Phage T4 endonuclease VII (Endo VII), the first enzyme shown to resolve Holliday junctions, recognizes a broad spectrum of DNA substrates ranging from branched DNAs to single base mismatches. We have determined the crystal structures of the Ca\(^{2+}\)-bound wild-type and the inactive N62D mutant enzymes at 2.4 and 2.1 Å, respectively. The Endo VII monomers form an elongated, highly intertwined molecular dimer exhibiting extreme domain swapping. The major dimerization elements are two pairs of antiparallel helices forming a novel ‘four-helix cross’ motif. The unique monomer fold, almost completely lacking β-sheet structure and containing a zinc ion tetrahedrally coordinated to four cysteines, does not resemble any of the known junction-resolving enzymes, including the Escherichia coli RuvC and λ integrase-type recombinases. The S-shaped dimer has two ‘binding bays’ separated by ~25 Å which are lined by positively charged residues and contain near their base residues known to be essential for activity. These include Asp40 and Asn62, which function as ligands for the bound calcium ions. A pronounced bipolar charge distribution suggests that branched DNA substrates bind to the positively charged face with the scissile phosphates located near the divalent cations. A model for the complex with a four-way DNA junction is presented.

Keywords: DNA junction resolvase/domain swapping/T4 endonuclease VII/X-ray structure

Introduction

The recognition and resolution of DNA junctions is an essential step in DNA recombination and repair, and proteins binding to and resolving these junctions have been found in all kingdoms of life (West, 1993; Kemper, 1997; White et al., 1997). Endonuclease VII (Endo VII), a 157 amino acid product of gene 49 of bacteriophage T4, expressed at early and late stages of infection from different promoters, is the prototypic junction-resolving enzyme, since it was the first enzyme shown to resolve Holliday junctions (Mizuuchi, 1982). Endo VII has been shown to be involved in mismatch repair; however, its major function in vivo, at least in the late stages of phage infection, appears to be the resolution of branchpoints prior to packaging of the DNA into the phage head (Kemper and Brown, 1976; Solaro et al., 1993; Grebenshchikova et al., 1994). Phages mutated in gene 49 are defective in packaging their newly synthesized DNA and accumulate highly branched DNA molecules.

In contrast to other resolvases, such as, for example, yeast CCE1 or Escherichia coli RuvC, Endo VII has a broad substrate specificity and recognizes a variety of branched DNA structures and/or structural perturbations in DNA. Besides Holliday junctions and cruciform DNA, Endo VII will not only cleave Y-junctions, heteroduplex loops, single-stranded overhangs and curved DNA, but also abasic sites and single base mismatches (Kemper et al., 1990; Kemper, 1997; Greger and Kemper, 1998). Mismatch cleavage by Endo VII has been exploited successfully for screening for mutations (Youil et al., 1995).

It has been suggested that the inclination of the DNA helix segments on either side of the branchpoint is important for recognition by Endo VII; however, the fact that single base mismatches and abasic sites are cleaved efficiently puts a question mark over this hypothesis (Bhattacharyya et al., 1991). Furthermore, comparative gel mobility experiments indicate a global structural change of four-way junctions on binding to Endo VII (Pöhler et al., 1996). Therefore, structural studies are required to understand the mechanism underlying the structural selectivity of Endo VII.

Endo VII is active as a dimer and nicks both strands in a divalent cation-catalyzed reaction 2–6 bp 3’ of the branchpoint in independent, but temporally closely correlated reactions (Pottmeyer and Kemper, 1992). For a supercoil-stabilized cruciform substrate, it was shown that the two strands of the junction are cleaved within the lifetime of the enzyme–junction complex (Giraud-Panis and Lilley, 1997). On the other hand, the time delay between nick and counternick of mismatches is such that it allows for the repair by DNA polymerase and ligase in vitro (Salaro et al., 1993). This has led to the conjecture that Endo VII primarily acts as a repair enzyme and represents a member of an ancient class of broad specificity repair enzymes (Kemper, 1997).

While Endo VII, like other resolvases, is active as a dimer, higher order complexes with cruciform DNA containing two or even three protein dimers per junction have been observed in band shift assays, electron micrographs and gel chromatography experiments (Golz et al., 1997; Kupfer et al., 1998; Törö and D.Suck, unpublished observations). At present, it is not known whether these complexes are in vitro artefacts or whether they have any biological meaning.
The cleavage pattern and the relative cleavage efficiencies of Endo VII are influenced by the local base sequence (Pottmeyer and Kemper, 1992); however, other resolvases display a more pronounced sequence specificity. For example, the yeast CCE1 protein shows a strong preference for cutting 3’ to a CT (Schofield et al., 1998).

Endo VII shows little sequence homology to other resolvases and, as a unique feature, contains one zinc ion per monomer, coordinated to four cysteines, which is essential for stabilizing the fold (Giraud-Panis et al., 1995). The relevance of significant sequence homology of the N-terminal 62 residues (containing the cysteines) with gp59 from mycobacteriophage L5 and of a 30 amino acid stretch near the C-terminus with endonuclease V from phage T4 is not clear.

We present here the high resolution crystal structures of Ca²⁺-bound wild-type and the inactive N62D mutant of phage T4 Endo VII. The implications of its unique fold and highly unusual domain-swapped dimer architecture for the catalytic mechanism and the interaction with branched DNA are discussed, and a model of a complex with a four-way DNA junction is presented.

Results and discussion
The crystal structure of the N62D mutant was determined first by MIRAS techniques using a seleno-methionine-labeled protein as the major source for phase determination at high resolution. The mercury and gold derivatives were used to locate nine out of 10 selenium atoms in an anomalous difference Fourier. The resulting electron density map was of excellent quality (see Figure 5), allowing chain tracing and model building of the complete Endo VII dimer present in the asymmetric unit. Hindering the X-ray analysis of the wild-type protein was severe non-isomorphism for many of the derivatives obtained by soaking, which often also led to a change in space group. In contrast to the catalytically inactive N62D mutant, the expression of the seleno-methionine-labeled wild-type Endo VII proved to be very difficult. Significant conformational changes in the mutant compared with the wild-type protein precluded the determination of the wild-type structure by simple molecular replacement and, therefore, SIRAS phasing based on a mercury derivative had to be applied. The wild-type and mutant Endo VII structures were refined to R-factors of 25.3 and 21.3% at 2.4 and 2.1 Å, respectively. Refinement statistics of the final models are summarized in Table I.

**Endo VII architecture: a domain-swapped dimer built from unstable monomers**

The fold of Endo VII revealed by our X-ray analysis is novel and shows several unusual features not resembling any of the known structures of junction-resolving enzymes. The structure of the individual 157 amino acid Endo VII monomer does not represent a stable fold, as it exposes many hydrophobic residues in its central region and contains two completely separated, flexibly linked domains. They consist of a 60 residue N-terminal domain and a 50 residue C-terminal domain, which are separated by ~50 Å and are connected by a central helix–loop–helix motif (residues 60–96) and an extended chain segment (residues 97–106) (Figures 1 and 2). The C-terminal domain comprising helices 4, 5 and 6 is a compact entity, while the N-terminal domain composed of helix 1, a loop region and a ‘β-finger’ displays a less compact structure whose integrity critically depends on the presence of a zinc ion. The ‘β-finger’ formed by residues 38–56 is a hairpin structure containing a four residue stretch of antiparallel β-sheet, the only β-sheet structure found in Endo VII. The zinc ion is tetrahedrally coordinated to
four cysteines located at the N-terminus of helix 2 (C58 and C61) and within the loop region (C23 and C26), which is thereby firmly tethered to the rest of the molecule (Figure 2). In keeping with the critical structural role of the zinc ions, cysteine mutants interfering with zinc binding were found to be inactive (Giraud-Panis et al., 1995).

Two extensively intertwined Endo VII monomers related by a non-crystallographic 2-fold axis form an elongated dimer of approximate dimensions $93 \times 45 \times 25 \, \text{Å}^2$ (Figure 2). Within the dimer, the C-terminal domain of one monomer interacts with the N-terminal domain of the other monomer, resulting in an extreme case of domain swapping. The C-terminal domain of the ‘top’ molecule (shown in red in Figure 2A) is located at the ‘bottom’ end of the dimer and vice versa. N- and C-terminal domains mutually stabilize each other and their orientation relative to the rest of the molecule. A lack of this interaction is presumably mainly responsible for the inactivity of N- and C-terminal deletion mutants (Golz et al., 1997; Birkenbihl and Kemper, 1998a; see discussion below). Only minor deviations from the local 2-fold symmetry are observed, and $\alpha_C$ positions of corresponding monomers can be superimposed with r.m.s.d. values of $\sim 0.5 \, \text{Å}$.

When viewed along the molecular dyad, the EndoVII dimer has an S-shaped structure with two ‘bays’ separated by $\sim 25 \, \text{Å}$ (the distance between the bound Ca$^{2+}$ ions in wild-type Endo VII is $24 \, \text{Å}$; Figure 2A). These bays or channels, which are lined by positively charged residues on one face of the dimer, contain a number of residues shown to be essential for activity, including N62, D40, H41 and E65 (Giraud-Panis and Lilley, 1996; Golz et al., 1997; R.P.Birkenbihl, unpublished results) as well as the bound calcium ions, clearly suggesting that they harbor the active sites (see discussion below).

In a view perpendicular to the molecular 2-fold axis, Endo VII appears as a very elongated, slightly curved object with a highly bipolar charge distribution (Figure 2A). Predominantly positively charged residues are exposed on the concave face of the dimer, in particular on helices 2, 4 and 6.

The major dimerization element of the Endo VII dimer is formed by the central part of the monomers comprising helices 2 and 3 and the 10 residue loop connecting them. The arrangement of these helices relative to the molecular 2-fold axis gives rise to the formation of a ‘four-helix cross’ with two pairs of antiparallel helices crossing each other at an angle of $\sim 70^\circ$ to $80^\circ$ (Figure 2). This highly unusual dimerization motif is reminiscent of the dimer interface in the recently determined structure of an E.coli RNA polymerase α subunit domain (Zhang and Darst, 1998). Numerous, mostly hydrophobic residues interacting with their symmetry-related counterparts form an extended hydrophobic core stabilizing the dimer (Figure 2B). The W87R mutant, which directly interferes with these inter-helical contacts, has lost its ability to dimerize and, as a consequence, shows neither DNA-binding nor cleavage activity (Birkenbihl and Kemper, 1998b).

**Differences between the wild-type and N62D mutant structures**

The structure of Endo VII was solved originally using crystals of the N62D mutant grown at acidic pH (4.5) in the presence of 20 mM Mg$^{2+}$ ions (see Materials and methods), while suitable crystals of the wild-type protein were obtained at pH 8.2 in the presence of high concentrations of Ca$^{2+}$ (200 mM). Both magnesium and calcium ions catalyze the resolution activity of the enzyme (S.Pottmeyer and B.Kemper, unpublished results); however, under the crystallization conditions used, no magnesium ions were found to be bound at the active site of the mutant. A comparison of the two structures reveals significant conformational changes (Figure 3) explaining the failure of attempts to solve the wild-type structure by molecular replacement using the refined mutant coordinates. While the overall protein architecture is the same in both structures, there is a significant change in the orientation of the N- and C-terminal domains relative to the central dimerization domain. This leads to a widening of the binding clefts in the mutant and a concomitant increase in the long dimension of the mutant dimer to $\sim 100 \, \text{Å}$.

As can be seen in Figure 3B, it is possible to superimpose fairly closely the upper, and also the lower halves of the dimers including the N-terminal domain of one subunit and the C-terminal domain from the other subunit, showing that the relative disposition of these domains is essentially conserved in the wild-type and the mutant protein. The position and orientation of helix 2, however, differ, and a closer inspection of the structures indicates that this helix is bent in wild-type Endo VII around position E65/C66 towards the β-finger and shifted roughly half a helical turn relative to its counterpart (Figure 3C). This narrows the binding cleft compared with the N62D mutant and also affects the packing of the helices, however, without disrupting the hydrophobic core of the dimerization interface. The Ca$^{2+}$ ion directly liganded to D40 and N62 in wild-type Endo VII, which neutralizes the negative charges of the closely spaced carboxylate groups of D40 and E65, could be responsible for these structural changes. The structural rearrangement could also be due to the widely different pH values of crystallization. The close approach of the D40 and E65 carboxylate groups ($\sim 3.0 \, \text{Å}$) in the N62D mutant suggests that at least one of them is protonated (Figure 3D). This may also account for the fact that no magnesium ions are bound at this site in the mutant. Finally, we cannot entirely exclude that the different packing environments in the two crystal structures are causing conformational changes in the protein.

The structural changes observed in the mutant indicate a certain degree of conformational flexibility of the Endo VII dimer which may be of importance for the interaction with substrates. On the other hand, a comparison of the two structures suggests that the N- and C-terminal domains are fairly rigid entities with a fixed spatial relationship.

**Mutants affecting activity and the location of the active site**

Extensive site-directed mutagenesis experiments have identified a number of residues affecting dimerization, DNA binding or cleavage (for a review see White et al., 1997; B.Kemper, unpublished results). The effects of most of the mutants can be rationalized in terms of the three-dimensional structures and, together with the Ca$^{2+}$-binding site found in the wild-type protein, allow the identification of the putative active site.

Residues D40, H41, N62 and E65 located in the cleft...
between the β-finger and helix 2 are essential for activity, but not for DNA binding (Giraud-Panis and Lilley, 1996; Golz et al., 1997; R.P.Birkenbihl, unpublished results). It is therefore highly likely that the Ca\(^{2+}\) ion directly coordinated to D40 and N62 and located close to E65 marks the position of the catalytic cation (Figure 2). A superposition of the putative active site regions in the wild-type and mutant proteins as shown in Figure 3D does not provide a straightforward explanation for the inactivity of the N62D mutant, since no metal ion is bound in the mutant structure and D40 and/or E65 appear to be protonated. In principle, D62 could also function as a ligand for the catalytic cation; however, the charges in the active site cleft may not be balanced properly in the mutant. Another acidic residue, E86, located on the active site cleft may not be balanced properly in the ligand for the catalytic cation; however, the charges in the active site cleft may not be balanced properly in the mutant. Another acidic residue, E86, located on the opposite face of the dimer and pointing into the solvent, has been reported to be catalytically essential (Pöhler et al., 1996). The reason for the apparent inactivity of an E86A mutant is not obvious from the three-dimensional structure (Figure 2), unless one assumes that the mutation somehow interferes with the proper folding of the protein.

Other critical residues including W87 and F72 are part of the hydrophobic four-helix cross dimerization interface (Figure 2B). The W87R and F72S mutants will clearly disrupt this interface and, as one would expect, show neither dimer formation nor activity (Birkenbihl and Kemper, 1998b).

As mentioned above, zinc ions are essential for the integrity of the fold by tethering the loop between helix 1 and the β-finger to the N-terminus of helix 2 (Figure 2) and, accordingly, the C23S and C61S mutants which no longer bind zinc are inactive (Giraud-Panis et al., 1995). The inner two zinc ligands (C26 and C58) seem to be redundant to a certain extent since they still retain some zinc-binding ability if either cysteine is mutated into a serine.

Deletions at either the C- or N-terminus severely affect DNA binding and activity, and mutants lacking residues 1–10 or 151–157 are inactive (Golz et al., 1997; Birkenbihl and Kemper, 1998a). The effect of these mutations will be two-fold: firstly, a disruption of the stabilizing interaction between the N-terminal domain of one monomer with the C-terminal domain of the other monomer, which will also affect their relative orientation to the rest of the molecule; and secondly, a loss of positively charged residues which may be important for DNA binding. In agreement with the three-dimensional structure is the observation that the individually inactive N62D and the C-terminal Δ151–157 deletion mutant can complement each other to form an active heterodimer (Golz et al., 1997). One half of this heterodimer will have a wild-type active site with correct interactions of intact N- and C-terminal domains.

Comparison with RuvC and other junction-resolving enzymes

A number of junction-resolving enzymes from various organisms have been studied in detail by biochemical and various biophysical techniques (for reviews see, for example, Kemper, 1997; White et al., 1997), but structural data are available for only very few of these enzymes including the E.coli RuvC protein and some λ integrase-type recombinases which form and resolve a Holliday junction intermediate during site-specific recombination (Ariyoshi et al., 1994; Guo et al., 1997; Hickman et al., 1997; Kwon et al., 1997; Subramanya et al., 1997).

There is rarely any sequence homology detectable between Endo VII and other junction resolvases. As noted before by Giraud-Panis et al. (1995), a database search reveals significant homology only for the N-terminal section (residues 1–62) containing the zinc-binding cysteines with some proteins including gp59 from mycobacteriophage L5, and for a 30 residue section near the C-terminus (residues 115–145) with T4 endonuclease V, a UV dimer excision-repair enzyme. The significance of these limited sequence homologies is not clear, particularly since in the case of endonuclease V whose three-dimensional structure is known (Vassylyev et al., 1995), no structural homology is present in this region. Likewise, the claimed moderate homology of a central 45 amino acid section with T7 endonuclease I, a functionally somewhat related resolvase which, however, cleaves 5’ to the junction in the continuous strands, may not be meaningful.

Escherichia coli RuvC resolves Holliday junctions, the central intermediates of genetic recombination, by cleaving the continuous strands to generate the recombinant DNA molecules (Iwasaki et al., 1991; Bennett and West, 1995a). Like Endo VII, it requires divalent cations and is active as a dimer, but otherwise the two proteins show very little homology, in terms of both sequence and structure. As revealed by the X-ray structure determination by Ariyoshi et al. (1994), the slightly larger RuvC protein (172 residues) has a fold very different from that of Endo VII at the monomer level. It is a typical αβ-protein with a central five-stranded β-sheet surrounded by helices displaying a topology closely related to that of E.coli RNase H1 (Katayanagi et al., 1990). In contrast, Endo VII almost completely lacks β-sheet structure and does not contain a stable fold at the monomer level.

The RuvC dimer with approximate dimensions 65×40×35 Å is far less elongated than the Endo VII dimer and has a far more conventional dimerization interface consisting of a pair of parallel helices without intertwining of the monomers. However, the relative disposition of the catalytic centers, which are separated by ~30 Å, and the bipolar charge distribution are similar, clearly suggesting binding of the Holliday junction to the

Fig. 2. Structure of the wild-type Endo VII dimer. (A) Ribbon plot representation produced with MOLSCRIPT (Kraulis, 1991) in two perpendicular views (one of them in stereo) with the individual monomers colored red and green. The bound zinc and calcium ions are shown as blue and yellow spheres, respectively. Indicated in ball and stick representation are the side chains of D40 and N62 liganding the calcium as well as the nearby E65. (B) Stereo representation of the four-helix cross region showing the hydrophobic residues. They form an extended hydrophobic core representing the major dimerization element of the Endo VII dimer. (C) Electrostatic surface representation of the Endo VII dimer. The two opposite faces of the dimer, viewed approximately along the dyad, show a distinctly different charge distribution. The predominantly positively charged face shown on the right (indicated by the blue color) also contains the bound calcium ion and residues known to be essential for activity. The figure was produced with GRASP (Nicholls et al., 1991) with the scale ranging from –10 (red) to +17 (blue).
Fig. 3. Superposition of the Endo VII wild-type and N62D mutant structures. (A) The orientation of the central four-helix cross relative to the N- and C-terminal domains has changed, leading to a wider cleft between the β-finger and helix 2 in the mutant (shown in black). (B) In contrast, the conformation of the N- and C-terminal domains as well as their relative orientation remains essentially the same. (C) Helix 2 in the mutant structure is shifted relative to its dyad-related counterpart by about half a helical turn compared with wild-type Endo VII. Cα positions 1–157, 100–157 and 62–96 were used for superposition in (A), (B) and (C), respectively, corresponding to r.m.s.d. values of 2.8, 0.5 and 1.0 Å, respectively. (D) Active site superposition of Endo VII wild-type and N62D mutant. The calcium ion bound in the wild-type protein (shown in gray) is liganded to N62 and D40. In the mutant (shown in black), E65 and D40 are in close contact (~3 Å), suggesting that at least one of the carboxylates is protonated. H105 is contributed by the other monomer.
predominantly positively charged face of the dimer (see discussion below). Four acidic residues essential for activity are found at the active site clefts, three of which are in a disposition closely related to that of the putative active site residues D40, E65 and N62 in Endo VII and the D10, E48 and D70 carboxylates in RNase H1. They presumably represent ligands of the catalytic metal ion and may suggest related catalytic mechanisms for these enzymes. Similar arrangements of acidic residues coordinating the catalytic cation have been found in other Mg$^{2+}$-dependent nucleases.

Three-dimensional structures are available for several $\lambda$ integrase-type recombinases including Cre, XerD, HP1, Int and the type I topoisomerase from vaccinia virus, which are all known to resolve Holliday junctions (Guo et al., 1997; Hickman et al., 1997; Kwon et al., 1997; Subramanya et al., 1997). However, no significant structural similarity to Endo VII is detectable with any of these enzymes.

Model of an Endo VII–four-way DNA junction complex

Endo VII was the first enzyme shown to resolve Holliday junctions in vitro (Mizuuchi et al., 1982) and later found to accept as substrates a range of branched DNAs with different structural perturbations. A wealth of data derived from various biophysical and biochemical techniques has since been accumulated concerning the interaction of Endo VII with four-way junctions and other branched DNA molecules (for recent reviews see Kemper, 1997; White et al., 1997).

In brief, the following picture has emerged for the binding of Endo VII to a four-way junction, based mainly on comparative gel electrophoresis, protection against hydroxyl radical attack and the analysis of cleavage sites in synthetic cruciform structures with tethered arms.

(i) Endo VII binds as a dimer to the minor groove side of a four-way junction, cleaving the exchanging strands in an antiparallel stacked X structure 3* to the point of strand exchange. (ii) The cleavage reactions of the two strands are independent, but temporally closely correlated, and do occur within the lifetime of the junction–protein complex. (iii) Experiments with a protein A–Endo VII fusion construct seem to indicate a change in the global structure of the junction. Other than in the free state, the presence or absence of Mg$^{2+}$ ions does not influence its structure in the bound state.

The pronounced bipolar charge distribution of the Endo VII dimer (Figure 2C) suggests that it binds with its predominantly positively charged face to the minor groove side of the four-way junction, such that the scissile phosphates reach into the ‘bays’ harboring the catalytic residues and the divalent metal ion, i.e. the Ca$^{2+}$ ion seen in the wild-type enzyme.

Given the 2-fold symmetry of the antiparallel stacked X structure of the junction (in the presence of Mg$^{2+}$ ions) and the 2-fold symmetric distribution of the preferred cutting sites relative to the point of strand exchange, we are assuming that in the complex the molecular dyad of the junction roughly coincides with the molecular 2-fold of the Endo VII dimer.

For docking, we used the coordinates of the antiparallel stacked X structure derived by von Kitzing et al. (1990) by computer modeling and extended the arms by regular B-DNA. The RuvA and Cre Holliday junction complexes whose structures have been reported recently (Gopaul et al., 1998; Roe et al., 1998) cannot be used as models, since they exhibit exact or near square-planar conformations with unstacked bases at the cross-over point which are not compatible with the Endo VII structure and biochemical results. The ending result of the four-way DNA junction complex is shown in Figure 4. This model has not been refined and thus cannot provide a detailed picture of the interactions, particularly since both the DNA junction and the protein may undergo conformational changes upon binding. Nevertheless, the overall features of the model are in agreement with the available experimental data.

Basic residues exposed on the central helices (residues H70, K71 and R74 on helix 2) could interact with the phosphates of the exchanging strands, and further contacts with the continuous strands are possible involving basic residues located on the helices in the C-terminal domain, in agreement with observed protection patterns (Parson et al., 1990). These latter contacts may also explain the observation by Seeman and coworkers (Mueller et al., 1990), that immobile junctions with arms shorter than 9 bp are not substrates for Endo VII. To avoid minor steric clashes with the C-terminal domains, the angle between the arms of the X structure junction model derived by von Kitzing et al. (1990) may have to be slightly adjusted. In contrast to RuvC and CCE1, where an open structure with unstacked bases at the cross-over point is indicated by hypersensitivity towards permanganate and hydroxyl radicals upon complex formation (Bennett and West, 1995b; White and Lilley, 1997), the footprinting patterns observed with Endo VII suggest a fully stacked conformation consistent with an antiparallel stacked X-structure of the junction.

The two putative active sites in the Endo VII dimer are ~24 Å apart as measured by the distance between the two Ca$^{2+}$ ions bound in the wild-type structure. This has to be compared with a distance of ~23 to ~35 Å between the scissile phosphates (highlighted in yellow in Figure 4) of the preferred cutting sites 2 or 3 bp on either side of the cross-over point. While there is a good match for cuts 2 bp from the cross-over point, a difference of >10 Å is encountered in the case of cutting 3 bp from the junction center. The model suggests that even in the optimal case, due to steric constraints imposed mainly by the central helices (2-fold-related, antiparallel helices 2), it may not be possible simultaneously to position both phosphates correctly with respect to the catalytic residues. This could possibly be achieved by a reorientation of the two arms of the junction contacting the protein, or by a corresponding conformational change in the Endo VII dimer. Alternatively, the junction would have to be reoriented somewhat after the first cut, possibly by a rocking motion, to bring the second phosphate into a productive binding position.

The latter scenario is attractive, since it would explain the asynchrony of bilateral cleavages and the time delay between nick and counter-nick allowing Endo VII to trigger mismatch repair (Solaro et al., 1993). It is also fully consistent with the notion that the two cuts do occur within the lifetime of the protein–DNA complex, as has
been shown for a supercoil-stabilized cruciform substrate (Giraud-Panis and Lilley, 1997).

The cleavage pattern and relative cleavage efficiencies of Endo VII are influenced by the local base sequence (Pottmeyer and Kemper, 1992), although other resolvases, such as CCE1, display a more pronounced sequence selectivity (Schofield et al., 1998). This sequence selectivity could be caused by indirect effects of the base sequence on structural parameters of the DNA and/or by direct contacts to the bases likely to occur according to our model in the minor groove around the cross-over point, as well as in the major groove some 8 bp away (Figure 4).

**What is the basis for the structural selectivity of Endo VII?**

A remarkable feature of Endo VII is its broad substrate specificity already mentioned above, which distinguishes it clearly from cellular resolvases such as RuvC or CCE1. Albeit with differing efficiencies, Endo VII cuts many different kinds of branched DNA species including Holliday junctions, but also DNA containing a single base mismatch or an abasic site. What are the common features of these structural perturbations recognized by the enzyme?

It has been suggested that Endo VII recognizes the inclination of DNA helices, i.e. the angle between the DNA segments on either side of a branchpoint or a structural perturbation (Bhattacharyya et al., 1991). This angle is expected to be ~120° in the stacked X conformation of a four-way junction, but also in a three-way junction or bulged DNA containing two additional adenines. Bent DNA, either intrinsically through the presence of A tracts, or by interaction with a compound, such as in a cis-Pt adduct, also appears to be a substrate for Endo VII (Bhattacharyya et al., 1991; Murchie and Lilley, 1993). Greger and Kemper (1998) recently have reported that an oligonucleotide with an abasic site, which was demonstrated to be extremely sensitive to cleavage by Endo VII, shows an unusual migration behavior on polyacrylamide gels typical for bent or kinked DNA.

Inspection of the Endo VII binding surface and the proposed model for the four-way junction complex (Figure 4) suggests that the X-ray structure is in general...
agreement with this hypothesis. However, it does not easily explain the recognition and cleavage of single base mismatches, which, as verified in several X-ray structures, do not display any significant bending. These experimental results seem to indicate that flexibility may be an important parameter in the recognition process, in line with the observation that the stability of a given mismatch (as measured by the melting temperature of the DNA) and cleavage efficiency are correlated reciprocally (Solaro et al., 1993). Possibly the lack of flexibility explains why UV-cross-linked thymine dimers containing DNA is not a substrate for Endo VII (B.Kemper, unpublished results).

Endo VII appears to accept as substrates branched DNA with intrinsically inclined helices or DNA which can easily adopt properly inclined configurations. This feature is somewhat reminiscent of the sequence-dependent cleavage by DNase I, which recognizes minor groove width and flexibility of the DNA and preferentially cleaves either where the DNA intrinsically has the proper groove geometry or where it needs the least energy to distort it for productive binding (Suck, 1994, 1998).

Materials and methods

Cloning and expression

The wild-type Endo VII was expressed and purified as described previously (Golz et al., 1995), while the N62D mutant (Golz et al., 1997) was recloned in vector pET24d to obtain better expression. The N62D mutant was expressed in E.coli BL21(DE3)pLys. Overnight precultures on LB plates (30 mg/l kanamycin, 20 mg/l chloramphenicol, 37°C) were used to inoculate one 2 l Erlenmeyer flask each, filled with 800 ml of LB medium (+ antibiotics). After 3 h shaking at 200 r.p.m., 37°C (OD600 μ0.8), the culture was induced with 0.25 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After another 3 h, the cells are harvested, spun down and stored at –80°C.

For the seleno labeling, the protein was expressed in E.coli B834 (DE3)pLys, and each pre-culture was used to inoculate 800 ml of M9 medium, supplemented with 50 mg/l of each of the following amino acids: Arg, His, Ile, Leu, Lys, Phe, Thr, Tyr, Val; 20 mg/ml of methionine, 2 mg/l biotin and thiamine, 30 mg/l kanamycin and 20 mg/ l chloramphenicol. At the required density (OD600 μ0.8), the cells were spun down, washed twice and dissolved in the same medium, but without methionine. After starvation for 8 h, 40 mg/l L-seleno-methionine was added. After another 2 h, the cells were induced with IPTG for 3 h, spun down and stored at –80°C.

Purification of EndoVII N62D

The frozen cells were thawed and suspended in lysis buffer consisting of 10% glycerol, 10 mM HEPES pH 7.5, 10 mM β-mercaptoethanol,
Wild-type Eno VII was crystallized in hanging drops at 4°C. A 1 μl aliquot of a 16 mg/ml protein solution was mixed with 1 μl of the reservoir solution containing 200 mM CaCl₂, 16.5–20% PEG 2K-MME, 100 mM Tris–HCl pH 8.2, 5 mM MgCl₂, 5 mM (NH₄)₂SO₄ and 10 mM β-mercaptoethanol. Mono-clinics crystals (space group C2, a = 145.0 Å, b = 39.4 Å, c = 75.7 Å, β = 106.2°, Z = 8).

### X-ray source (λ)
- **Wild-type**
  - Cu Kα: 2.33 Å
  - Cu Kβ: 2.33 Å
- **N62D**
  - Cu Kα: 2.1 Å
  - Cu Kβ: 2.48 Å

### dmin used (Å)
- **Wild-type**
  - 2.33 Å
- **N62D**
  - 2.1 Å

### Observed reflections
- **Wild-type**
  - 59,911
- **N62D**
  - 67,931

### X-ray source (λ)
- **Wild-type**
  - Cu Kα: 2.33 Å
  - Cu Kβ: 2.33 Å
- **N62D**
  - Cu Kα: 2.1 Å
  - Cu Kβ: 2.48 Å

### dmin used (Å)
- **Wild-type**
  - 2.33 Å
- **N62D**
  - 2.1 Å

### Observed reflections
- **Wild-type**
  - 59,911
- **N62D**
  - 67,931

Table I. Data collection and phasing statistics for the wild-type and N62D mutant crystals

<table>
<thead>
<tr>
<th>Crystallographic Data</th>
<th>Wild-type</th>
<th>N62D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>X-ray source (λ)</strong></td>
<td>Cu Kα</td>
<td>Cu Kα</td>
</tr>
<tr>
<td>dmin used (Å)</td>
<td>2.33 Å</td>
<td>2.1 Å</td>
</tr>
<tr>
<td>Observed reflections</td>
<td>59,911</td>
<td>67,931</td>
</tr>
<tr>
<td>Completeness (outer shell (%))</td>
<td>99.3 (98.2)</td>
<td>99.9 (99.0)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>6.0</td>
<td>7.5</td>
</tr>
<tr>
<td>No. of sites found</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Phasing power (centric/anomalous)</td>
<td>9/2.1/1.5</td>
<td>2.9/3.4/1.7</td>
</tr>
<tr>
<td>Figure of merit (solvent-flattened)</td>
<td>0.42 (0.87)</td>
<td>0.55 (0.97)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Space group C2 a = 145.0 Å, b = 39.4 Å, c = 75.7 Å, β = 106.2°, Z = 8.

<sup>b</sup>Space group P2₁ a = 58.0 Å, b = 35.9 Å, c = 92.0 Å, β = 103.9°, Z = 4.

<sup>c</sup>EMTS = ethyl mercury thiosalicylate.

<sup>d</sup>At EMBL outstation Hamburg, DESY.

<sup>e</sup>Excluding the protein-bound Zn ions.

<sup>f</sup>Phasing power = |Fobs|/|Fcalc| - 1.

<sup>g</sup>H.Raaijmakers et al.

The crystals kept growing for ~5 days. The crystals were not successful.

Phasing

The EMTS anomalous difference Patterson function and the EMTS–KAu(CN)₂ derivatives were obtained by soaking the wild-type crystals for 1 h in their mother liquor without β-mercaptoethanol and 10% PEG 400 plus 1 mM of the heavy atom compound. In the case of the wild-type protein, only crystals >2 months old survived this treatment. Many of the heavy atom compounds tried caused serious physical damage and/or non-isomorphism. The space groups and cell dimensions were determined and oscillation data were processed using HKL (Otwinowski, 1997) except for the wild-type EMTS derivative data which were processed with XDS98 (Kabsch, 1988) and Scala (CCP4, 1994).

Model building and refinement

The electron density maps were displayed using the program ‘O’ (Jones et al., 1991). The sigmaA (CCP4, 1994) weighted 2m |Fcalc| - D |Fc| map diffracted to 1.6 Å at the BW7B beamline but, due to time constraints, only 2.0 Å data were collected.

All Cu Kα diffraction data were collected using a MAR345 area detector system mounted on an Enraf-Nonius rotating anode generator operating at 40 kV and 90 mA and equipped with nickel-coated focusing mirrors (in-house design). The seleno-methionine derivative data were collected at the EMBL BW7B beamline at the DORIS storage ring, DESY, Hamburg. All data were collected at 100 K using the mother liquor + 20% PEG 400 as a cryoprotectant.

The EMTS and KAu(CN)₂ derivatives were obtained by soaking crystals for 1 h in their mother liquor without β-mercaptoethanol and additionally 10% PEG 400 plus 1 mM of the heavy atom compound. In the case of the wild-type protein, only crystals >2 months old survived this treatment. Many of the heavy atom compounds tried caused serious physical damage and/or non-isomorphism. The space groups and cell dimensions were determined and oscillation data were processed using HKL (Otwinowski, 1997) except for the wild-type EMTS derivative data which were processed with XDS98 (Kabsch, 1988) and Scala (CCP4, 1994).
of N62D was of excellent quality (Figure 5), allowing construction of a complete model of Endo VII N62D, including side chains. The selenium positions were very helpful for tracing the chain. Two clusters containing three closely spaced methionines each (M124, M125, M128) were easily recognized. Another cluster of two selenium atoms near the NCS axis was soon identified as M68, leaving the other selenium site to correspond to one of the N-termini. Tracing the molecule from this N-terminus to the C-terminus was straightforward. The second molecule was traced by applying the NCS operator and was adjusted manually.

The wild-type structure was traced by manually fitting whole helices or strands, from the N62D model, as rigid bodies into the signalA-weighted 2m|Fo|−D|Fc| electron density map. A second round of fitting optimized the individual amino acids. The Cα trace was clearly visible, with only a few chain breaks at the 1σ level. Most side chains were also visible.

The structures were refined using the CCP4 suite of programs DM (Cowtan, 1994), REFMAC (Mushovod et al., 1997) and ARPP (Lamzin and Wilson, 1993) and validated with PROCHECK (Laskowski et al., 1993) and WHATIF (Friend, 1990). A total of 5% of the reflections have been selected to calculate Rmerge. The starting R-factor for the N62D mutant was 0.396 (σ = 2.1 Å). Individual B-factors were restrained but no NCS restraints were applied; nevertheless the monomers differ only marginally (r.m.s.d. = 0.476 Å for the Cα atoms). Arpp was run to find water sites automatically in the later stages of refinement; 262 waters and 12 sulfates were included in the N62D structure.

The starting R-factor in the case of the wild-type structure was 0.395 (15–2.5 Å). Individual B-factors and NCS were restrained, but residues involved in crystal contacts that were clearly different in omit maps were excluded from NCS restraints. Three calcium ions were included and, additionally, 63 waters were found by Arpp. Seven side chains remained completely disordered. The coordinates have been submitted to the Brookhaven Protein Data Bank.

Acknowledgements

We would like to thank our colleagues Anja Christoph for providing the original clone for the N62D mutant, Hiang Dreher and Gunter Stier for help and advice with the cloning, expression and purification of the proteins, and Yorgo Modis and Joachim Meyer for advice in various computational aspects. We thank Ned Seeman for providing the model coordinates of the four-way DNA junction. We gratefully acknowledge the help of the staff of the EMBL outstation in Hamburg, in particular of Paul Tucker, with data collection at the BWTB beam line. This work was supported by the European Commission through an institutional fellowship to O.V. (contract no. ERBCBIOCT940535). Work in the group of N62D was of excellent quality (Figure 5), allowing construction of a complete model of Endo VII N62D, including side chains. The selenium positions were very helpful for tracing the chain. Two clusters containing three closely spaced methionines each (M124, M125, M128) were easily recognized. Another cluster of two selenium atoms near the NCS axis was soon identified as M68, leaving the other selenium site to correspond to one of the N-termini. Tracing the molecule from this N-terminus to the C-terminus was straightforward. The second molecule was traced by applying the NCS operator and was adjusted manually.

The wild-type structure was traced by manually fitting whole helices or strands, from the N62D model, as rigid bodies into the signalA-weighted 2m|Fo|−D|Fc| electron density map. A second round of fitting optimized the individual amino acids. The Cα trace was clearly visible, with only a few chain breaks at the 1σ level. Most side chains were also visible.

The structures were refined using the CCP4 suite of programs DM (Cowtan, 1994), REFMAC (Mushovod et al., 1997) and ARPP (Lamzin and Wilson, 1993) and validated with PROCHECK (Laskowski et al., 1993) and WHATIF (Friend, 1990). A total of 5% of the reflections have been selected to calculate Rmerge. The starting R-factor for the N62D mutant was 0.396 (σ = 2.1 Å). Individual B-factors were restrained but no NCS restraints were applied; nevertheless the monomers differ only marginally (r.m.s.d. = 0.476 Å for the Cα atoms). Arpp was run to find water sites automatically in the later stages of refinement; 262 waters and 12 sulfates were included in the N62D structure.

The starting R-factor in the case of the wild-type structure was 0.395 (15–2.5 Å). Individual B-factors and NCS were restrained, but residues involved in crystal contacts that were clearly different in omit maps were excluded from NCS restraints. Three calcium ions were included and, additionally, 63 waters were found by Arpp. Seven side chains remained completely disordered. The coordinates have been submitted to the Brookhaven Protein Data Bank.

Acknowledgements

We would like to thank our colleagues Anja Christoph for providing the original clone for the N62D mutant, Hiang Dreher and Gunter Stier for help and advice with the cloning, expression and purification of the proteins, and Yorgo Modis and Joachim Meyer for advice in various computational aspects. We thank Ned Seeman for providing the model coordinates of the four-way DNA junction. We gratefully acknowledge the help of the staff of the EMBL outstation in Hamburg, in particular of Paul Tucker, with data collection at the BWTB beam line. This work was supported by the European Commission through an institutional fellowship to O.V. (contract no. ERBCBIOCT940535). Work in the laboratory of B.K. was supported in part by the Deutsche Forschungsgemeinschaft (DFG Normalverfahren Ke188/12-3) and the Fonds der Chemischen Industrie.

We thank Ned Seeman for providing the model coordinates of the four-way DNA junction. We gratefully acknowledge the help of the staff of the EMBL outstation in Hamburg, in particular of Paul Tucker, with data collection at the BWTB beam line. This work was supported by the European Commission through an institutional fellowship to O.V. (contract no. ERBCBIOCT940535). Work in the laboratory of B.K. was supported in part by the Deutsche Forschungsgemeinschaft (DFG Normalverfahren Ke188/12-3) and the Fonds der Chemischen Industrie.
H. Raaijmakers et al.


Received December 7, 1998; revised and accepted January 19, 1999