly-repeated *psi* sites. A map of this plasmid, and its derivation, is shown in Figure 5.2. Mutant strains that fail to resolve this plasmid can be detected by virtue of the kanamycin resistance marker, which is deleted in Xer^+ strains to give the product plasmid p751. In order to select for *xer* mutants derived from a Tn*5* mutagenesis using NK467, vectors were required with alternative deletable markers, since the resistance encoded by NK467 is also for kanamycin. The marker chosen for this purpose was a gentamycin resistance cassette. Two plasmids with this deletable resistance marker were constructed, with different origins of replication, pJEN81 and pJEN96. The plasmids, and their method of construction, are shown in Figures 5.3 and 5.4.

Pilot mutageneses made use of pJEN81. However, it was found that selection using this plasmid did not give clean results. After mutagenesis with Tn*5*, the Km^r cells were transformed with pJEN81, Ap^rKm^r colonies were selected, and were replica-plated onto plates containing Ap, Km and Gm. The majority of the colonies were found to be resistant to gentamycin, however when their plasmid content was analysed on single colony gels, the plasmid was found to be mainly resolved. This was taken to indicate either that recombination between the *psi* sites in pJEN81 might be at such a low rate that unresolved plasmid persisted in the cell for long enough to allow the gentamycin resistance phenotype to be expressed, or perhaps that the two resolution products are able to recombine together sufficiently well to allow a small number of gentamycin resistant plasmids to be present in the cell. If the latter is true, it might be through homologous recombination, since the *psi* sites in pJEN81, at 363 bp, are large enough to permit homologous recombination and being a pUC18-based plasmid, the number of copies of each product within the cell would be expected to be high. For whatever reason, selection with pJEN81 was not clean enough to allow mutants to be isolated through their gentamycin resistance.

It was because of the inability to select mutants using pJEN81 that pJEN96 was constructed. There were three reasons for making this λ*dv*-based plasmid. Firstly, the copy number of this plasmid is much lower than pJEN81, so any problems resulting from homologous recombination between resolution products should be minimised by reducing the number of sites available to recombine together. Secondly, if the selection difficulty of pJEN81

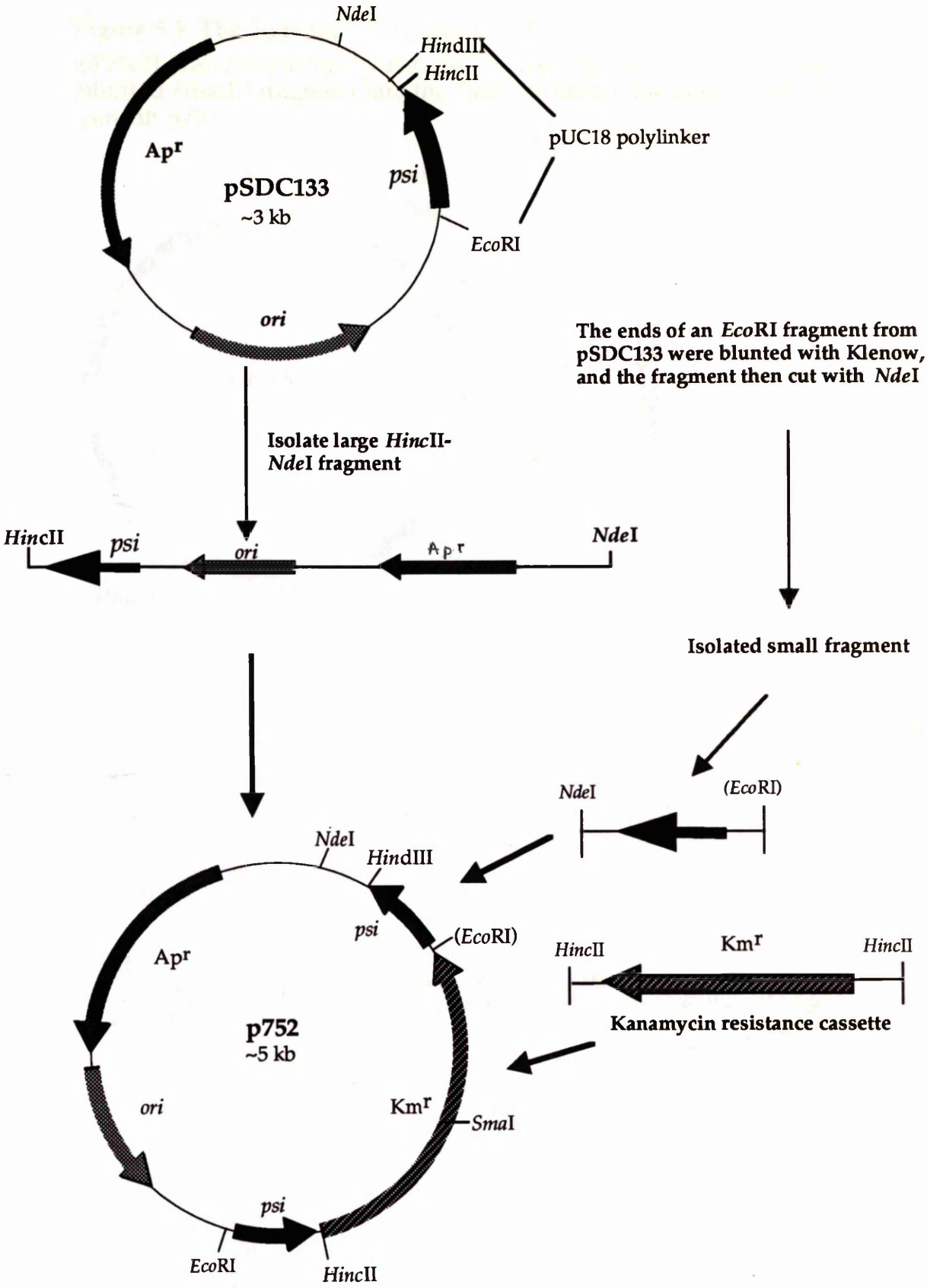


Figure 5.2: Derivation of the *psi* reporter plasmid p752 from pSDC133. At the top of this page is a map of pSDC133, showing the single *psi* site in the polylinker of pUC18. With the exception of the *Sma*I site at which the 363 bp *psi* fragment was cloned, all of the pUC18 restriction sites remain. Below pSDC133 is shown the cloning strategy used by M. Burke to produce p752, from which pJEN81 and pJEN96 (shown in Figures 5.3 & 5.4) were derived. In all three Figures, only the relevant restriction sites used in cloning are shown, and where sites are removed by cloning, they are shown in brackets.

Figure 5.3: The 2-*psi* reporter plasmid pJEN81.

pJEN81 was constructed by sub-cloning a gentamycin resistance cassette as a blunted *Hin*dIII fragment into the *Sma*I site within the kanamycin resistance gene in p752.

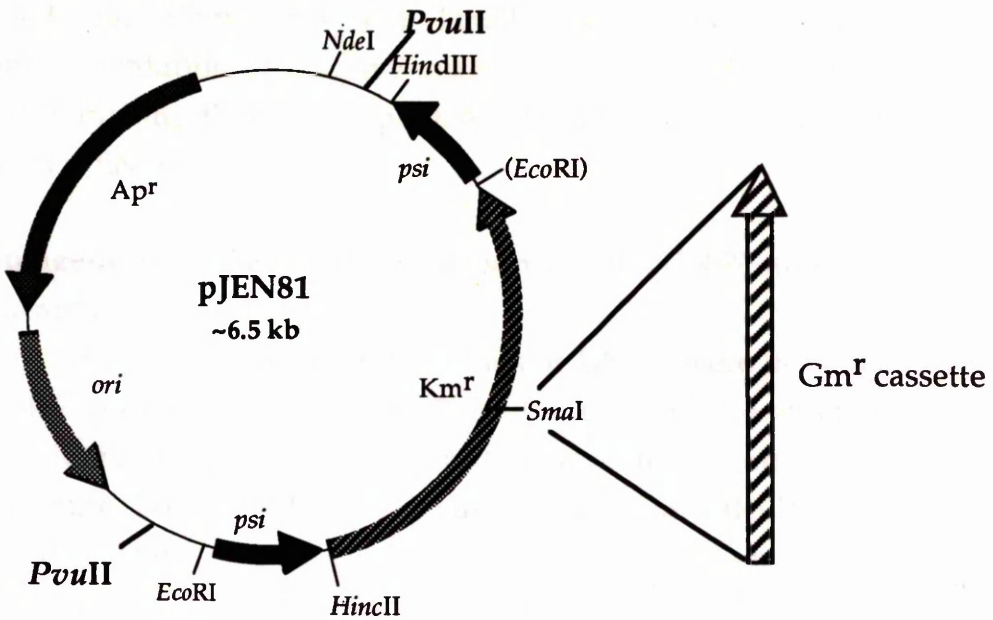
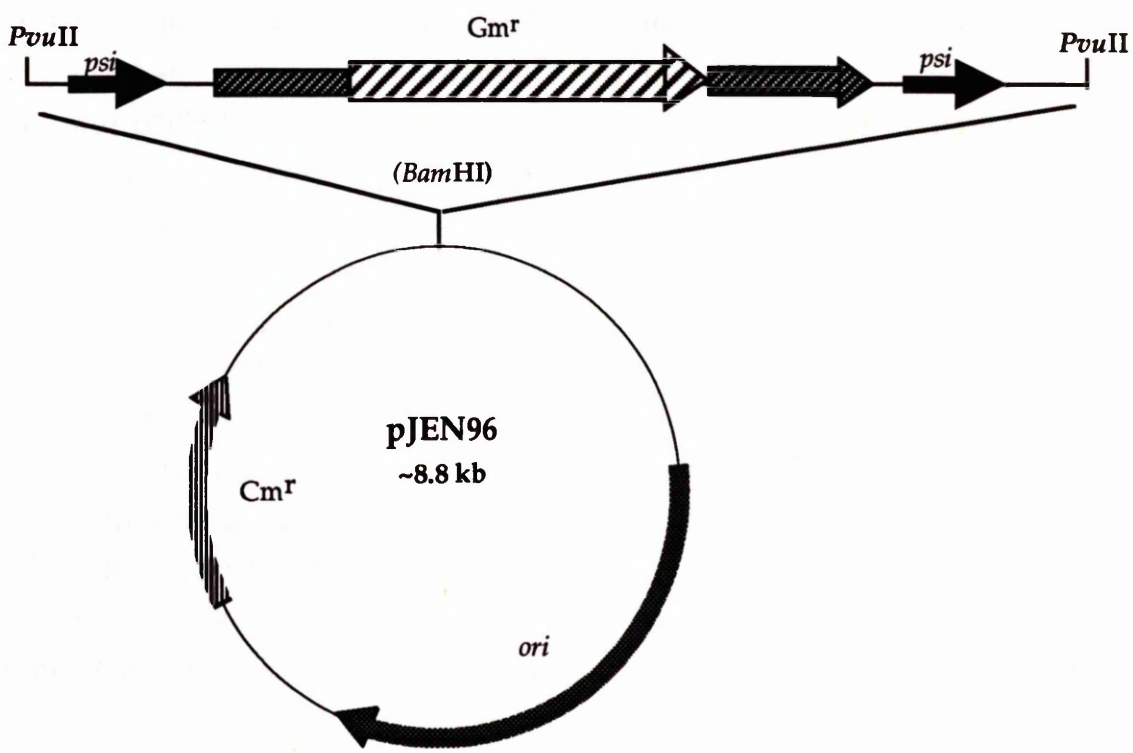


Figure 5.4: Derivation of pJEN96 from pJEN81.

The large *Pvu*II fragment from pJEN81 was sub-cloned into pCB104 at the *Bam*HI site, which had first been blunted with Klenow. This resulted in a 2-*psi* reporter plasmid with low-copy number, as this vector has a λ-*dv* origin of replication.



was due to residual unresolved plasmid remaining in the cell as a result of a low rate of recombination, the reduction in the number of plasmid copies by using pJEN96 should allow better resolution by not overwhelming the site-specific recombination system. Finally, the plasmids selected for subsequent complementation of putative mutants, pMAY5 and pCS112, are both pMB1-based, making them incompatible with pJEN81, but compatible with the λ*dv*-based pJEN96.

Mutagenesis of the *E. coli* chromosome with NK467, and selection of putative mutants

Four 2.5 ml overnight cultures of DS953 were mutagenised with NK467 as described in Chapter 2. Half of each culture was plated on LB agar containing 50 $\mu g.ml^{-1}$ kanamycin to determine the number of mutants per ml of culture. For each of the four cultures, this was calculated to be approximately 2000 mutants. ml^{-1} (these should be mainly independent mutants since NK467 was used to infect cultures that were already at stationary phase, and infection time was kept to two hours). The remainder of the cultures was used to inoculate 20 ml cultures in 2 x YT to prepare competent cells. Competent cells were made by growing cultures to A_{600} = 0.6, which were then harvested and resuspended in 5 ml ice-cold 50 mM $CaCl_2$, then left on ice for 60 minutes. The cells were then harvested, resuspended in 2.5 ml 50 mM $CaCl_2$ and kept on ice for a further 30 minutes, before the addition of approximately 0.5 µg pJEN96 to 250 µl aliquots of cells. A two minute heat-shock at 42°C was followed by the addition of 1 ml 2 x YT, and 60 minutes expression time at 37°C. Transformants were plated on LB agar containing kanamycin and chloramphenicol, to select for mutants (Km^r) containing pJEN96 (Cm^r). After overnight growth, each plate contained approximately 1000 colonies.

The transformants were replica-plated onto fresh $LBKm_{50}Cm_{12.5}$ plates and left to grow overnight. This was repeated three times, before replica-plating onto $LBKm_{50}Cm_{12.5}Gm_5$ and $LBKm_{50}Cm_{12.5}$. After overnight growth, the number of colonies on the $LBKm_{50}Cm_{12.5}Gm_5$ plates was found to be substantially reduced compared with the plates with no gentamycin. Several hundred colonies were able to grow on each plate in the presence of gentamycin, which appeared to be a relatively high number compared with previous mutageneses to select for *xer* mutants (however the protocol described here was different from those used previously: Stirling, 1987; Stirling *et al*, 1988b; Colloms, 1990; Colloms *et al.*, 1990). Fifty colonies from

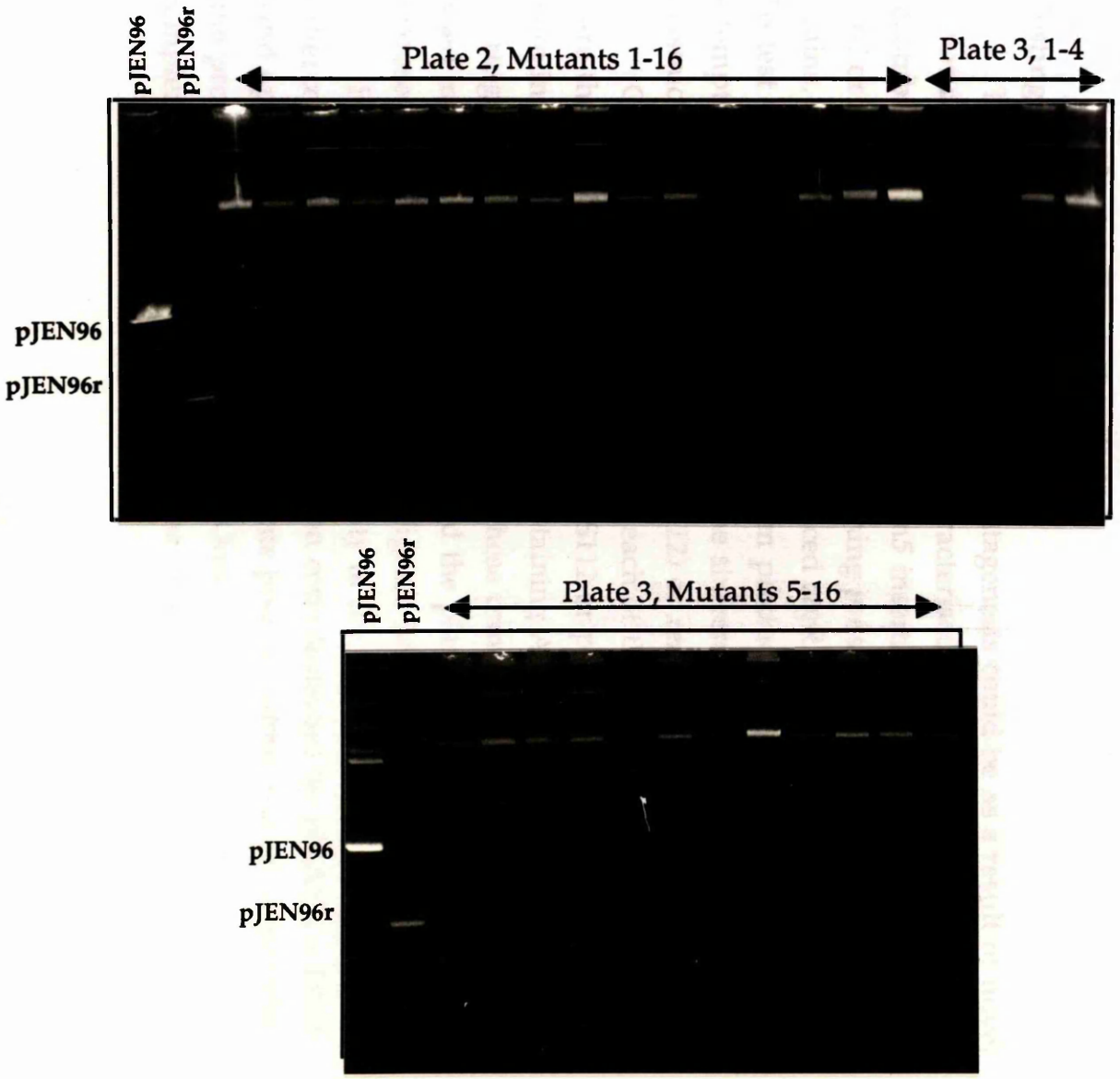


Figure 5.5: Putative Xer mutants failing to resolve pJEN96.
Sixteen colonies from each of two of the four plates containing putative mutants were selected at random, and their plasmid content analysed on single colony agarose gels. These 1% gels are shown above. pJEN96 and pJEN96r markers are shown, to indicate the resolution status of the plasmid in each of these 32 potential mutants.

each of the four $LBKm_{50}Cm_{12.5}Gm_5$ plates were replica-streaked onto $LBKm_{50}Cm_{12.5}Gm_5$ and $LBKm_{50}Cm_{12.5}$ plates, and all grew equally well on both plates, confirming that these colonies were resistant to gentamycin. This suggested that they were *xer* mutants which were unable to delete the gentamycin resistance cassette of pJEN96.

To confirm that these colonies contained unresolved pJEN96, sixteen colonies from each of two plates (plates 2 & 3) were chosen at random and their plasmid content analysed on single colony gels, which are shown in Figure 5.5. One problem with using pJEN96 is clear from this gel: since the plasmid is λ-*dv* based, the amount of DNA recovered was very low, therefore it was difficult to assess by ethidium bromide staining whether a low proportion of resolution had occurred. However, all of the colonies tested appeared to contain unresolved pJEN96, indicating that they may be *xer* mutants. This further suggested that the other $Km^rCm^rGm^r$ colonies on plates 1-4 might also be chromosomal mutants unable to recombine at *psi* (not tested).

Testing for novel *xer* mutants

The mutants derived by this mutagenesis could be as a result of novel mutations in genes previously un-characterised as being involved in *xer* recombination, or alternatively, the Tn*5* insertions could be within *pepA*, *xerC* or *xerD*; because plasmids containing *psi* sites can recombine in *argR* strains, the possibility of having produced *argR* mutants could be excluded. To test whether the mutations had been produced in *pepA*, or *xerC* or *xerD*, attempts were made to complement the sixteen mutants from plate 2 with plasmids expressing either *pepA* (pCS112) or *xerC* and *xerD* (pMAY5).

Competent cells were made for each of the sixteen mutants, which were then transformed with either pCS112 or pMAY5, and transformants were inoculated into 2.5 ml 2 x YT containing $Ap_{100}Cm_{12.5}Km_{50}$ to make overnight cultures. 100 µl aliquots of these transformants were harvested and lysed in 100 µl single colony buffer, and the plasmid content analysed on 1% agarose gels. These gels are shown in Figures 5.6 and 5.7.

Figure 5.6 shows that the majority of the mutants being tested are either *xerC* or *xerD* mutants, since when complemented by pMAY5, a DNA band appeared that migrated to the same position within the gel as pJEN96r (the product of resolution of pJEN96). Only mutant 9 failed to be complemented by pMAY5, as this is the only mutant which had no DNA

Figure 5.6

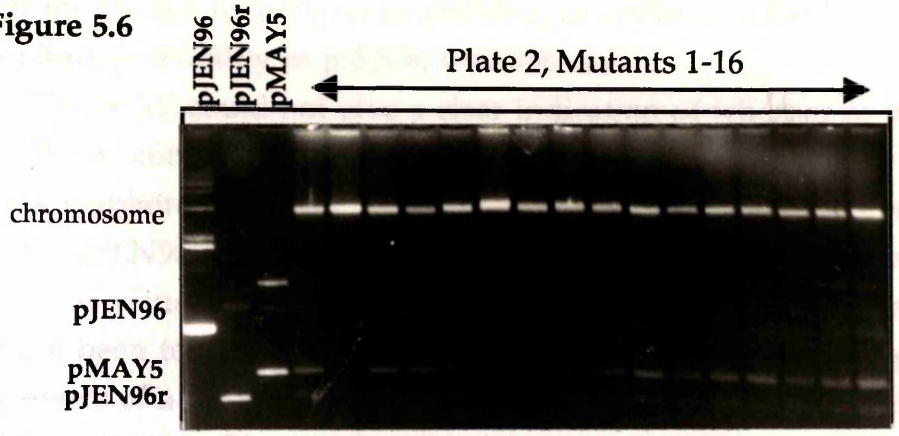


Figure 5.6: Complementation of putative mutants with XerC and/or XerD. An attempt was made to complement the sixteen mutants from plate two that had been demonstrated to fail to resolve pJEN96. Competent cells were prepared from the pJEN96 transformants (shown on Figure 5.5), which were then transformed with pMAY5. pMAY5 expresses *xerC* and *xerD* from a *lac* promoter, so that chromosomal *xerC* or *xerD* mutants can be complemented. The plasmid content of transformants was analysed on the single colony agarose gel shown above. pJEN96, pJEN96r and pMAY5 size markers are shown, to indicate whether the presence of pMAY5 within the cells allows resolution of pJEN96 by these sixteen mutants.

Figure 5.7

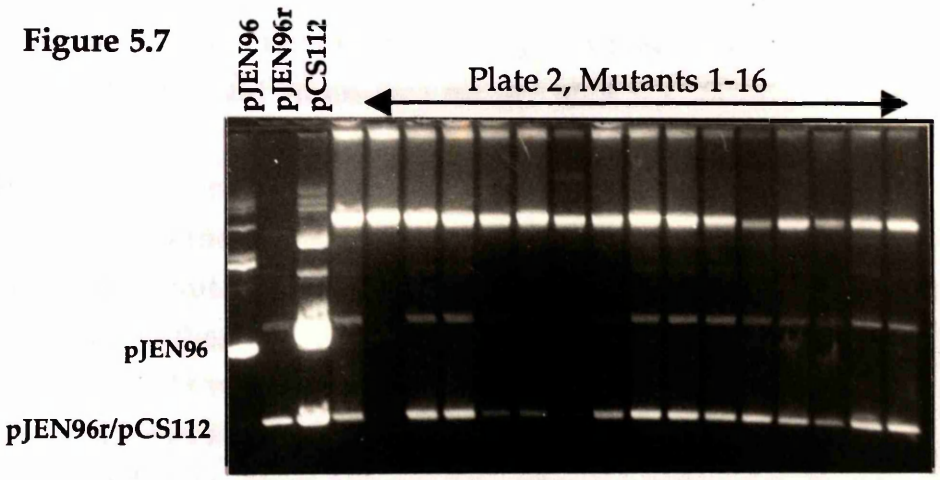


Figure 5.7: Complementation of putative mutants with PepA. Complementation of the same sixteen mutants as in Figure 5.6 was attempted with the *pepA* expression plasmid, pCS112. Competent cells of mutants 1-16 from plate 2 containing pJEN96 were transformed with pCS112. The plasmid content of transformants was analysed on the single colony agarose gel shown above. Markers of pJEN96, pJEN96r and pCS112 are indicated, however analysis is complicated as pCS112 and pJEN96r co-migrate in the gel.

band of the correct mobility to be pJEN96r; in addition, a band of the same electrophoretic mobility as pJEN96 was detected.

Figure 5.7 could not give a clear indication of whether or not mutant 9 was complemented by PepA, since pCS112 and pJEN96r have the same electrophoretic mobility, making it impossible to decide if resolution of pJEN96 to pJEN96r had occurred. However, a faint band corresponding to pJEN96 was visualised in this track indicating that if resolution had occurred, it had not been to completion. Further confusion arose from this gel due to the existence of a DNA species that could correspond with a dimer of either pCS112 or pJEN96r. This species did not appear in all tracks; in lane 10, with mutant 7, this band was absent, and, in addition, a band corresponding in size to pJEN96 could be detected. This might indicate that this was dimeric pJEN96r, and if so, that all of the mutants except mutant 7 were complemented by pCS112, and are therefore *pepA* mutants. This theory is unattractive, since, with the exception of mutant 9, they were all additionally complemented by pMAY5, and there is no simple explanation for how this could occur. In addition, no band of similar mobility could be detected in Figure 5.6 for those mutants that were complemented by pMAY5, indicating that it was unlikely to be a dimer of pJEN96r. These observations led to the conclusion that this band was more likely to be a dimer of pCS112 than a dimer of pJEN96r. Consequently, mutant 9 was the only candidate mutant pursued, as this was the one not clearly complemented either by PepA, XerC or XerD. This mutant strain was named JR29.

JR29 is mutant for recombination at *cer* and *psi*, but not at *dif*

In vivo recombination assays were used to confirm that JR29 was a chromosomal mutant unable to carry out site-specific recombination at *psi*, and to test whether this was also true for *cer*. DS941, DS956, DS957, DS981, DS9008 and JR29 were transformed with p752 and with pKS455, to assay the ability of JR29 to resolve these two plasmids compared with previously characterised *xer* mutant strains. The plasmid content of the resulting transformants was analysed on a single colony gel, which is shown in Figure 5.8.

Figure 5.8a shows that p752 was resolved in DS941 and DS956, but not in DS957, DS981, DS9008 or JR29. This confirms that JR29 is a mutant for recombination at *psi*, and illustrates the independence of this site from ArgR, since in the *argR* mutant strain, DS956, p752 was resolved to completion. Figure 5.8b shows that the 2-*cer* reporter plasmid pKS455 was also unable to

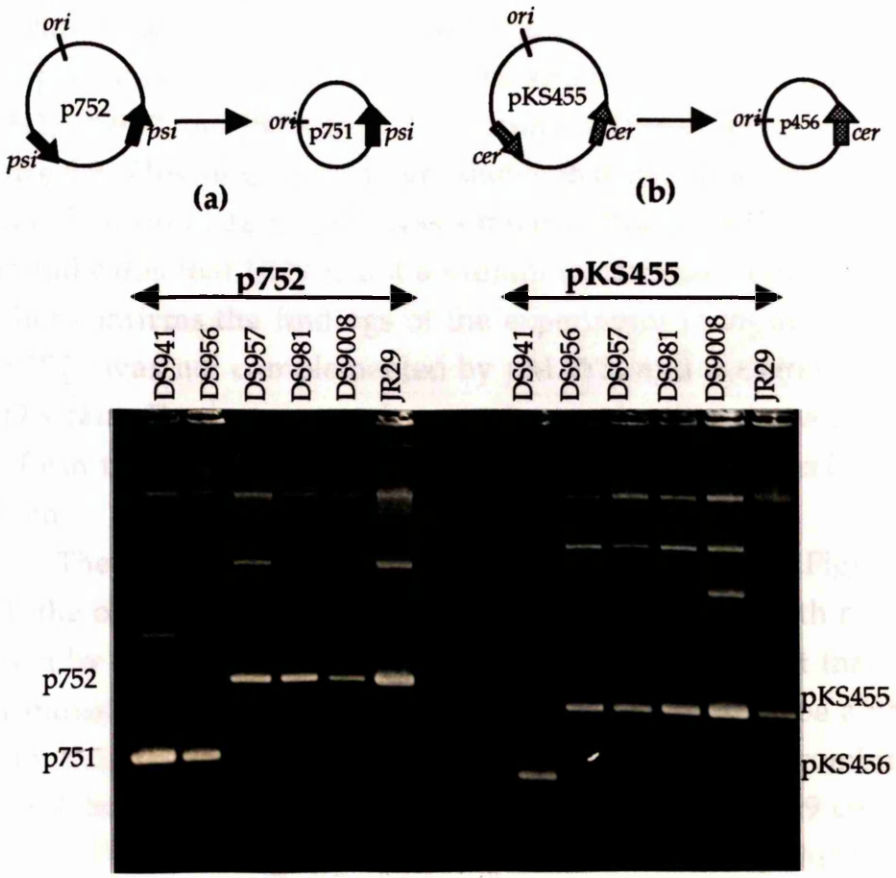


Figure 5.8: Assessing the ability of JR29 to recombine at *psi* and *cer*. The 2-*psi* reporter plasmid p752 (a) and the 2-*cer* reporter plasmid pKS455 (b) were transformed into DS941, and its *argR*, *pepA*, *xerC* and *xerD* derivatives, and into JR29. The ability of these six strains to resolve the two plasmids was analysed by assessing the plasmid content on a single colony agarose gel, which is shown here.

resolve in JR29, as a DNA band of the same size was recovered from the JR29 pKS455 transformant as from transformants of the other four *xer* mutant strains. Only DS941 was able to resolve pKS455 into its product, pKS456.

The results shown in Figure 5.8 indicated that the mutation in JR29 abolishes site-specific recombination at *cer* as well as at *psi*. A similar experiment was carried out to establish whether JR29 was also a mutant in site-specific recombination at *dif*. The strains assayed in this experiment were DS941, DS957 and JR29 only. The results of this experiment are shown in Figure 5.9. This single colony gel shows that the extent of site-specific recombination at *dif* in JR29 was similar to that in both DS941 and DS957. This indicates that JR29 is not a mutant in terms of recombination at *dif*, and further confirms the findings of the experiment in Figure 5.6 which showed that JR29 was not complemented by pMAY5, and is therefore not an *xerC* or *xerD* strain. Furthermore, microscopic examination showed that JR29 does not form the filamentous cells characteristic of *xerC* or *xerD* mutants (data not shown).

The uncertainty in the interpretation of the gel in Figure 5.7, coupled with the observation that JR29 is also a mutant strain with respect to *cer* (as shown by the failure of pKS455 to resolve in JR29) meant that the data could be rationalised in two possible ways. Firstly, JR29 could be a *pepA* mutant strain: this would explain the fact that it is mutant in recombination at *cer* and *psi*, but not at *dif*. Alternatively, the mutation in JR29 could be a novel *xer* mutation affecting both *cer* and *psi*; if so, it is most likely to be a mutation abolishing another accessory factor, since recombination at *dif* was unaffected in JR29.

P1 transduction of the *xer* mutation from JR29 into DS941

The simplest method to characterise the gene which has been mutated in JR29 is to exploit the Tn*5* insertion that created the mutation, and identify the sequences flanking Tn*5* in the JR29 chromosome. To do this, it was essential to ensure that the Tn*5* insertion was responsible for the *xer* phenotype being characterised. This could be verified by P1 transduction of the Tn*5* into another strain; if the insertion is responsible for the mutant phenotype, the *xer* mutation should co-transduce with the kanamycin resistance marker of the Tn*5* insertion. Consequently, the Tn*5* insertion from JR29 was P1 transduced into DS941. One kanamycin resistant colony was obtained from this P1 transduction, and this was named JR30.

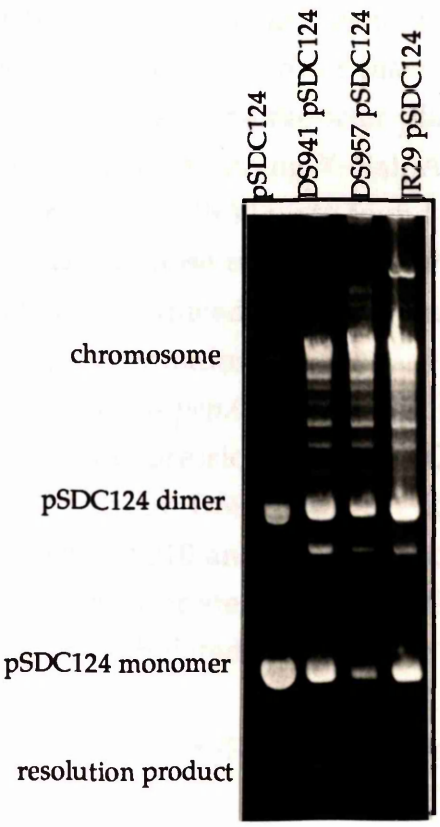


Figure 5.9: The effect of JR29 on recombination at *dif in vivo*. The 2-*dif* reporter plasmid pSDC124 was transformed into DS941, DS957 and JR29. DS957 was included because the possibility of JR29 being a *pepA* mutant had not been eliminated. Intermolecular and intramolecular recombination at *dif* was assessed by analysing the plasmid content of pooled transformants on a single colony agarose gel, shown here. pSDC124 is included as a marker; comparison of this marker track with the other three lanes allows comparison of the three strains with respect to their ability to carry out Xer-mediated recombination at *dif*.

To ascertain the *xer* phenotype of JR30, JR30 cells were transformed with the 2-*cer* reporter plasmid pCS210. This plasmid was selected for further analysis of JR29 and JR30 since it had already been established that the mutation affects *cer*, and this plasmid gives a colour assay for Xer-mediated recombination. pCS210 contains a *lacZ* gene, flanked by two *cer* sites in direct repeat, so that resolution of the reporter plasmid results in white colonies (in derivatives of DS941), while if the reporter plasmid is unresolved, the colonies are blue on plates containing X-Gal. A map of pCS210 is shown in Figure 5.10. Transformants of JR30 were found to be blue on X-Gal plates, as were those of JR29, while those of DS941 were white. This implied that the *xer* phenotype had co-transduced with the Km^r allele.

To verify that the mutation in JR29 and in JR30 which resulted in the *xer* phenotype is not within *pepA*, an attempt was made to complement the mutation with the *pepA* expression plasmid pCS112 using pCS210 as a reporter plasmid. Competent DS941, DS957, JR29 and JR30 were transformed with pCS210 alone, or pCS210 and pCS112, and the transformants selected on plates containing the appropriate antibiotics, X-Gal and IPTG. The results of the transformations are tabulated in Table 5.1:

Table 5.1: Complementation of JR29 and JR30 by pCS112

Strain	**Plasmid(s)**	**Colour on X-Gal plates**
DS941	pCS210	white
DS941	pCS210 pCS112	white
DS957	pCS210	blue
DS957	pCS210 pCS112	white, occasional blue
JR29	pCS210	blue
JR29	pCS210 pCS112	white, occasional blue
JR30	pCS210	blue
JR30	pCS210 pCS112	blue

Several conclusions were derived from these results. Firstly, the *xer* mutation in JR29 appeared to be complemented by pCS112 to a similar extent to DS957, which implies that JR29 is a *pepA* mutant. Secondly, despite the fact that JR30 appears to be an *xer* mutant, it was not complemented by pCS112, and so it seemed unlikely that JR30 is a *pepA* mutant. This result was confusing, since the *xer* allele in JR30 was derived by P1 transduction from JR29. To ascertain whether the plasmid had resolved in these four strains, the plasmid content

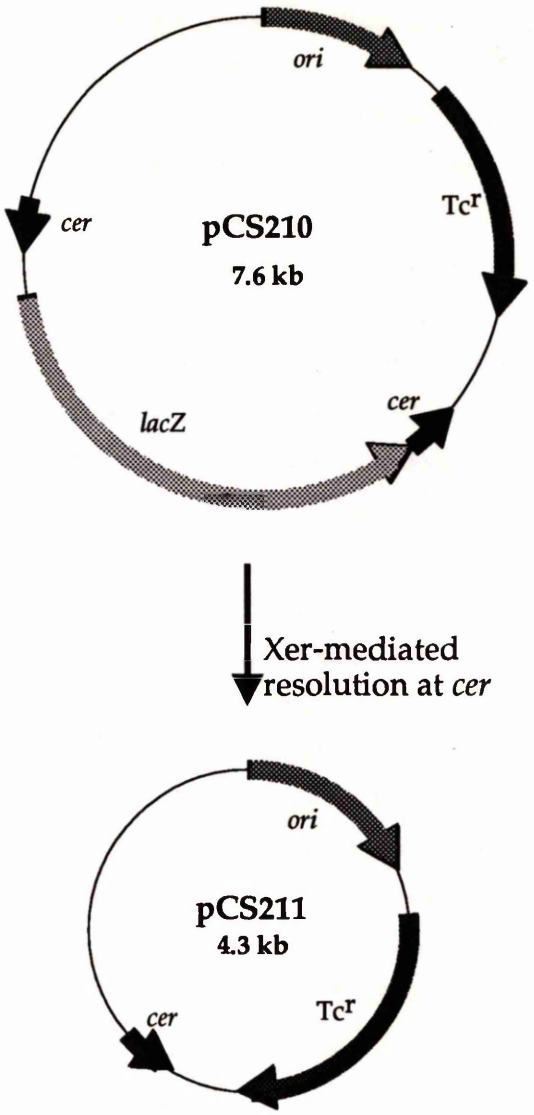


Figure 5.10: Map of pCS210.
This plasmid was constructed by C. Stirling, to detect recombination at *cer* by the blue-white colour assay which results from deletion of the *lacZ* gene between the two *cer* sites in an Xer$^+$ host. Only one product of resolution, pCS211, possesses an origin of replication, which derives from pACYC184. The other resolution product, containing the *lacZ* gene, is not maintained in the cell as it has no origin and is therefore not shown here. *xer* derivatives of DS941 produce blue colonies on agar containing X-Gal when transformed with pCS210, while their wild-type counterparts produce white colonies.

of a population of transformants was analysed, by recovering a large number of colonies from the transformation plate, then lysing the cells in single colony final sample buffer. The plasmids were then electrophoresed on a 1% agarose gel, which is shown in Figure 5.11.

Figure 5.11 shows that in DS941, as expected, no DNA band of the correct size to be pCS210 could be detected, either with or without pCS112 (lanes 4 & 5). However, in the presence of pCS112, it was impossible to detect the pCS211 band, as this co-migrated with pCS112. With both DS957 (lane 6) and JR29 (lane 8), a band of the correct mobility to be pCS210 could be detected in the absence of pCS112, indicating that no resolution had occurred. However, when these strains are co-transformed with pCS210 and pCS112, no band of the correct size to be pCS210 was detected (lanes 7 & 9), implying that complementation of the *xer* mutation had occurred. These results are consistent with the X-Gal plates, which showed that these co-transformants were mainly white, while the single pCS210 transformants were blue. By contrast, the plasmid DNA recovered from JR30 pCS210 (lane 10) appeared to be totally resolved to pCS211, even in the absence of pCS112. This was inconsistent with the X-Gal plates, on which the colonies were blue.

To further investigate the extent of resolution of pCS210 in each of these four strains with and without pCS112, mini-prep plasmid DNA was prepared from a population of cells recovered from the transformation plates, and this was used to transform the *xerD* mutant strain DS9008, to gain an estimate of the proportion of pCS210 which had been resolved in populations of DS941, DS957, JR29 and JR30 cells. The transformants were plated on X-Gal plates containing tetracycline to select for the resistance marker of pCS210. The percentage of blue transformants was scored, and is tabulated in Table 5.2:

Table 5.2: Resolution of pCS210 in JR29 and JR30 in presence and absence of pCS112

DS9008 transformed with DNA isolated from..	**% blue colonies**
DS941 pCS210	0
DS941 pCS210 pCS112	0
DS957 pCS210	100
DS957 pCS210 pCS112	0.5
JR29 pCS210	100
JR29 pCS210 pCS112	17.5

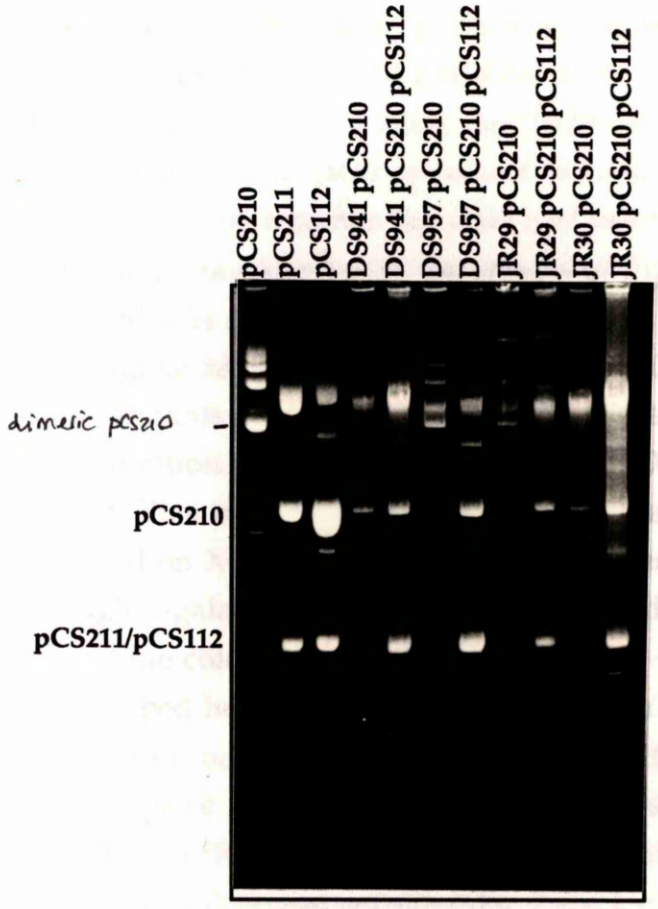


Figure 5.11: Complementation of recombination at pCS210 by pCS112. DS941, DS957, JR29 and JR30 competent cells were transformed with pCS210, or co-transformed with pCS210 and pCS112, and the plasmid content of a population of transformants was analysed by lysis in single colony final sample buffer followed by electrophoresis on a 1% agarose gel. Markers of pCS210, pCS211 and pCS112 are run on this gel, and their monomer positions indicated alongside the photograph. pCS211 and pCS112 co-migrate, complicating analysis; disappearance of pCS210 rather than appearance of pCS211 must be assessed.

On this gel, the pCS210 DNA preparation that was used as a marker on this gel, and was used for transformation was mainly in dimeric form. Consequently, the resolution status of this reporter plasmid was followed by the disappearance of dimeric pCS210 in preference to the monomeric species.

JR30 pCS210	0
JR30 pCS210 pCS112	0

These results indicated that JR29 was not totally complemented by pCS112, in contrast to DS957, since the number of blue colonies obtained from this transformation was much greater when the source of the mini-prep DNA was JR29 than when it was DS957. This infers that JR29 differs from DS957. The mini-prep DNA obtained from JR30 appeared to contain completely resolved pCS211 irrespective of whether the cells had been complemented with pCS112. While this is consistent with the results of Figure 5.11, which showed that the plasmid was completely resolved on a gel, it cannot be reconciled with the original results which showed that JR30 pCS210 and JR30 pCS210 pCS112 transformants were blue on X-Gal. One possible explanation is that JR30 contains a mutation which decreases the rate of Xer recombination, but does not abolish it. This might explain the blue colonies formed when the transformants are plated on X-Gal: if pCS210 is not resolved immediately after transformation, enough β-galactosidase might be expressed to allow the colonies to produce a blue colour on X-Gal.

The results described here suggest that the mutation that was P1 transduced into DS941 to produce JR30 is not the same as the mutation causing the *xer* phenotype of JR29. This unfortunately means that the insertion position of Tn*5* in JR29 could not be used to identify the gene that is mutated in JR29. One possibility, not excluded by these experiments, is that there are two Tn*5* insertions in JR29, since only one transductant was obtained. It might therefore be possible to determine whether another insertion exists in JR29, and, if so, this could be investigated further. To determine whether more than one Tn*5* insertion is responsible for the mutation in JR29, either the P1 transduction could be repeated to establish whether a different mutant from JR30 can also be obtained from JR29, or alternatively Southern hybridisation could be used to determine how many Tn*5* insertions are within the JR29 chromosome. However, the current results imply that the Xer mutation is likely to be affecting *pepA*, since JR29 can be complemented almost completely by the PepA over-expressing plasmid pCS112.

Assaying peptidase activity of JR29 and JR30

To further determine whether or not JR29 and JR30 are *pepA* mutants, the peptidase activity in cell extracts of those two strains was assayed *in vitro*

Strain	A_{400}.20µl cell extract^{-1}. 30 min^{-1}	A_{400}. µg total protein^{-1}. min^{-1}	units peptidase activity. mg total protein^{-1}. min^{-1}
DS941	0.025	5.34×10^{-5}	3.44
DS957	0.007	1.86×10^{-5}	1.20
JR29	0.002	1.44×10^{-5}	0.93
JR30	0.010	5.14×10^{-5}	3.32

Figure 5.12: Peptidase assays *in vitro*.
Peptidase levels in DS941, DS957, JR29 and JR30 were assayed *in vitro* using the substrate leucine-*p*-nitroanalide, which is cleaved to give the coloured compound *p*-nitroanaline by peptidases. Cell extracts were made from the four strains by growing 100 ml overnight cultures in LB broth, resuspending the cells in 1 ml 2 x assay buffer, and releasing cellular proteins by the addition of 50 µl 25 mg.ml^{-1} CTAB. Extracts were incubated with CTAB for 10 minutes at 37°C, then at 70°C for 10 minutes to denature other peptidases (PepA is not denatured at 70°C).The extracts were cleared by centrifugation, and 20 µl of each supernatent was assayed for peptidase activity. Assays were performed for 30 minutes at 37°C, as described in the text. The amount of *p*-nitroanaline released was assayed spectrophotometrically, as a change in A_{400}. To calculate peptidase activity, the total amount of protein in each extract was assayed using Bradford's reagent, and a standard curve (not shown); total protein. ml^{-1} extract was found to be 780 mg, 626 mg, 230 mg, 324 mg for DS941, DS957, JR29 and JR30 respectively. Peptidase units were calculated using the extinction co-efficient for *p*-nitroanaline, and are recorded in the final column of the table, where units of peptidase activity is expressed as µmoles *p*-nitroanaline produced per mg total protein per minute. This experiment was performed once only, therefore no attempt was made to determine whether the results are statistically significant.

using the substrate leucine-*p*-nitroanalide, which is cleaved by peptidases to release the coloured product *p*-nitroanaline which can be detected spectrophotometrically. 100 ml overnight cultures of DS941, DS957, JR29 and JR30 were grown in LB broth, and the cells harvested and resuspended in 1 ml 2 x peptidase assay buffer (see Chapter 2). The protein content of the cells was released by adding 50 µl 25 $mg.ml^{-1}$ CTAB, and incubating at 37°C for ten minutes. The extracts were then heated to 70°C for ten minutes to denature the other cellular peptidases. The cell extracts were cleared by centrifugation, and the supernatant was used as a protein extract. The protein content of each of the four extracts was calculated using Bradford assays (as described in Chapter 2). Peptidase assays were performed by adding 20 µl of each protein extract to a reaction mixture containing 1 mM leucine-*p*-nitroanalide in 1 x peptidase assay buffer, then incubating for 30 minutes at 37°C, and were stopped by the addition of EDTA to a final concentration of 10 mM. The release of *p*-nitroanaline was assayed spectrophotometrically by measuring A_{400}, and the units of peptidase activity (as expressed as µmoles *p*-nitroanaline produced per mg protein per minute) was calculated. The results of these assays are shown in Figure 5.12.

Figure 5.12 suggests that the peptidase activity of JR30 was similar to that of DS941, while JR29 was more similar to DS957. The peptidase activities of all four strains is very low, presumably because chromosomal levels of expression were being assayed. These assays were performed only once, therefore the results are not conclusive, but they can perhaps be taken as confirmation of previous results described in this Chapter: JR29 appears to have a mutation which could be at least partially complemented by PepA.

In addition to the *in vitro* peptidase assays using cell extracts, peptidase activity was assayed *in vivo* using minimal media plates containing the dipeptide *L*-valyl-leucine-amide (val-leu). This peptide is cleaved by PepA (a leucine aminopeptidase) to release valine, which is toxic to cells growing on minimal medium, by feedback inhibition of the genes involved in valine, leucine and isoleucine biosynthesis. DS941, DS957, JR29 and JR30 colonies were streaked onto a plate containing val-leu at a final concentration of 0.5 mM. After overnight incubation, DS941 and JR30 were unable to grow on this medium indicating that they were able to cleave the val-leu and are therefore not *pepA* mutants. By contrast, DS957 grew well on this medium, as a result of its *pepA* mutation. JR29 also grew on val-leu, further indicating that JR29 has a mutation impairing *pepA* activity. A photograph of this plate is shown in Figure 5.13. This assay was kindly performed by R. McCulloch.

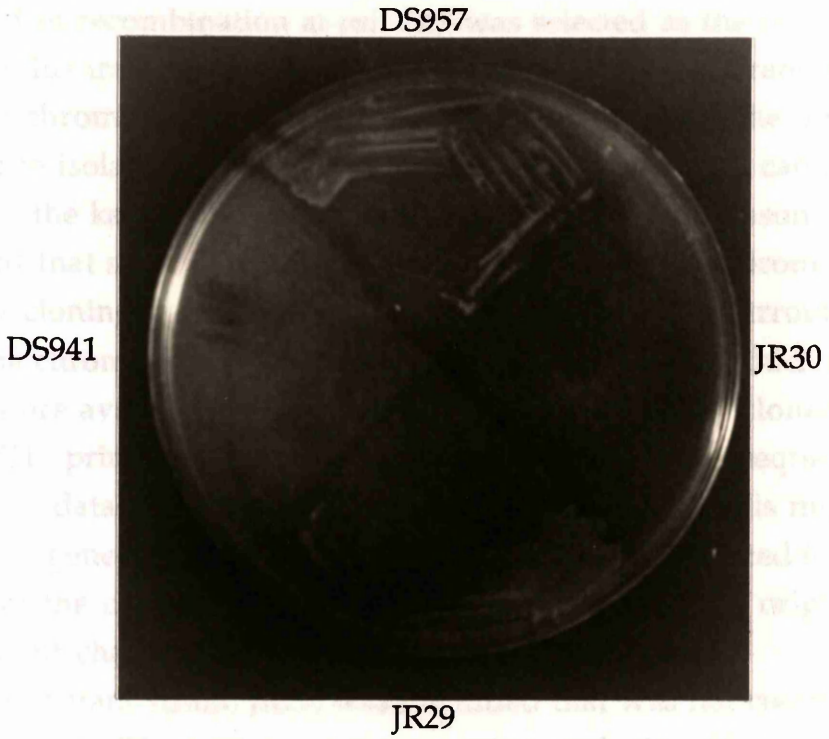


Figure 5.13: Peptidase assays *in vivo*.
Peptidase assays were carried out *in vivo* by Richard McCulloch for DS941, DS957, JR29 and JR30 using minimal media containing the dipeptide amide val-leu-NH_2. This peptide is cleaved by PepA to release valine, which is toxic to cells grown on minimal media. Consequently, *pepA* mutants should grow on plates containing this peptide, while *pepA*$^+$ cells will not. This photograph shows such a val-leu plate. DS941 and JR30 are unable to grow, while DS957 grows well, and JR29 is able to grow, but to a lesser extent.

Discussion

The results discussed in this Chapter describe an attempt to characterise further *xer* genes, but specifically aimed at identifying a putative factor involved in recombination at *psi*. Tn*5* was selected as the mutagenic agent because this transposon is believed to insert at relatively random positions into the chromosome, causing insertional inactivation. The inserted Tn*5* can be used to isolate the surrounding DNA sequences, which can be cloned by virtue of the kanamycin resistance encoded by the transposon. It was intended that some flanking sequence could be obtained from the inactivated gene by cloning the kanamycin resistance marker with surrounding DNA from the chromosome into a vector such as pUC18, for which sequencing primers are available commercially; sequencing into the cloned DNA using the pUC18 primers might be expected to yield sufficient sequence information to allow a data-base search of known *E. coli* sequences. This might enable the wild-type gene to be identified, or would at least be expected to pinpoint the region of the chromosome from which the mutated gene originated, allowing subsequent characterisation.

A mutant strain, JR29, was identified that was not complemented by XerC or XerD. The initial complementation analysis with the PepA expression plasmid pCS112 was complicated since this plasmid has the same electrophoretic mobility as the resolution product of the 2-*psi* reporter, pJEN96. However, on agarose gel electrophoresis, unresolved pJEN96 could be visualised, indicating that if complementation by PepA was occurring, it was not to completion. JR29 was also found to be a mutant for Xer recombination at *cer*, but not at *dif*. This observation could be taken to indicate that the mutation is not within the gene encoding the putative factor required for recombination at *psi* that carries out a similar role to ArgR during recombination at *cer*. However, it was considered possible that this mutation was within a previously unidentified *xer* gene encoding a further accessory factor required for recombination at *cer* and *psi*. In addition, the fact that *cer* was affected by this mutation as well as *psi* meant that additional reporter plasmids became available to analyse JR29 *in vivo*. To further characterise the mutation in JR29, the 2-*cer* reporter plasmid pCS210 was utilised, as this allows the resolution state of the plasmid to be detected by a blue/white colour assay.

To clone and sequence the mutation by exploiting the position of Tn*5* insertion in JR29, it was first essential to ensure that the mutation had arisen

as a result of the Tn*5* insertion. This was attempted by P1 transduction of the kanamycin resistance encoded by Tn5 from the original mutant, JR29, into DS941, to produce a strain called JR30. Analysis of JR30 revealed that the mutation in this strain was not the same as in JR29, as JR30 was able to resolve the 2-*cer* reporter plasmid pCS210. Resolution of pCS210 by JR30 differed from DS941, as JR30 pCS210 colonies were blue on X-Gal while those of DS941 pCS210 were white. JR29 was found to be partially complemented by PepA, but the complementation was not to the same extent as with the previously characterised *pepA* strain DS957.

Peptidase assays were carried out with DS941, DS957, JR29 and JR30, and the results suggested that JR30 is unlikely to be a *pepA* mutant strain as the peptidase activity measured from extracts of this strain was found to be similar to that of DS941, and it was unable to grow on minimal medium containing val-leu. JR29 was found to have lower peptidase activity, similar to that of DS957, and grew on val-leu plates, indicating that this strain might be a *pepA* mutant. The *in vivo* and *in vitro* peptidase assays, in conjunction with the complementation analysis, indicated that JR29 is likely to be a *pepA* strain.

Not all of the results can be explained if JR29 is a *pepA* mutant. JR29 is not totally complemented by pCS112, which totally complements the previously characterised *pepA* strain DS957. It was found that in JR29 pCS112, about 17% of pCS210 was unresolved compared with DS957 pCS112, which resolved 99.5% of the pCS210 to product. If this result is taken in conjunction with the observed inconsistency between colony colour on X-Gal plates and plasmid size in JR30 pCS210, a tentative hypothesis can be proposed: JR29 might have two mutations affecting Xer-mediated recombination at *cer* and *psi*. One of these mutations is affecting PepA, the other is an unknown mutation affecting the rate of Xer-mediated recombination. The mutation that co-transduced with the kanamycin resistance marker of JR29 to make JR30 might affect rate of Xer-mediated recombination, but does not abolish it; the assumption that the rate or efficiency of recombination is affected is derived from the observations that colonies of JR30 pCS210 were blue on X-Gal plates, while the plasmid appeared totally resolved on a gel, and if mini-prep DNA was prepared from JR30 pCS210 and transformed into DS9008, it resulted in white colonies. It should be possible to identify the gene which is mutated in JR30 by its linkage to the Tn*5* insertion if this is considered worthwhile by others in the future. Unfortunately, I had too little time to analyse these mutants further.

It is disappointing that this mutagenesis and mutant screen did not uncover either a new *xer* gene, or the gene involved in recombination at *psi*. The most promising aspect of these experiments for future investigations of *psi* is that the mutagenesis strategy described in this Chapter, along with the plasmid pJEN96, were able to isolate *xer* genes. The majority of the mutants that were screened were mutants at *xerC* or *xerD*, as their lesions were complemented by the XerC/D over-expressing plasmid, pMAY5. More difficulty arose when attempting to complement with *pepA* expression plasmids: by bad luck the two *pepA* expression plasmids pCS112 and pCS126 used in this laboratory are similar in size, and are also too close in size to be distinguished from pJEN96r or pCS211, making analysis of complementation assays on agarose gels extremely difficult. The blue/white colour assay of pCS210 helped overcome this technical difficulty, although the results with this plasmid were not always clear-cut, as illustrated by the findings with JR30 where the colour assay and plasmid content as analysed on gels did not tally. However, once these technical problems can be overcome (for example by constructing a plasmid which expresses *pepA* in addition to *xerC* and *xerD*) it should be possible to repeat this mutagenesis, and hopefully reveal the putative factor which acts in *psi* recombination.

Chapter 6

Concluding Remarks

Many questions concerning Xer recombination still await answers. Three major topics were studied during the course of this work, with the global aim of furthering the current understanding of Xer-mediated recombination.

One of the intriguing features of XerC/D-mediated recombination is the selectivity that is shown at some sites, but not at others. The experiments described in Chapter 3 gave some insight into the mechanisms that might be used by XerC and XerD to ensure resolution selectivity at some sequences, and a freedom from the constraints of selectivity at others. From the experimental results, it appeared that XerC/D-mediated recombination shows resolution selectivity at sites that have reduced affinities for the recombinases. For example, XerC/D bind with reduced overall affinity to the *parB* core recombination site compared with that of the type II hybrid. Reduced affinity for XerC/D is correlated with the formation of protein-DNA complexes with different conformations. The reduced affinity for *parB* (as well as the conformational change) was particularly associated with a single base pair substitution in the left arm of the core recombination site.

Since *cer* has the same left arm sequence as the type II hybrid, resolution selectivity during recombination at *cer* must be attributed to the length of the *cer* central region (8 bp; compared with the type II hybrid's 6 bp), which would be expected to affect the co-operative interaction between XerC and XerD bound at the left and right arms respectively. This is exemplified by the resolution selectivity shown during recombination at the type I hybrid, which differs in sequence from the type II hybrid only in that it has a 2 bp insertion in the central region, and by the failure of *dif*-8 to recombine without accessory factors. Therefore, while the sequence of the XerC binding site contributes to the resolution selectivity mechanism, it is one of a number of mechanisms used to prevent recombination between sites that are inappropriately arranged for a resolution reaction.

For those sites that have evolved to resolve plasmid multimers, resolution selectivity ensures that site-specific recombination maintains plasmids in a monomeric form to increase their heritable stability. These sites cannot recombine as core recombination sites, and therefore require accessory factors to allow recombination to proceed. The pSC101 *psi* site is a substrate for XerC/D-mediated recombination, and acts as a plasmid resolution site. One of the most interesting features of this site is that it requires PepA, and accessory DNA sequences for recombination, but is independent of ArgR. The identification of a potential protein-binding site in the *psi* accessory DNA led

to the suggestion that another DNA binding protein might be able to bind to this site to "substitute" for ArgR during recombination at *psi*. While the putative factor binding to this site remained elusive after the mutagenesis described in Chapter 5, work is currently in progress to identify it, and therefore to gain some insight into the role of the accessory factors during Xer recombination. That PepA is required for recombination at *psi* indicates that either ArgR and PepA may not interact during recombination, or that PepA may be able to interact with another protein to achieve the same end-result. The identification of this protein might further understanding of the role of ArgR during recombination at *cer*, and might also shed light on the function of PepA, which is currently unknown.

The second reason for the chromosomal mutagenesis was to attempt to identify additional *xer* genes not belonging to the existing four complementation groups. It is possible that we have already identified all of the genes involved in Xer recombination. The main reason for suspecting that there may be additional proteins involved in recombination is that we have to date been unable to produce resolution products *in vitro*. The present *in vitro* conditions efficiently produce Holliday junctions, implying that one strand-exchange is catalysed *in vitro*. As a consequence of the inability to produce resolution products *in vitro*, we do not currently have direct evidence that Xer recombination directs both strand-exchanges; it is possible that the role of this system is to form Holliday junctions that are then resolved by other cellular Holliday junction-resolving enzymes, such as RuvC or RecG (Connolly *et al.*, 1991; Lloyd, 1991).

Intermolecular recombination might occur at *dif* due to its high affinity for both recombinases. Chapter 4 gave preliminary evidence that in a *fis* mutant strain, expression of XerC is decreased, and that recombination at *dif* is reduced. This might be further evidence for the importance of XerC binding as a control-step in Xer recombination. More experiments are required to further investigate these findings. It would be desirable to study expression of XerC more directly in the presence and absence of Fis, and also to examine the effect of a *fis* mutation on expression of XerD. At present, little is known about the expression of XerC and XerD: Fis is the first protein that has been proposed to affect expression of either protein. To my knowledge, this is the first example of a site-specific recombination system that is affected indirectly by Fis: in other systems Fis interacts at the recombination site. The potentially novel involvement of Fis during Xer recombination therefore surely warrants further investigation.

References

References

Abdel-Meguid S. S., Grindley N. D. F., Templeton N. S., and Steitz T. A. (1984). Cleavage of the site-specific recombination protein γδ resolvase: the smaller of the two fragments binds DNA specifically. Proc. Natl. Acad. Sci. USA **81**: 2001-2005.

Abremski K., Frommer B. and Hoess R. H. (1986). Linking-number changes in the DNA substrate during Cre-mediated *loxP* site-specific recombination. J. Mol. Biol. **192**: 17-26.

Abremski K. E. and Hoess R. H. (1992). Evidence for a second conserved arginine residue in the integrase family of recombination proteins. Protein Engineering **5**: 87-91.

Abremski K. E., Hoess R. H. and Sternberg N. (1983). Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. Cell **32**: 1301-1311.

Andrews B. J., Proteau G. A., Beatty L. G. and Sadowski P. D. (1985). The FLP recombinase of the 2-μm circle DNA of yeast: interaction with its target sequences. Cell **40**: 795-803.

Argos P., Landy A., Abremski K., Egan J. B., Haggard-Ljungquist E., Hoess R. H., Kahn M. L., Kalionis B., Narayana S., Pierson L. S., Sternberg N. and Leong J. M. (1986). The integrase family of site-specific recombinases: regional similarities and global diversity. EMBO J. **5**: 433-440.

Austin S., Ziese M., and Sternberg N. (1981). A novel role for site-specific recombination in maintenance of bacterial replicons. Cell **25**: 729-736

Bachmann B.J. (1972). Pedigrees of some mutant strains of *Escherichia coli* K12. Bacteriol. Reviews **36**: 525-557

Ball C. A. and Johnson R. C. (1991a). Efficient excision of phage λ from the *Escherichia coli* chromosome requires the Fis protein. J. Bact. **173**: 4027-4031.

Ball C. A. and Johnson R. C. (1991b). Multiple effects of Fis on integration and control of lysogeny in phage λ. J. Bact. **173**: 4032-4038.

Bednarz A. L. , Boocock M. R. and Sherratt D. J. (1990). Determinants of correct *res* site alignment in site-specific recombination by Tn3 resolvase. Genes & Dev. **4**: 2366-2375.

Berg D. E. and Berg C. M. (1983). The transposable element Tn5. Biotechnology **1**: 417-435.

Berg D. E., Weiss A. and Crossland L. (1980). The polarity of Tn5 insertion mutants in *Escherichia coli*. J. Bact. **412**: 439-446.

Bernardi A. and Bernardi F. (1984). Complete sequence of pSC101. Nucl. Acids Res. **12**: 9415-9426.

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

Blake D. G. (1993). Interaction of Tn3 resolvase with its recombination site, *res*. PhD thesis, University of Glasgow.

Blakely G., Colloms S.D., May G., Burke M. and Sherratt D.J. (1991). *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. New Biologist **3**: 789-798

Blakely G., May G., McCulloch R., Arcizweska L. K., Burke M., Lovett S. T., and Sherratt D.J. (1993). Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. Cell **75**: 351-361.

Boocock M.R., Brown J.L. and Sherratt D.J. (1987). Topological specificity in Tn3 resolvase catalysis. UCLA Symposia on Molecular and Cellular Biology **47**: 703-718.

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

Bruist M.F., Glasgow A.C., Johnson R.C. and Simon M.I. (1987). Fis binding to the recombinational enhancer of the Hin DNA inversion system. Genes and Development **1**: 762-772

Carey J. (1988). Gel retardation at low pH resolves Trp repressor-DNA complexes for quantitative study. Proc. Natl. Acad. Sci. USA. **85**: 975-979.

Carey J. (1991). Gel retardation. Meth. Enzymol. **208**: 103-117.

Chen J. W., Lee J. and Jayaram M. (1992). DNA cleavage in *trans* by the active site tyrosine during FLP recombination: switching partners before exchanging strands. Cell. **69**: 647-658.

Cohen S. N. and Chang A. S. N. (1973). Recircularisation and autonomous replication of a sheared R-factor segment in *Escherichia coli* transformants. Proc. Natl. Acad. Sci. USA **70**: 1293-1297.

Cohen S. N. and Chang A. S. N. (1977). Revised interpretation of the origin of the pSC101 plasmid. J. Bact. **132**: 734-737.

Colloms S.D. (1990). Isolation and characterisation of *E.coli xerC*, a gene required for *cer* site-specific recombination. PhD thesis: University of Glasgow.

Colloms S.D., Sykora P., Szatmari G. and Sherratt D.J. (1990). Recombination at ColE1 *cer* requires the *Escherichia coli xerC* gene product, a member of the lambda integrase family of site-specific recombinases. J. Bact. **172**: 6973-6980.

Connolly B., Parsons C.A., Benson F.A., Dunderdale H.J., Sharples G.J., Lloyd R.G. and West S.C. (1991). Resolution of Holliday junctions *in vitro* requires the *Escherichia coli ruvC* gene product. Proc. Natl. Acad. Sci USA. **88**: 6063-6067.

Craig N. L. (1988). The mechanism of conservative site-specific recombination. Ann Rev Genet. **22**: 77-105.

Crothers D. M., Gartenburg M. R. and Schrader T.E. (1991). DNA bending in protein-DNA complexes. Methods in Enzymology **208**: 118-146.

de Bruijn F. J. and Lupski J. R. (1984). The use of transposon Tn5 mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multi-copy plasmids - a review. Gene **27**: 131-149.

de Gennes P. -G. (1971). Reptation of a polymer chain in the presence of fixed obstacles. J. Chem. Phys. **55**: 572-579.

Deutsch J. M. (1988). Theoretical studies of DNA during gel electrophoresis. Science **240**: 922-924.

Durkacz B. and Sherratt D. J. (1973). Segregation of the colicinogenic factor ColE1 from a bacterial population temperature sensitive for DNA polymerase I. Mol. Gen. Genet. **121**: 71-75.

Filutowicz M., Ross W., Wild J. and Gourse R. L. (1992). Involvement of Fis protein in replication of the *Escherichia coli* chromosome. J. Bact. **174**: 398-407.

Finkel S. E. and Johnson R. C. (1992). The Fis protein: it's not just for DNA inversion anymore. Mol. Micro. **6**: 3257-3265

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

Futcher A. B. (1986). Copy number amplification of the 2-μm plasmid of *Saccharomyces cerevisiae*. J. Theor. Biol. **119**: 197-204.

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

Gellert N. and Nash H. (1987). Communication between segments of DNA during site-specific recombination. Nature **325**: 401-404.

Gille H., Egan J. B., Roth A. and Messer W. (1991). The Fis protein binds and bends the origin of chromosomal DNA replication, *oriC*, of *Escherichia coli*. Nucl. Acids Res. **19**: 4167-4172.

Glasgow A.C., Kelly H.T., and Simon M. (1989). Bacterial DNA inversion systems. In: Mobile DNA. Ed. Berg D., and Howe M. pp 637-660

Grindley N.D.F., Lauth M.R., Wells R.G., Wityk R.J., Salvo J.J., and Reed R.R. (1982). Transposon-mediated site-specific recombination: identification of three binding sites for resolvase at the *res* sites of γδ and Tn3. Cell **30**: 19-27.

Gronostajski R. M. and Sadowski P. D. (1985a). Determination of DNA sequences essential for FLP-mediated recombination by a novel method. J. Biol. Chem. **260**: 2320-2327.

Gronostajski R. M. and Sadowski P. D. (1985b). The FLP recombinase of the *Saccharomyces cerevisiae* 2-μm plasmid attaches covalently to DNA *via* a phosphotyrosyl linkage. Mol. Cell. Biol. **5**: 3274-3279.

Gronostajski R. M. and Sadowski P. D. (1985c). The FLP protein of the 2-μm plasmid of yeast - intermolecular and intramolecular reactions. J. Biol. Chem. **260**: 2328-2335.

Haffter B. and Bickel T. A. (1987): Purification and DNA binding properties of Fis and Cin, two proteins required for the bacteriophage P1 site-specific recombination system, *cin*. J. Mol. Biol. **198**: 579-587.

Hagerman P. J. (1985). Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. Biochemistry **24**: 7033-7037.

Hakkaart M.J., Van Den Elzen P., Veltkamp E., and Nijkamp H.J.J. (1984). Maintenance of multicopy plasmid CloDF13 in *Escherichia coli* cells: evidence for site-specific recombination at *parB*. Cell **36**: 203-209

Hatfull G.F., and Grindley N.D.F. (1986). Analysis of γδ resolvase mutants *in vitro*: evidence for an interaction between serine 10 of resolvase and site 1 of *res*. Proc. Natl. Acad. Sci. USA **83**:5429-5433.

Hatfull G.F. and Grindley N.D.F. (1988). Resolvases and DNA invertases: a family of enzymes active in site-specific recombination. In: Genetic Recombination. Ed. Kucherlapati R. and Smith G.R. 857-867.

Haykinson M. J. and Johnson R. C. (1993). DNA looping and the helical repeat *in vitro* and *in vivo*: effect of HU protein and enhancer location on Hin invertasome assembly. EMBO J. **12**: 2503-2512.

Heery D.M., Gannon F. and Powell R. (1990). A simple method for sub-cloning DNA fragments from gel slices. TIG **6**: 173.

Heichman K. A. and Johnson R. C. (1990). The Hin invertasome: protein-mediated joining of distant recombination sites at the enhancer. Science **249**: 511-517.

Heichman K. A., Moskowitz I. P. G. and Johnson R. C. (1991). Configuration of DNA strands and mechanism of strand-exchange by the Hin invertasome as revealed by analysis of recombinant knots. Genes & Dev. **5**: 1622-1634.

Hoess R. H. and Abremski K. (1984). Interaction of the bacteriophage P1 recombinase Cre with the recombining site *loxP*. Proc. Natl. Acad. Sci. USA **81**: 1026-1029.

Hoess R. H. and Abremski K. (1985). Mechanism of strand cleavage and exchange in the Cre-*lox* site-specific recombination system. J. Mol. Biol. **181**: 351-362.

Hoess R. H., Wierzbicki A. and Abremski K. (1986). The role of the *loxP* spacer region in P1 site-specific recombination. Nucl. Acids Res. **14**: 2287-2300.

Hoess R.H., Wierzbicki A. and Abremski K. (1987). Isolation and characterisation of intermediates in site-specific recombination. Proc. Natl. Acad. Sci. USA **84**: 6840-6844.

Holliday R. (1964). A mechanism for gene conversion in fungi. Genet. Res (Cambridge). **5**: 282-304.

Hubner P. and Arber W. (1989). Mutational analysis of a prokaryotic recombinational enhancer element with two functions. EMBO J. **8**: 577-585.

Hughes R. E., Hatfull G. F., Rice P., Steitz T. and Grindley N. D. F. (1990). Co-operativity mutants of γδ resolvase identify an essential interdomain interaction. Cell **63**: 1331-1338.

Iwasaki H., Takahagi M., Shiba T., Nakatu A. and Shinagawa H. (1991). *Escherichia coli* RuvC is an endonuclease that resolves the Holliday structures. EMBO J. **10**: 4381-4389.

Jayaram M., Crain K. L., Parsons R. L. and Harshey R. M. (1988). Holliday junctions in FLP recombination: resolution by step-arrest mutants of FLP protein. Proc. Natl. Acad. Sci. USA **85**: 7902-7906.

Johnson R.C. (1991). Mechanism of site-specific DNA inversion in bacteria. Curr. Opin. Genet. Dev. **1**: 404-411.

Johnson R.C., and Bruist M.F. (1989). Intermediates in Hin-mediated DNA inversion: a role for Fis and the recombination enhancer in the strand exchange reaction. EMBO J. **8**: 1581-1590

Johnson R. C. and Simon M. I. (1985). Hin-mediated site-specific recombination requires two 26 bp recombination sites and a 60 bp recombinational enhancer. Cell **41**: 781-791.

Kahmann R., Rudt F., Koch C, and Mertens G. (1985). G inversion in bacteriophage Mu DNA is stimulated by a site within the invertase gene and a host factor. Cell **41**: 771-780.

Kanaar R., Klippel A., Shektman E., Dungan J. M., Kahmann R. and Cozzarelli N. R. (1990). Processive recombination by the phage Mu Gin system: Implications for the mechanisms of DNA strand-exchange, DNA site alignment and enhancer action. Cell **62**: 353-366.

Kerppola T. K. and Curran T. (1993). DNA bending by Fos and Jun: structural and functional implications. In: Nucleic Acids and Molecular Biology. Eds: Eckstein F. and Lilley D. M. J. (Published by Springer-Verlag). pp 70-105.

Kim J., Zweib C., Wu C. and Adhya S. (1989). Bending of DNA by gene regulatory proteins: construction and use of a DNA bending vector. Gene **85**: 15-23.

Klippel A., Mertens G., Patschinsky T., and Kahmann R. (1988a). The DNA invertase Gin of phage Mu: formation of a covalent complex with DNA *via* a phosphoserine at amino acid position 9. EMBO J. **7**: 1229-1237.

Klippel A., Cloppenborg K. and Kahmann R. (1988b). Isolation and characterisation of unusual *gin* mutants. EMBO J. 3983-3989.

Koch C., Ninnemann O., Fuss H. and Kahmann R. (1991). The N-terminal part of the *E. coli* DNA binding protein Fis is essential for stimulating site-specific DNA inversion but is not required for specific DNA binding. Nucl. Acids Res. **19**: 5915-5822.

Kolb A., Spassky A., Chapon C., Blazy B. and Buc H. (1983). On the different binding affinities of CRP at the *lac, gal* and *malT* promoter regions. Nucl. Acids Res. **11**: 7833-7852.

Kolot M. (1990). Par site of the ColN plasmid: Structural and functional organisation. Mol. Gen. Genet. **222**: 77-80.

Koo H. -S., Wu H. -M. and Crothers D. M. (1986). DNA bending at adenine-thymine tracts. Nature **320**: 501-506.

Kostrewa D. , Granzin J., Koch C., Choe H. -W., Raghunathan S., Wolf W., Labahn J., Kahmann R. and Saenger W. (1991). Three-dimensional structure of the *E. coli* DNA-binding protein Fis. Nature **349**: 178-180.

Koudelka G. B. (1993). Indirect effects of DNA sequence on 434 repressor-DNA interactions. In: Nucleic Acids and Molecular Biology. Eds: Eckstein F. and Lilley D. M. J. (Published by Springer-Verlag): pp 16-27.

Kremer W., Klenin K., Diekmenn S. and Langowski J. (1993). DNA curvature influences the internal motions of supercoiled DNA. EMBO J. **12**: 4407-4412.

Kuempel P.A., Henson J.M., Dircks L., Tecklenburg M. and Lim D.F. (1991). *dif*, a recA-independent recombination site in the terminus region of the chromosome of *Escherichia coli*. New Biologist **3**: 799-949

Landy A. (1989). Dynamic, structural, and regulatory aspects of lambda site-specific recombination. Ann. Rev. Biochem. **58**: 913-949.

Lane D., Prentki P. and Chandler M. (1992). Use of gel retardation to analyse protein-nucleic acid interactions. Micro. Rev. **56**: 509-528.

Lazarus L. T. and Travers A. A. (1993). The *Escherichia coli* Fis protein is not required for the activation of *tyrT* transcription on entry into exponential growth. EMBO J. **12**: 2483-2494.

Lee J. and Jayaram M. (1993). Mechanism of site-specific recombination: logic of assembly of recombinase active site from fractional active sites. J. Biol. Chem. **268**: 19152-19159.

Levene S. D. and Zimm B. H. (1989). Understanding the anomalous electrophoresis of bent DNA molecules: a reptation model. Science **245**: 396-399.

Lim D., Oppenheim J.D., Eckhart T. and Maas W.K. (1987). Nucleotide sequence of the *argR* gene of *Escherichia coli* K12 and isolation of its product, the arginine repressor. Proc. Natl. Acad Sci USA. **84**: 6697-6701.

Lloyd R.G. (1991). Conjugational recombination in resolvase-deficient *ruvC* mutants of *Escherichia coli* K12 depends on *recG*. J. Bact. **173**: 5414-5418.

Lovett S.T. and Clark A.J. (1985). Cloning of the *Escherichia coli recJ* chromosomal region and identification of its encoded proteins. J. Bact. **162**: 280-285.

Lovett S.T. and Kolodner R.D. (1991). Nucleotide sequence of the *Escherichia coli recJ* chromosomal region and construction of RecJ-overexpression plasmids. J. Bact. **173**: 353-364.

Lumpkin O. J., Dejardin P. and Zimm B. H. (1985). Theory of gel electrophoresis of DNA. Biopolymers **24**: 1573-1593.

Lumpkin O. J. and Zimm B. H. (1982). Mobility of DNA in gel electrophoresis. Biopolymers **21**: 2315-2316.

McClelland M., Hanish J., Nelson M. and Patel Y. (1988). KGB: a single buffer for all restriction endonucleases. Nucl. Acids Res. **16**: 364.

McCulloch R. (1992). Analysis of Xer site-specific recombination *in vivo*. PhD thesis, University of Glasgow.

McCulloch R., Colloms S., Coggins L., Sherratt D. J. (1994a). Xer-mediated site-specific recombination at *cer* generates Holliday junctions *in vivo*. EMBO J., in press.

McCulloch R., Burke M. and Sherratt D. J. (1994b). Peptidase activity of *E. coli* aminopeptidase A is not required for its role in site-specific recombination. Mol. Micro., in press.

Meyer-Leon L., Huang L. C., Umlauf S. W., Cox M. M. and Inman R. B. (1988). Holliday intermediates and reaction by-products in FLP protein-promoted site-specific recombination. Mol. Cell. Biol. **8**: 3784-3796.

Meyer-Leon L., Inman R. B. and Cox M. M. (1990). Characterisation of Holliday structures in FLP protein-promoted site-specific recombination. Mol. Cell. Biol. **10**: 235-242.

Miller C. G. and Schwartz G. (1978). Peptidase-deficient mutants of *Escherichia coli*. J. Bact. **135**: 603-611.

Moitoso de Vargas L., Pargellis C.A., Hasan N.M., Bushman E.W. and Landy A. (1988). Autonomous DNA binding domains of λ-integrase recognise two different sequence families. Cell **54**: 923-929.

Morlon J., Chartier M., Bidaud M. and Lazdunski C. (1988). The complete nucleotide sequence of the colicinogenic plasmid ColA. High extent of homology with ColE1. Mol. Gen. Genet. **211**: 231-243.

Moskowitz I. P. G., Heichman K. A. and Johnson R. C. (1991). Alignment of recombination sites in Hin-mediated site-specific DNA recombination. Genes & Dev. **5**: 1635-1645.

Newman B. J. and Grindley N. D. F. (1984). Mutants of γδ resolvase: a genetic analysis of the recombination function. Cell **38**: 463-469.

Nilsson L., Vanet A., Vijgenboom E. and Bosch L. (1990). The role of Fis in *trans*-activation of stringent RNA operators of *E. coli*. EMBO J. **9**: 727-734.

Ninnemann O., Koch C. and Kahmann R. (1992). The *E. coli fis* promoter is subject to stringent control and auto-regulation. EMBO J. **11**: 1075-1083.

Nunes-Duby S. E., Matsumoto L. and Landy A. (1987). Site-specific recombination intermediates trapped with suicide substrates. Cell **50**: 779-788.

Osuna R., Finkel S. E. and Johnson R. C. (1991). Identification of two functional domains in Fis: the N-terminus is required to promote Hin-mediated DNA inversion but not λ excision. EMBO J. **10**: 1593-1603.

Pargellis C. A., Nunes-Duby S. E., Devargas L. M. and Landy A. (1988). Suicide recombination substrates yield covalent λ integrase-DNA complexes and lead to identification of the active-site tyrosine. J. Biol. Chem. **263**: 7678-7685.

Patient M. E. and Summers D. K. (1993). ColE1 multimer formation triggers inhibition of *Escherichia coli* cell division. Mol. Micro. **9**: 1089-1095.

Qian X-H., Inman R. and Cox M. M. (1990). Protein-based asymmetry and protein-protein interactions in FLP recombinase-mediated site-specific recombination. J. Biol. Chem. **265**: 21779-21788.

Richaud C., Higgins W., Mengin-Lecreulx D. and Stragier P. (1987). Molecular cloning, characterisation and chromosomal localisation of *dapF*, the *E. coli* gene for diaminopimelate epimerase. J. Bact. **169**: 1454-1459

Richaud C. and Printz C. (1988). Nucleotide sequence of the *dapF* gene and flanking regions from *Escherichia coli* K12. Nucl. Acids Res. **16**: 10367

Ross W., Thompson J. F., Newlands J.T., Gourse R. L. (1990). *E. coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. EMBO J. **9**: 3733-3742.

Sadowski P. D. (1986). Site-specific recombinases: changing partners and doing the twist. J. Bact. **165**: 341-347.

Saenger W., Sandmann C., Theis K., Starikov, E. B., Kostrewa, D., Labahn, J. and Granzin, J. (1993). Structural and functional aspects of the DNA binding protein Fis. In: Nucleic Acids and Molecular Biology. Eds: Eckstein F. and Lilley D. M. J. (Published by Springer-Verlag): pp 158-169.

Salvo N. J. and Grindley N. D. F. (1987). Helical phasing between DNA bends and the determination of bend direction. Nucl. Acids. Res. **15**: 9771-9779.

Sambrook J., Fritsch E. F. and Maniatis T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition (Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press).

Sanger F., Nicklen S. and Coulson A. R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**: 5463-5467.

Sato T., Samori Y. and Kabayashi Y (1990). The *cisA* cistron of *Bacillus subtilis* sporulation gene *spoIVC* encodes a protein homologous to a site-specific recombinase. J. Bact. **172**: 1092-1098.

Scott T. N. and Simon M. I. (1982). Genetic analysis of the mechanism of *Salmonella* phase variation site-specific recombination system. Mol. Gen. Genet. **188**: 313-321.

Senecoff J. F. and Cox M. M. (1986). Directionality in FLP protein-promoted site-specific recombination is mediated by DNA-DNA pairing. J. Biol. Chem. **261**: 7380-7386.

Shaw K. J. and Berg C. M. (1979). *Escherichia coli* K12 auxotrophs induced by the insertion of the transposable element Tn5. Genetics **92**: 741-747.

Sherratt D. J. (1989). Tn3 and related transposable elements. In: Mobile DNA. Ed. Berg D., and Howe M. pp163-184.

Sherratt D. J. (1993). Site-specific recombination and segregation of circular chromosomes. In: Nucleic Acids and Molecular Biology. Eds: Eckstein F. and Lilley D. M. J. (Published by Springer-Verlag). pp 202-216.

Sherratt D. J., Blakely G., Burke M., Colloms C., Leslie N., McCulloch R., May G., and Roberts J. (1993). Site-specific recombination and the partition of bacterial chromosomes. In: The Chromosome, J. S. Heslop-Harrison and R. B. Flavell, eds. (Oxford: Bios Scientific Publishers Ltd), pp 25-41.

Sigman D. S., Kuwabara M. D., Chen C.-H. B. and Bruce T. W. (1991). Nuclease activity of 1, 10-phenanthroline-copper in study of protein-DNA interactions. Meth. Enzymol. **208**: 414-433.

Stark W.M., Sherratt D.J., and Boocock M.R. (1989a). Site-specific recombination by Tn3 resolvase: topological changes in the forward and reverse reactions. Cell. **58**: 779-790

Stark W.M., Boocock M.R., and Sherratt D.J. (1989b). Site-specific recombination by Tn3 resolvase. TIG. **5**: 304-309

Stark W.M., Boocock M.R., and Sherratt D.J. (1992). Catalysis by site-specific recombinases. TIG. **8**: 432-439

Sternberg N., Hoess R. and Abremski K. (1983). The P1 *lox*-Cre site-specific recombination system: properties of *lox* sites and biochemistry of *lox*-Cre interactions. In: Mechanisms of DNA replication and recombination. (Alan Liss Inc. publ): pp 671-684.

Stirling C. J. (1987). Cloning and characterisation of *Escherichia coli xer* genes. PhD thesis: University of Glasgow.

Stirling C. J., Stewart G. and Sherratt D. J. (1988a). Multi-copy plasmid stability in *Escherichia coli* requires host-encoded functions that lead to plasmid site-specific recombination Mol. Gen. Genet. **214**: 80-84.

Stirling C.J., Szarmari G., Stewart G., Smith M.C.M. and Sherratt D.J. (1988b). The arginine repressor is essential for plasmid-stabilising site-specific recombination at the ColE1 *cer* locus. EMBO J. **7**: 4389-4395.

Stirling C.J., Colloms S.D., Collins J.F., Szarmari G., and Sherratt D.J. (1989). *xerB*, an *Escherichia coli xer* gene required for ColE1 site-specific recombination, is identical to *pepA*, encoding aminopeptidase A, a protein with substantial similarity to bovine lens leucine aminopeptidase. EMBO J. **8**: 1623-1627.

Summers D.K. and Sherratt D.J. (1984). Multimerisation of high copy number plasmids causes instability: ColE1 encodes a determinant for plasmid monomerisation and stability. Cell. **36**: 1097-1103.

Summers D.K., Yaish S., Archer J. and Sherratt D.J. (1985). Multimer resolution systems of ColK and ColE1: localisation of the crossover site. Mol. Gen. Genet. **201**: 334-338.

Summers D.K. and Sherratt D.J. (1988). Resolution of ColE1 dimers requires DNA sequences implicated in the three dimensional organisation of the *cer* site. EMBO J. **7**: 851-858.

Summers D.K. (1989). Derivatives of ColE1 *cer* show altered topological specificity in site-specific recombination. EMBO J. **8**: 309-315.

Summers D. K., Beton C. W. H. and Withers H. L. (1993). Multicopy plasmid instability: the dimer catastrophe hypothesis. Mol. Micro. **8**: 1031-1038.

Thompson J. F. and Landy A. (1988). Empirical estimation of protein-induced bending angles: applications to λ site-specific recombination complexes. Nucl. Acids Res. **16**:9687-9705.

Vetter D., Andrews B.J., Roberts-Beatty L. and Sadowski P.D. (1983). Site-specific recombination of yeast 2-μm DNA *in vitro*. Proc. Natl. Acad. Sci USA. **80**: 7284-7288

Wasserman S. A., Dungan J. A. and Cozzarelli N. R. (1985). Discovery of a predicted DNA knot substantiates a model for site-specific recombination. Science **229**: 171-174.

Winkler F. K., Banner D. W., Oefner, C., Tsernoglou D., Brown, R. S., Heathman, S. P., Bryan R. K., Martin., P. D., Petratos K. and Wilson K. S. (1993). The crystal structure of *Eco*RV endonuclease and of its complexes with cognate and non-cognate DNA fragments. EMBO J. **12**: 1781-1795,

Wu H. -M. and Crothers D. M. (1984). The locus of sequence-directed and protein-induced DNA bending. Nature **308**: 509-513.

Yanisch-Perron C., Vieira J. and Messing J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**: 103-119.

Zinkel S. S. and Crothers D. M. (1987). DNA bend direction by phase sensitive detection. Nature **328**: 178-181.

Zinkel S. S. and Crothers D. M. (1991). Catabolite activator protein-induced DNA bending in transcriptional activation. J. Mol. Biol. **219**: 201-215.

Tea time!