The EMBO Journal Vol.16 No.17 pp.5178–5187, 1997

Crystal structure of the site-specific recombinase, XerD

Hosahalli S.Subramanya[1], Lidia K.Arciszewska[2], Rachel A.Baker[2], Louise E.Bird[1], David J.Sherratt[1,3] and Dale B.Wigley[1]

[1]Laboratory of Molecular Biophysics and [2]Microbiology Unit, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK

[3]Corresponding author
e-mail: Sherratt@bioch.ox.ac.uk

The structure of the site-specific recombinase, XerD, that functions in circular chromosome separation, has been solved at 2.5 Å resolution and reveals that the protein comprises two domains. The C-terminal domain contains two conserved sequence motifs that are located in similar positions in the structures of XerD, λ and HP1 integrases. However, the extreme C-terminal regions of the three proteins, containing the active site tyrosine, are very different. In XerD, the arrangement of active site residues supports a *cis* cleavage mechanism. Biochemical evidence for DNA bending is encompassed in a model that accommodates extensive biochemical and genetic data, and in which the DNA is wrapped around an α-helix in a manner similar to that observed for CAP complexed with DNA.

Keywords: DNA binding/recombination mechanism/site-specific recombinase/XerD structure

Introduction

In site-specific recombination, DNA molecules are cleaved in both strands at two separate recombination sites, and the ends are rejoined to new partners, without any synthesis or degradation of DNA or hydrolysis of phosphodiesters. The reactions are catalysed by specialized recombinase proteins and may involve other protein accessory factors that have structural and modulatory roles in the nucleoprotein complex that contains synapsed recombination sites. Site-specific recombinases mediate a wide range of microbial programmed DNA rearrangements that include the integration and excision of bacteriophages from bacterial chromosomes, the control of circular replicon inheritance, the processing of the initial products of genetic transposition and the mediation of genetic 'switches' through inversion or deletion of specific DNA segments (reviewed in Stark *et al.*, 1992; Landy, 1993; Stark and Boocock, 1995a; Nash, 1996). There are two families of site-specific recombinases; the resolvase/invertase family use a serine nucleophile to mediate a concerted double strand cleavage and rejoining reaction at nucleotide phosphates separated by 2 bp, while the λ integrase family enzymes use a tyrosine nucleophile to mediate sequential pairs of strand exchanges that are positioned 6–8 bp apart. In site-specific recombination reactions mediated by both families, four recombinase molecules bound to two ~30 bp recombination core sites catalyse the breaking and rejoining of four DNA phosphodiester bonds.

The 298 amino acid site-specific recombinase XerD, along with its related partner recombinase, XerC, belong to the λ integrase family of site-specific recombinases. They function in the stable inheritance of the *Escherichia coli* chromosome and multicopy circular plasmids, apparently by converting to monomers the circular multimers that can arise by homologous recombination (Blakely *et al.*, 1993; Sherratt *et al.*, 1995). Recombinases of the integrase family are highly diverged in primary amino acid sequence, with only four completely conserved amino acids (RHRY) (Argos *et al.*, 1986; Abremski and Hoess, 1992). All four of the conserved residues have been implicated in catalysis (Pargellis *et al.*, 1988; Evans *et al.*, 1990; Chen *et al.*, 1992a; Friesen and Sadowski, 1992; Lee *et al.*, 1992). In XerD and XerC, the two conserved arginines and the tyrosine are required for DNA cleavage, while the conserved histidine is required for DNA rejoining (Blakely *et al.*, 1993; Arciszewska and Sherratt, 1995; Arciszewska *et al.*, 1997; L.K.Arciszewska, R.A.Baker, P.A.Wigge and D.J.Sherratt, unpublished data). Recombination is initiated when the conserved tyrosine hydroxyl attacks the scissile phosphate, forming a 3′ phosphotyrosyl–DNA complex and a free 5′ hydroxyl. In the second step, a 5′ hydroxyl from the adjacent partner duplex attacks the phosphotyrosine to form a Holliday junction intermediate. The recombination reaction is completed by the exchange of the second pair of strands, using the same mechanism, 6–8 bp away from the site of the initial strand exchanges.

Xer site-specific recombination exhibits three features that distinguish it from other well characterized members of the family. First, it uses two related recombinases, XerC and XerD, each of which catalyses one specific pair of strand exchanges (Blakely *et al.*, 1993, 1997; Arciszewska and Sherratt, 1995; Colloms *et al.*, 1996, 1997; Arciszewska *et al.*, 1997). This is proving to be a powerful tool in establishing the roles of the two recombinases and for defining the determinants for recombinase binding to DNA. Analysis of recombination site function has also been facilitated by the availability of a wide variety of naturally occuring Xer recombination sites that contain a range of related recombinase-binding DNA sequences.

Second, the recombination reaction has different requirements and outcomes depending on whether it occurs at plasmid or chromosomal recombination sites. Recombination at natural plasmid sites is preferentially intramolecular and requires, in addition to the two recombinases and the 28–30 bp recombination core site,

© Oxford University Press

additional accessory proteins and ~200 bp of adjacent accessory DNA sequences. Interaction of the accessory proteins and accessory sequences promotes the formation of a synaptic complex of precise topology, that can only form efficiently on directly repeated recombination sites in the same molecule (Colloms *et al.*, 1996, 1997). In contrast, recombination at the *E.coli* chromosomal site, *dif*, requires only a 28 bp recombination core site at which the two recombinases act. Recombination *in vivo* at *dif*, present in multicopy plasmids, occurs intermolecularly and intramolecularly (Blakely *et al.*, 1991, Leslie and Sherratt, 1995; Tecklenberg *et al.*, 1995).

Third, despite the sequence divergence of integrase family recombinases, conserved Xer-like recombinase sequences are present in the chromosomes of almost all bacteria examined (including the archaebacterium, *Methanococcus janaschii*), suggesting that there is a strong constraint on how Xer recombination functions in chromosome segregation. These enzymes may be the progenitors of the many integrase-like enzymes found in different microbes.

In this study, the substantial body of information that has accumulated on how XerC and XerD interact with their recombination site DNA and mediate recombination is used in order to relate the XerD structure to function. XerC and XerD each bind cooperatively to related 11 bp sites that are separated by a 6–8 bp 'central' region; binding of the recombinases to DNA leads to substantial DNA bending (Blakely *et al.*, 1993, 1997; Blakely and Sherratt, 1994, 1996a). XerC-mediated strand exchange of 'top' strands occurs at the border of its binding site and the central region, while XerD exchange of 'bottom' strands occurs 6–8 bp away at the border of its binding site and the central region. Genetic and biochemical analysis of the recombination core site has identified the nucleotides that provide specificity for recombinase binding (Blake *et al.*, 1997; Hayes and Sherratt, 1997), and has indicated which backbone and base contacts are involved in this interaction (Blakely and Sherratt, 1994, 1996a; Blakely *et al.*, 1997). A deletion and pentapeptide insertion analysis of XerD has revealed parts of the protein involved in DNA binding and interaction with XerC (Spiers and Sherratt, 1997; Y.Cao, B.Hallet, D.J.Sherratt and F.Hayes, unpublished data). XerC and XerD are catalytically autonomous as judged by the demonstration that normal strand exchange by either XerC or XerD does not require the tyrosine nucleophile of the partner recombinase (Arciszewska and Sherratt, 1995; Arciszewska *et al.*, 1997), and that when either XerC or XerD are incubated with supercoiled plasmid containing *dif*, a XerC or XerD site-specific type I topoisomerase activity is detectable (Cornet *et al.*, 1997; Spiers and Sherratt, 1997). Finally, the substrate requirements for XerC- and XerD-mediated catalysis have been compared and the topological parameters of the recombination reaction determined (Colloms *et al.*, 1996, 1997; Arciszewska *et al.*, 1997).

Results and discussion

Structure of the protein

Details of the structure determination and refinement are presented in Table I and in Materials and methods. Of the 298 amino acids of the protein, 271 were defined in our final model. The missing residues are located at the N- and C-termini and in three disordered surface loops. The enzyme comprises two domains: domain 1 consists of residues 1–107, while domain 2 comprises residues 108–298 (Figure 1). Domain 1 contains four α-helices, arranged such that there are two parallel helix hairpins arranged at 90° to each other. Domain 2 is also mainly α-helical, but with a three-stranded antiparallel β-sheet along one edge. The fold of this domain is similar to that determined recently for λ and HP1 integrases (λ Int and HP1 Int, respectively), and is at present unique to this family of proteins (Hickman *et al.*, 1997; Kwon *et al.*, 1997). Domains 1 and 2 of XerD correspond to domains of λ Int, HP1 Int and FLP identified by limited proteolysis, although in FLP the C-terminal domain has been divided further into three further sub-domains (Moitoso de Vargas *et al.*, 1988; Evans *et al.*, 1990; Chen *et al.*, 1991; Pan and Sadowski, 1993; Sadowski, 1995; Hickman *et al.*, 1997; Kwon *et al.*, 1997).

The region of structural homology within the C-terminal domains of the XerD, λ Int and HP1 Int spans ~170 residues (Figure 2). Two conserved sequence motifs, that have been proposed to indicate a conservation of structure across a wide family of integrases that includes XerD and XerC as well as λ Int and HP1 Int (Argos *et al.*, 1986; Abremski and Hoess, 1992; Blakely and Sherratt, 1996b), are located in domain 2 of XerD. The locations of motif I and the N-terminal portion of motif II are similar in the structure of XerD (residues 145–159 and 244–281, respectively) and those of λ and HP1 integrases (Hickman *et al.*, 1997; Kwon *et al.*, 1997). However, the extreme C-terminal portions of the proteins, which include the C-terminal portion of motif II, could hardly be more different (Figure 2). This region of the proteins is of particular interest because it contains the active site tyrosine residue to which the DNA becomes attached covalently during the recombination reaction.

In λ Int, these C-terminal residues (334–356) form a flexible loop that is disordered in one of the two molecules in the asymmetric unit, but is more ordered in the other, where the final 15 residues form two additional β-strands along one edge of the antiparallel sheet. In HP1 Int, this region (residues 307–337) forms an extended structure which protrudes from the surface of the protein molecule and contains two short helices. This region is involved in crystal contacts which the authors propose to be representative of one of the protein dimer interfaces during the recombination reaction. By contrast, in XerD, this region (residues 271–298) forms a turn followed by a long α-helix, containing the active site tyrosine, that extends almost to the C-terminus (the last six residues of the protein are disordered in the crystal structure).

It is intriguing that a region of such vital importance should be so different in the three enzymes. Furthermore, it is likely that this region will differ further in other members of the λ integrase family. For example, in FLP and other yeast recombinases, this region is longer than in the bacterial members of the family, a difference that has been proposed to play a role in determining whether the tyrosine nucleophile of a given enzyme molecule attacks the scissile phosphate bond and is activated by that molecule (cleavage *in cis*), or whether it attacks the

Table I. Summary of crystallographic structure analysis

Data collection

	λ (Å)	Resolution (Å)	R_{sym} (%)	Redundancy of data	Completeness (%)
Native	0.89	2.5	4.8	3.0	98.9
$NaAuCl_4$ (10 mM, 3 h)	0.89	3.0	7.8	3.5	96.1
Trimethyl lead acetate (100 mM, 18 h)	0.89	2.5	5.6	2.6	95.4
Ethylmercury phosphate (5 mM, 5 days)	1.54	2.5	7.4	2.9	99.1

Phasing statistics

Derivative	Anomalous data	No. of sites	MFID	R_{cullis}	Phasing power	Mean FOM
$NaAuCl_4$	yes	8	0.28	0.73	1.4	0.57
Trimethy lead acetate	yes	9	0.21	0.77	1.3	
Ethylmercury phosphate	yes	6	0.23	0.70	1.4	

Refinement of native crystal

Resolution (Å)	2.5
Final *R*-factor (all data, 10–2.5Å)	22.4
R_{free} (5% of data)	28.7
r.m.s.d. bonds (Å)	0.013
r.m.s.d. angles (Å)	0.041
No. of residues	271
No. of water molecules	105
Ramachandran analysis (most favoured/additional allowed/generous/disallowed)[a] (%)	89.4/10.2/0.4/0.0
B-factors (lowest/highest/mean)	6.8/134.8/50.4
protein only	6.8/134.8/49.9
water molecules	20.8/97.8/59.9

[a]Definitions according to Laskowski *et al.* (1993).

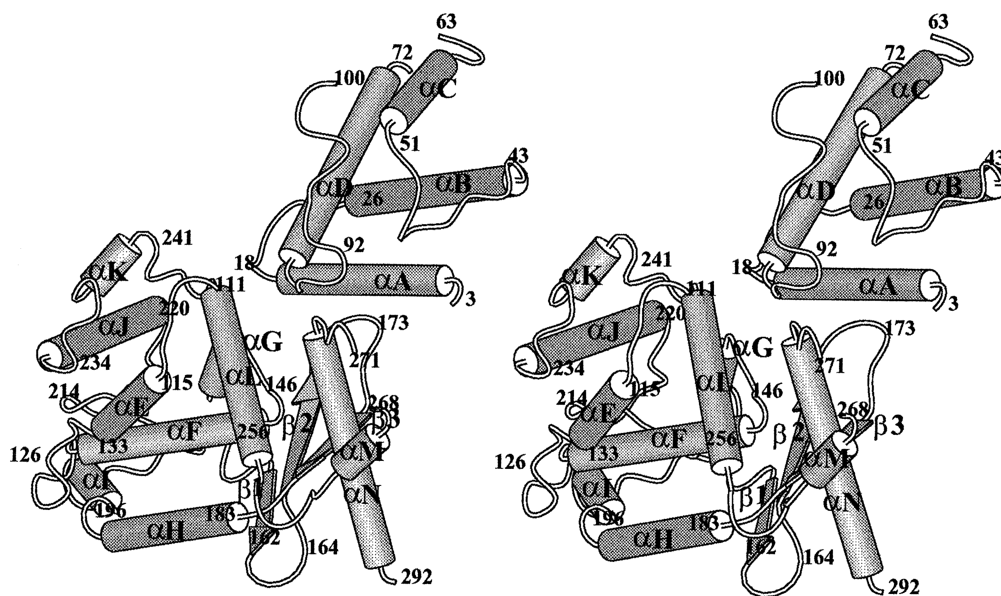

Fig. 1. Overall structure of the protein. This stereo figure was prepared using PREPPI. The numbering refers to the beginning and end of secondary structural elements. Residues that are not defined are located at the N- and C-termini and in three disordered loops (residues 64–70, 101–110 and 269–270).

scissile phosphate bond and is activated by a different molecule (cleavage *in trans*) (discussed in Landy, 1993; Jayaram and Lee, 1995; Stark and Boocock, 1995b; Blakely and Sherratt, 1996b; Jayaram, 1997).

Interactions between XerD and DNA

The structures of the catalytic domains of λ and HP1 integrases suggested how DNA might interact with the C-terminal region of these proteins (Hickman *et al.*, 1997;

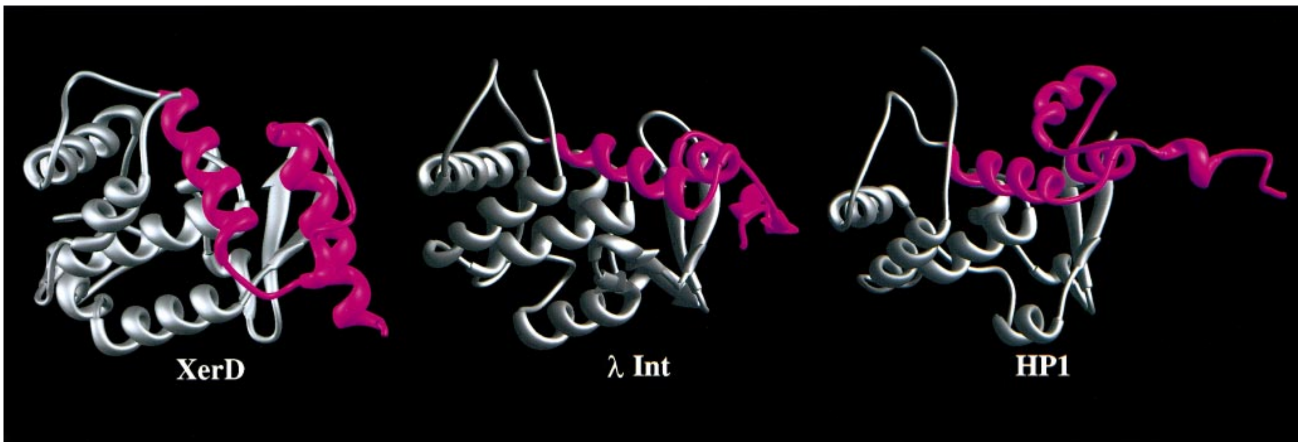

Fig. 2. Comparison of the structures of the C-terminal domains of XerD, λ Int and HP1 Int. Regions of the C-terminal domains of the proteins that show the greatest structural similarity are shown in grey. The major structural differences (shown in magenta) are located in the polypeptide segments that extend from conserved motif II (Argos *et al.*, 1986) to the C-terminus of the proteins.

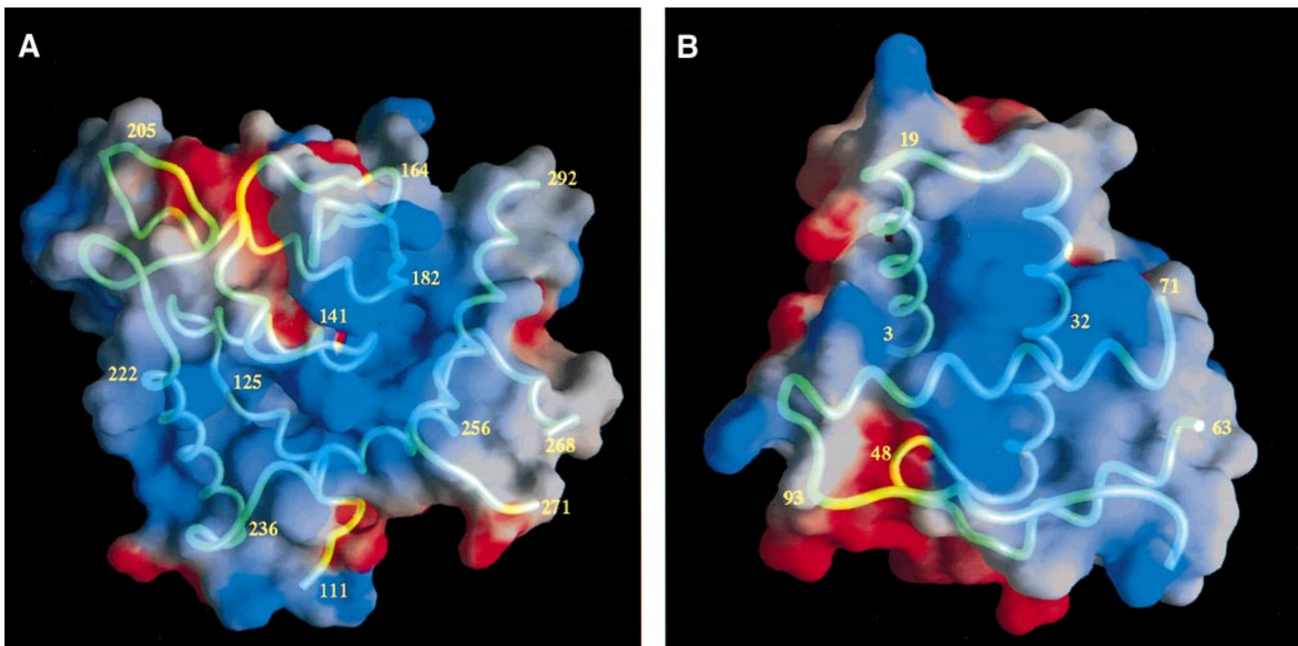

Fig. 3. Electrostatic surface potential of domains of XerD. (**A**) Domain 2 and (**B**) domain 1. Regions of negative potential are coloured in red and positive potential in blue. The surface is transparent to reveal the underlying C_{α} backbone of the protein (shown in green). This figure was prepared using GRASP (Nicholls and Honig, 1991).

Kwon *et al.*, 1997). Calculations of the electrostatic potential of the surface of domain 2 of XerD [using GRASP (Nicholls and Honig, 1991)] reveal an obvious site for interaction with DNA (Figure 3A) that is consistent with those proposed for λ Int and HP1 Int. However, when combined with the extensive biochemical information that is available for the Xer proteins, the structure of XerD provides a more detailed view of this interaction. The footprinting data indicate clearly that there are contacts between the protein and DNA that encircle the DNA duplex in the region around the cleavage site, an observation that is not explained by the interactions between domain 2 and DNA alone. Consequently, it is likely that domain 1 also contributes to the DNA-binding site, a view supported by genetic and biochemical studies of XerD and other recombinases (Hoess *et al.*, 1990; Panigrahi *et al.*, 1992; Panigrahi and Sadowski, 1994; Sadowski, 1995; Spiers and Sherratt, 1997). While the proposed DNA-binding sites in the structures of the catalytic core fragments of λ and HP1 integrases are exposed, in the XerD structure, access to the active site is blocked by the positioning of domain 1 over this region (Figure 1). Hence, in order for XerD to bind to DNA, there has to be a large conformational change to allow access of the DNA to the active site region. The electrostatic potential of the surface of domain 1 also reveals a likely candidate for a DNA-binding surface (Figure 3B), supporting the view that both domain 1 and 2 contribute to the DNA-binding surface of the protein. Because of the conformational changes of the protein required to bind to DNA, we have restricted our

detailed discussion of the DNA binding and recombinase–recombinase interactions to the C-terminal catalytic domain of XerD, which we shall refer to as domain 2.

Initial attempts to model the XerD domain 2–DNA complex were based upon those proposed for the interactions of catalytic core fragments of λ and HP1 integrases with linear B-form DNA (Hickman *et al.*, 1997; Kwon *et al.*, 1997). However, it was evident immediately that the application of the λ Int and HP1 models to XerD did not take into account the biochemical evidence which shows that the DNA binding of XerD induces an ~40° bend in the DNA (Blakely and Sherratt, 1996a). Furthermore, these simple models were inconsistent with the footprinting and binding interference data which show that XerD binding to its DNA target involves major and minor groove interactions that encircle the DNA (Blakely and Sherratt, 1994, 1996a; Blakely *et al.*, 1997). For example, the DNA footprinted by XerD extends over a much larger region than would be possible by docking the XerD structure onto linear duplex B-form DNA, but could be accommodated by wrapping the DNA around the protein. The degree of bending of the DNA in the XerD–DNA complex has been shown to be comparable with that induced by each of the molecules of the catabolite activator protein (CAP) dimer, as observed in the crystal structure of the CAP–DNA complex and measured in biochemical experiments (Schultz *et al.*, 1991). This led us to compare the structures to seek any similarities that might help to understand the XerD–DNA complex. In the CAP–DNA complex, there is a sharp 40° kink in the bound DNA as it wraps around a helix–turn–helix motif on the protein surface. Comparison of the DNA-binding surface of CAP with that of XerD revealed a striking similarity in the spatial arrangement of the helix–turn–helix motif in CAP and the positions of two helices (helices αG and αJ) in the XerD structure (Figure 4A), although in the XerD structure these two helices are separated by ~65 residues rather than the tight turn found in CAP. Furthermore, the positions of residues which are involved in contacts with the phosphodiester backbone of the DNA, in the CAP structure, are conserved in the XerD structure (Figure 4). These residues also occur in other helix–turn–helix motifs (Brennan and Matthews, 1989) and, perhaps more importantly, the sequences and structures of the comparable regions of the λ and HP1 integrases are also consistent with them being DNA recognition helices (Figure 4B and C; Hickman *et al.*, 1997; Kwon *et al.*, 1997). This latter point may be particularly significant given the low level of sequence similarity in these regions of the proteins.

Taken together, these data allow us to construct a model for the interaction between domain 2 of XerD and its recognition sequence based upon the CAP–DNA complex (Schultz *et al.*, 1991). The two helices of the helix–turn–helix motif of one subunit of the CAP–DNA complex were superimposed upon helices αG and αJ of XerD (Figure 4A) and allowed us to position the DNA relative to these helices in XerD. The resulting model is presented in Figure 5, and is in remarkably good agreement with the DNA bending, footprinting and binding interference data (Blakely *et al.*, 1993, 1997; Blakely and Sherratt, 1994, 1996a; summarized in Figure 5). The proposed backbone contacts (Figure 4B and C) compare well with the phosphates whose ethylation prevents binding (Figure 5). The proposed base-specific contacts between the XerD recognition helix, αJ and recombination site DNA are very similar to the comparable CAP–DNA contacts, with both the CAP and the XerD recognition helices being oriented in the same way (Figure 4B and C). In the model, XerD residues 220R and 221Q could make base-specific contacts at precisely the positions that we have identified as being important for XerD binding and XerD–XerC binding specificity. For example, oxidation of any of three adjacent Ts (positions 11–13; Figures 4C and 5) prevents XerD binding to its site, while nucleotides at positions 10, 11 and 13 contribute to XerC–XerD binding discrimination, and the presence of a T or G, but not a C, at position 9 correlates with tight XerD binding (Blake *et al.*, 1997; Hayes and Sherratt, 1997). Moreover, examination of known XerD and XerC recombinases shows that all XerD recombinases have the equivalent of 220R and 221Q, whereas XerC recombinases have a conserved R in place of Q at the equivalent of position 221 and a non-conserved residue in the preceding position. Other putative Xer recombinase sequences present in the databases have either RQ at the positions corresponding to 220 and 221, respectively, or a conserved R at the position corresponding to 221, preceded by a non-conserved residue. This indicates that these presumptive recombinases can be classified as either XerC or XerD proteins on the basis of the amino acid sequence at positions corresponding to 220 and 221, and that these amino acids may provide much of the discrimination that directs XerC and XerD to their specific DNA-binding sites. The weaker binding of XerC, and the reduced bending it appears to induce, may be a consequence of fewer base-specific contacts. Furthermore, the high conservation of amino acid residues at these two positions in XerD recombinases from different bacteria (and at the one position in different XerC enzymes) suggests a very strong functional selection for the maintenance of specific recombinase–DNA contacts in these enzymes. This is supported by our demonstration that the *Bacillus subtilis* XerC and XerD homologues mediate strand exchange on an *E.coli dif*-containing Holliday junction (G.B.Blakely and D.J.Sherratt, unpublished data).

The DNA-binding properties of deletion and pentapeptide insertion mutants also agree well with the model in which helix αJ of domain 2 is the recognition helix that interacts with DNA. A truncated XerD derivative containing residues 1–233 is proficient in DNA binding and retains all but the last residue of helix αJ. In contrast, an even shorter XerD derivative, that is deleted for the six C-terminal residues of helix αJ, is binding-deficient (Spiers and Sherratt, 1997). Insertions of proline-containing pentapeptides into this same helix also abolish detectable DNA binding (Y.Cao, B.Hallet, D.J.Sherratt and F.Hayes, unpublished data). The positioning of the scissile phosphate adjacent to the active site residues adds further credence to the model (Figure 6).

The 11 bp XerD- and XerC-binding sites can be subdivided into two regions; the inner four nucleotides, that are dyad symmetrical in the XerC- and XerD-binding sites, and the outer seven nucleotides, at least four of which contribute to specific XerD binding (Figure 4C; Blakely and Sherratt, 1994, 1996a; Blake *et al.*, 1997;

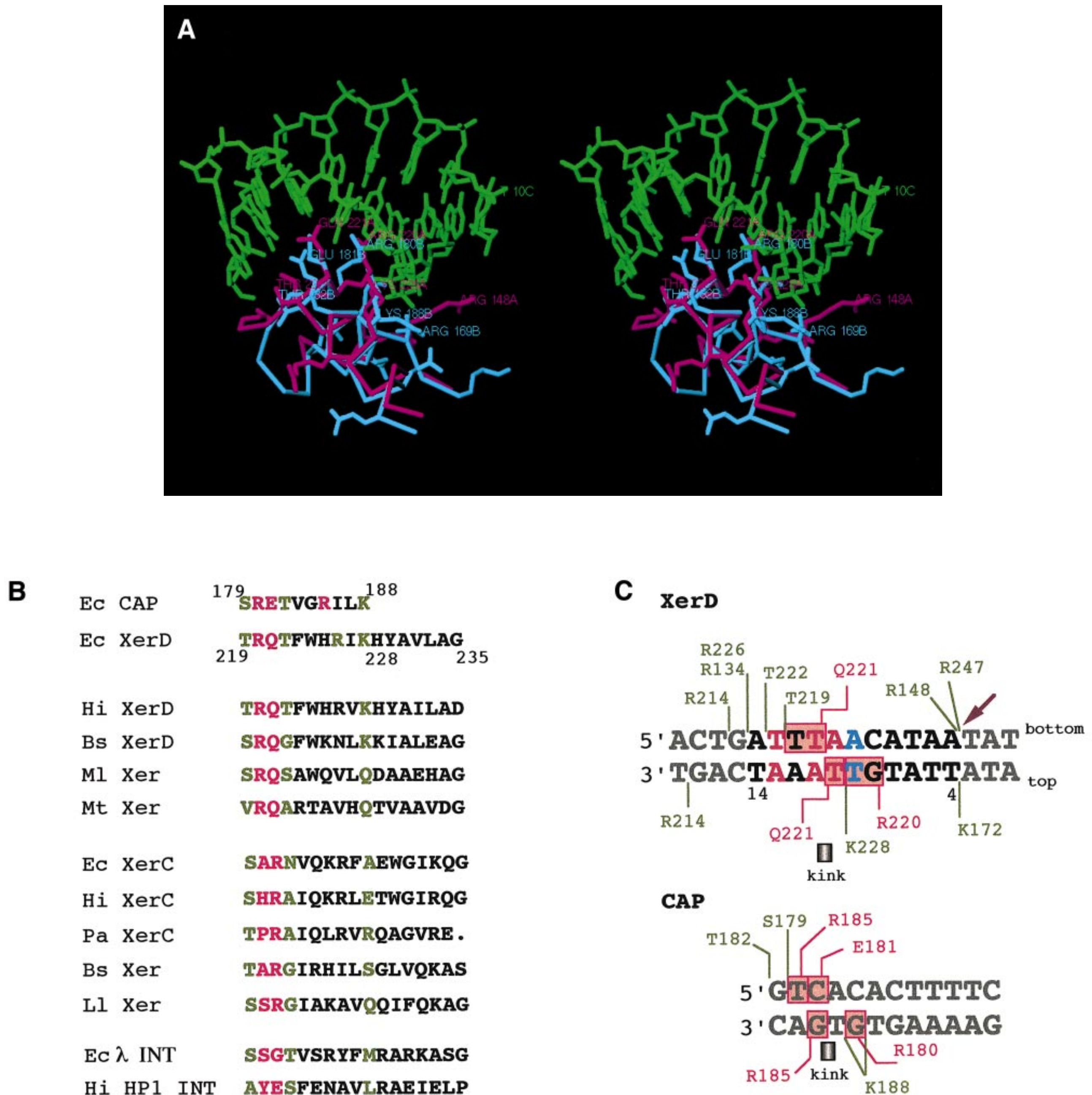

Fig. 4. Model for interactions between XerD and DNA. (**A**) The helix–turn–helix motif (residues 166B–194B) in the CAP–DNA complex is shown in cyan, with the DNA around the kink site shown in green. The corresponding helices of XerD (residues 146A–155A and 217A–234A) are overlaid in magenta. Only the Cα positions of the main chain are shown, together with the side chains that interact directly with the DNA. (**B**) Amino acid sequence alignment of the helix–turn–helix DNA recognition α-helix of *E.coli* CAP with the putative recognition helix of XerD and the comparable regions present in Xer recombinase sequences of other bacteria, λ Int and HP1 Int. The three CAP residues that make base-specific contacts in the major groove are shown in magenta (Schulz *et al.*, 1991). Other helix–turn–helix proteins also often make DNA contacts through these positions (Brennan and Matthews, 1989). Two amino acids of XerD helix αJ that could make base-specific contacts are also coloured magenta. In XerC enzymes, we propose that the conserved R at the position corresponding to 221 of XerD is important for sequence-specific binding. Amino acids of CAP that make DNA backbone contacts are coloured green. In XerD, these equivalent amino acids may contact the DNA backbone, as may 226R (all coloured green). Recombinases described as XerC or XerD have been shown to possess that specific function (G.Blakely, L.Neilson and D.J.Sherratt, unpublished data). Amino acid sequences in this region of *Salmonella typhimurium* XerD and XerC are identical to those in *E.coli*. (Hayes *et al.*, 1997). Recombinases described as 'Xer' are presumed to be XerD and XerC homologues based on sequence comparison; residues at position 220 and 221 are used for classification. The accession numbers for the recombinase sequences are as follows: *E.coli* (Ec) XerD, P21891[a]; *S.typhimurium* (St) XerD 492525[b]; *Haemophilus influenzae* (Hi) XerD P44630[a]; *Mycobacterium tuberculosis* (Mt) Xer Q10815[a]; *M. leprae* (Ml) Xer 467161[b]; *Bacillus subtilis* (Bs) RipX (XerD), P46352[a]; *E.coli* XerC, P22885[a]; *S.typhimurium* XerC, 492524[b]; *H.influenzae* XerC, P44818[a]; *Pseudomonas aeruginosa* Sss (XerC), X78478[c]; *B.subtilis* CodV (XerC), P39776[a]; *Lactobacillus leichmannii* Xer, X84261[c]; *E.coli* λ Int, P03700[a]; *H.influenzae* HP1 Int, P21442[a]; *Saccharomyces cerevisiae* FLP, P03870[a]. [a]SWISS-PROT, [b]NCIB, [c]GenEMBL. (**C**) Comparison of CAP recognition helix interactions with DNA (Schultz *et al.*, 1991), and those that our model predicts will be involved in XerD–DNA interactions. Base-specific contacts are in magenta and backbone contacts in green. The sequences are oriented so that the kinks induced by recombinase binding are in the same direction. The relative positions of the contacts are remarkably similar. The XerD-binding site consists of residues 4–14. The nucleotides in deep magenta are implicated in binding specificity, whereas those in blue are important for binding (see text). The position of the scissile phosphate bond is indicated with an arrow.

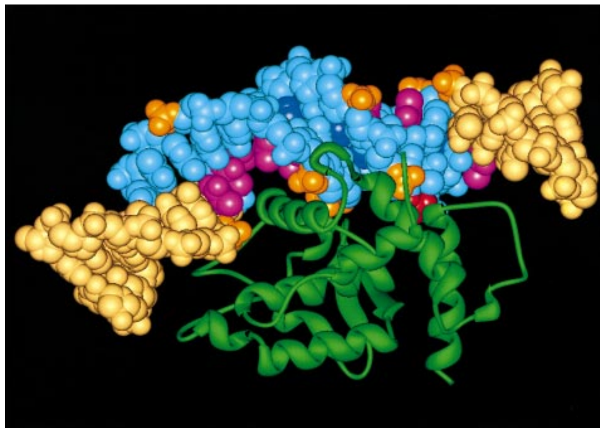

Fig. 5. Model of XerD bound to DNA. Model of XerD domain 2 bound at its recognition sequence, derived from the CAP–DNA complex. The protein is shown in green as a ribbon, while the DNA is shown as a space-filling representation. Residues of the DNA within the OP-Cu footprint are shown in cyan, and those outside of the footprint in beige. Specific contacts, as shown by interference binding analysis, are overlaid in orange (phosphates), blue (adenine minor groove contacts) and magenta (thymine, major groove). The scissile phosphate is shown in red. The view of the active site residues in relation to the scissile phosphate is as in Figure 6. This figure was prepared using RIBBONS (Carson, 1991).

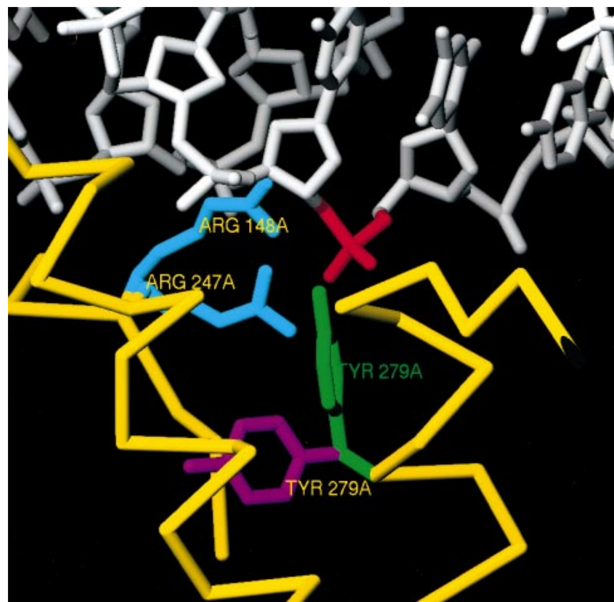

Fig. 6. The active site region of XerD. The positions of residues in the active site of XerD, that have been shown to be important for the cleavage reaction, are consistent with a *cis* cleavage mechanism for the enzyme. The Cα backbone is shown in yellow, the side chains of Arg148 and Arg247 in blue and the side chain of Tyr279 in purple. A simple rotation about the Cα–Cβ bond of Tyr279 (now depicted in green) would position the side chain appropriately for in-line attack of the scissile phosphate (coloured red).

Blakely *et al.*, 1997; Hayes and Sherratt, 1997). 1,10-Phenanthroline-copper (OP-Cu) intercalates into DNA through the minor groove, from where it can cleave the DNA backbone. The OP-Cu 'footprint' made by XerD covers the whole of the XerD DNA-binding site (Figure 5). The resistance of the outer part of the site to cleavage by OP-Cu could be the consequence of a widening of the minor groove because of the proposed kink in this region, thus preventing intercalation of the footprinting reagent. The resistance of the minor groove, on the inner part of the site and the proximal part of the central region, to OP-Cu cleavage could result from the interactions of residues 236–245 with the minor groove on the 'front' face of the DNA as viewed in Figure 5, as well as interactions from 'behind' the DNA by the antiparallel strands β2 and β3. The sites of N-3 adenine methylation that interfere with XerD binding are also consistent with the same minor groove interactions in the inner region of the binding site (Figure 5). Furthermore, the lysine residues at positions 172 and 175, which are in the turn of the β2–β3 hairpin, could make the DNA backbone contacts on each of the DNA strands in the vicinity of the scissile phosphate that have been revealed by ethylation interference footprinting (Blakely *et al.*, 1997). Two pentapeptide insertions into the β2–β3 hairpin turn result in XerD proteins that are defective in XerD catalysis, despite binding DNA cooperatively (Y.Cao, B.Hallet, D.J.Sherratt and F.Hayes, unpublished data). This suggests that this region of the protein, that is structurally conserved in λ and HP1 integrases and which shows primary amino acid sequence conservation in other integrases, may have an important function in the strand exchange reaction.

Binding interference data (Figure 5) also implicate the major groove on the inner part of the XerD-binding site in recombinase binding. This interaction cannot be accounted for by the proposed domain 2 binding, and therefore probably arises from a domain 1 interaction. We expect such an interaction to be similar in XerC and XerD because of the dyad symmetry in this region of the DNA. Moreover, the same interaction with the major groove in the inner part of the recombinase-binding site may occur in other integrase family enzymes, since the regions of FLP and Cre, corresponding to XerD domain 1, have been implicated in binding to the inner four nucleotides of the recombinase-binding site, whereas their C-terminal domains bind to the outer 9 bp of the target site (Hoess *et al.*, 1990; Panigrahi *et al.*, 1992; Panigrahi and Sadowski, 1994). Helices αB and αD are major contributors to the basic surface of domain 1 (Figure 3B), and mutations that disrupt either helix are impaired in DNA binding (Spiers and Sherratt, 1997; Y.Cao, B.Hallet, D.J.Sherratt and F.Hayes, unpublished data). Therefore, one of these helices, and the equivalent region of XerC, is a candidate for a domain 1 interaction with the major groove of the inner part of the recombinase-binding site.

Mechanistic implications

The model we present for the interaction between XerD and DNA is consistent also with data concerning the catalytic mechanism of the enzyme. There has been considerable debate concerning the mechanism of enzymes of the integrase family (discussed in Landy, 1993; Jayaram and Lee, 1995; Stark and Boocock, 1995b; Blakely and Sherratt, 1996b; Jayaram, 1997). The debate focuses upon whether the residues involved in catalysis at a particular scissile phosphate within a recombination site arise from the same molecule (a *cis* cleavage mechanism) or can belong to a partner molecule (*trans* cleavage). For most systems, it is difficult to distinguish between these alternatives, although the evidence for FLP strongly supports a

trans cleavage mechanism in which one of the two protomers bound to a given recombination site provides the tyrosine that attacks the scissile phosphate that is contacted and activated by the other protomer (Chen *et al.*, 1992b; Lee *et al.*, 1994). With λ Int, some evidence supports *cis* cleavage (Nunes-Düby *et al.*, 1994), while other data indicate a *trans* mechanism (Han *et al.*, 1993). It is important to note that both modes of cleavage could occur at different stages of the reaction. One advantage of studying the Xer recombinases is that the partner molecules bound to a duplex are different and it is therefore easier to distinguish the roles played by each of the partners during catalysis. By using combinations of wild-type and mutant proteins, in which the active site tyrosine of each recombinase was replaced by phenylalanine, it has been demonstrated conclusively that XerC- and XerD-mediated cleavage of a Holliday junction substrate does not involve the participation of an active site tyrosine from the partner recombinase and therefore occurs by a *cis* mechanism (Arciszewska and Sherratt, 1995; Arciszewska *et al.*, 1997). The structure of XerD explains this observation, since the active site tyrosine, Y279, is in close proximity to other residues known to be at the active site (Figure 6). In λ Int, the structure has been interpreted as being consistent with both *cis* and *trans* cleavage, while the HP1 structure supports *cis* cleavage (Hickman *et al.*, 1997; Kwon *et al.*, 1997). The mechanism for formation of the covalent DNA–protein adduct is likely to involve in-line nucleophilic attack of the scissile phosphate by the hydroxyl of the active site tyrosine. One feature of the model for the XerD–DNA complex is that although Y279 is positioned appropriately in the active site, it is buried in the protein in a conformation that would not allow attack of the scissile phosphate (Figure 6). However, a simple rotation of the side chain of the tyrosine about the Cα–Cβ bond could place the residue in a position that would be ideal for in-line attack. This rotation would be hindered sterically in the present structure, but would require only a small alteration in the position of helix αN to allow rotation of the tyrosine side chain. Although XerD can bind to a recombination site in the absence of XerC, both proteins have to be bound to catalyse efficient cleavage or strand exchange of the DNA, showing that interaction between the recombinases is required to promote efficient cleavage of the DNA. The properties of recombinase mutants are also consistent with recombinase–recombinase interactions being necessary to 'activate' catalysis (see below).

A model for interactions between XerC and XerD at a recombination site

The model for DNA binding to XerD presented above may also have important implications for the cooperative interactions that occur between XerC and XerD on DNA binding (Blakely *et al.*, 1993; Blakely and Sherratt, 1996a; Spiers and Sherratt, 1997). Since the XerC and XerD amino acid sequences, DNA-binding sites and footprints are similar, the complex of both recombinases bound to DNA will contain a pseudo 2-fold symmetry axis. Consequently, we are able to construct a model to suggest how the two proteins might interact at a single recombination site. This model was prepared by rotating the XerD–DNA model described above by an ~2-fold axis and then superimposing the central region of bases to produce a pseudo-continuous helical duplex. The DNA model required small adjustments in the central region to make it more similar to duplex B-form DNA, because of disturbances in the CAP–DNA structure as a result of crystal contacts. The resulting model, despite being derived by modelling of the DNA substrate alone, and without reference to the proteins, shows how XerC and XerD might interact when bound together at a *dif* site (Figure 7). This arrangement is consistent with the biochemical properties of mutant enzymes. A truncated XerD protein containing residues 1–268 is able to bind to DNA and to interact cooperatively with XerC, while a protein containing residues 1–262 lacks cooperativity, although it binds DNA normally (Spiers and Sherratt, 1997). Pentapeptide insertions into XerD also define this region as being important for cooperative interactions with XerC (Y.Cao, B.Hallet, D.J.Sherratt and F.Hayes, unpublished data). A second region likely to be involved in XerC–XerD interactions is defined by a XerD mutant containing a tripeptide substitution at residues 256–258 (B.Hallet and D.J.Sherratt, unpublished data). The mutant protein is proficient in XerD cleavage and strand exchange, and can undergo cooperative interactions with XerC. Nevertheless, it is unable to promote efficient catalysis by XerC, thus identifying a region of XerD involved in activation of XerC catalysis (see below). These two regions are highlighted in Figure 7 and are at the subunit interface in the model. We should stress that this model can only provide an approximation of an XerC–XerD interaction on DNA since XerD–XerD interactions have not been observed, even on recombination sites having two XerD-binding sites. Moreover, XerC appears to induce a smaller bend than XerD (Blakely and Sherratt, 1996a). It is possible that the XerD–XerC interactions utilize amino acid differences between the recombinases in the interface region. Despite the limitations of the model for XerC–XerD interactions, we are confident that it provides the basis for explaining much experimental data and for designing new experiments.

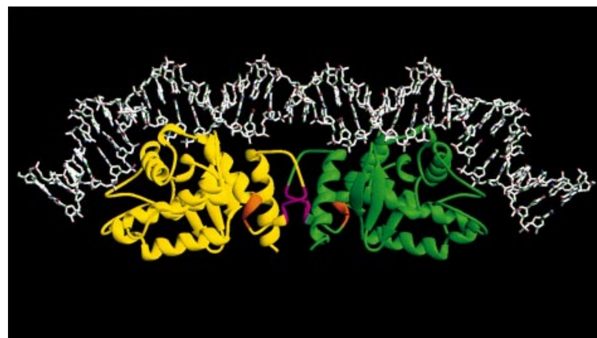

Fig. 7. A model for the complex between XerC, XerD and DNA. Ribbon representation of the catalytic domains of the two recombinase proteins at a *dif* site. Regions of the proteins implicated in XerC–XerD interactions are coloured in orange (residues 256–258) and magenta (residues 263–267).

It is evident from this model that the C-terminal helix, which contains the active site tyrosine at one end, forms a major part of the interaction of XerD with XerC. This interaction suggests a simple mechanism for how the binding of both proteins at a recombination site may

promote cleavage of the DNA by the active site tyrosine (see above), whereby a small shift in the position of the C-terminal helix, αN, could allow rotation of the side chain of Y279 into a position suitable for attack of the scissile phosphate. It is plausible that the mutants of XerD which do not support XerC catalysis fail to facilitate the conformational change that would allow the active site tyrosine to initiate an in-line attack. The 'inactive' buried position of Tyr279 is stabilized by two H-bonds with main-chain amide and carbonyl groups. We note that the λ Int structure is of a protein containing an active site tyrosine to phenylalanine substitution (Kwon *et al.*, 1997). The failure of the phenylalanine to make these H-bonds could lead to an inability to stabilize the 'inactive' position of this residue, thereby explaining the looped out position of the λ Int active site 'tyrosine' in the structure. In the structure of the dimer of HP1 Int, the active site tyrosines are in the active position ready for in-line attack (Hickman *et al.*, 1997). We believe that this provides additional evidence for the involvement of protein–protein interactions in the activation of catalysis.

The XerD–XerC pseudodimer proposed here is organized similarly to a dimer observed in the structure of HP1 Int (Hickman *et al.*, 1997). The two active site tyrosines in the HP1 dimer are positioned 28 Å apart, consistent with a model in which the dimer binds to a single recombination site, as proposed here for XerC–XerD. Although Hickman *et al.* (1997) considered this possibility, they preferred an alternative model in which the dimer is derived from monomers bound to separate recombination sites. The model for λ Int bound to DNA has a single protomer bound to B-form DNA, the protein molecule having the same overall orientation with respect to the recombination site as that proposed here for the XerD–DNA complex (Kwon *et al.*, 1997).

The three different structures of λ integrase family recombinases reveal many similarities, though the differences in the region of the protein containing the active site tyrosine may indicate significant differences in details of the catalytic mechanism and its control. Alternatively, the differences between the structures may represent 'snapshots' of different conformational states of enzymes that have essentially identical catalytic mechanisms. For example, the differences in the C-terminal regions of the proteins (Figure 2) could represent inactive (XerD) and activated forms of the enzymes (λ Int and HP1 Int). The structures provide an important starting point for further structural, biochemical and genetic experiments that should give new insight into how recombination sites come together and undergo the two sequential pairs of strand exchanges that constitute a normal recombination reaction.

Materials and methods

Purification and crystallization of XerD

XerD was purified from *E.coli* strain DS9009 (AB1157 *recF lacI^q^ lacZ*ΔM15 *xerD2*:Tn10-9 *xerC*Y17) containing plasmid pRM132 which expresses XerD from the *plac* promoter (Blakely *et al.*, 1993). Cells (3 l) were grown at 37°C in LB supplemented with ampicillin (125 μg/ml) to $A_{600} = 0.4$, and expression of XerD was induced with isopropyl-β-D-thiogalactopyranoside (IPTG; 0.5 mM) for 4 h, after which the cells were harvested, washed and frozen. Cells were resuspended (3 ml/g) in a buffer containing 50 mM Tris–HCl, pH 8.0, 0.3 M NaCl and 10% glycerol, incubated with lysozyme (400 μg/ml) for 15 min on ice, sonicated and particulate debris removed by centrifugation. Because XerD binds avidly to Ni resins, the cell lysate was fractionated on a 15 ml Ni-NTA Superflow column (Qiagen); bound proteins were eluted with the above buffer containing 200 mM imidazole. Fractions enriched with XerD were loaded directly onto a 5 ml HiTrap Heparin column (Pharmacia) equilibrated with a buffer containing 50 mM Tris–HCl, pH 8.0, 0.3 M NaCl, 1 mM EDTA, 1 mM dithiothrietol (DTT) and 10% glycerol. After washing with a similar buffer containing 0.4 M NaCl, the column was eluted with a NaCl gradient (0.4–1.5 M). Fractions containing XerD were loaded directly onto a 5 ml HiTrap Blue column (Pharmacia) equilibrated with a buffer containing 50 mM Tris–HCl, pH 8.0, 1 M NaCl, 1 mM EDTA, 1 mM DTT and 10% glycerol, and XerD was eluted by increasing the NaCl concentration to 2.25 M. The XerD-containing fraction was made 0.33 M in $(NH_4)_2SO_4$ before being applied to a 1 ml phenyl Superose (Pharmacia) column which was washed with a buffer containing 50 mM Tris–HCl, pH 8.0, 1 M NaCl, 0.33 M $(NH_4)_2SO_4$, 1 mM EDTA, 1 mM DTT and 10% glycerol. XerD was eluted from the column with a NaCl/$(NH_4)_2SO4$ gradient (1–0.1 M and 0.33–0 M respectively). XerD (~10 mg/3 l) was >95% pure, with most of the contaminants being products of proteolytic degradation of XerD. The identity of XerD and its degradation products was confirmed by N-terminal sequencing. Mixtures of protease inhibitors were used wherever possible to minimize XerD degradation. Crystals were grown at 20°C by the hanging drop vapour diffusion method. XerD (2.5–5 mg/ml) in 10 mM Tris (pH 8.5), 500 mM NaCl or KCl, 1 mM DTT and 1 mM $(NH_4)_2SO_4$ was mixed at a 1:1 ratio with well solution containing 100 mM Tris (pH 8.0) and 6–14% polyethylene glycol (PEG) 400. Crystals formed after 24–48 h incubation.

Structure determination

The crystals belong to the space group $P6_5$, with unit cell dimensions $a = b = 102.9$ Å, $c = 56.0$ Å. Crystals were collected into harvest solution (0.5 M KCl, 12% PEG 400, 100 mM Tris pH 7.5) prior to soaking in harvest solution containing heavy atom compounds as indicated in Table I. All data were collected from crystals that were flash frozen in rayon loops at 100 K. X-ray data were collected on a MarResearch image plate detector, either at the Synchrotron Radiation Source (Daresbury) or using a rotating anode source. Integrated intensities were calculated and scaled together with the programs DENZO and SCALEPACK (Otwinowski, 1993). Initially, two heavy atom sites in the lead derivative were identified by direct methods using the program SHELXS (Sheldrick, 1993). Further heavy atom sites in the lead derivative, together with the sites in the gold and mercury derivatives, were determined by difference Fourier calculations. A 2.5 Å map was calculated using the MIR phases. The CCP4 program suite (Collaborative Computing Project No. 4, 1994) was used unless stated otherwise. The initial map was solvent flattened using DM and an initial model was built into the solvent-flattened map using the graphics program TURBO FRODO (Roussel and Cambillau, 1989). The model was refined using positional refinement in REFMAC, with restrained temperature factors. Manual rebuilding was carried out between the refinement cycles. The crystallographic free *R*-factor (Brünger, 1992) was monitored at each stage to prevent model bias. Statistics on the final model are presented in Table I.

Acknowledgements

We wish to thank SRS (Daresbury) for access to synchrotron radiation facilities, and S.Islam and M.Sternberg for providing their graphics programme PREPPI. We are grateful to Tom Ellenberger, Art Landy and Fred Dyda for communicating their results prior to publication and for providing coordinates for their structures of λ Int and HP1 Int. We thank Garry Blakely and Bernard Hallet for helpful discussions and for collating much relevant information on recombinase–DNA interactions. This work was supported by Wellcome Trust grants to D.B.W and D.J.S.

References

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

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

Arciszewska,L.K. and Sherratt,D.J. (1995) Xer site-specific recombination *in vitro. EMBO J.*, **14**, 2112–2120.

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.
Blake,J.A.R., Ganguly,N. and Sherratt,D.J. (1997) DNA sequence of recombinase-binding sites can determine Xer site-specific recombination outcome. *Mol. Microbiol*., **23**, 387–398.
Blakely,G.W. and Sherratt,D.J. (1994) Interactions of the site-specific recombinases XerC and XerD with the recombination site *dif. Nucleic Acids Res*., **22**, 5613–5620.
Blakely,G.W. and Sherratt,D.J. (1996a) Determinants of selectivity in Xer site-specific recombination. *Genes Dev*., **10**, 762–773.
Blakely,G.W. and Sherratt,D.J. (1996b) *Cis* and *trans* in site-specific recombination. *Mol. Microbiol*., **20**, 234–237.
Blakely,G.W., 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 Biol*., **3**, 789–798.
Blakely,G.W., May,G., McCulloch,R., Arciszewska,L., Burke,M., Lovett,S. and Sherratt,D.J. (1993) Two related recombinases are required for site-specific recombination at *dif* and *cer* in *Escherichia coli* K12. *Cell*, **75**, 351–361.
Blakely,G.W., Davidson,A.O. and Sherratt,D.J. (1997) Binding and cleavage of nicked substrates by site-specific recombinases XerC and XerD. *J. Mol. Biol*., **265**, 30–39.
Brennan,R.G. and Matthews,B.W. (1989) The helix–turn–helix DNA binding motif. *J. Biol. Chem*., **264**, 1903–1906.
Brünger,A.T. (1992) The free *R*-value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–474.
Carson,M. (1991) Ribbons 2.0. *J. Appl. Crystallogr*., **24**, 958–961.
Chen,J.-W., Evans,B.R., Yang,S.-H., Teplow,D.B. and Jayaram,M.(1991) Domain of a yeast site-specific recombinase (Flp) that recognizes its target site. *Proc. Natl Acad. Sci. USA*, **88**, 5944–5948.
Chen,J.-W., Evans,B.R., Yang,S.-H., Araki,H., Oshima,Y. and Jayaram,M. (1992a) Functional analysis of box I mutations in yeast site-specific recombinases Flp and R: pairwise complementation with recombinase variants lacking the active-site tyrosine. *Mol. Cell Biol*., **12**, 3757–3765.
Chen,J.-W., Lee,J. and Jayaram,M. (1992b) DNA cleavage *in trans* by the active site tyrosine during Flp recombination: switching protein partners before exchanging strands. *Cell*, **69**, 647–658.
Collaborative Computing Project No. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr*., **D50**, 760–763.
Colloms,D.S., McCulloch,R., Grant,K., Neilson,L. and Sherratt,D.J. (1996) Xer-mediated site-specific recombination *in vitro*. *EMBO J*., **15**, 1172–1181.
Colloms,S.D., Bath,J. and Sherratt,D.J. (1997) Topology of Xer mediated site-specific recombination at *psi* and *cer. Cell*, **88**, 855–864.
Cornet,F., Hallet,B. and Sherratt,D.J. (1997) Xer recombination in *Escherichia coli*: site-specific DNA topoisomerase activity of the XerC and XerD recombinases. *J. Biol. Chem.*, in press.
Evans,B.R., Chen,J.W., Parsons,R.L., Bauer,T.K., Teplow,D.B. and Jayaram,M. (1990) Identification of the active-site tyrosine of FLP recombinase. *J. Biol. Chem*., **265**, 18504–18510.
Friesen,H. and Sadowski,P.D. (1992) Mutagenesis of a conserved region of the gene encoding the FLP recombinase of *Saccharomyces cerevisiae*. *J. Mol. Biol*., **225**, 313–326.
Han,Y.W., Gumport,R.I. and Gardner,J.F. (1993) Complementation of bacteriophage lambda integrase mutants: evidence for an intersubunit active site. *EMBO J*., **12**, 4577–4584.
Hayes,F. and Sherratt,D.J. (1997) Recombinase binding specificity at the chromosome dimer resolution site *dif* of *Escherichia coli. J. Mol. Biol*., **266**, 525–537.
Hayes,F., Lubetzki,S. and Sherratt,D.J. (1997) *Salmonella typhimurium* specifies a circular chromosome dimer resolution system which is homologous to the Xer site-specific recombination system in *Escerichia coli. Gene*, in press.
Hickman,A.B., Waninger,W., Scocca,J. and Dyda,F. (1997) Molecular organisation in site-specific recombination: the catalytic domain of bacteriophage HP1 integrase at 2.7 Å resolution. *Cell*, **89**, 227–237
Hoess,R., Abremski,K., Irwin,S., Kendall,M. and Mack,A. (1990) DNA specificity of the Cre recombinase resides in the 25 kDa carboxyl domain of the protein. *J. Mol. Biol*., **216**, 873–882.
Jayaram,M. (1997) The *cis–trans* paradox of integrase. *Science*, **276**, 49–51.
Jayaram,M. and Lee,J. (1995) Return to sobriety after the catalytic party. *Trends Genet*., **11**, 432–433.
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. (1993) Mechanistic and structural complexity in the site-specific recombination pathways of Int and Flp. *Curr. Opin. Genet. Dev.*, **3**, 699–707.
Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*., **26**, 283–291.
Lee,J., Serre,M.-C., Yang,S.-W., Whang,I., Araki,H., Oshima,Y. and Jayaram,M. (1992) Functional analysis of Box II mutations in yeast site-specific recombinase Flp and R. *J. Mol. Biol.*, **228**, 1091–1103.
Lee,J., Whang,I., Lee,J. and Jayaram,M. (1994) Directed protein replacement in recombination full sites reveals *trans*-horizontal DNA cleavage by Flp recombinase. *EMBO J*., **13**, 5346–5354.
Leslie,N. and Sherratt,D.J. (1995) Site-specific recombination in the replication terminus region of *Escherichia coli*: functional replacement of *dif*. *EMBO J*., **14**, 1561–1570.
Moitoso de Vargas,L., Pargellis,C.A., Hasan,N.M., Bushman,E.W. and Landy,A. (1988) Autonomous DNA binding domains of λ integrase recognize two different sequence families. *Cell*, **54**, 923–929.
Nash,H.A. (1996) Site-specific recombination: integration, excision, resolution and inversion of defined DNA segments. In Neidhardt,F.C. *et al.* (eds), *Escherichia coli and Salmonella*: *Cellular and Molecular Biology.* ASM Press, Washington, DC, pp. 2363–2376.
Nicholls,A. and Honig,B.J. (1991) A rapid finite-difference algoritim, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation. *J. Comput. Chem*., **12**, 435–445.
Nunes-Düby,S.E., Tirumalai,R.S., Dorgai,L., Yagil,E., Weisberg,R.A. and Landy,A. (1994) λ Integrase cleaves DNA in *cis*. *EMBO J*., **13**, 4421–4430.
Otwinowski,Z. (1993) *Data Collection and Processing*. SERC Daresbury DL/SC1/R34, 56–62.
Pan,G. and Sadowski,P.D. (1992) Ligation activity of Flp recombinase. *J. Biol. Chem*., **267**, 12397–12399.
Pan,G. and Sadowski,P.D (1993) Identification of the functional domains of the FLP recombinase. *J. Biol. Chem*., **268**, 22546–22551.
Panigrahi,G.B. and Sadowski,P.D. (1994) Interaction of the NH_2- and COOH-terminal domains of the FLP recombinase with the FLP recognition target sequence. *J. Biol. Chem*., **269**, 10940–10945.
Panigrahi,G.B., Beatty,L.G. and Sadowski,P.D. (1992) The FLP protein contacts both major and minor grooves of its recognition target sequence. *Nucleic Acids Res.*, **20**, 5927–5935.
Pargellis,C.A., Nunes-Düby,S.E., Moitoso de Vargas,L. 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.
Roussel,A. and Cambillau,C. (1989) TURBO-FRODO. In *Silicon Graphics Geometry Partner Directory.* Silicon Graphics, Mountain View, CA, pp. 77–78.
Sadowski,P.D. (1995) The Flp recombinase of the 2μm plasmid of *Saccharomyces cerevisiae. Prog. Nucleic Acid Res. Mol. Biol*., **51**, 53–89.
Schultz,S.C., Shields,G.C. and Steitz,T.A. (1991) Crystal structure of a CAP–DNA complex: the DNA is bent by 90°. *Science*, **253**, 1001–1007.
Sheldrick,G.M., Dauter,Z., Wilson,K.S., Hope,H. and Sieker,L.C. (1993) The application of direct methods and Patterson interpretation to high resolution protein data. *Acta Crystallogr*., **D49**, 18–23.
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*. Philos. Trans. R. Soc. Lond. B.*, **347**, 37–42.
Spiers,A. and Sherratt,D.J. (1997) Relating primary structure to function in the *Escherichia coli* XerD site-specific recombinase. *Mol. Microbiol*., **24**, 1071–1082.
Stark,W.M. and Boocock,M.R. (1995a) Topological selectivity in site-specific recombination. In Sherratt,D.J. (ed.), *Mobile Genetic Elements*. IRL Press at Oxford University Press, pp. 101–129.
Stark,W.M. and Boocock,M.R. (1995b) Gatecrashers at the catalytic party. *Trends Genet*., **11**, 121–123.
Stark,W.M., Boocock,M.R. and Sherratt,D.J. (1992) Catalysis by site-specific recombinases. *Trends Genet*., **8**, 432–439.
Tecklenburg,M., Naumer,A., Nagapan,O. and Kuempel,P.L. (1995) The *dif* resolvase locus of the *Escherichia coli* chromosome can be replaced by a 33-bp sequence, but function depends on location. *Proc. Natl Acad. Sci. USA*, **92**, 1352–1356.

Received on May 21, 1997; revised on June 26, 1997