The EMBO Journal Vol.17 No.14 pp.4175–4187, 1998

Structure of the Holliday junction intermediate in Cre–*loxP* site-specific recombination

Deshmukh N.Gopaul, Feng Guo and Gregory D.Van Duyne[1]

Department of Biochemistry and Biophysics and Johnson Research Foundation, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA

[1]Corresponding author

We have determined the X-ray crystal structures of two DNA Holliday junctions (HJs) bound by Cre recombinase. The HJ is a four-way branched structure that occurs as an intermediate in genetic recombination pathways, including site-specific recombination by the λ-integrase family. Cre recombinase is an integrase family member that recombines 34 bp *loxP* sites in the absence of accessory proteins or auxiliary DNA sequences. The 2.7 Å structure of Cre recombinase bound to an immobile HJ and the 2.5 Å structure of Cre recombinase bound to a symmetric, nicked HJ reveal a nearly planar, twofold-symmetric DNA intermediate that shares features with both the stacked-X and the square conformations of the HJ that exist in the unbound state. The structures support a protein-mediated crossover isomerization of the junction that acts as the switch responsible for activation and deactivation of recombinase active sites. In this model, a subtle isomerization of the Cre recombinase–HJ quaternary structure dictates which strands are cleaved during resolution of the junction via a mechanism that involves neither branch migration nor helical restacking.

Keywords: Cre recombinase/Holliday junction/ recombination

Introduction

Genetic recombination results in the exchange of information between DNA segments and plays a central role in the maintenance and propagation of genomes. In eukaryotes, for example, meiotic recombination generates diversity by exchanging regions of homologous chromosomes and serves as a mechanical organizer in preparation for chromosome segregation (Kleckner, 1996). In bacteria, homologous recombination is a key component of conjugation and DNA-repair pathways (Kowalczykowski *et al*., 1994). A site-specific recombination process is responsible for integration and excision of bacteriophage genomes into and out of their bacterial host chromosomes (Campbell, 1962; Weisberg and Landy, 1983). Although the enzymatic details of the early steps involved in these recombination reactions differ, and are in some cases not yet well understood, each case involves a common four-way branched DNA intermediate known as the Holliday junction (Holliday, 1964).

A Holliday junction (HJ) is formed when a single strand from each of two duplex DNA segments involved in recombination is exchanged in a region of sequence homology to yield a branched structure that has four duplex arms. An important property of the HJ is its ability to undergo branch migration through regions of sequence homology; this isomerization process involves the simultaneous melting and annealing of base pairs and results in movement of the branch point along the arms of the junction (Meselson, 1972). The ultimate fate of HJs *in vivo* is cleavage of two of the DNA single strands to yield unbranched products. Depending on the extent of branch migration and which pairs of strands are cleaved, the nature of the duplex products can be quite different. A number of HJ-resolving enzymes that differ in sequence and strand specificity of cleavage have been identified from bacteria (Dunderdale *et al*., 1991; Iwasaki *et al*., 1991; Sharples *et al*., 1994), bacteriophages (Mizuuchi *et al*., 1982; de Massey *et al*., 1984), yeast (Symington and Kolodner, 1985; West and Korner, 1985; Kleff *et al*., 1992) and higher eukaryotes (Elborough and West, 1990; Sekiguchi *et al*., 1996). It has recently become clear that the cleavage specificities of these enzymes, and therefore the types of recombinant products generated, are intimately related to the structures adopted by the branched DNA substrates (Parsons *et al*., 1990; Bennett and West, 1995b; Duckett *et al*., 1995; Arciszewska *et al*., 1997; Azaro and Landy, 1997).

We have been studying the structures of intermediates in the site-specific recombination pathway of bacteriophage P1 Cre recombinase (Cre) (Abremski *et al*., 1983). Cre is a member of the bacteriophage λ-integrase family of recombinases, which also includes the Flp recombinase from yeast (Sadowski, 1995), the XerC/XerD recombinases from *Escherichia coli* (Blakely *et al*., 1993) and >100 other members identified from sequence analyses (Argos *et al*., 1986; Nunes-Düby *et al*., 1998). Members of this large family of enzymes perform a variety of genetic manipulations. λ-integrase catalyzes the integration and excision of the phage λ genome to and from the host chromosome (Weisberg and Landy, 1983; Landy, 1989); Cre and the XerC/XerD recombinases maintain circular DNA replicons in a monomeric state (Sternberg *et al*., 1981; Sherratt *et al*., 1995) and Flp recombinase mediates the amplification of 2μ plasmid copy number by inverting a DNA segment during replication (Futcher, 1986).

The λ-integrase (Int) family shares a common mechanism for carrying out site-specific recombination that involves the formation of a HJ intermediate (Craig, 1988; Stark *et al*., 1989). The reaction is shown schematically in Figure 1A. A nick is produced in each of the two DNA substrates when conserved tyrosine side chains from two recombinase subunits cleave at specific sites, forming 3′-

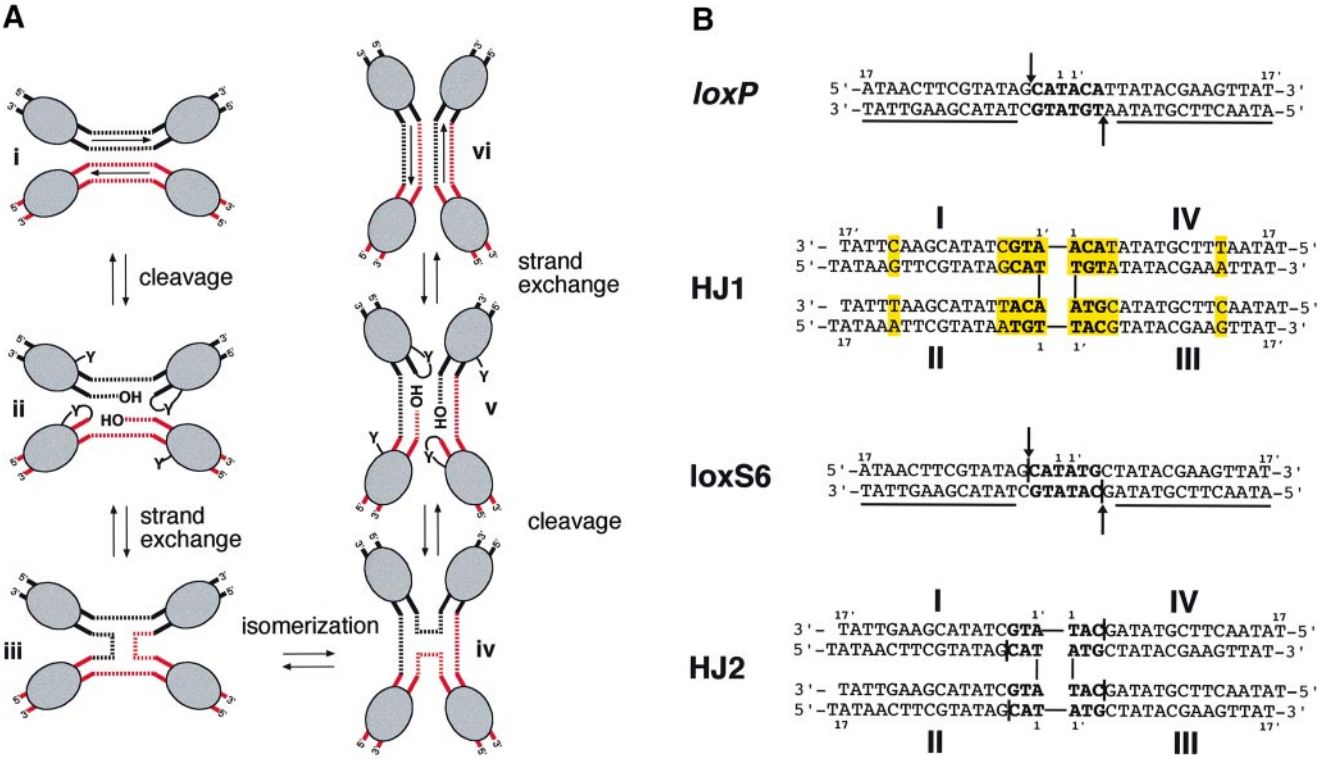

Fig. 1. (**A**) Schematic drawing of the Cre–*loxP* site-specific recombination pathway, based on the strand-swapping model (Nunes-Düby *et al*., 1995) and on Cre–*loxP* structural models (Guo *et al*., 1997; this work). Conserved tyrosine residues from two of the four recombinases in a synaptic tetramer cleave the DNA backbones of the recombining segments to form transient 3′-phosphotyrosine linkages. The released 5′-hydroxyl ends of the cleaved DNA undergo intermolecular nucleophilic attack of the partner phosphotyrosine linkages to complete the exchange of one pair of strands and form a Holliday intermediate. A second round of cleavages and strand exchanges using the remaining two recombinase subunits and the complementary DNA strands gives recombinant products. (**B**) Sequences of *loxP* and *loxS6* DNA duplexes used to design the HJs HJ1 and HJ2. HJ1 and HJ2 are shown in the same orientation as in Figure 3. For HJ1, the strand-bridging arms I and II contain the wild-type *loxP* sequence and the strand-bridging arms III and IV contain the *loxP* complementary sequence. Bases that are not related by twofold symmetry and prevent branch migration are highlighted in yellow. Vertical lines indicate missing phosphates in the DNA backbone. The 13 bp inverted repeat binding sites for Cre recombinase are underlined and the 6 bp crossover region between cleavage sites are in bold. For both the Cre–HJ1 and Cre–HJ2 complexes, an additional 5′-overhanging thymidine residue was found to facilitate crystallization.

phosphotyrosine linkages with the substrates and releasing free 5′-hydroxyl DNA ends. The 5′-hydroxyl groups carry out intermolecular nucleophilic attack of the partner phosphotyrosine linkages to complete the exchange of one set of strands and produce a Holliday intermediate. The Holliday structure is then resolved to the recombinant products by cleavage and exchange of the remaining pair of strands by the second pair of recombinase subunits. Structural and mechanistic similarities between the Int family and the eukaryotic topoisomerase I family of enzymes have recently been demonstrated (Cheng *et al*., 1998).

Cre recombinase is a 38.5 kDa protein that catalyzes site-specific recombination between 34 bp *loxP* sites. The *loxP* sequence (Figure 1B) is composed of two 13 bp recombinase-binding elements arranged as inverted repeats which flank an 8 bp central region where cleavage and ligation reactions occur (Hoess *et al*., 1982). This core recombination site organization is common to the Int family members that have been well-studied, although the length of the binding sites and central segments vary and some enzymes require auxiliary DNA sequences outside of this core region. Recent investigations into the mechanism of Int family site-specific recombination have focused on the nature of the recombinase-bound Holliday intermediate, and in particular on the ability of the DNA junction conformation to influence the direction of resolution to products (Arciszewska *et al*., 1997; Azaro and Landy, 1997; Lee *et al*., 1998).

Previously we described the structure of a covalent intermediate in the Cre–*loxP* site-specific recombination pathway (Figure 1A, ii and v), where two of the four catalytic tyrosine residues had cleaved the DNA substrates and formed 3′-phosphotyrosine linkages (Guo *et al*., 1997). This synaptic complex was trapped at the strand exchange step of the reaction. Here we report two X-ray crystal structures of Cre recombinase–HJ (Cre–HJ) complexes, representing the pivotal intermediate in the Cre–*loxP* site-specific recombination pathway (Figure 1A, iii and iv). The first, determined at 2.7 Å resolution, contains an immobile junction with the branch point fixed at the center of the crossover region. The second, determined at 2.5 Å, contains a partially mobile junction formed from nicked DNA substrates. Both structures indicate that the Cre-bound HJ adopts a pseudo-planar, dyad-symmetric conformation distinct from either the stacked-X or the extended square models that have been deduced from studies of isolated junctions (Duckett *et al*., 1988). The Cre–HJ structures also suggest a model for isomerization of the recombinase–HJ complex that involves a more subtle rearrangement of protein and DNA quaternary structure than previously proposed.

Results

Construction of Cre–HJ complexes

To construct a HJ substrate for Cre recombinase, four 35-mer oligodeoxynucleotides were annealed to give the junction HJ1 shown in Figure 1B. The sequences of the junction arms are modifications of the wild-type *loxP* site. Because a HJ generated from two 34 bp *loxP* sites contains dyad symmetry with respect to the sequences of the arms and would spontaneously branch-migrate into duplex molecules in solution, it was necessary to introduce asymmetry into the arm sequences of HJ1 (Kallenbach *et al*., 1983). The HJ1 arms were designed by using the wild-type *loxP*-crossover sequence for arms I and II, but swapping top and bottom strands of the *loxP*-crossover sequence in arms III and IV. In addition, 1 bp was changed in each of the inverted repeat regions of the arms. This site was chosen by inspection of the Cre recombinase–DNA (Cre–DNA) interface observed in the structure of a synaptic Cre–DNA covalent intermediate (Guo *et al*., 1997) and was predicted to tolerate change to any of the four DNA base pairs. Together, the asymmetric crossover region and the unique base pair located in each of the inverted repeat arms of the junction effectively destabilizes formation of alternative secondary structures from the single-stranded components and yields an immobile junction.

In order to prevent cleavage of HJ1 by the recombinase, a Cre R173K mutant was used in the structure analysis. Arginine-173 is a strictly conserved active-site residue in the Int family (Abremski and Hoess, 1992) that participates in activation of the scissile phosphate. Mutations at this site render Cre inactive in the complete recombination reaction (Wierzbicki *et al*., 1987) and unable to cleave the *loxA* suicide DNA substrate (Guo *et al*., 1997; data not shown). We chose not to use inactive Cre mutants involving changes to the conserved nucleophile Tyr324 for these structural studies. In our previous covalent Cre–DNA complex structure and in the Cre–HJ structures described here, Tyr324 makes a hydrogen bond adjacent to the scissile phosphate and this interaction may be an important component of active-site assembly prior to cleavage.

The symmetric junction HJ2 was derived from a pair of symmetrized *loxP* sites containing nicks at the positions of the scissile phosphates. To symmetrize the *loxP* site, three positions in the crossover region were changed to give the sequence shown in Figure 1B. This sequence is identical to the *loxA* suicide substrate used to prepare covalent Cre–DNA intermediates (Guo *et al*., 1997), except the nick has been moved to the position of the scissile phosphate. The *loxS6* HJ precursor was formed by annealing two oligodeoxynucleotides to form duplexes with 6 bp self-complementary overhangs which dimerize to give the full site. Following Cre binding and synapsis of the *loxS6* sites, strand exchange between duplexes occurs in the absence of ligation, leading to formation of a HJ complex (see Discussion). Wild-type Cre recombinase was used in complex with HJ2, since the phosphodiester linkages normally attacked by Cre are missing in this substrate.

Table I. Summary of crystallographic refinements

	Cre–HJ1	Cre–HJ2
Resolution	48–2.7 Å	27–2.5 Å
R_{symm} overall/R_{symm} high res shell	0.072/0.33	0.053/0.25
Completeness/redundancy	99%/4.7	94%/2.8
R-working/R_{free} ($F \geqslant 2\sigma$)	0.22/0.29	0.20/0.27
Refined model:		
Protein residues	622/686	614/686
DNA bases	68/70	68/70
Solvent	239	361
r.m.s. bonds	0.005 Å	0.007 Å
r.m.s. angles	0.95°	1.3°

Structure of the Cre–HJ complex

Structures of the Cre–HJ1 and Cre–HJ2 complexes were determined at 3 Å resolution and refined to final resolutions of 2.7 and 2.5 Å, respectively (see Materials and methods; Table I). The two complexes have nearly identical structures, despite sequence differences near the branch point of the DNA junctions and the presence of four nicks in the phosphate backbone of HJ2. The root mean square (r.m.s.) difference in α-carbon positions between the two different complexes is 0.63 and 0.60 Å for the subunits bound to DNA arms I and II, respectively. The structure of the Cre–HJ2 complex is shown in Figure 2. The Cre–HJ1 complex is indistinguishable from Cre–HJ2 in these drawings (not shown). One Cre recombinase protein binds to each of the four arms of the HJ, forming a 'C-clamp' around the inverted repeat binding sites (base pairs 5–17). The N-terminal domain covers one face of the duplex and the C-terminal domain covers the opposite face (Figure 2D). These Cre–DNA arm complexes form a pseudo-fourfold-symmetric tetramer through an extensive network of protein–protein interactions involving both the N- and C-terminal domains. The Cre–DNA interfaces in the arms are all similar to one another and to that seen in the covalent Cre–*loxA* complex (Guo *et al*., 1997). The protein–protein interactions in the Cre–HJ complexes are also similar to those observed in the Cre–*loxA* structure, including a cyclic exchange of C-terminal α-helices (helix N) between adjacent subunits (Figure 2C). The Cre–HJ1 and Cre–HJ2 complexes have a crystallographic dyad axis that relates opposite halves of the structures. The r.m.s. difference between well-ordered α-carbon atoms in the two unique recombinase subunits is 1.7 Å for the Cre–HJ1 complex and 1.5 Å for Cre–HJ2.

Structure of the Cre recombinase-bound HJ

The most conspicuous features of the Cre recombinase-bound HJ are the relative dispositions of the junction arms and the unstacking of base pairs at the branch point (Figure 3). The complexes have exact twofold symmetry, but only approximate fourfold symmetry. For example, the angle between arms I and II is 76° and the angle between arms II and III is 101°. This gives rise to a pronounced distortion from a fourfold-symmetric structure that would have equivalent inter-arm angles (see Materials and methods for definitions of geometric parameters). The Cre–HJ arms do not form pairs of coaxially stacked helices as shown for unbound junctions in the presence of physiological salt concentrations (Duckett *et al*., 1988),

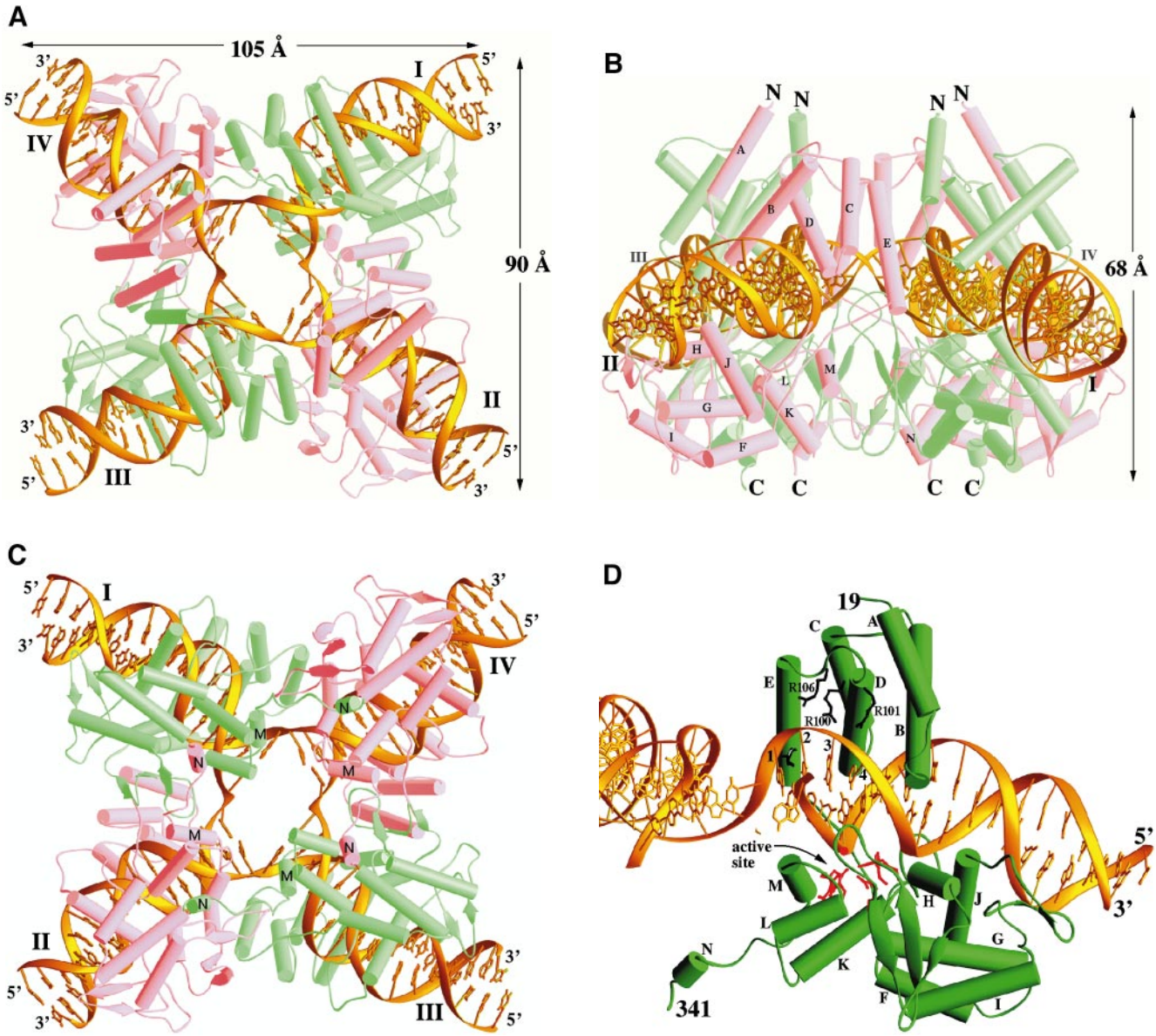

Fig. 2. Structures of the Cre–HJ1 and Cre–HJ2 complexes. Because the complexes are very similar, only the Cre–HJ2 structure is shown, with continuous DNA ribbons drawn through the missing phosphates. (**A**) View from the N-terminal recombinase domain side and the minor groove face of the junction. (**B**) Side view illustrating the near planar arrangement of the DNA arms. Helices of one recombinase protomer are labeled. (**C**) View from the C-terminal recombinase domain side and the major groove face of the junction. C-terminal helices M and N, which are involved in a cyclic exchange of α-helices between recombinase protomers, are labeled. (**D**) One recombinase protomer bound to a single junction arm. The active site residues are in red and the position of the scissile phosphate is represented by a red sphere. Arginine residues that contact phosphates near the branch point are shown in black.

but instead have sharp discontinuities in stacking at the branch point. The DNA bases located at the branch point all make Watson–Crick base pairs, but have become unstacked, with local tilt angles of –113° between base pairs bridging arms I and II and –65° for base pairs between arms II and III. The same tilt angles with respect to a global helicoidal axis through neighboring arms are –110 and –60°, respectively.

Although the HJ adopts an approximately planar conformation, a small out-of-plane bend in the junction arms creates a concave surface on the face of the junction that is contacted by the C-terminal catalytic domains of the recombinase tetramer (Figure 2B). This is also the face of the DNA junction where the major grooves of the four arms converge (Figure 3). The cleavage and ligation steps in the Cre recombinase site-specific recombination reaction therefore occur on the major groove face of the four-way junction substrate. The opposite minor groove face of the junction is bound by the N-terminal recombinase domains. The extent of bending with respect to the best plane through the DNA junction differs for the two unique arms in the Cre–HJ complexes. Arm I makes an angle of 13° with the plane and an angle of 153° with its dyad-symmetric arm III, whereas arm II is 8° from the plane and makes an angle of 164° with arm IV. The distortion from a fourfold-symmetric junction is therefore not limited to rotational differences about an axis perpendicular to the plane of the junction, but also includes small out-of-plane bends that differ for neighboring arms. Within the arms of the junction, the helicoidal DNA parameters

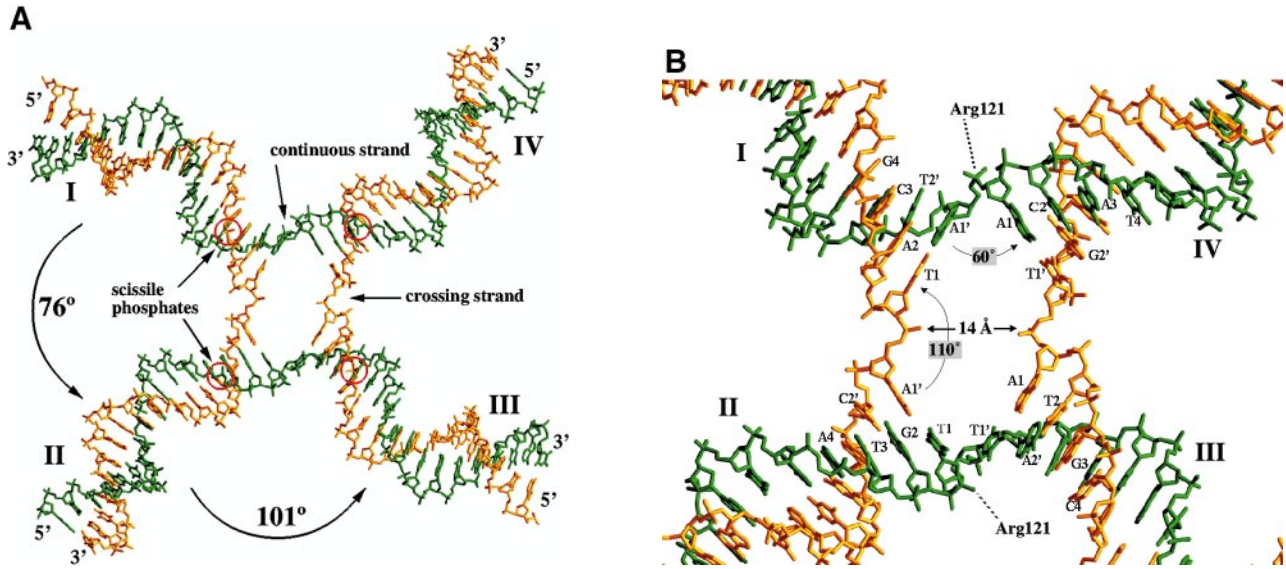

Fig. 3. Structure of the Cre recombinase-bound HJ. The HJ1 structure is shown, with the arm sequences as in Figure 1 (see Materials and methods). (**A**) Overall geometry of the four-way junction. (**B**) Close-up of the center of the junction, where angles refer to helicoidal tilt angles as described in the text.

are consistent with those found in moderately bent DNA duplexes in high resolution crystal structures of protein–DNA complexes. A bend of ~20° is smoothly distributed throughout the 13 bp binding sites of each arm and is responsible for the concave DNA surface on the major groove face of the junction. DNA arms I and II have similar structures, with r.m.s. differences of 1.6 Å within both the Cre–HJ1 and Cre–HJ2 complexes, and a maximum of 0.8 Å for pairwise comparison of the same arms between the two complexes.

The source of the sharp bends between junction arms I and II is different than for arms II and III. For arms II and III, the torsion angle about the C3′–O3′ bond (ε) for the 1–1′ phosphodiester linkage is –87° instead of the *trans* conformation typical of a B-DNA duplex. For arms I and II, the ε torsion angle is *trans*, but the P–O3′ torsion angle (ζ) and the P–O5′ torsion angle (α) are +44 and +58°, respectively. Both are close to a *gauche*$^+$ conformation, whereas the same torsion angles in standard B-DNA are approximately –90 and –60°, respectively. In addition to creating a sharp bend at the junction branch point, the g^+/g^+ conformation for torsion angles ζ and α causes the phosphate group to swivel 180° towards the center of the junction. This phosphate inversion is evident in Figure 3, which shows the two pairs of dyad-related non-bridging phosphate oxygen atoms facing one another across the center of the junction with a separation of about 14 Å.

The unstacking of bases at the junction branch point causes a dramatic increase in solvent accessibility to the planar purine and pyrimidine surfaces. Both the Cre N-terminal domain tetramer on the minor groove face of the junction and the C-terminal domain tetramer on the major groove face form channels that are large enough to allow solvent access to the junction branch point (Figure 4). An average of ~170 Å^2 of base pair surface is accessible to solvent for the four AT base pairs at the branch point in the Cre–HJ complexes. For comparison, the same base

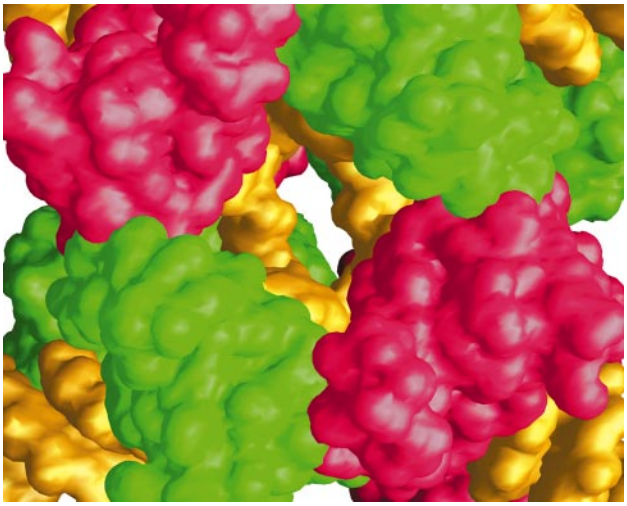

Fig. 4. Solvent accessible surface at the center of the Cre–HJ complex, as viewed from the N-terminal, minor groove side of the junction. The surfaces of adenine and thymine bases at the branch point as well as the inverted phosphates are clearly visible through a solvent-accessible tunnel formed by the N-terminal domains. The orientation of the complex is the same as in Figure 2A.

pair type near the center of the *loxP*-inverted repeat region has only ~80 Å^2 of base-pair surface accessible to solvent (ignoring the bound recombinase). Chemical probing of solvent accessibility of bases at or near the branch point of four-way DNA junctions has proved to be a revealing tool in the investigation of both free junctions and protein–junction complexes (Duckett *et al*., 1988; Bennett and West, 1995b; Arciszewska *et al*., 1997). As discussed later, the arm geometry and solvent accessibility of bases near the branch point of the Cre–HJ complexes allow us to draw close parallels with models resulting from some studies and sharp contrasts with others.

Given the requirement for large movements of DNA

strands within the *loxP* crossover region during the strand exchange step of the site-specific recombination pathway, it is perhaps not surprising that we observe few protein–DNA contacts in this region of the Cre–HJ structures or in the covalent Cre–*loxA* complex previously reported. Within the 4 bp adjacent to the branch point in each arm of the junction, there are no direct contacts of the protein to bases in the major or minor grooves. Each arm of the junction is, however, contacted by recombinase side chains on the phosphate backbone in the crossover region. Three arginine side chains from the N-terminal recombinase helix D (Arg100, Arg101 and Arg106) interact with the phosphates between bases 1–4 in each arm (Figure 2D). The contacts are slightly closer for the recombinase subunits bound to arms I and III than for the subunits bound to arms II and IV. In addition, the phosphate backbone of DNA residues 1′, 2′ and 3′ is closely approached by a poorly ordered loop that includes residues Thr200–Thr206. The electron density for these residues is of insufficient quality to reliably identify DNA contacts. This loop was better-ordered in the covalent Cre–*loxA* complex and we were able to identify Lys201 in the minor groove of the DNA, interacting with the base pair adjacent to the scissile phosphate (Guo *et al*., 1997). Lys201 is conserved in the Int family, except for a small number of proteins with arginine in this position (Nunes-Düby *et al*., 1998).

The central phosphate between arms II and III forms the most conspicuous backbone contact within the crossover region, hydrogen bonding with Arg121 in the Cre–HJ1 complex, and interacting from a slightly larger distance (~4 Å) in the Cre–HJ2 complex (Figure 3B). An equivalent interaction is not made with the central phosphate between arms I and II, which has turned towards the center of the junction away from Arg121. As a consequence of the lack of hydrophobic and major–minor groove interactions within the crossover region, the mobility of DNA in the center of the junction is increased as reflected in larger average thermal parameters (<B> = 49 Å^2 in the crossover region versus 35 Å^2 in the 13 bp inverted repeat region of the Cre–HJ1 complex) following crystallographic refinement.

The active sites in the Cre–HJ intermediate

The Cre recombinase active site contains four amino acids that are strictly conserved in the Int family (Abremski and Hoess, 1992). Three of these residues, Arg173, His289 and Arg292, make up the 'RHR triad' and form hydrogen bonds to the scissile phosphate during strand cleavage (Guo *et al*., 1997). The fourth conserved residue is the nucleophile Tyr324. The side chain of this residue is poised to attack the scissile phosphate, hydrogen bonding to the neighboring phosphate that lies at the boundary of the 13 bp inverted repeat of the *loxP* site. Trp315, an amino acid that is more often represented by histidine in the Int family, also contributes to formation of the active site and forms a hydrogen bond to the scissile phosphate. A similar active site geometry was observed in the recently described structures of two eukaryotic topoisomerases (Cheng *et al*., 1998; Stewart *et al*., 1998).

The Cre–HJ1 and Cre–HJ2 complexes are only twofold-symmetric overall, therefore the active sites formed by each of the two independent recombinases (i.e. the green and red subunits in Figure 2) might be expected to show some stereochemical differences. Indeed, during cleavage of the HJ substrate, only one pair of dyad-related Tyr324 side chains will cleave the DNA to initiate strand exchange. The geometric arrangement of residues in both active sites of the Cre–HJ1 and both active sites of the Cre–HJ2 complex are nevertheless quite similar, with r.m.s. deviations of 0.5 Å for α carbon atoms. There are modest differences in side chain torsion angles among the active sites in the two complexes, but it is not yet clear whether these have functional relevance, given the missing scissile phosphates in the Cre–HJ2 structure and the altered DNA sequence used in Cre–HJ1.

The largest difference between the Cre–HJ1 active sites and the active site observed in a covalent Cre–*loxA* complex where cleavage had not occurred (Guo *et al*., 1997) involves the scissile phosphates in arms II and IV. The trajectories of the DNA strands are slightly different in the HJ complex and result in an ~1 Å shift in the position of the scissile phosphates. The consequences of this shift are an increase in the Tyr324-OH to phosphate distance from 5 Å observed in arms I and III to 6 Å in arms II and IV, and the removal of this phosphate group from ideal hydrogen bonding position. It will be important to verify this observation with further studies using alternative DNA sequences and other inactive Cre mutants in order to establish this phosphate relocation as a component of the stereochemical switch used to deactivate cleavage of a given pair of strands in the Holliday intermediate.

Although the active site geometries are similar for the two unique arms in the Cre–HJ complexes, the conformations of the linker peptides between the two C-terminal α-helices M (residues 319–326) and N (residues 333–340) adjacent to the active sites are quite different. Helix N is buried in a hydrophobic pocket of a neighboring recombinase subunit, forming one of the four cyclic exchanges of helices that link the arms of the recombinase-junction tetramer on the C-terminal side of the junction (Figure 2C). Catalytic Tyr324 is located at the end of helix M and this entire helix shifts by ~3 Å during cleavage of the DNA substrate relative to the same helix in the non-cleaving active site (Guo *et al*., 1997). The conformation and flexibility of the peptide segment tethered to helix M could therefore play a crucial role in regulating cleavage activity. This linker adopts a compact and well-ordered conformation in the connection between arms I and II of the junction and a more extended and poorly ordered conformation between arms II and III.

Discussion

Intermediates in Cre–loxP recombination have similar architectures

In this report, we have described the structures of two Cre–HJ complexes that together provide a model for this intermediate in the Cre–*loxP* site-specific recombination pathway. Previously, we described the structure of a covalent Cre–DNA intermediate (Figure 1A, ii and v) in which two of the four recombinases in a synaptic tetramer of symmetrized *loxP* sites had cleaved a suicide DNA substrate to form 3′ phosphotyrosine linkages (Guo *et al*., 1997). Surprisingly, the 5′ ends of the cleaved DNA

strands were melted away from their complementary strands and directed towards the center of that synaptic complex, which appeared to be caught in the act of exchanging strands. Because a 5′ cytidine had been removed in forming the phosphotyrosine bond between Cre and DNA, the 5′ strands were one base too short to reach the active sites of the partner substrates. In the Cre–HJ2 structure reported here, this idea has been taken one step further. The *loxS6* sites used to form HJ2 (Figure 1B) contain nicks where the phosphotyrosine linkage would be formed. In addition to melting and directing the 5′-ends of the crossing strands towards the middle of the synapse, Cre has mediated formation of a HJ without covalent ligation to the partner substrate. Such a structure was in fact suggested by Burgin and Nash (1995) as a possible product of Int recombination using 5′-bridging phosphorothioate-substituted suicide substrates.

Since the sequences of the two *loxS6* sites that form HJ2 are identical, this protein–HJ complex could, in principle, branch-migrate by 1–2 bp. The branch point of the observed crystal structure, however, is located at the center of the sites. In contrast, the immobile HJ1 DNA substrate was pre-formed before binding by Cre and was designed to have its branch point located at the center of the crossover region. The fact that the Cre–HJ1 and Cre–HJ2 structures are similar to one another and to the structure of the covalent intermediate (ii) in Figure 1A indicates that on the level of quaternary structure, the recombinase tetramer pre-organizes the DNA arms for HJ formation and stabilizes a single branch-point isomer of the DNA intermediate. We cannot rule out the possibility that other branch-point isomers exist for the Cre–HJ complex during site-specific recombination and that crystallization has selected a single low energy structure for Cre–HJ2. However, the architectures of Cre–HJ1 and Cre–HJ2 argue strongly against stable recombinase tetramers with alternative branch-point positions. In order to branch-migrate a *loxP*-derived HJ substrate with Cre recombinase-bound, the extensive protein–protein contacts (~5000 Å^2 of buried protein surface) that stabilize the tetramer would have to be disrupted. Alternatively, if the DNA arms were allowed to rotate within the bound recombinase domains in order to accomplish branch migration, the specific Cre–DNA interface would be disrupted. Since these possibilities seem unlikely, we would argue that energy barriers to branch migration are too high in the Cre–*loxP* system. However, restricted branch migration may occur more readily in the Int and Flp systems, where the spacings between cleavage sites are 1–2 bp longer and the nature of the intersubunit contacts may be somewhat different.

Recent studies in the λ-integrase, XerC/XerD and Flp systems support a model of limited branch migration in the site-specific recombination pathway. Experiments performed using immobilized HJ substrates (Dixon and Sadowski, 1993; Arciszewska *et al*., 1995; Lee *et al*., 1995; Nunes-Düby *et al*., 1995) led to a model that does not require branch migration of the Holliday intermediate between cleavage sites for resolution of the junction. Experiments using nicked DNA substrates (Lee and Jayaram, 1995; Zhu *et al*., 1995), suicide phosphorothioate-containing substrates (Burgin and Nash, 1995), and suicide half-site substrates (Nunes-Düby *et al*., 1997) provided evidence that homology between recombining sites is first sensed following initial strand cleavage but preceding the ligation step that completes the first round of strand exchange. We believe that the Cre–HJ2 structure represents what Burgin and Nash (1995) refer to as the 'communication' step that precedes strand ligation. The Cre–HJ2 complex also represents an intermediate in the 'strand-swapping model' of site-specific recombination proposed by Landy and coworkers (Nunes-Düby *et al*., 1995) where the initial swapping step has just occurred (but without ligation) and homology between three out of the six base pairs between cleavage sites has been successfully tested.

The Cre recombinase-bound versus free HJ

The properties and behavior of free HJs have been extensively studied over the past fifteen years using a number of experimental techniques, including electrophoresis, NMR, fluorescence resonance energy transfer and electric dichroism (reviewed in Lilley and Clegg, 1993; Seeman and Kallenbach, 1994). As a result of these studies, a consensus view of the solution structure of four-way junctions has emerged (Figure 5). In the presence of physiological concentrations of magnesium ions (≥100 μM), the arms of the junction are organized into two quasi-continuous duplexes that are fully stacked at the junction branch point. The helical stacks form a right-handed cross with an acute angle of roughly 60°, leading to an overall dyad-symmetric shape termed the 'stacked-X' (Churchill *et al*., 1988; Duckett *et al*., 1988). The pairs of arms in this four-way junction are distinctly non-coplanar (Figure 5B). Within each helical stack, a continuous strand forms an uninterrupted helix through the branch-point region. The opposite crossing strand has a sharp discontinuity in its helicoidal path at the junction branch point, where it crosses from one stack to the other. The two strands in a given arm therefore have distinctly different roles at the branch point. There are several ways of constructing such a junction; the continuous strands can be in parallel or antiparallel orientation with respect to alignment of the 5′–3′ strand direction and the roles of crossing and continuous strands can be switched for each orientation, leading to strand equivalence during recombination (Sigal and Alberts, 1972). Experimental and theoretical studies indicate that the antiparallel alignment is favored (von Kitzing *et al*., 1990; Lu *et al*., 1991), but a given antiparallel junction can exist as one of two crossover isomers depending on which strand crosses between helical stacks. These crossover isomers differ by relatively small free energies that depend on the sequences present at and near the branch point and both isomers can be populated for some junction sequences (Miick *et al*., 1997).

In the presence of EDTA and low salt, the free HJ adopts a fourfold-symmetric square conformation, where the arms of the junction point towards the vertices of a square (Figure 5A; Duckett *et al*., 1988; Clegg *et al*., 1994). In this case, there is an unstacking of bases at the branch point which is detectable by chemical probing with osmium tetroxide. In contrast, bases at the branch point of the stacked-X junction do not show this sensitivity to oxidation. In addition to direct experimental analyses of synthetic four-way junctions, analyses of packing in crystal structures of DNA duplexes (Goodsell *et al*., 1995; Wood

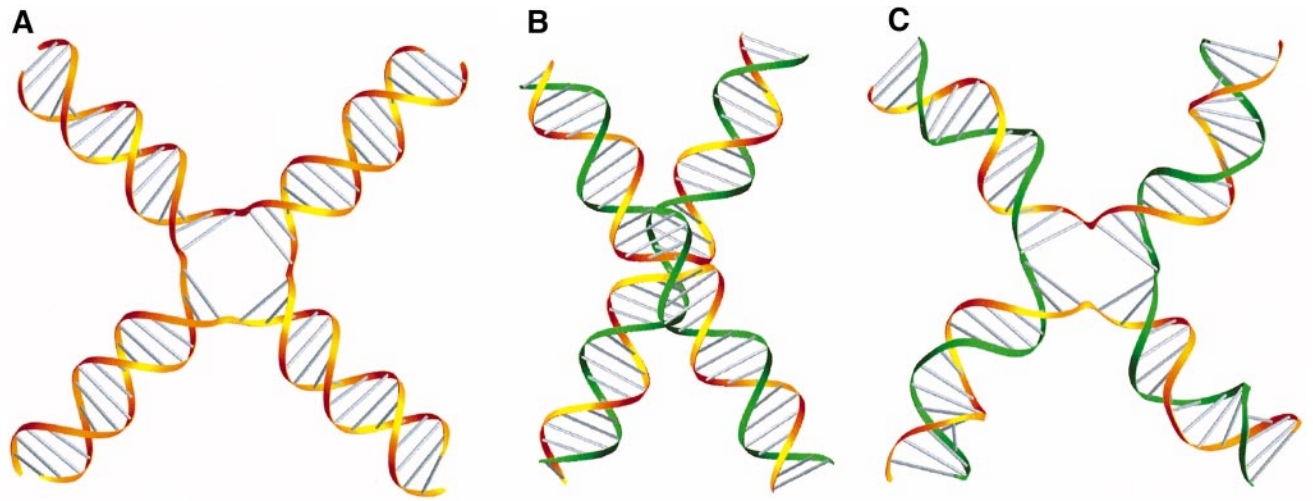

Fig. 5. Comparison of the (**A**) extended square, (**B**) stacked-X, and (**C**) Cre–HJ structures. Structures (A) and (B) were obtained by model-building with approximate geometries taken from theoretical models (von Kitzing *et al*., 1990). Continuous strands are green and crossing strands are gold for (B) and (C). The junction in (C) has been rotated 90° with respect to Figure 3.

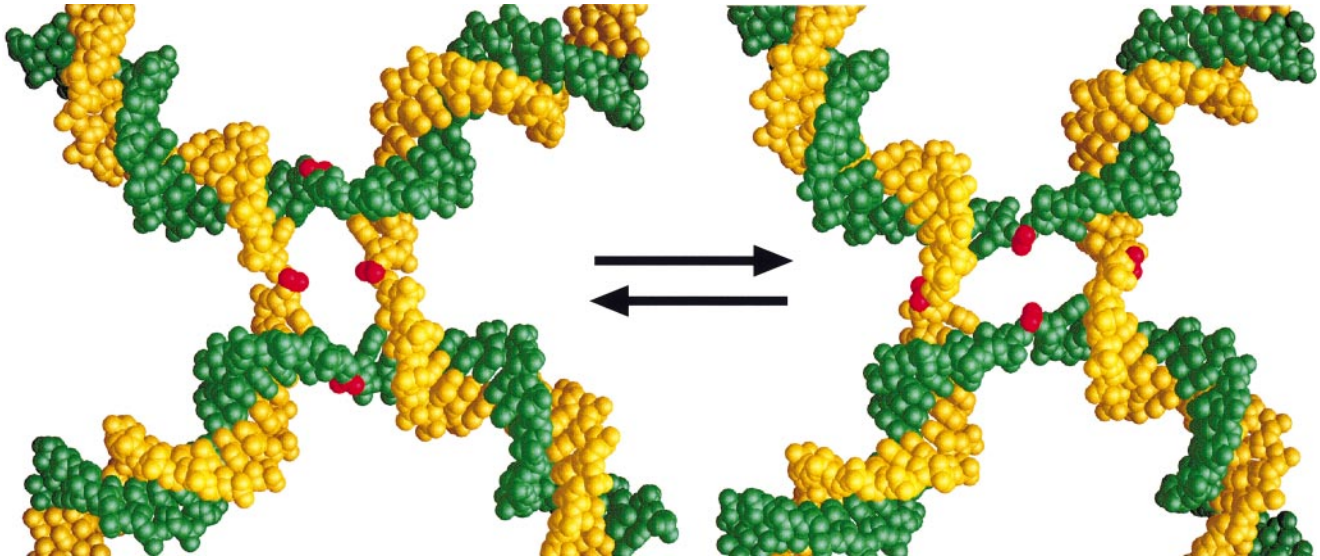

Fig. 6. Structural model for crossover isomerization of the Cre–HJ intermediate, based on the Cre–HJ1 and Cre–HJ2 complexes described here. Isomerization results in an exchange in the roles of the two pairs of 'continuous' and 'crossing' strands. This model provides a structural basis for the central isomerization step of the 'strand-swapping model' proposed by Nunes-Düby *et al.* (1995).

et al., 1997) and theoretical studies (von Kitzing *et al*., 1990) have contributed to the current models for the unbound HJ.

The Cre–HJ1 and Cre–HJ2 complexes have features in common with both the stacked-X and the square HJ models (Figure 5). Like the stacked-X model, the Cre–HJ structures are dyad-symmetric with two acute and two obtuse inter-arm angles. The assignment of continuous and crossing strands is less obvious for the Cre-bound junction, however, since both strands are discontinuous at the branch point. Based on the differences in tilt angles, phosphodiester torsion angles, and degree of unstacking, we refer to the strands bridging arms II–III and I–IV as 'continuous' and the strands bridging arms I–II and III–IV as 'crossing', in keeping with current terminology. It is also clear that, like the stacked-X model, two crossover isomers are possible that correspond to swapping the roles between continuous and crossing strands. With respect to the overall Cre–DNA assembly, crossover isomerization could be readily accomplished by compressing the obtuse arm II–III and I–IV angles, opening the acute arm I–II and III–IV angles, and rotating the central phosphate linkages into opposite orientations (Figure 6). Because the Cre–Cre contacts in the HJ-bound tetramer are nearly fourfold-symmetric, only a small adjustment in the network of protein–protein and protein–DNA interactions is required to accomplish this rather subtle transition. The relationship between crossover isomer and direction of junction resolution is of central importance in understanding the function of HJ-resolving enzymes and this has

been a central theme in recent studies of the Int family site-specific recombination pathway (Arciszewska *et al*., 1997; Azaro and Landy, 1997; Lee *et al*., 1998).

The Cre–HJ1 and Cre–HJ2 complexes share more gross architectural features with the square, low-salt form of the free HJ. Like the square junction, the arms of the Cre–HJ complexes lie approximately in the same plane and the bases at the branch point have become distinctly unstacked. In the case of the Cre–HJ2 complex, the DNA sequence of the junction that is formed is compatible with formation of a strictly fourfold-symmetric Cre–HJ complex. The fact that we observe a dyad-symmetric distortion from a fourfold-symmetric complex suggests that a fourfold-symmetric structure may exist only as a transient intermediate during crossover isomerization and that the lower symmetry of both the covalent Cre–DNA intermediate (Guo *et al*., 1997) and the Cre–Holliday intermediates described here may have functional significance.

Comparison with other protein–HJ complex studies

The flexible nature of the HJ in solution is reflected in its ability to adopt different conformations when bound by junction-specific proteins (reviewed in White *et al*., 1997). RuvC, an enzyme that cleaves HJs in conjunction with RuvAB-mediated branch migration in bacteria (Dunderdale *et al*., 1991; Iwasaki *et al*., 1991, 1992; Tsaneva *et al*., 1992), binds to the junction as a dimer and introduces symmetry-related nicks in the pair of continuous strands (Bennett and West, 1995a). West and colleagues (Bennett and West, 1995b) have shown using electrophoresis of junctions with pairs of short and long arms that the RuvC–HJ complex is twofold-symmetric. Chemical probing of thymine bases with potassium permanganate also indicated that the crossover point forms an unstacked, open conformation. We find the pattern of electrophoretic mobilities and the evidence of solvent accessibility at the branch point in the RuvC–HJ complex to be remarkably consistent with the geometric and stereochemical features in the Cre–HJ1 and Cre–HJ2 complexes. Given the close relationship between the nearly planar dyad-symmetric Cre-bound HJ with the fourfold-symmetric junction thought to exist in the RuvAB complex (Rice *et al*., 1997; Yu *et al*., 1997), it seems plausible that RuvC may bind to a HJ with a similar architecture as seen with Cre recombinase. Such a structure was in fact proposed in schematic form by Bennett and West (1995b).

In contrast to the RuvC–HJ complex, studies of the complexes of RuvA and phage T7 endonuclease I with synthetic HJs suggest that their structures differ from what we observe with Cre recombinase. RuvA binds HJs as a tetramer and recruits the hexameric ring helicase RuvB to drive branch migration (Tsaneva *et al*., 1992). In studies of RuvA–HJ complexes, Parsons *et al.* (1995) identified a fourfold-symmetric complex that was unstacked at the branch point, as expected for a square junction conformation. A still different complex was observed by Lilley and coworkers (Duckett *et al*., 1995) in a study of T7 endonuclease I, which binds and cleaves HJs as a dimer. In this case, the protein–junction complex is dyad-symmetric, but with apparent coaxial stacking of pairs of helical arms in a structure that appears more closely related to the stacked-X conformation of unbound junctions.

Within the λ-integrase family, XerC/XerD binding to synthetic HJs has been shown using permanganate probing to induce a pronounced unstacking of bases at the branch point relative to that observed in free junctions, indicating that these recombinases also change the unbound stacked-X structure (Arciszewska *et al*., 1997). The same studies also showed that the branch point of the XerC/XerD-bound junction is located at the center of the crossover region, as we observe in the Cre–HJ complexes. Recent studies in the Flp recombinase system have also shown that the recombinase unfolds HJ substrates and exposes bases at the branch point to solvent (Lee *et al*., 1998). Flp recombinase is the only Int family system for which electrophoretic mobilities and deduced symmetries of the recombinase–HJ complexes have been reported (Lee *et al*., 1998). Flp–HJ complexes show approximately fourfold symmetry, indicating a nearly planar arrangement of junction arms and either a complex with exact fourfold symmetry or one that readily converts between dyad-symmetric forms. Taken together with the recently established structural similarity of the proteins in the Int family (Guo *et al*., 1997; Hickman *et al*., 1997; Kwon *et al*., 1997; Subramanya *et al*., 1997), it seems likely that the XerC/XerD, λ-integrase, and Flp–HJ intermediate complexes are closely related to the Cre–HJ1 and Cre–HJ2 structures described here.

A model for strand cleavage-coupled HJ isomerization

A central question in the Int family site-specific recombination pathway involves the asymmetry of cleavage of both the initial duplex substrates and of the Holliday intermediate. Only two of the four tyrosine nucleophiles are permitted to cleave the DNA substrates at each step. How is the appropriate pair of tyrosine residues activated and how is the alternative pair prevented from cleaving? In the resolution of the Holliday intermediate, the choice of which pair of tyrosine residues cleaves (or equivalently, which pair of DNA strands are cleaved) defines whether recombinant products or starting substrates will be formed. In early models of Int family site-specific recombination, branch migration of the Holliday intermediate between cleavage sites was suggested to be the switch responsible for deactivation of one pair of recombinases and activation of the other pair (Weisberg *et al*., 1983; Kitts and Nash, 1987; Nunes-Düby *et al*., 1987). This model was based on the requirement for homology between cleavage sites and it was assumed that branch migration would bring the second pair of scissile phosphates into mutual proximity to allow a second strand exchange and formation of recombinant products. In a more recent model of the Int family pathway, the optimal location of the Holliday intermediate branch point is limited to the central 1–2 bases of the crossover region, and homology between recombining sites is instead tested during the strand transfer step (Nunes-Düby *et al*., 1995). In this 'strand-swapping' model, the switch between top and bottom strand cleavage required for resolution of the Holliday intermediate to form recombinants is linked to an isomerization of the HJ structure that results in the roles of the DNA strands. Both Int and the XerC/XerD

recombinases have been shown to preferentially cleave the crossing strands in HJ substrates that have been biased to exist predominantly as a single crossover isomer in the unbound state. Using tethered junction arms (Arciszewska *et al*., 1997) and branch-point sequences designed to bias a single crossover isomer (Azaro and Landy, 1997), strand cleavage preference has been programmed by selecting the desired crossover isomer.

A related experiment was reported by Hoess *et al*. (1987) in a study of Cre mutants that accumulate Holliday intermediates but are unable to resolve them efficiently. Using a linear substrate containing two *loxP* sites, α-structures were formed by the Cre mutants that could then be efficiently resolved to products by wild-type Cre. Resolution of the α-structure gave primarily recombinant products, whereas resolution of a χ-structure obtained by digestion of the loop connecting adjacent junction arms in the α-structure gave an equal mixture of recombinants and starting substrates. In the context of the Int and XerC/XerD studies, it seems likely that the Cre-catalyzed α-structure effectively biased the Holliday intermediate towards a single crossover isomer. The χ-structure in this case was presumably free to isomerize and populate both isomeric forms. Flp recombinase shares this apparent lack of resolution bias in studies with immobile Holliday substrates (Lee *et al*., 1998), yet the existence of two distinct kinetic forms of the Flp–HJ intermediate has been observed (Meyer *et al*., 1990).

The Cre–HJ1 and Cre–HJ2 complexes described here strongly support the idea that strand cleavage specificity is coupled to a HJ crossover isomer. We propose that crossover isomerization in the Cre–*loxP* pathway also involves an exchange in the roles of the DNA strands, but requires only subtle changes in quaternary structure and no restacking of the helical arms (Figure 6). An important feature of this mechanism is that the extensive interface (~5000 Å^2 of buried protein surface) formed between the recombinase subunits in the synaptic tetramer would be preserved during isomerization. The energetic barrier to crossover isomerization may in fact be quite small for Cre recombinase. Assuming that the isomeric forms have roughly similar free energies, this would explain the lack of bias observed in the resolution of Holliday intermediates for this system and for the Flp system. In a kinetic analysis of Flp site-specific recombination, Meyer *et al.* (1990) were able to observe this isomerization process and showed that the interconversion of two Flp–HJ kinetic intermediates becomes rate-limiting at 0°C.

The stereochemical coupling between strand cleavage and HJ isomer is likely to involve the peptide linker connecting the C-terminal helix N to the catalytic tyrosine-containing helix M. Where helix N is exchanged between Cre protomers on arms that make an acute angle, the peptide linking helices M and N is well-ordered and adopts a compact conformation. In contrast, the peptide linking helices M and N across two arms that make an obtuse angle adopts a more extended, unrestrained conformation that may provide the required mobility for helix M to initiate cleavage. This 'activated' active site is positioned on what we have termed the crossing strand in the Cre–HJ complex and corresponds to the same active site where a covalent 3′-phosphotyrosine linkage to a suicide DNA substrate was formed in a previous study (Guo *et al*., 1997). This arrangement is consistent with reports that Int and XerC/XerD both cleave the strand that is crossed in free Holliday substrates. We have assumed in these cases that the crossed strands in the protein-bound junctions maintain their identities as those bridging the arms that are closest in space, as in the stacked-X model and the Cre–HJ complex. The details of the stereochemical coupling between strand cleavage and HJ isomer are the subject of a separate study in this laboratory (F.Guo and G.Van Duyne, in preparation).

Although the isomerization model discussed above suggests a potential link between the stereochemical activation of a given pair of catalytic tyrosine residues and the isomeric state of the Cre–HJ complex, it does not address how cleavage preferences are related to the DNA sequences of the substrates. Int and Cre, for example, have been shown to proceed with a defined order of strand exchanges (Hoess *et al*., 1987; Kitts and Nash, 1988), which for Cre must be dictated initially by the DNA sequences in the recombining sites. Similarly, Landy and coworkers (Azaro and Landy, 1997) have demonstrated that specific HJ substrate isomers for Int can be dictated by the sequences present at the branch point. It is not clear from the Cre–HJ complexes why certain bases might prefer to adopt the crossing rather than continuous strand positions. Structural models of the recombinase proteins bound to their duplex substrates prior to cleavage and further studies of the Holliday intermediate complexes with wild-type sequences and at higher resolution should provide yet more pieces to the complex mechanistic puzzle in the Int family.

Materials and methods

Preparation of HJs

The HJ1 junction was formed from four 35-mer oligonucleotides using the sequences shown in Figure 1B. Oligonucleotides were synthesized by conventional phosphoramidite chemistry, and purified on a Hamilton PRP-1 reversed phase HPLC column with the dimethoxytrityl protecting group attached. Purified oligonucleotides were concentrated on a 2 ml Q-sepharose column (Pharmacia), deprotected with 0.5% trifluoroacetic acid, and eluted with 20 mM Tris–HCl pH 7.8, 1 M NaCl, 0.5 mM EDTA. Following concentration and buffer exchange into 20 mM Tris pH 7.8, 50 mM NaCl, 0.5 mM EDTA by Centricon-3 concentrators (Amicon), the concentrations of individual strands were determined by absorbance at 260 nm and used to prepare a stoichiometric mixture at ~1 mM concentration. The strands were annealed in a Perkin-Elmer 2400 thermal cycler using a linear 1°/min gradient from 80–20°C. The yield of four-way junction was >95% as judged by the formation of a single band with low mobility on 22% non-denaturing polyacrylamide gels. The *loxS6* sites were constructed by annealing the 15-mer 5′-TATAACTTCGTATAG-3′ and the 20-mer 5′-CATATGCTATACGAA-GTTAT-3′ following purification as described above.

Protein purification and crystallization of Cre–HJ1 and Cre–HJ2

The CreR173K mutant was generated by PCR mutagenesis using the wild-type expression construct in pET21a (Novagen) and the single-site mutation was verified by sequencing the entire coding region. Wild-type Cre and Cre R173K were purified, concentrated and stored as described (Guo *et al*., 1997). Crystals of both the HJ1 and HJ2 complexes were obtained by the hanging-drop vapor diffusion method. For Cre R173K-HJ1, 4 μl drops containing 50 μM protein, 20 μM HJ1, 50 mM sodium acetate pH 5.6, 80 mM $MgCl_2$, 2 mM dithiothreitol (DTT) and 22% 2-methyl-2,4-pentanediol (MPD) were equilibrated against 1 ml reservoirs containing 50 mM sodium acetate pH 5.0, 26% MPD, 80 mM $MgCl_2$ at 18°C. Thin plates (150×150×20 μm) were obtained after two weeks. For Cre–HJ2, 6 μl drops containing 65 μM Cre, 81 μM *loxS6*

half-site, 25 mM sodium acetate pH 5.7, 12.5% MPD, 5 mM $CaCl_2$, 120 mM NaCl, 2.5 mM DTT and 0.5 mM spermine were equilibrated against 1 ml reservoirs containing 50 mM sodium acetate pH 5.0, 24% MPD, 20 mM $CaCl_2$, 200 mM NaCl. Thin plates grew to a maximum size of 200×200×30 μm after one month at 4°C.

Data collection and structure determination

Crystals were flash frozen in propane and stored in liquid nitrogen prior to data collection as described (Van Duyne *et al*., 1996). All diffraction data were measured at 100 K at the NSLS X25 beamline using a MAR research detector and processed using DENZO (Otwinowski, 1994). Cre–HJ1 crystals are orthorhombic, space group $C222_1$, a = 106.3, b = 122.7, c = 179.5 Å, with two Cre recombinase molecules and one half of the HJ in the asymmetric unit. Data are 99% complete to 2.7 Å, with R_{sym}= 0.072 and average redundancy = 4.7. Cre–HJ2 crystals are orthorhombic, space group $C222_1$, a = 106.9, b = 122.5, c = 180.0 Å, with two Cre molecules and one *loxS6* site in the asymmetric unit. Data are 94% complete to 2.5 Å, with R_{sym} = 0.053 and average redundancy = 2.8. Because Cre–HJ1 and Cre–HJ2 are nearly isomorphous with the Cre–*loxA* structure (Guo *et al*., 1997), both were determined by rigid body refinement of the Cre–*loxA* domains at 3 Å, where each recombinase domain and the 13 bp inverted repeat segment of the DNA sites were treated as rigid bodies. The C-terminal residues (316–341) of the recombinases and base pairs 1–4 and 1′–4′ were excluded from the rigid body refinement and were not included in the structure refinement until they could be unambiguously fit into unbiased σ_A-weighted $2F_o–F_c$ electron density maps (Read, 1986).

Refinement of Cre–HJ1 and Cre–HJ2

The Cre–HJ1 and Cre–HJ2 structures were refined using torsion angle dynamics followed by conjugate gradients minimization in XPLOR (Rice and Brunger, 1996) at 3 and 2.7 Å, and for Cre–HJ2, at 2.5 Å resolution. Refinement followed by model fitting to σ_A-weighted $2F_o–F_c$ maps using O (Jones *et al*., 1991) was repeated at each resolution.

A crystallographic dyad relates the two halves of the Cre–HJ1 complex, therefore the DNA in this structure is at least twofold disordered in the 4 bp adjacent to the branch point in each arm and in a single base pair located within the Cre binding site (Figure 1). Two models for disorder were considered. In model 1, only one isomer of the HJ is present in the crystal: either the one shown in Figure 3 or the alternative isomer with arms II and III forming an acute angle. In either case, the Cre–HJ complex could pack in the crystal in two identical ways that are not influenced by the asymmetric base pairs and these bases would therefore be twofold disordered. In model 2, both isomers are equally populated; there are four equivalent ways to assign the arms of the junction, and the asymmetric bases are fourfold disordered. Because the minimal requirement of twofold disorder superimposes a purine and a pyrimidine at each asymmetric site, it was not possible to distinguish between these possibilities based on simple inspection of the electron density at 2.7 Å. The difficulty was compounded by the higher than average thermal parameters for DNA residues near the branch point. The models could also not be distinguished by comparing cross-validated R-factors following refinement of the structures while imposing the appropriate non-crystallographic symmetry-constrained disorder model. The final refinement of the Cre–HJ1 complex was performed using a twofold disorder model where each of the asymmetric bases in the HJ1 arms are present at 50% occupancy, using the arm and sequence assignments shown in Figures 1 and 3. The final refinement model consists of residues 19–197 and 207–341 for Cre subunit A (bound to arm II), residues 19–197, 209–329 and 334–341 for subunit B (arm I), all DNA residues except the 5′ thymidine on arms II and IV, 239 water molecules with good hydrogen bonding stereochemistry at unit occupancy and grouped thermal parameters (two per residue). The refinement for the Cre–HJ1 complex converged at R = 0.22 and R_{free} = 0.29 (5% of data) using data to 2.7 Å ($F \geqslant 2\sigma$) following a correction for the contribution of bulk solvent (Jiang and Brunger, 1994). R.m.s. deviations from ideal bond lengths and bond angles are 0.005 Å and 0.95°, respectively.

The refinement of the Cre–HJ2 complex did not require a model for disorder, since the DNA sequence is twofold-symmetric. The final refinement model included residues 19–198, 210–275 and 283–341 for Cre subunit A, residues 19–198, 207–326 and 333–341 for subunit B, all DNA residues except the 5′-thymidine on arm II, 361 water molecules with good hydrogen bonding stereochemistry at unit occupancy, and tightly restrained individual thermal parameters. The refinement for the Cre–HJ2 complex converged at R = 0.20 and R_{free} = 0.27 ($F \geqslant 2\sigma$) using data to 2.5 Å following a bulk solvent correction. r.m.s. deviations from ideal bond lengths and bond angles are 0.007 Å and 1.3°, respectively.

Geometric calculations and illustrations

DNA helicoidal parameters, torsion angles and curvature were computed using the program CURVES (Lavery and Sklenar, 1988). The angles between junction arms were defined as the angles between the principle moments of inertia of the duplex arm segments and were calculated using local programs. Figures 2, 3 and 5 were produced using RIBBONS (Carson, 1991); Figures 4 and 6 were produced with GRASP (Nicholls and Honig, 1993).

Cre–HJ1 and Cre–HJ2 structures

Coordinates for the Cre–HJ1 and Cre–HJ2 structures will be deposited with the Brookhaven Protein Data Bank under accession codes 3CRX and 2CRX, respectively.

Acknowledgements

We thank Lonny Berman for access to and help with the NSLS X25 beamline, Jianping Ni for assistance with data collection and Mark Lemmon, Hillary Nelson, Art Landy and Simone Nünes-Duby for helpful suggestions with the manuscript. This work was supported by NIH grant GM-55041 and an American Cancer Society Institutional Research Grant to the University of Pennsylvania.

References

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

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

Arciszewska,L., Grainge,I. and Sherratt,D. (1995) Effects of Holliday junction position on Xer-mediated recombination *in vitro*. *EMBO J.*, **14**, 2651–2660.

Arciszewska,L.K., Grainge,I. and Sherratt,D.J. (1997) Action of site-specific recombinases XerC and XerD on tethered Holliday junctions. *EMBO J.*, **16**, 3731–3743.

Argos,P. *et al.* (1986) The integrase family of site-specific recombinases: regional similarities and global diversity. *EMBO J.*, **5**, 433–440.

Azaro,M.A. and Landy,A. (1997) The isomeric preference of Holliday junctions influences resolution bias by lambda integrase. *EMBO J.*, **16**, 3744–3755.

Bennett,R.J. and West,S.C. (1995a) RuvC protein resolves Holliday junctions via cleavage of the continuous (noncrossover) strands. *Proc. Natl Acad. Sci. USA*, **92**, 5635–5639.

Bennett,R.J. and West,S.C. (1995b) Structural analysis of the RuvC–Holliday junction complex reveals an unfolded junction. *J. Mol. Biol.*, **252**, 213–226.

Blakely,G., May,G., McCulloch,R., Arciszewska,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.

Burgin,A.B. and Nash,H.A. (1995) Suicide substrates reveal properties of the homology-dependent steps during integrative recombination of bacteriophage lambda. *Curr. Biol.*, **5**, 1312–1321.

Campbell,A. (1962) Episomes. *Adv. Genet*., **11**, 101–145.

Carson,M. (1991) Ribbons 2.0. *J. Appl. Crystallogr.*, **24**, 958–961.

Cheng,C., Kussie,P., Pavletich,N. and Shuman,S. (1998) Conservation of structure and mechanism between eukaryotic topoisomerase I and site-specific recombinases. *Cell*, **92**, 841–850.

Churchill,M.E.A., Tullius,T.D., Kallenbach,N.R. and Seeman,N.C. (1988) A Holliday recombination intermediate is twofold symmetric. *Proc. Natl Acad. Sci. USA*, **85**, 4653–4656.

Clegg,R.M., Murchie,A.I.H. and Lilley,D.M.J. (1994) The solution structure of the four-way DNA junction at low-salt conditions: a fluorescence resonance energy transfer analysis. *Biophys. J.*, **66**, 99–109.

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

de Massey,B., Studier,F.W., Dorgai,L., Appelbaum,F. and Weisberg,R.A. (1984) Enzymes and the sites of genetic recombination: studies with gene-3 endonuclease of phage T7 and with site-affinity mutants of phage lambda. *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 715–726.

Dixon,J.E. and Sadowski,P.D. (1993) Resolution of synthetic chi structures by the FLP site-specific recombinase. *J. Mol. Biol.*, **234**, 522–533.

Duckett,D.R., Murchie,A.I., Diekmann,S., von,K.E., Kemper,B. and Lilley,D.M. (1988) The structure of the Holliday junction, and its resolution. *Cell*, **55**, 79–89.

Duckett,D.R., Panis,M.E.J. and Lilley,D.M.J. (1995) Binding of the junction-resolving enzyme bacteriophage T7 endonuclease I to DNA: separation of binding and catalysis by mutation. *J. Mol. Biol.*, **246**, 95–107.

Dunderdale,H.J., Benson,F.E., Parsons,C.A., Sharples,G.J., Lloyd,R.G. and West,S.C. (1991) Formation and resolution of recombination intermediates by *E.coli* RecA and RuvC proteins. *Nature*, **354**, 506–510.

Elborough,K.M. and West,S.C. (1990) Resolution of synthetic Holliday junctions in DNA by an endonuclease from calf thymus. *EMBO J.*, **9**, 2931–2936.

Futcher,A.B. (1986) Copy number amplification of the 2 micron circle plasmid of *Saccharomyces cerevisiae*. *J. Theor. Biol.*, **119**, 197–204.

Goodsell,D.S., Grzeskowiak,K. and Dickerson,R.E. (1995) Crystal structure of C-T-C-T-C-G-A-G-A-G. Implications for the structure of the Holliday junction. *Biochemistry*, **34**, 1022–1029.

Guo,F., Gopaul,D.N. and Van Duyne,G.D. (1997) Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature*, **389**, 40–46.

Hickman,A.B., Waninger,S., Scocca,J.J. and Dyda,F. (1997) Molecular organization in site-specific recombination: the catalytic domain of bacteriophage HP1 integrase at 2.7 Å resolution. *Cell*, **89**, 227–237.

Hoess,R.H., Ziese,M. and Sternberg,N. (1982) P1 site-specific recombination: nucleotide sequence of the recombining sites. *Proc. Natl Acad. Sci. USA*, **79**, 3398–3402.

Hoess,R., Wierzbicki,A. and Abremski,K. (1987) Isolation and characterization 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.*, **5**, 282–304.

Iwasaki,H., Takahagi,M., Shiba,T., Nakata,A. and Shinagawa,H. (1991) *Escherichia coli* RuvC protein is an endonuclease that resolves the Holliday structure. *EMBO J.*, **10**, 4381–4389.

Iwasaki,H., Takahagi,M., Nakata,A. and Shinagawa,H. (1992) *Escherichia coli* RuvA and RuvB proteins specifically interact with Holliday junctions and promote branch migration. *Genes Dev.*, **6**, 2214–2220.

Jiang,J.S. and Brunger,A.T. (1994) Protein hydration observed by X-ray diffraction: solvation properties of penicillopepsion and neuraminidase crystal structures. *J. Mol. Biol.*, **243**, 100–115.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst.*, **A47**, 110–119.

Kallenbach,N.R., Ma,R. and Seeman,N.C. (1983) An immobile nucleic acid junction constructed from oligonucleotides. *Nature*, **305**, 829–831.

Kitts,P.A. and Nash,H.A. (1987) Homology-dependent interactions in phage lambda site-specific recombination. *Nature*, **329**, 346–348.

Kitts,P.A. and Nash,H.A. (1988) Bacteriophage lambda site-specific recombination proceeds with a defined order of strand exchanges. *J. Mol. Biol.*, **204**, 95–107.

Kleckner,N. (1996) Meiosis: how could it work? *Proc. Natl Acad. Sci. USA*, **93**, 8167–8174.

Kleff,S., Kemper,B. and Sternglanz,R. (1992) Identification and characterization of yeast mutants and the gene for a cruciform cutting endonuclease. *EMBO J.*, **11**, 699–704.

Kowalczykowski,S.C., Dixon,D.A., Eggleston,A.K., Lauder,S.D. and Rehrauer,W.M. (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.*, **58**, 401–465.

Kwon,H.J., Tirumalai,R., Landy,A. and Ellenberger,T. (1997) Flexibility in DNA recombination: structure of the lambda integrase catalytic core. *Science*, **276**, 126–131.

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

Lavery,R. and Sklenar,H. (1988) The definition of generalised helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dynam.*, **6**, 63–91.

Lee,J. and Jayaram,M. (1995) Role of partner homology in DNA recombination. Complementary base pairing orients the 5′-hydroxyl for strand joining during Flp site-specific recombination. *J. Biol. Chem.*, **270**, 4042–4052.

Lee,J., Lee,J. and Jayaram,M. (1995) Junction mobility and resolution of Holliday structures by Flp site-specific recombinase. Testing partner compatibility during recombination. *J. Biol. Chem.*, **270**, 19086–19092.

Lee,J., Voziyanov,Y., Pathania,S. and Jayaram,M. (1998) Structural alterations and conformational dynamics in Holliday junctions induced by binding of a site-specific recombinase. *Mol. Cell.*, **1**, 483–494.

Lilley,D.M.J. and Clegg,R.M. (1993) The structure of the four-way junction in DNA. *Annu. Rev. Biophys. Biomol. Struct.*, **22**, 299–328.

Lu,M., Guo,Q., Seeman,N.C. and Kallenbach,N.R. (1991) Parallel and antiparallel Holliday junctions differ in structure and stability. *J. Mol. Biol.*, **221**, 1419–1432.

Meselson,M. (1972) Formation of hybrid DNA by rotary diffusion during genetic recombination. *J. Mol. Biol.*, **71**, 795–798.

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

Miick,S.M., Fee,R.S., Millar,D.P. and Chazin,W.J. (1997) Crossover isomer bias is the primary sequence-dependent property of immobilized Holliday junctions. *Proc. Natl Acad. Sci. USA*, **94**, 9080–9084.

Mizuuchi,K., Kemper,B., Hays,J. and Weisberg,R.A. (1982) T4 endonuclease VII cleaves Holliday structures. *Cell*, **29**, 357–365.

Nicholls,A. and Honig,B. (1993) GRASP. Columbia University, New York.

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

Nunes-Düby,S., Azaro,M.A. and Landy,A. (1995) Swapping DNA strands and sensing homology without branch migration in lambda site-specific recombination. *Curr. Biol.*, **5**, 139–148.

Nunes-Düby,S.E., Yu,D. and Landy,A. (1997) Sensing homology at the strand-swapping step in lambda excisive recombination. *J. Mol. Biol.*, **272**, 493–508.

Nunes-Düby,S.E., Tirumalai,R.S., Kwon,H.J., Ellenberger,T. and Landy,A. (1998) Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucleic Acids Res.*, **26**, 391–406.

Otwinowski,Z. (1994) DENZO and SCALEPACK. Yale University, New Haven, CT.

Parsons,C.A., Kemper,B. and West,S.C. (1990) Interaction of a four-way junction in DNA with T4 endonuclease VII. *J. Biol. Chem.*, **265**, 9285–9289.

Parsons,C.A., Stasiak,A., Bennett,R.J. and West,S.C. (1995) Structure of a multisubunit complex that promotes DNA branch migration. *Nature*, **374**, 375–378.

Read,R.J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst.*, **A42**, 140–149.

Rice,L.M. and Brunger,A.T. (1996) Torsion angle dynamics: reduced variable confomational sampling enhances crystallographic structure refinement. *Proteins Struct. Func. Genet.*, **19**, 277–290.

Rice,D.W., Rafferty,J.B., Artymiuk,P.J. and Lloyd,R.G. (1997) Insights into the mechanisms of homologous recombination from the structure of RuvA. *Curr. Opin. Struct. Biol.*, **7**, 798–803.

Sadowski,P.D. (1995) The Flp recombinase of the 2-microns plasmid of *Saccharomyces cerevisiae*. *Prog. Nucleic Acid Res. Mol. Biol.*, **51**, 53–91.

Seeman,N.C. and Kallenbach,N.R. (1994) DNA branched junctions. *Annu. Rev. Biophys. Biomol. Struct.*, **23**, 53–86.

Sekiguchi,J., Seeman,N.C. and Shuman,S. (1996) Resolution of Holliday junctions by eukaryotic DNA topoisomerase I. *Proc. Natl Acad. Sci. USA*, **93**, 785–789.

Sharples,G.J., Chan,S.N., Mahdi,A.A., Whitby,M.C. and Lloyd,R.G. (1994) Processing of intermediates in recombination and DNA repair: identification of a new endonuclease that specifically cleaves Holliday junctions. *EMBO J.*, **13**, 6133–6142.

Sherratt,D.J., Arciszewska,L.K., Blakely,G., Colloms,S., Grant,K., Leslie,N. and McCulloch,R. (1995) Site-specific recombination and circular chromosome segregation. *Phil. Trans. R. Soc. Lond. Ser. B*, **347**, 37–42.

Sigal,N. and Alberts,B. (1972) Genetic recombination: the nature of crossed strand-exchange between two homologous DNA molecules. *J. Mol. Biol.*, **71**, 789–793.

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

Sternberg,N., Hamilton,D., Austin,S., Yarmolinsky,M. and Hoess,R. (1981) Site-specific recombination and its role in the life cycle of bacteriophage P1. *Cold Spring Harbor Symp. Quant. Biol.*, **1**, 297–309.

Stewart,L., Redinbo,M.R., Qiu,X., Hol,W.G.J. and Champoux,J.J. (1998) A model for the mechanism of human topoisomerase I. *Science*, **279**, 1534–1540.

Subramanya,H.S., Arciszewska,L.K., Baker,R.A., Bird,L.E., Sherratt,D.J. and Wigley,D.B. (1997) Crystal structure of the site-specific recombinase, XerD. *EMBO J.*, **16**, 5178–5187.

Symington,L. and Kolodner,R. (1985) Partial purification of an endonuclease from *Saccharomyces cerevisiae* that cleaves Holliday junctions. *Proc. Natl Acad. Sci. USA*, **82**, 7247–7251.

Tsaneva,I.R., Muller,B. and West,S.C. (1992) ATP-dependent branch migration of Holliday junctions promoted by RuvA and RuvB proteins of *E.coli*. *Cell*, **69**, 1171–1180.

Van Duyne,G.D., Ghosh,S., Maas,W.K. and Sigler,P.B. (1996) Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *J. Mol. Biol*., **256**, 377–391.

von Kitzing,E., Lilley,D.M.J. and Diekmann,S. (1990) The stereochemistry of a four-way DNA junction: a theoretical study. *Nucleic Acids Res*., **18**, 2671–2683.

Weisberg,R. and Landy,A. (1983) Site-specific recombination in phage lambda. In Hendrix, R.W., Roberts,J.W., Stahl,F.W. and Weisberg,R.A. (eds), *Lambda II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 211–250.

Weisberg,R.A., Enquist,L.W., Foeller,C. and Landy,A. (1983) A role for DNA homology in site-specific recombination: the isolation and characterization of a site affinity mutant of coliphage lambda. *J. Mol. Biol*., **170**, 319–342.

West,S.C. and Korner,A. (1985) Cleavage of cruciform DNA structures by an activity from *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA*, **82**, 6445–6449.

White,M.F., Giraud-Panis,M.J., Pohler,J.R. and Lilley,D.M. (1997) Recognition and manipulation of branched DNA structure by junction-resolving enzymes. *J. Mol. Biol*., **269**, 647–664.

Wierzbicki,A., Kendall,M., Abremski,K. and Hoess,R. (1987) A mutational analysis of the bacteriophage P1 recombinase Cre. *J. Mol. Biol*., **195**, 785–794.

Wood,A.A., Nunn,C.M., Trent,J.O. and Neidle,S. (1997) Sequence-dependent crossed helix packing in the crystal structure of a B-DNA dodecamer yields a detailed model for the Holliday junction. *J. Mol. Biol*., **269**, 827–841.

Yu,X., West,S.C. and Egelman,E.H. (1997) Structure and subunit composition of the RuvAB-Holliday junction complex. *J. Mol. Biol*., **266**, 217–222.

Zhu,X.D., Pan,G., Luetke,K. and Sadowski,P.D. (1995) Homology requirements for ligation and strand exchange by the FLP recombinase. *J. Biol. Chem*., **270**, 11646–11653.

Received May 18, 1998; accepted May 22, 1998