The EMBO Journal vol.13 no.7 pp.1514–1524, 1994

Model for a DNA-mediated synaptic complex suggested by crystal packing of γδ resolvase subunits

Phoebe A.Rice[1] and Thomas A.Steitz[2]

Departments of Molecular Biophysics and Biochemistry and Chemistry, Howard Hughes Medical Institute, Yale University, New Haven, CT, USA
[1]Present address: NIDDK/NIH, Building 5, Room 237, Bethesda, MD 20892, USA
[2]Corresponding author at: Yale University, Department of Molecular Biophysics and Biochemistry, 266 Whitney Avenue, Bass Center—Room 418, PO Box 208114, New Haven, CT 06520-8114, USA

Communicated by J.A.Steitz

The packing arrangement of the 12 subunits of intact γδ resolvase in the unit cell of a hexagonal crystal form suggests a model for site-specific recombination that involves a DNA-mediated synaptic intermediate. The crystal structure has been determined by molecular replacement and partially refined at 2.8/3.5 Å resolution. Although the small DNA-binding domain is disordered in these crystals, packing considerations show that only a small region of space in the crystal could accommodate a domain of its size. A family of related models for a synaptic complex between two DNA duplexes and 12 monomers that are arranged as situated in the crystal is consistent with the known topology of the complex and the distances between the three resolvase dimer-binding sites per DNA; further, these models place the two DNA recombination sites in contact with each other between two resolvase dimers, implying that strand exchange is accomplished through direct DNA–DNA interaction. A major role postulated, then, for the resolvase protein assembly is to stabilize a *res* DNA structure that is close to the topological transition state of the reaction.

Key words: DNA–protein complex/resolvase/site-specific recombination

Introduction

γδ resolvase, a member of the Tn3 family of site-specific recombinases, is a 20.5 kDa enzyme responsible for the second step in the transposition of the γδ transposon (reviewed in Grindley and Reed, 1985; Hatfull and Grindley, 1988; Hatfull *et al.*, 1988; Sherratt, 1989). This enzyme resolves a co-integrate DNA molecule that contains both the donor and target plasmids bounded by directly oriented copies of the transposon into two product circles each containing one copy of the transposon (Figure 1A). Resolvase acts at a 115 bp site, termed *res*, within the transposon. *res* contains three binding sites for resolvase dimers, each consisting of an inverted repeat of a 12 bp binding site with a variable internal spacer (Grindley *et al.*, 1982; Kitts *et al.*, 1983) (Figure 1B). As this site overlaps the promoters for both the *tnpR* (resolvase) and the *tnpA* (transposase) genes, resolvase also acts as a transcriptional repressor.

Strand exchange occurs via staggered double-strand cleavage at site I, creating a two base 3′ overhang (Reed and Grindley, 1981). All three sites must be present for the reaction to proceed (Wells and Grindley, 1984). The reaction can be performed *in vitro* and requires only supercoiled substrate, resolvase and Mg^{2+} (Reed, 1981). If Mg^{2+} is not included, resolvase will cut but not religate the DNA, and an intermediate can be isolated in which resolvase is covalently linked to the 5′ phosphate of the DNA via serine 10 (Reed and Grindley, 1981; Reed and Moser, 1984; Hatfull and Grindley, 1986). Thus, serine 10 is expected to mark the active site of this recombinase and should be able to contact the DNA phosphates at the cleavage sites.

Resolvase exhibits remarkable specificity in two major respects. First, it only cuts the DNA at the center of site I, although it is bound to all three inverted sequence repeats in the *res* site. Second, recombination only occurs if two copies of the *res* site are found on the same molecule, and in direct (head-to-tail) orientation. Products of both intermolecular reactions and reactions between *res* sites in inverted orientation are nearly undetectable (Reed, 1981; Kitts *et al.*, 1983; Krasnow and Cozzarelli, 1983). An individual monomer of resolvase is, therefore, sensitive to

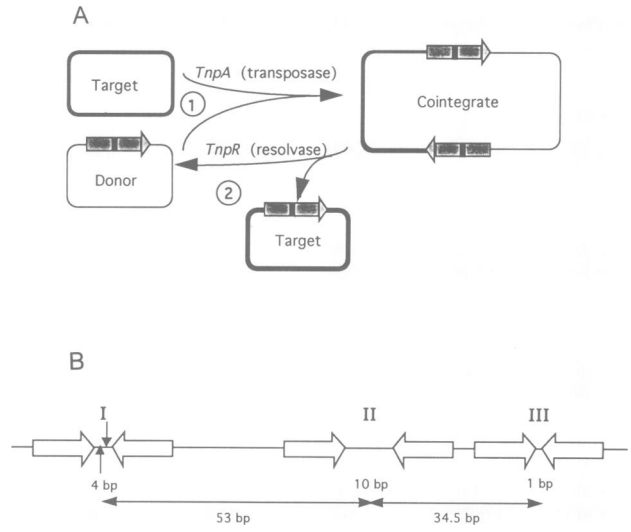

Fig. 1. Cartoon of transposition by the γδ transposon and diagram of the *res* site. (**A**) γδ is a member of the Tn3 family of transposons, which transpose by a two-step process. The first step, catalyzed by transposase, is fusion of the donor and target plasmids to form a co-integrate molecule in which the transposon (shaded box) is duplicated. The second step, catalyzed by resolvase, is a conservative recombination event at a site within the transposon termed *res* (black bar). (**B**) The *res* site DNA contains three inverted repeat sequences (large open arrows) to which resolvase dimers bind. Recombination occurs at the center of site I at positions indicated by arrows. The numbers of base pairs between the centers of the three sites are shown below.

both the details of the particular inverted repeat to which it is bound and to the topological context of the site to which it is bound.

The complex formed by resolvase and a single linear *res* site (resolvosome) appears to be compact (Salvo, 1987). Resolvase binds to *res* in a highly co-operative manner and bends the DNA at each inverted repeat (Salvo and Grindley, 1988). An unstacking or kinking of the bases at the center of the cross-over site is indicated by enhanced intercalation of MPE-Fe(II) and hydrolysis of the DNA (Hatfull *et al.*, 1987). The spacing between sites I and II can only be varied by an integral number of helical turns. DNase I cleavage patterns indicate that the DNA in this segment is curved. This looping out of the DNA between sites I and II is dependent on the spacing of the sites, and on the presence of site III (Salvo and Grindley, 1988).

The synaptic complex formed by the interaction of two *res* sites (the 'synaptosome') and the reaction product have tightly defined topologies. Formation of the synaptosome traps three negative interdomainal nodes, while strand exchange creates a single positive node. This results in the formation of product circles that are singly linked catenanes joined by two negative nodes (Cozzarelli *et al.*, 1985; Wasserman and Cozzarelli, 1985; Wasserman *et al.*, 1985).

The resolvase protein can be cleaved by α-chymotrypsin into two domains: a 140 residue N-terminal catalytic domain, or large fragment, which contains the active site residue serine 10 and is responsible for protein−protein contacts, and a 43 residue C-terminal domain containing a helix−turn−helix motif. The small domain binds to each of the six half-sites of *res* DNA independently without the co-operativity observed with the intact protein (Abdel-Meguid *et al.*, 1984).

The structure of the catalytic domain of resolvase has been determined (Hatfull *et al.*, 1989; Sanderson *et al.*, 1990) and refined at 2.3 Å resolution (P.A.Rice and T.A.Steitz, in preparation). These crystals are in space group $C222_1$, with three monomers in the asymmetric unit. The dimer formed by two of these monomers in the crystal is now known to represent the solution dimer (Hughes *et al.*, 1993), while another contact between monomers in the crystal has been shown to represent an important interdimer interaction in the resolvosome (Hughes *et al.*, 1990).

Crystals of intact resolvase are hexagonal bipyramids of space group $P6_422$, with a single monomer in the asymmetric unit (Weber *et al.*, 1982). Both the intact protein and the large fragment can crystallize isomorphously in this crystal form, although it was clear from the small size of the intensity differences between data sets that the small domain was most probably disordered (Abdel-Meguid *et al.*, 1986).

The structure of the hexagonal crystal form reported here shows conservation of most intersubunit contacts seen in the orthorhombic crystal form, has only a small locus of points where the disordered, 60 residue DNA-binding domain could lie and suggests a possible structure for the synaptosome. Using the arrangement of the 12 subunits in the unit cell of this hexagonal crystal and information about the roles of these subunit interactions derived from molecular genetic studies (Hughes *et al.*, 1990, 1993), we propose a model of the synaptic complex. In this model, the two duplex DNAs at site I are in contact with each other at a point of 222 symmetry, rather than being on the outside of a 222 tetramer of protein subunits. This model implies that exchange of resolvase subunits is not necessary for the strand-exchange reaction, as has often been proposed. Rather, this protein assembly is postulated to stabilize a distorted DNA structure that is close to the topological transition state and requires only modest changes in its structure during the subsequent covalent steps of recombination.

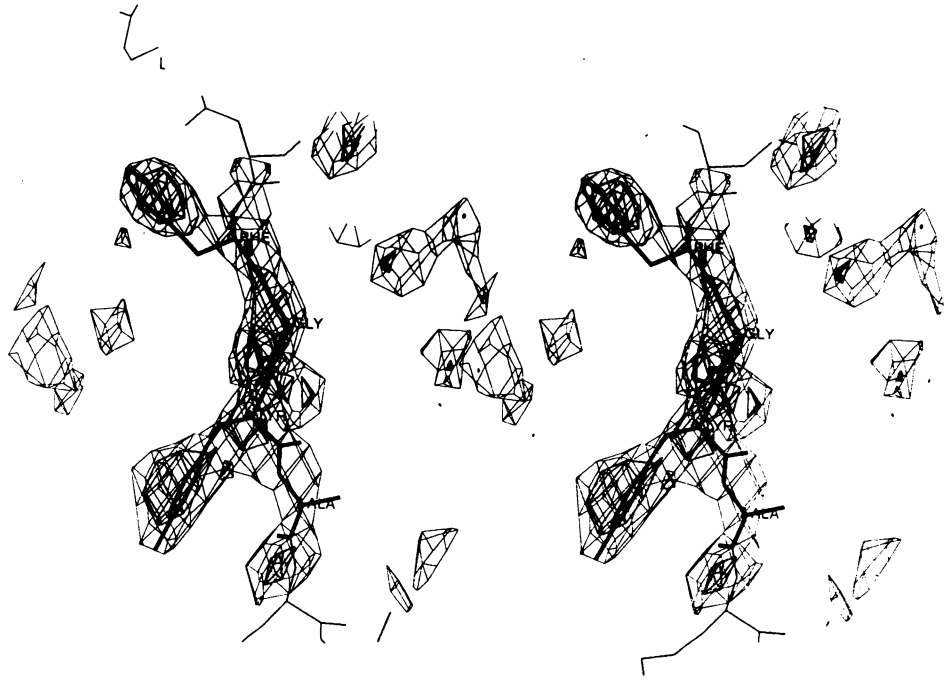

Fig. 2. Omit map of residues 4−7. A stereo view of a difference electron density map calculated using $F_{obs}-F_{calc}$ as coefficients, large fragment data from 8 to 3.3 Å resolution and phases calculated from co-ordinates which had been subjected only to rigid body refinement after molecular replacement, and from which residues 4−7 (shown in bold) had been deleted. The map is contoured at 1 and 2 σ.

Results

Structure determination and quality of the model

The structure was solved by molecular replacement and analysis of crystal packing. Several tests were performed to check the validity of the molecular replacement solution. First, when the search model was refined as nine individual rigid pieces, one corresponding to each α-helix or individual strand of β-sheet, the separate pieces did not wander in different directions. However, upon refinement of monomer 1 of the orthorhombic form, whose central β-sheet has an overall twist 10.5° less (P.A.Rice and T.A.Steitz, in preparation), as individual rigid pieces of secondary structure against data from the hexagonal form, the molecule regained the twist of monomer 3. Thus, the overall twist of the single monomer in the asymmetric unit of the hexagonal form resembles that of monomers 2 and 3 rather than monomer 1 of the orthorhombic form. Finally, in a difference map calculated using $F_o - F_c$ as coefficients and phases derived from a rigid-body refined model with part of the first β-strand deleted, density for the omitted residues was clearly visible (Figure 2). This is not likely to result from model bias, as only rigid-body refinement had been done. The current model includes residues 1–37 and 45–115, and has been partially refined to an R-factor of 31%.

Subunit packing

The packing of subunits in this crystal form is remarkably similar to that seen in the previously reported $C222_1$ form. The non-crystallographic dyad axes that relate the three monomers in the asymmetric unit of the orthorhombic form all superimpose approximately on crystallographic dyads in the hexagonal form and, in fact, all 2-fold-related intersubunit interactions seen in the orthorhombic crystal form are present here as well. Both crystals can be described as arrays of dimers held together by the formation of tetramers with local 222 symmetry (Figure 3). There are two major differences between the two crystal forms. First, in the orthorhombic crystal form, monomer 1 does not participate in a tetramer. Second, an additional non-2-fold-related packing is generated

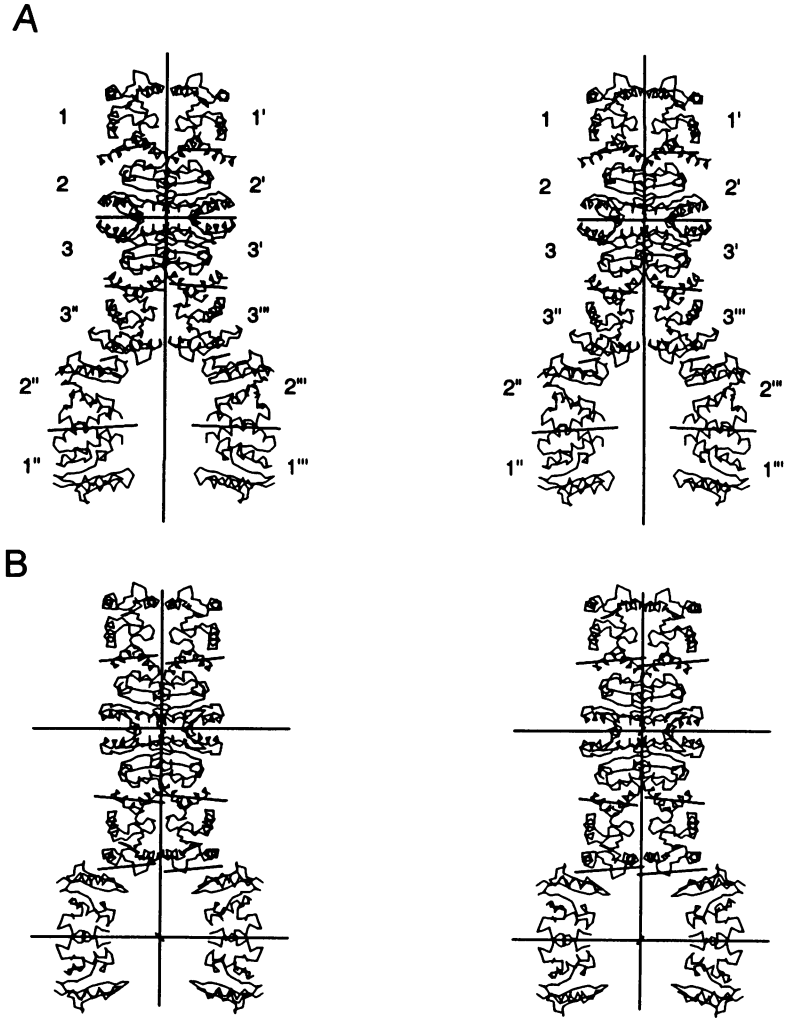

Fig. 3. Diagram of the packing of resolvase monomers found in both crystal forms. The α-carbon backbones of 12 monomers as they are packed in the previously reported $C222_1$ crystal form (A) and in the $P6_422$ crystal form (B) are shown. (**A**) shows four asymmetric units (with three independent monomers in each), or half the contents of a full unit cell of the $C222_1$ crystal form. Individual monomers are numbered as in Sanderson *et al.* (1990), with primes denoting monomers related by crystallographic symmetry operations. The vertical axis and the axes related monomers 3 to 3'' and 3' to 3''' are crystallographic; all others are local, non-crystallographic dyads. (**B**) shows the full unit cell of the $P6_422$ crystal. These crystals contain only one monomer in the asymmetric unit and thus all the dyad axes are crystallographic in this case. Both crystals can be described as assemblies of dimers of the type formed by monomers 1 and 2, and by 3 and 3'', held together by the formation of tetramers with 222 symmetry and formed by monomers 2,3,2' and 3', as numbered in (**A**).

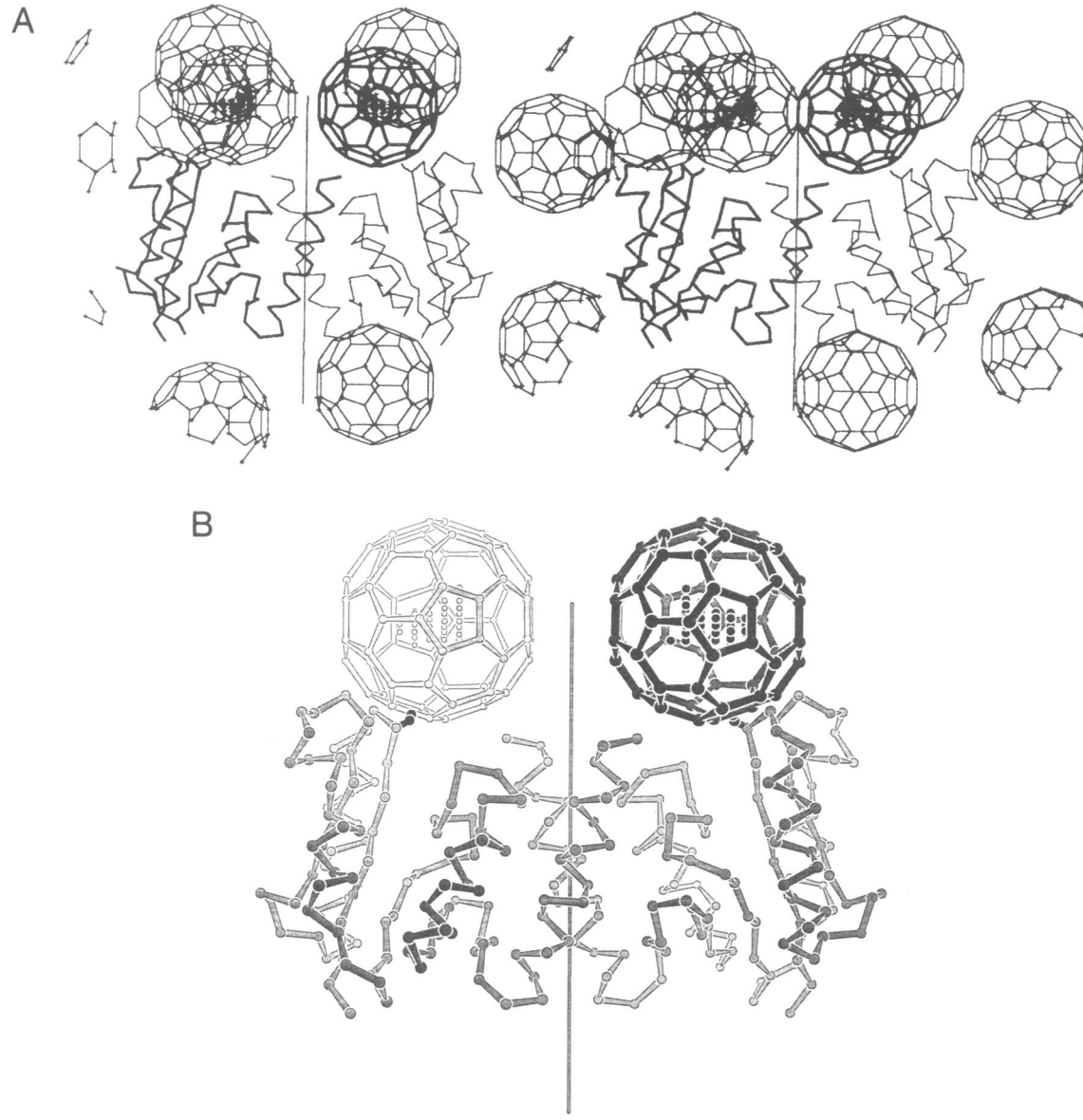

Fig. 4. Model of the intact protein. (**A**) A stereo α-carbon drawing showing a dimer of resolvase. The single monomer drawn in bold represents the contents of one asymmetric unit, and the sphere drawn in bold represents the probable location of the DNA-binding domain belonging to this monomer. The small crosses within the sphere represent the points found in a packing search where the center of a sphere 20 Å in diameter could fit. All other symmetry-related model-built DNA-binding domains within 50 Å of the end of the C-terminal helix of this monomer are drawn as lighter spheres. The bold sphere has been assigned to the bold monomer as it is the one that lies the closest to the C-terminus of that monomer. (**B**) A dimer of resolvase is shown with spheres representing the probable location of the DNA-binding domain. The small dots inside the spheres represent those points in the hexagonal unit cell where a sphere 20 Å in diameter could fit. The side chain of the active site serine 10 is shown in black. This figure was made by the program MAXIMAGE, written by Mark Rould at Yale.

by the crystallographic 2_1 screw axis in the orthorhombic form. In the hexagonal form, there is only one monomer in the asymmetric unit, so all monomers make both dimer- and tetramer-type contacts.

Location of the DNA-binding domain

No interpretable experimental electron density for the small domain was found. In fact, the molecule appears to become disordered at the same place as does the large fragment in the orthorhombic crystal form: although interpretable density extends only to about residue 115, there is some density suggesting another turn of α-helix beyond this. The previously reported large fragment model extends to residue 122.

Analysis of the crystal packing, however, shows that there is only a small region of space where a domain of its size could lie. This is not due to the small size of the solvent channels in the crystal, but rather to the large number of 2-fold axes that cannot run through protein. Figure 4 shows a dimer of resolvase with the DNA-binding domains represented as 20 Å diameter spheres centered over this region.

The DNA-binding domain lies just past the end of the C-terminal helix of the large fragment, and near the active site serine 10 of the opposite monomer. It also lies near the solvent-exposed face of the β-sheet of another monomer. While we cannot be certain to which of these neighboring monomers the DNA-binding domain should be attached, the

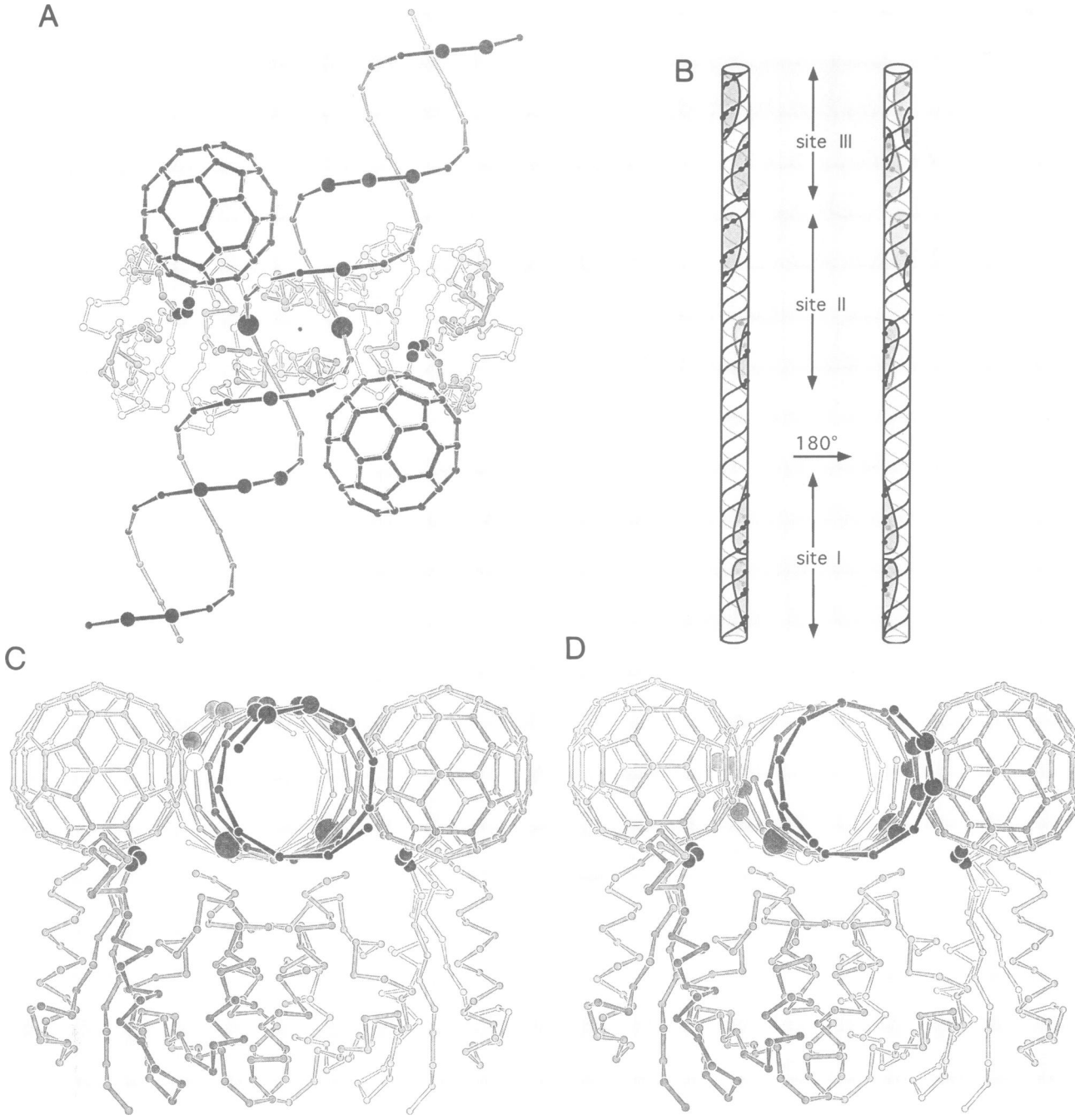

Fig. 5. (**A**) Model of the resolvase–site I DNA complex. One possible orientation of the DNA across the resolvase dimer is shown. The protein is drawn as a Cα model with a sphere to represent the possible location of the DNA-binding domain. Phosphates whose ethylation interferes with binding of the small domain and the intact protein are shown enlarged and in black, phosphates whose ethylation interferes with binding of the intact protein but not the small domain are shown enlarged and in white, and the phosphates that are attacked during the recombination reaction are shown greatly enlarged and in black (Falvey and Grindley, 1987; Rimphanitchayakit *et al.*, 1989). The side chain of the active site serine 10 is shown enlarged and in black. This particular orientation of the DNA across the protein closely resembles that proposed by Mazzarelli *et al.* (1993). (**B**) The *res* site diagrammed on B-form DNA. Two views of the *res* site are shown with those phosphates whose ethylation interferes with binding by the small domain of resolvase highlighted, and approximate 'footprints' on each half-site shown as shaded regions. While the two halves of sites II and III lie on opposing faces of the DNA, the two halves of site I lie on the same face. If the DNA at the center of site I were to be unwound by half a turn, it would much more closely resemble the other two sites. (**C**) Model for the resolvase–site I DNA complex as in (A), but viewed with the dyad axis lying in the plane of the page. (**D**) The phosphate backbone has been broken at the cross-over site and each half-site rotated by 90° to show the consequences of untwisting the DNA by half a helical turn at the center of the site. The phosphates whose ethylation interferes with binding of the small domain now lie on the same faces of the DNA helix that the small domains contact. These drawings (A, C and D) were made with the program MAXIMAGE written by Mark Rould.

most logical choice is the one whose C-terminal helix points directly to it, as suggested in Figure 4. Any other choice requires an extensive linker region to connect the end of this helix to the small domain. This is technically possible, as there are ~20 disordered residues between the last clearly defined residue in the electron density and the chymotrypsin cleavage site, but we cannot justify such an assumption without further data.

Discussion

Structure of intact resolvase

The similarity in subunit packing between the hexagonal and orthorhombic crystal forms underscores the significance of the dimer and the tetramer contacts seen in both crystals. Recent experiments show that the dimer formed by the interaction between the two C-terminal helices of the large fragment of monomers 1 and 2 (Figure 4) is the dimer in solution that binds to an individual DNA inverted repeat (Hughes *et al.*, 1993). Mutations of residues lying at a second interface between monomers 2 and 3′ of the tetramer eliminate not only recombination activity, but also the co-operativity among dimers normally seen when resolvase binds to the *res* site (Hughes *et al.*, 1990). Thus, we believe that the tetramer seen in both crystal forms arises from an important interdimer interaction, implying that the arrangement of resolvase monomers in the crystal probably reflects the organization of higher-order complexes formed in solution during recombination. There are also other examples in which the packing of monomers in a crystal reflects higher-order complexes formed in solution, including the tobacco mosaic virus coat protein (Bloomer *et al.*, 1978) and recA (Story *et al.*, 1992).

The location of the DNA-binding domain in the intact protein (Figure 4) is based on packing considerations and provides a useful framework for considering the nature of the DNA–protein complex. However, the assignment of a particular large fragment to a particular model DNA-binding domain is based solely on logical considerations.

Models of resolvase dimer complexed with a single DNA binding site

We can simply dock canonical B-form DNA onto the resolvase dimer and note the correspondence with footprinting data, and the relative positions of Ser10 and the cleavage site at site I (Figure 5A). In spite of the fact that there are several possible ways in which site I DNA might lie on the protein and detailed models of the complex cannot be built, important conclusions can be drawn about the possible nature of the DNA and protein distortions that might occur in such a complex at site I. These conclusions arise from addressing the several discrepancies that exist between the biochemical data and even the best docking of undistorted DNA and protein.

First, the two serine 10 residues of the dimer are further apart than the phosphates which are attacked during the reaction and the DNA-binding domains are closer together than the region of DNA they protect from chemical modification, suggesting a protein conformational change. Although much of this discrepancy can be explained by distortions in the DNA that are known to occur (Hatfull *et al.*, 1987), the discrepancy is so large that some distortions in the protein structure are likely. The large fragment in the orthorhombic crystal form shows an unusually flexible molecule (P.A.Rice and T.A.Steitz, in preparation), which may be important in its recognizing the three different inverted repeat sequences of the *res* site. Not only does the twist of the central β-sheet vary, but also the arrangement of the two monomers forming the catalytic dimer differs. The two versions of this dimer have differences in relative subunit positions that can be described as a scissors-like motion a few degrees in magnitude, with the two C-terminal helices representing the blades of the scissors. Opening the scissors simultaneously moves the serine 10 residues closer together, and the ends of the C-terminal helices and the attached DNA-binding domains farther apart. A larger structural change of this sort may occur when the protein binds to DNA.

Second, the sidedness of DNA contacts with resolvase at the three sites suggests that site I may be unwound by one half-turn upon binding resolvase. Chemical modification protection and interference studies show that the regions of site I contacted by the DNA-binding domains lie on one face of B-form DNA, whereas the contacting half-sites lie on opposing faces of the DNA in sites II and III (Figure 5B) (Falvey and Grindley, 1987; Rimphanitchayakit *et al.*, 1989; Graham and Dervan, 1990). In our model, the DNA-binding domains contact opposing faces of the DNA, consistent with the protection data on sites II and III. The protection data on site I require that either the DNA-binding domains are positioned differently at site I as compared with sites II and III, or the DNA is unwound (or overwound) by half a turn (Figure 5C and D). Negative supercoiling favors unwinding the DNA which would trap at the cross-over site one of the negative supercoils that is subsequently lost during recombination. Klippel *et al.* (1993) have recently reported experimental evidence supporting such unwinding by the related recombinase Gin.

Models for synapsis

Two general classes of site-specific recombination models exist (Figure 6): one has the two DNA duplexes located on the outside of a protein complex, and strand exchange is mediated by a 180° rotation of the protein subunits to which the cleaved DNA duplexes are bound; the second model involves direct DNA–DNA interactions in which the protein complex stabilizes a DNA structure that is the intermediate of the recombination reaction. The first model has generally been assumed for the resolvase/invertase family of site-specific recombinases, while the second model has been demonstrated for the λ integrase family of recombinases, which proceed through a Holliday-junction intermediate (for reviews see Craig, 1988; Landy, 1989).

The first model (Figure 6A), assumed to apply to resolvase, involves two dimers, each bound to a single site I, which form a tetramer, presumably with 222 symmetry (three mutually perpendicular 2-fold axes). The two DNA duplexes lie on the outside of the complex. After strand breakage, two of the protein subunits, each covalently linked to a 5′ phosphate of the DNA, rotate by 180° relative to the other two subunits to accomplish strand exchange, exactly accounting for the topological changes known to occur (Stark *et al.*, 1989). The primary problem with this mechanism is that there is nothing to prevent the subunits and cleaved DNA from coming apart while rotating. The tightly defined topological changes that occur during resolvase-mediated recombination require that the subunits must maintain some contact with one another. Although one might imagine that two relatively flat surfaces could rotate 180° relative to one another while maintaining close contact, the interface of the resolvase dimer is ridged and notched rather than flat. It is difficult to imagine how a pair of dimers could exchange subunits without separating far enough to risk permanent dissociation.

The alternative model (Figure 6B) places the two DNA duplexes in direct contact with one another and postulates that strand exchange takes place via relatively small motions

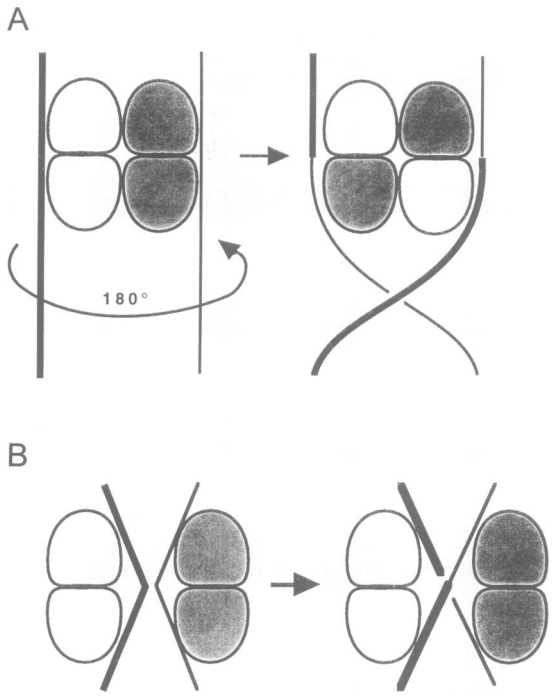

Fig. 6. Models for strand exchange. (**A**) The subunit-exchange model for strand exchange. Two dimers of resolvase, each bound to site I, come together to form a tetramer. After double-strand breaks are made and each half-site is covalently linked to the protein via serine 10, the lower two monomers of the tetramer switch positions by rotating clockwise about one another. The DNA can then be religated in the recombinant configuration. (**B**) Recombination without subunit exchange. This model supposes that the two site I DNAs to be recombined are bound to resolvase dimers and brought together such that DNA strands at the cross-over point are in close proximity. Strand exchange could then occur by a relatively modest motion of only the free 3′ OH of the intermediate from one strand to another, and without dissociation of the resolvase dimers.

of the cleaved DNA ends. This may be achievable, we suggest, because the protein uses binding energy to distort the DNA substrate and stabilize DNA structures that are the topological transition states of this recombination reaction. This class of model appears most consistent with the present crystal structures.

A family of related models of the synaptosome can be constructed using the observed arrangement of subunits; they show the appropriate spacing of sites I, II and III, and bring two site I DNA duplexes into close proximity at a point of 222 symmetry. Starting with the array of 12 subunits as found in the crystallographic unit cell (six 1,2 dimers), two *res* site DNA duplexes can be built onto this protein assembly to form models of the synaptic complex or 'synaptosome' (Figure 7) that account for its known properties. This family of models provides a framework for thinking about a structural basis for the reaction and makes several testable predictions. They are based on the following assumptions. (i) The '1,2' dimer formed by interaction between the C-terminal helices of the large fragment and found in both crystal forms (Figure 4) is the solution dimer that binds to each inverted repeat (Hughes *et al.*, 1993). (ii) The contacts between the dimers that mediate tetramer formation are important higher-order interactions in the synaptosome (Hughes *et al.*, 1990). (iii) The DNA should be arranged such that three negative nodes are trapped in the synaptic complex (Wasserman and Cozzarelli, 1985; Wasserman *et al.*, 1985).

In these models, the two DNA duplexes are held together in synapsis by the interaction between the dimers bound at sites II and III to form a tetramer, and perhaps also by interactions between the DNA-binding domains at site I. The dimers bound at the two site Is interact with proteins at site III via the same type of interactions as between monomers 2 and 3 of the tetramer (as numbered in Figure 3). The four models drawn differ only in how the site I DNA bound to the two lower dimers is arranged. That this tetramer-of-dimers may form the 'core' of a synaptic complex between sites II and III has also been proposed by Grindley and colleagues, and is supported by complementation experiments between mutant resolvases that lack the active site serine 10 and ones that cannot make proper tetramer contacts (Grindley, 1993; Hughes *et al.*, 1993). The path of the DNA in these models closely resembles that predicted from topological analysis of minor products resulting from multiple rounds of recombination (Cozzarelli *et al.*, 1985; Wasserman *et al.*, 1985).

The distance of 178 Å, or slightly less, between the centers of site I and site II in our model is in good agreement with the 180 Å calculated from the separation of 53 bp at 3.4 Å/bp. This model is consistent with the observations that the distance between sites I and II can be increased by integral numbers of helical repeats and that the bend between these sites is dependent on the presence of site III (Salvo and Grindley, 1988). It does not easily accommodate the Tn2501 *res* site, where the site I–II spacer is shorter by one helical turn (Hatfull and Grindley, 1988), although small individual twists of the monomers relative to one another, summed along the length of the synaptosome, might accommodate the difference. The 74 Å distance between the centers of sites II and III in our model is shorter than the calculated distance of 117 Å for 34.5 bp. However, the path we have drawn for the DNA between these sites is the most direct path allowable and makes the tightest possible contact with the protein, and therefore represents a minimum length.

Although the distance between the active sites of the two dimers hypothesized to bind site I in this model is somewhat small to accommodate two undistorted DNA duplexes, several explanations are apparent. At the closest approach, the Cα atoms of residue 120 of two symmetry-related molecules in the hexagonal crystal form are separated by 32.0 Å. In the orthorhombic crystal, the equivalent distance is 3.7 Å larger due to the cumulative effect of small shifts in the relative orientation of the monomers. Such effects may be even larger in solution, especially when the protein is bound to the DNA. Additionally, or alternatively, untwisting and bending of the DNA at the center of site I might flatten the DNA somewhat, as occurs with TATA DNA bound to TATA-binding protein (Kim *et al.*, 1993a,b).

Our model also makes the predictions that the spacer DNA between sites II and III may come into close contact with a set of residues previously identified for their involvement in tetramer-forming contacts: R2, R32 and E56 (Hughes *et al.*, 1990), and that the site I–II spacer of one *res* site passes directly over the site II–III spacer of the other.

Our models are consistent with the observation of Mazzarelli *et al.* (1993) that resolvase binding to sites II and III is significantly asymmetric. The center of the bend in the DNA is probably offset from the protein dyad at both sites II and III. Further, hydroxyl radicals generated at a chemically modified cysteine at position 102 (near the N-terminus of the C-terminal helix) cleaved the 'outside'

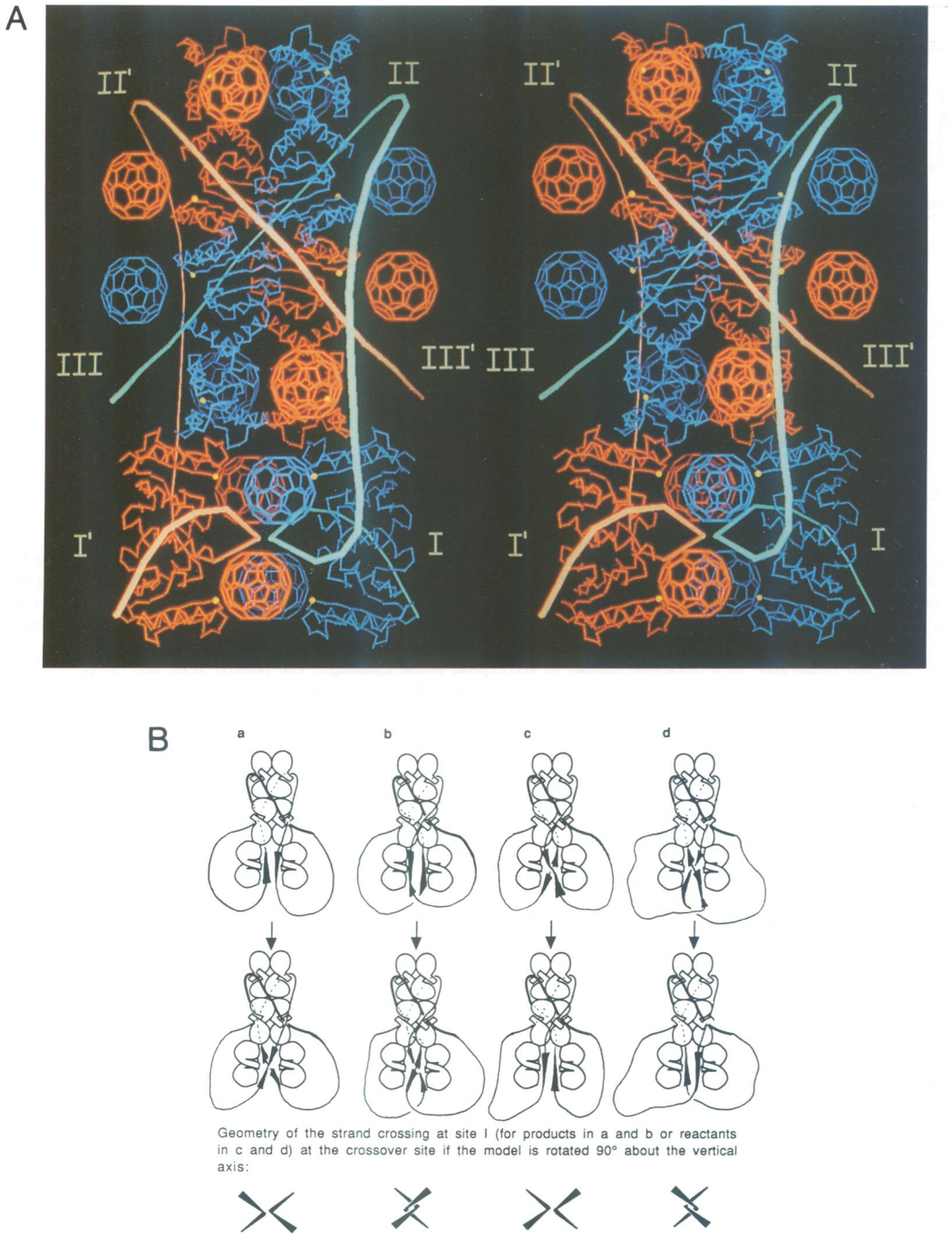

Fig. 7. Model of the synaptosome. (**A**) A stereo Cα representation of 12 monomers of resolvase as they are arranged in the crystallographic unit cell, with pink and blue lines representing how the DNA might wrap around them to form a synaptosome. The three inverted repeats of the two *res* sites are numbered I, II and III, and I′, II′ and III′. The dimers of resolvase are colored to match the *res* sites to which they are bound and the possible locations of the DNA-binding domains are marked by spheres. The α-carbon of the active site serine 10 on each monomer is highlighted in yellow. The line representing the DNA is drawn at a minimum distance from the protein equivalent to the Van der Waals radius of the DNA plus that of a carbon atom (~13 Å). Formation of the tetramer brings sites II and III from the two strands together. The dimer bound to site I interacts with the protein bound to site III via a subset of the contacts involved in formation of the tetramer. The two site I DNAs approach one another quite closely at a second point where there is 222 symmetry in the crystal. The orientation of the DNA at site I is equivalent to cartoon (a) below. The figure was made using the programs MAXIMAGE and SHAZAM written by Mark Rould and Art Perlo, respectively. (**B**) Cartoon diagrams of the model above. The resolvase monomers are represented as spheres with the C-terminal helix (at the dimer interface) protruding. Possible paths for the two *res* sites are drawn as thick and thin lines. The presumed path of the remainder of the plasmid on which the two *res* sites would be found is also drawn in (not to scale, and without supercoiling). The four models diagrammed here differ only in the arrangement of the DNA at site I, and all can give rise to a singly linked catenane product, as shown below. Models (a) and (b) assume that in the reactant complex a single site I is bound across a single resolvase dimer, while models (c) and (d) address the possibility that the DNA may rearrange during synapsis (before strand cleavage), such that in the product complex a single (recombinant) site I is bound across a single dimer. However, if this is the case, there is then no advantage to be gained from unwinding the DNA by half a helical turn at the center of site I when it is bound across a single dimer, as is suggested by the models in Figure 5 and as discussed in the text. Furthermore, these two models require an intertwining of the two site I DNAs in either the product (b) or reactant (d) complex, as shown in the lowermost portion of the figure, which seems rather improbable. Therefore, we believe model (a) to be the most probable representation of the synaptic complex.

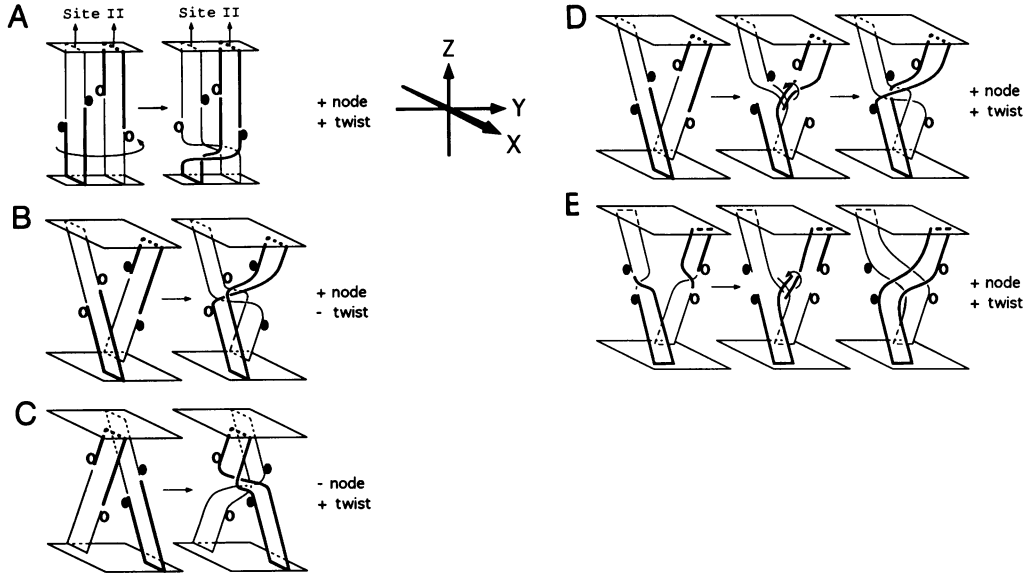

Fig. 8. Models for strand exchange. Relaxed DNA is drawn as a flat ribbon, and the resolvase protein is represented as an open or filled circle attached to the recessed 5′ end of the DNA. The 'plates' attached to both ends of each piece of DNA are kept immobile in all diagrams; thus, the topological consequences of each model can be seen by the twists and crossings of the 'product' ribbons. The two site Is are always assumed to be aligned in essentially parallel orientation, as shown explicitly in (A). (**A**) The subunit exchange model. Each site I is bound by a dimer of resolvase, double-strand breaks are introduced into the DNA, and resolvase becomes covalently linked to the 5′ phosphates via serine 10. The two right-hand monomers then dissociate and rotate 180° in a clockwise manner to accomplish strand exchange. The positive node and the positive twist of one-half introduced to each product circle by this mechanism are in complete agreement with experimental data (see the text). The remaining diagrams all assume that the protein is immobile and strand exchange is accomplished by movement of DNA only. (**B**) The DNA is drawn with the minor grooves facing away from the protein at the center of the cross-over site, as indicated by the MPE-Fe(II) intercalation data (Hatfull and Grindley, 1987). Placing the strands at an angle to one another brings the appropriate free 3′ OH quite close to the proper 5′ phosphate for strand exchange, but simple religation from this configuration results in products with the wrong twist. The crossing angle of the two strands is the same as for models (a) and (c) in Figure 7B (if viewed from different angles). This model for strand exchange is equivalent to Figure 10a of Stark *et al.* (1989). (**C**) Major grooves facing away from the protein at the center. In this case, angling the strands the opposite way relative to one another brings the appropriate 3′ hydroxyls and serine 10s near one another, but simple religation creates a negative rather than a positive node. The crossing angle of the strands here is the same as in models (b) and (d) of Figure 7B. The incorrect sign of the node introduced by simple religation in this case explains the odd intertwining of the two DNAs in Figure 7B if the correct node sign is forced. (**D**) The same orientation of the strands as (B). Here we show that if the overhanging 3′ ends rotate about one another before finishing strand exchange, the topology of the product is correct [and is the same as for the subunit exchange model shown in (A)]. (**E**) The same as (D), except the DNA is assumed to be unwound by half a turn before strand exchange begins. The negative supercoiling thus trapped by the binding of resolvase dimers to site I is destroyed by rotating the overhanging ends about one another, and the final product has no local topological strain at all. Note that this is the case only for this figure and for the subunit-exchange case.

half-sites of sites II and III, but not the half-sites adjacent to the II−III spacer region. In our model, only the outer two half-sites lie near residue 102.

The resolvases are homologous to another family of site-specific recombinases, the invertases, which exhibit specificity for sites in inverted, rather than direct repeats; invertases thus catalyze an inversion rather than a deletion of the DNA segment between the cross-over points. The simpler invertase sites consist of only one inverted repeat, but require an additional site known as the enhancer, which is bound by the accessory protein FIS. The resolvase dimers bound to sites II and III and FIS may serve similar functions in holding the two DNA strands together in a synaptic complex with the correct topology. However, the invertases are capable of synapsing two duplexes of DNA in the absence of FIS and the enhancer (Haffter and Bickle, 1988; Klippel *et al.*, 1988; Heichman and Johnson, 1990; Kanaar *et al.*, 1990). This implies that two invertase dimers at site I must interact directly with one another. In our model, this interaction must be largely via the DNA-binding domains.

Models for the mechanism of strand exchange

Any model for the mechanism of the strand-exchange reaction must account for the known DNA topological change which entails the loss of four negative supercoils (Boocock *et al.*, 1987; Stark *et al.*, 1989). In addition to the two negative nodes trapped by the formation of the singly linked catenane product, a positive interdomainal node must be introduced, and a twist of +1/2 or one additional single-strand crossing, the topological equivalent of half of a double-strand node, must be added to each product circle.

The subunit exchange model predicts this topology directly (Figure 8A), while our strand-exchange model can account for the known topology changes if a change in the local twist of each substrate duplex of half a turn is postulated. Models for strand exchange without subunit exchange can be drawn, placing the phosphates and hydroxyls that must exchange partners close to one another by crossing the two DNA strands at the appropriate angle (Figure 8B and C). However, as pointed out by Stark *et al.* (1989), the simplest reconnections of the strands predict either the wrong sign for the introduced node, or the wrong sign for the twist, depending on the details of how the two strands cross. For the case drawn in Figure 8B, this may be resolved by assuming that the overhanging 3′ ends rotate 180° about one another before religation (Figure 8D). Such rotation could be driven by strain introduced by unwinding the DNA at site I (Figure 8E). Note that if the reactant DNA is indeed unwound by half a turn at each site I, then introducing a twist of +1/2 to each product circle relieves the initial

underwinding and results in product molecules that have no local supercoiling strain at all (i.e. no under- or over-winding within site I itself). For the subunit-exchange model, the degree of local supercoiling strain does not change during the reaction, and there is no obvious advantage to unwinding the DNA before the reaction. If our model is correct, however, then unwinding the DNA by half a turn at each site might be an elegant way to trap one of the negative supercoils that is destroyed during the recombination reaction.

Conclusions

We have constructed a model for the resolvase synaptosome that accounts for most of the known biochemical and topological observations by assuming that the arrangements and interactions of the 12 subunits in the unit cell of the hexagonal crystal are close to those seen during recombination. Although the 60 residue DNA-binding domains are disordered in these crystals of intact resolvase, their approximate location in the hexagonal crystal has been established by a packing analysis. This model for intact resolvase places DNA-binding domains on opposite sides of docked DNA, consistent with DNA protection and interference data for *res* sites II and III, but inconsistent with that on site I. To account for the sidedness of the interaction at site I, either the DNA could be untwisted by half a turn or the DNA-binding domains could be differently arranged.

By wrapping two 115 bp *res* site DNAs around the six dimers of resolvase found in the crystal, a model of the synaptosome results that places the two site I DNAs in contact with each other between two resolvase dimers at a point of 222 symmetry. We propose that strand exchange occurs within this complex without the exchange of protein subunits, and that by unwinding the DNA by half a turn at the cross-over site, resolvase utilizes one of the supercoils that is subsequently released to drive the reaction. Thus, in our model of the synaptic complex, resolvase brings the DNA into a conformation close to that of the transition state of the reaction.

Materials and methods

Protein purification and X-ray data collection

Resolvase large fragment was purified as previously described (Hatfull *et al.*, 1989). Intact resolvase was purified by the same protocol, but with an additional chromatography step: before the final low-salt precipitation, resolvase in 6 M urea/20 mM Tris−HCl (pH 7.5)/1 mM EDTA was loaded onto an FPLC monoS cation-exchange column, then eluted with a NaCl gradient, appearing at ~150 mM NaCl. Crystals were grown under essentially the conditions described previously (Weber *et al.*, 1982; Abdel-Meguid *et al.*, 1986) by the hanging-drop method. A total of 10−15 μl of protein at ~10 mg/ml for the intact protein and 24 mg/ml for the large fragment in 10% saturated $(NH_4)_2SO_4$ was mixed with an equal volume of 25−35% saturated $(NH_4)_2SO_4$ buffered with 50 mM Na acetate (pH 5−6) and allowed to equilibrate at 30°C against 1 ml of the latter solution. Although resolvase will crystallize under a variety of conditions, crystals seemed to form more readily below pH 7; at temperatures <30°C the protein tends to form showers of microcrystals. The crystals are hexagonal bipyramids in space group $P6_422$, with one monomer in the asymmetric unit and cell dimensions a = b = 60.2, c = 170.1 Å.

Data were collected on a Xuong−Hamlin multiwire area detector (Hamlin, 1985), mounted on a Rigaku RU300 rotating anode X-ray generator, and were reduced and initially scaled using the programs written by Anderson and Nielsen (Anderson, 1986). Two crystals were used for each data set. The crystals of the large fragment were ~0.75 × 0.75 × 1 mm, while crystals of the intact protein were ~0.25 mm in the small dimension. This size difference is reflected in the resolution of the data. Both data sets are severely anisotropic (Table I).

The program ANISOB (Sheriff and Hendrickson, 1987) was used to correct for the anisotropy of the diffraction intensities. Despite the application of anisotropic B-factors to the solved structure, there were still systematic differences between F_{obs} and F_{calc} along the different crystal axes. These were eliminated by local scaling structure factor amplitudes calculated from the model to the unmerged data using a version of the program SCALAR (written by M.Rould). Scaling was done so as to reach a compromise between scaling up inaccurate weak reflections along a* and b*, and scaling down the more intense and accurate reflections along c*, and so as to give the model a reasonable overall B-factor of ~30 Å^2.

Structure determination

The model used for molecular replacement was a 'trimmed' version of monomer 3 from the previously published orthorhombic form (Sanderson *et al.*, 1990). Residues 37−44 were deleted from the search model because they compose a very flexible turn, and the last four residues (116−119) were deleted from the C-terminal helix because this helix appears to bend slightly at this point on some monomers.

The rotation function was solved using the Crowther function in the package MERLOT (version 2.1) (Crowther, 1972; Fitzgerald, 1988). Over the several different resolution ranges tried, one peak consistently appeared as the highest peak at 4−4.75 σ. Attempts to solve the translation function using the Crowther and Blow (1967) function in MERLOT failed, perhaps because of the complexity of Patterson space for Laue group 6/mmm.

The translation position was finally solved by analysis of the crystal packing using the program PACK, written by J.Warwicker, and the modeling program FRODO (Jones, 1987). After applying the rotation function solution, the search model was translated in a $P6_422$ unit cell to a point at which all of the 2-fold related interactions seen in the orthorhombic crystal were repeated across crystallographic axes in this crystal form.

The translation solution established by visual inspection was initially refined in X-PLOR (Brünger, 1990) as a single rigid body, then as three rigid bodies corresponding to residues 1−36, 45−65 and 66−115. Unlike several incorrect translation solutions, the crystallographic R-factor ($\Sigma F_o - F_c / \Sigma F_c$) dropped quickly from 49.2% to 43.2%. For consistency, all R-factors quoted were recalculated against the final scaled version of the data. The initial R-factor against data which had only an overall anisotropic B factor correction was slightly higher. The root mean square (r.m.s.) deviation between co-ordinates before and after this refinement was 0.9 Å.

The structure has been partially refined by alternating cycles of energy minimization using the program X-PLOR and manual rebuilding using the

Table I. Statistics of the data

	Unique reflections	Merging R	% complete	Resolution limit along c*/a* (Å)
Intact[a]	3455 (3.2)	4.9% (3.2)	85 (3.2)/99 (3.8)	3.3/3.9
RLF[b]	5173 (2.7)	4.1% (2.7)	93 (2.7)/94 (3.4)	2.75/3.55

The cross R-factor between the two data sets is 53% over 1828 common reflections from 99.0 Å to 4.0 Å. The resolution limit used for each calculation is given in parentheses. The resolution limit along each axis was defined as the point where the average value of I/sigma(I) drops below 2. This was calculated using a program written by Jonathan Friedman that bins reflections radially as well as by resolution. The merging factor is defined as: $R = \Sigma|I - \langle I \rangle| I/\Sigma(I)$ where I is the observed intensity.

[a]Data collected form crystals of intact resolvase.

[b]An abbreviation for resolvase large fragment.

program FRODO (Jones, 1991), resulting in an R-factor of 31%. A total of 3553 unique reflections between 8.0 and 2.7 Å resolution were used in the refinement after removal of the following reflections from the data set: all those with $F < 1\ \sigma(F)$ lying outside an ellipsoid defined by resolution limits of 4 Å along a* and b*, and 3 Å along c*; all those with $F < 2.5\ \sigma(F)$ beyond 3.7 Å/2.8 Å; and all those beyond 3.4 Å/2.6 Å. The r.m.s. deviation from ideality for bond length and angles was 0.015 Å and 3.55°, respectively. Owing to the limited resolution of the diffraction data, simulated annealing was not used, and further refinement was not deemed worthwhile.

Location of the small domain

No identifiable electron density corresponding to the small domain could be found in any of a large variety of electron density maps, including a difference map calculated using experimentally measured diffraction amplitudes from the intact and large fragment crystals.

In order to identify the possible location of the DNA-binding domain, a 20 Å diameter sphere consisting of 162 neutral atoms was used in a packing search performed with X-PLOR. The calculated diameter of a spherical protein of 63 amino acids is 25.6 Å, assuming 115 Da/residue and a partial specific volume of 0.73 cm^3/g. We chose a slightly smaller diameter of 20 Å because the domain is likely to be ellipsoidal rather than spherical, and it is the smallest dimension that is relevant in this search. The large-fragment protein co-ordinates used for the packing search were those of monomer 2 from the orthorhombic crystal form which extends to residue 120 rather than 115. The sphere was translated on a 4 Å grid from 0 to 60, 0 to 60 and 0 to 30 Å, thus covering two asymmetric units. At each grid point, the intermolecular Van der Waals energy between the sphere and resolvase, and between symmetry-related spheres, was calculated. This packing energy was small or negative at only a very small cluster of grid points. The search was repeated over this region on a 1.5 Å grid to better define its boundaries. Using a 15 Å diameter sphere resulted in a slightly broader cluster of points, but no additional locus. These points define the only possible locations for the center of a 20 Å sphere in the hexagonal lattice.

Acknowledgements

We thank Nigel Grindley and Robert Hughes for helpful discussions and comments. This work was supported by NIH grant GM-22778 to T.A.S. P.A.R. was supported in part by a NSF graduate fellowship. Full co-ordinates and structure factors for the $P6_422$ and C222 crystal forms have been deposited in the Brookhaven Protein Data Bank.

References

Abdel-Meguid,S.S., Grindley,N.D.F., Templeton,N.S. and Steitz,T.A. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 2001–2005.

Abdel-Meguid,S.S., Krishna Murthy,H.M. and Steitz,T.A. (1986) *J. Biol. Chem.*, **261**, 15934–15935.

Anderson,D.H. (1986) PhD Thesis, University of San Diego, San Diego, CA.

Bloomer,A.C., Champness,J.N., Bricogne,G., Staden,R. and Klug,A. (1978) *Nature*, **276**, 362–368.

Boocock,M.R., Brown,J.L. and Sherratt,D.J. (1987) Topological specificity in Tn3 resolvase catalysis. In Kelley,T.J. and McMacken,R. (eds), *DNA Replication and Recombination*. Alan R.Liss, Inc., New York, pp. 703–718.

Brünger,A.T. (1990) *X-PLOR (Version 2.1) Manual*. Yale University, New Haven, CT.

Cozzarelli,N.R., Krasnow,M.A., Gerrard,S.P. and White,J.J. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 383–400.

Craig,N.L. (1988) *Annu. Rev. Genet.*, **22**, 77–105.

Crowther,R.A. (1972) In Rossmann,M.G. (ed.), *The Molecular Replacement Method*. Gordon and Breach, New York.

Crowther,R.A. and Blow,D.W. (1967) *Acta Crystallogr.*, **23**, 544–548.

Falvey,E. and Grindley,N.D.F. (1987) *EMBO J.*, **6**, 815–821.

Fitzgerald,P.M.D. (1988) *J. Appl. Crystallogr.*, **21**, 273–278.

Graham,K.S. and Dervan,P.B. (1990) *J. Biol. Chem.*, **265**, 16534–16540.

Grindley,N.D.F. (1993) *Science*, **262**, 738–740.

Grindley,N.D.F. and Reed,R.R. (1985) *Annu. Rev. Biochem.*, **54**, 863–896.

Grindley,N.D.F., Lauth,M.R., Wells,R.G., Wityk,R.J., Salvo,J.J. and Reed,R.R. (1982) *Cell*, **30**, 19–27.

Haffter,P. and Bickle,T.A. (1988) *EMBO J.*, **7**, 3911–3996.

Hamlin,R. (1985) *Methods Enzymol.*, **114**, 416–452.

Hatfull,F.G. and Grindley,N.D.F. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 5429–5433.

Hatfull,G.F. and Grindley,N.D.F. (1988) In Smith,G. and Kucherlapati,R. (eds), *Genetic Recombination*. American Society for Microbiology, Washington, DC, pp. 357–396.

Hatfull,G.F., Noble,S.M. and Grindley,N.D.F. (1987) *Cell*, **49**, 103–110.

Hatfull,G.F., Salvo,J.J., Falvey,E.E., Rimphanitchayakit,V. and Grindley,N.D.F. (1988) *Symp. Soc. Gen. Microbiol.*, **43**, 149–181.

Hatfull,G.F., Sanderson,M.R., Freemont,P.S., Raccuia,P.R., Grindley,N.D.F. and Steitz,T.A. (1989) *J. Mol. Biol.*, **208**, 661–667.

Heichman,K.A. and Johnson,R.C. (1990) *Science*, **249**, 511–517.

Hughes,R.E., Hatfull,G.F., Rice,P., Steitz,T.A. and Grindley,N.D.F. (1990) *Cell*, **63**, 1331–1338.

Hughes,R.E., Rice,P.A., Steitz,T.A. and Grindley,N.D.F. (1993) *EMBO J.*, **12**, 1447–1458.

Jones,T.A. (1987) *J. Appl. Crystallogr.*, **11**, 1–10.

Jones,T.A. (1991) *O Version 5.7 User's Manual*. 1991, Uppsala, Sweden.

Kanaar,R., Klippel,A., Shekhtman,E., Dungan,J.M., Kahmann,R. and Cozzarelli,N.R. (1990) *Cell*, **62**, 353–366.

Kim,J.L., Nikolov,D.B. and Burley,S.K. (1993a) *Nature*, **365**, 520–527.

Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993b) *Nature*, **365**, 512–520.

Kitts,P.A., Symington,L.S., Dyson,P. and Sherratt,D.J. (1983) *EMBO J.*, **329**, 346–348.

Klippel,A., Cloppenberg,K. and Kahmann,R. (1988) *EMBO J.*, **7**, 3983–3989.

Klippel,A., Kanaar,R., Kahmann,R. and Cozzarelli,N.R. (1993) *EMBO J.*, **12**, 1047–1057.

Krasnow,M.A. and Cozzarelli,N.R. (1983) *Cell*, **32**, 1313–1324.

Landy,A. (1989) *Annu. Rev. Biochem.*, **58**, 913–949.

Mazzarelli,J.M., Ermacora,M.R., Fox,R.O. and Grindley,N.D.F. (1993) *Biochemistry*, **32**, 2979–2986.

Reed,R.R. (1981) *Cell*, **25**, 713–719.

Reed,R.R. and Grindley,N.D.F. (1981) *Cell*, **25**, 721–728.

Reed,R.R. and Moser,C.D. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 245–249.

Rimphanitchayakit,V., Hatfull,G.F. and Grindley,N.D.F. (1989) *Nucleic Acids Res.*, **17**, 1035–1050.

Salvo,J.J. (1987) PhD Thesis, Yale University.

Salvo,J.J. and Grindley,N.D.F. (1988) *EMBO J.*, **7**, 3609–3616.

Sanderson,M.R., Freemont,P.S., Rice,P.A., Goldman,A. Hatfull,G.F. and Grindley,N.D.F. (1990) *Cell*, **63**, 1323–1329.

Sheriff,S. and Hendrickson,W.A. (1987) *Acta Crystallogr.*, **A43**, 118–121.

Sherratt,D. (1989) In Berg,D.E. and Howe,M.N. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 163–184.

Stark,W.M., Sherratt,D.J. and Boocock,M.R. (1989) *Cell*, **58**, 779–790.

Story,R.M., Weber,I.T. and Steitz,T.A. (1992) *Nature*, **355**, 318–325.

Wasserman,S.A. and Cozzarelli,N.R. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 1079–1083.

Wasserman,S.A., Dungan,J.M. and Cozzarelli,N.R. (1985) *Science*, **229**, 171–174.

Weber,P.C., Ollis,D.L., Bebrin,W.R., Abdel-Meguid,S.S. and Steitz,T.A. (1982) *J. Mol. Biol.*, **157**, 689–690.

Wells,R.G. and Grindley,N.D.F. (1984) *J. Mol. Biol.*, **179**, 667–687.

Received on October 11, 1993; revised on January 11, 1994