Published online October 12, 2004

5442–5451 *Nucleic Acids Research, 2004, Vol. 32, No. 18*
doi:10.1093/nar/gkh869

Substrate recognition and catalysis by the Holliday junction resolving enzyme Hje

Claire L. Middleton, Joanne L. Parker[1], Derek J. Richard[1], Malcolm F. White[1] and Charles S. Bond*

Division of Biological Chemistry and Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK and [1]Centre for Biomolecular Sciences, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, UK

Received August 31, 2004; Revised and Accepted September 16, 2004

PDB accession nos 1OB8 and 1OB9

ABSTRACT

Two archaeal Holliday junction resolving enzymes, Holliday junction cleavage (Hjc) and Holliday junction endonuclease (Hje), have been characterized. Both are members of a nuclease superfamily that includes the type II restriction enzymes, although their DNA cleaving activity is highly specific for four-way junction structure and not nucleic acid sequence. Despite 28% sequence identity, Hje and Hjc cleave junctions with distinct cutting patterns—they cut different strands of a four-way junction, at different distances from the junction centre. We report the high-resolution crystal structure of Hje from *Sulfolobus solfataricus*. The structure provides a basis to explain the differences in substrate specificity of Hje and Hjc, which result from changes in dimer organization, and suggests a viral origin for the *Hje* gene. Structural and biochemical data support the modelling of an Hje:DNA junction complex, highlighting a flexible loop that interacts intimately with the junction centre. A highly conserved serine residue on this loop is shown to be essential for the enzyme's activity, suggesting a novel variation of the nuclease active site. The loop may act as a conformational switch, ensuring that the active site is completed only on binding a four-way junction, thus explaining the exquisite specificity of these enzymes.

INTRODUCTION

The Holliday junction resolving enzymes bind and cleave the four-way Holliday junctions in DNA created during repair and rearrangement by the ubiquitous process of homologous recombination (1). These junctions are formed by strand exchange between homologous duplex DNA molecules. Subsequent branch migration of the Holliday junction generates stretches of heteroduplex recombinant DNA. The introduction of paired nicks in opposing strands by a structure-specific endonuclease, or junction resolving enzyme, and subsequent ligation ends the recombination process. Unresolved and partially resolved junctions are potent mutagens, therefore junction resolution must occur with high structure-specificity, and cleave both opposing strands during the lifetime of the enzyme–DNA complex.

Nature has invented this resolving activity many times in the form of highly basic, dimeric, metal-dependent enzymes. The archaeal resolving enzymes belong to a family represented by Hjc [Holliday junction cleavage (2,3)] which, like the type II restriction enzymes (TIIREs) (4,5), is a member of the nuclease superfamily bearing the sequence motif **E**(X)$_m$p**D**(X)$_n$**EXK**. Structural studies of Hjc from *Sulfolobus solfataricus* (*Sso*Hjc) (6) and *Pyrococcus furiosus* (*Pfu*Hjc) (7,8) have been reported, in addition to resolving enzymes from bacteria [RuvC (9), RusA (10)], yeast mitochondria [Ydc2 (11)], bacteriophages T4 [endonuclease VII (12)] and T7 [endonuclease I (13,14); T7eI]. The latter enzyme is also a member of the nuclease superfamily, with domain swapping resulting in an active site contributed by parts of both subunits in the dimer (13). The nuclease domain identified in Hjc and T7eI is also present in the related structure-specific nucleases XPF-ERCC1 (Rad1-Rad10 in yeast) and Mus81-Eme1. XPF-ERCC1 plays a role in nucleotide excision repair and in other repair processes, and cleaves at junctions between double-stranded and single-stranded DNA (15). Mus81-Eme1 was initially identified as a putative Holliday junction resolving enzyme (16), but is now generally considered to have a function in the rescue of stalled replication forks or cleavage of nicked Holliday junctions in meiotic recombination (17). The archaeal homologue of XPF/Mus81, known as Hef in *Pyrococcus* (18) and *Sso*XPF in *Sulfolobus* (19), is a homodimer with two identical active sites that have the same core structure as the nuclease site of Hjc (20). Thus, the nuclease domain found in Hjc is widespread in a variety of DNA repair endonucleases found in archaea and eukarya.

*To whom correspondence should be addressed: Tel: +44 1382 348325; Fax: +44 1382 345764; Email: C.S.Bond@dundee.ac.uk
Correspondence may also be addressed to Malcolm White. Tel: +44 1334 463432; Fax: +44 1334 462595; Email: mfw2@st-and.ac.uk
Present address:
Derek J. Richard, Queensland Institute of Medical Research, PO Royal Brisbane Hospital, Queensland 4029, Australia

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

Amongst the archaea, *S.solfataricus* appears unique in encoding a second resolving enzyme with 28% amino acid sequence identity to *Sso*Hjc, namely *Sso*Hje [Holliday junction endonuclease (3,21)]. Both enzymes are highly specific for Holliday junction structures but show no detectable sequence specificity (3,21). These properties distinguish Hje and Hjc from the other cellular junction resolving enzymes which are all sequence dependent, and the phage enzymes which are sequence independent and much more promiscuous in their substrate specificity (1). Despite their sequence similarity, *Sso*Hje and *Sso*Hjc differ in both their strand preference and the positioning of the nicks that resolve Holliday junction substrates. Hje shows an exquisite specificity for cleavage of strands that are continuous in the folded, stacked-X Holliday junction structure, while Hjc is specific for the exchanging strand (3,21).

No crystal structures are available for resolving enzymes bound to DNA substrates, although models have been proposed for such complexes (6,10,22). Major questions remain regarding the detail of molecular recognition of junctions by resolving enzymes, the manipulation of both the local and global structures of four-way junctions on binding, and the mechanisms by which catalysis is coupled at two independent sites to ensure productive resolution (1).

In this work, we describe the three-dimensional structure of *Sso*Hje and a model for the enzyme four-way junction complex. Comparisons with *Sso*Hjc explain the differences in Holliday junction binding and cutting patterns, and are extended to include T7eI, highlighting the observation that the DNA-cutting specificity of resolving enzymes within the nuclease superfamily, like that of the related TIIREs, is determined by the variation of interactions at the dimer interface. A conserved serine was identified and confirmed as a catalytic residue in addition to the well-characterized nuclease active site residues, and our structural model suggests a role in ensuring productive junction resolution.

MATERIALS AND METHODS

Protein expression, purification and site directed mutagenesis

The expression and purification of Hjc and Hje were carried out as described previously (3,23). Site directed mutants of Hjc and Hje were produced using the QuikChange protocol (Stratagene) and checked by DNA sequencing. The mutant proteins were purified as for the wild-type enzymes.

Equilibrium DNA-binding affinities

These were assessed by gel electrophoretic retardation analysis, as described previously, using junction Jbm5, a mobile four-way junction with 25 bp arms (3).

Single turnover kinetic assays

The assays were carried out using 1 μM purified recombinant Hjc or Hje protein in reaction buffer (20 mM Tris–HCl, pH 7.5, 50 mM NaCl, 15 mM $MgCl_2$) using 80 nM radioactively 5′-^{32}P-labelled junction 1 as a substrate (3). Calf thymus DNA (0.2 mg/ml) was added as a competitor to minimize the non-specific endonuclease activity. The reactions were initiated by the addition of Mg^{2+} to the assay mixture in 5 μl total volume, and incubated at 35 or 65°C. At set time points, the aliquots were removed and the reactions were stopped by the addition of 4 μl formamide/EDTA loading mixture and heating to 95°C, and the products were analysed by denaturing gel electrophoresis and phosphorimaging as described previously (24). All kinetic assays were carried out in triplicate, and standard errors were calculated. The four-way junction used for the kinetic studies was constructed by annealing the following four oligonucleotides, yielding a fixed four-way junction with 25 bp arms.

b-strand (5′–3′): CCTCGAGGGATCCGTCCTAGCAAGCCGCTGCTACCGGAAGCTTCTGGACC
h-strand (5′–3′): GGTCCAGAAGCTTCCGGTAGCAGCGAGAGCGGTGGTTGAATTCCTCGACG
r-strand (5′–3′): CGTCGAGGAATTCAACCACCGCTCTTCTCAACTGCAGTCTAGACTCGAGC
x-strand (5′–3′): GCTCGAGTCTAGACTGCAGTTGAGAGCTTGCTAGGACGGATCCCTCGAGG

Crystallization. Native and selenomethione *Sso*Hje were crystallized in space group $P\ 6_5$ as described previously (23). Furthermore, the tetragonal bipyramidal crystals were grown to a typical size of 0.3 mm using the hanging-drop method at 20°C and a reservoir containing 0.1 M HEPES–HCl, pH 7.5 and 2.0 M ammonium formate and drops composed of 1μl of protein solution (10 mg/ml in 0.1 M Tris–HCl, pH 8.5 and 0.2 M NaCl) and 1 μl of reservoir.

X-ray analysis. For data collection, the crystals were flash cooled in a stream of nitrogen gas maintained at 100 K either directly from the drop (hexagonal form) or after brief (5–10 s) washing in a mixture of reservoir solution with 20% (v/v) ethylene glycol. Single-wavelength anomalous dispersion (SAD) data were collected on a selenomethionine-derivative hexagonal crystal at beam line ID14-EH4 at the European Synchrotron Radiation Facility, Grenoble. An X-ray fluorescence scan spanning the Se K-edge was performed to locate the absorption peak (0.9793 Å). The diffraction data were collected and processed in space group $P\ 6_{1/5}$ with unit cell dimensions a = 92.03 Å, c = 72.39 Å using DENZO and SCALEPACK (25) and details are provided in Table 1. SAD phasing was performed using the CCP4 software suite (26). The analysis of anomalous difference Patterson maps (FFT, RSPS) yielded strong peaks corresponding to four Se positions. Phase refinement (MLPHARE) with these peaks and their enantiomer in $P\ 6_5$, followed by solvent flattening and histogram matching (DM) resulted in excellent quality interpretable electron density (figure of merit from MLPHARE and DM were 0.25 and 0.86, respectively). The initial model of an *Sso*Hje dimer was built using ARPWARP (27).

Higher resolution (1.8 Å) data, collected at SRS, Daresbury (Table 1), were used for subsequent analysis. Cycles of restrained refinement (REFMAC), addition of waters (ARP_WATERS) and graphical manipulation [O (28)] resulted in a model consisting of residues A6–A29, A35–A129 and B4–B29, B35–B135. Inspection of difference maps indicated the locations of 251 ordered water molecules. Dual conformations were modelled for the sidechains of 14 and 9 residues in subunits A and B, respectively with occupancies of

Table 1. Diffraction data and model refinement statistics

	Hje – $P\ 6_5$ SeMet	Native	Hje – $I\ 4_1$ Native
Cell constants (Å)	a = 92.03	a = 90.58	a = 76.49
	c = 72.35	c = 70.92	c = 88.94
Wavelength (Å)	0.97926	0.860	1.5418
Resolution (Å)	20.0–2.2	25.0–1.8	30.0–2.0
Observations	155702	128043	148592
Unique reflexions	23745	30127	50200
R_{sym}	0.041 (0.179)	0.033 (0.652)	0.079 (0.544)
R_{ano}	0.060 (0.129)		
Completeness (%)	90.7 (53.7)	98.0 (94.1)	97.8 (95.9)
<I/σI>	19.2 (4.6)	41.1 (1.75)	11.9 (2.46)
Refinement			
R_{cryst}		0.199	0.199
R_{free} (5%)		0.247	0.239
Protein atoms		2072 (Dimer)	1006 (Monomer)
Solvent atoms		251 (Water)	151 (Water)
		30 (Sulfate)	21 (Formate)
		12 (Ethylene glycol)	4 (Ethylene glycol)
Cruickshank's DPI		0.130	0.133

Values in parentheses refer to the highest resolution shell.

0.33:0.67 or 0.50:0.50 based on omit map electron density levels and behaviour in refinement. Model quality was monitored with PROCHECK (29), indicating no Ramachandran outliers. Other significant electron density features were modelled as two ethylene glycol (cryoprotectant) molecules and six sulfate ions, two of which are modelled at half occupancy.

Diffraction data to 2.0 Å collected on tetragonal crystals using a Rigaku rotating anode CuKα X-ray source and RAXIS-IV image plate were phased by molecular replacement (MOLREP) using a monomer from the $P\ 6_5$ structure, producing a significant solution for one molecule per a.u. in space group $I\ 4_1$ with unit cell dimensions a = 76.49 Å, c = 88.94 Å. The dimer is formed by a crystallographic 2-fold rotation axis. The model was successfully rebuilt and refined using similar protocols to the hexagonal form, producing a model consisting of residues 6–129 (four dual sidechain conformers), one ethylene glycol molecule, seven formate ions and 151 ordered waters.

Molecular modelling. The X3DNA program (30) was used to create segments of idealized B-DNA. The programs O, LSQMAN (31) and PDB-MODE (32) were used for the manipulation of coordinates. CNS program (33) was used for energy minimization.

RESULTS AND DISCUSSION

The structure of *Sso*Hje was solved by Se-targeted SAD methods from a sulfate-containing hexagonal crystal form which diffracts to 1.8 Å, and subsequently a tetragonal form (2.0 Å) The structure of the Hje monomer (Figure 1A) is, similar to its homologue Hjc, characteristic of the nuclease superfamily. In summary, it is an α/β protein with a central doubly wound six-stranded mixed β-sheet flanked by three α-helices, α1 and α3 on one face and α2 on the other, that run approximately parallel to the β-strands. The C-terminal ends of the third and fourth strands of this sheet (βC and βE) pack face-on against a short two-stranded antiparallel β-sheet (βD and βG). The catalytic site, identified by the presence of the conserved residues, Glu-10, Pro-38, Asp-39, Glu-52 and Lys-54 is positioned close to the N-terminus of strand βB and a bend in strand βC. The N-termini of the α-helices are oriented towards the DNA-binding face of the molecule.

The sequence identity between *Sso*Hje and *Sso*Hjc (28%) or *Pfu*Hjc (29%) is reflected in a structural comparison of the monomers. The superpositions of representative monomers of *Sso*Hje, *Sso*Hjc and *Pfu*Hjc yield root-mean-squared deviations (r.m.s.d.) of 1.5 Å for 90 and 106 equivalent Cα atoms. (For reference, an r.m.s.d. of 0.5 Å for 115 atoms is observed between subunits of the hexagonal and tetragonal forms of Hje). The clearest differences in superposition occur in residues 58–64 and 108–112 of *Sso*Hjc. These residues are involved in crystal contacts in two *Sso*Hjc crystal forms and may be related to the concentration-dependent auto-inhibition observed for *Sso*Hjc (34) but not *Sso*Hje.

Hje and Hjc have different quaternary structures

Like the other characterized members of the Hjc family (and all known Holliday junction resolving enzymes), Hje is dimeric both in solution (running as a dimer during gel filtration; data not shown) and in the crystal (Figure 1B): congruent dimers (r.m.s.d. 0.8 Å for 232 Cα atoms) occur in both crystal forms, reflected by crystallographic (tetragonal form) or non-crystallographic (hexagonal form) symmetry. However, when *Sso*Hje and either *Sso*Hjc or *Pfu*Hjc are superposed using only the structurally analogous Cα atoms of one subunit, the relative orientations of the partner subunit are startlingly different (Figure 2A). The operation required to transform subunit B of Hjc onto subunit B of *Sso*Hje can be described by a screw axis running approximately perpendicular to the molecular dyad. There is a translation of 6 Å along and a rotation of 30° about this axis [DYNDOM (35)].

The cause of this large structural difference is clear from the analysis of the interfaces of *Sso*Hje and *Sso*Hjc (Figure 2A and C). Of the 17 residues forming the interface in *Sso*Hje, only five are conserved in sequence in *Sso*Hjc and five conservatively substituted, and even the conserved residues play different roles in each interface. The Cα positions of the interface residues of Hje and Hjc are structurally well conserved within the monomer, except for the region 77–82 (*Sso*Hje) which corresponds to a relative insertion of two residues in the Hje sequence. This difference results in the sidechains of three large hydrophobic residues, Met-77, Phe-78 and Met-80, being inserted into the dimer interface and prising the subunits apart (Figure 2B). Residues Pro-28 (Hje) and Pro-30 (*Sso*Hjc) have comparable conformations relative to the monomer and contribute to the interface by interacting with the dyad-related proline, yet the arrangement at the interface is such that the proline imide rings interact with opposite faces in the two structures (Figure 2C, Hje Pro-28 is to the left of the dyad, while Hjc Pro-30 is to the right), corresponding to a relative shift of over 9 Å between the partner prolines in the two structures. While Phe-74 is conserved in sequence, its intersubunit interactions are also totally different: in *Sso*Hjc, it interacts with the sidechain of Ala-25' (apostrophe indicates the other subunit) and the backbone of Val-26',

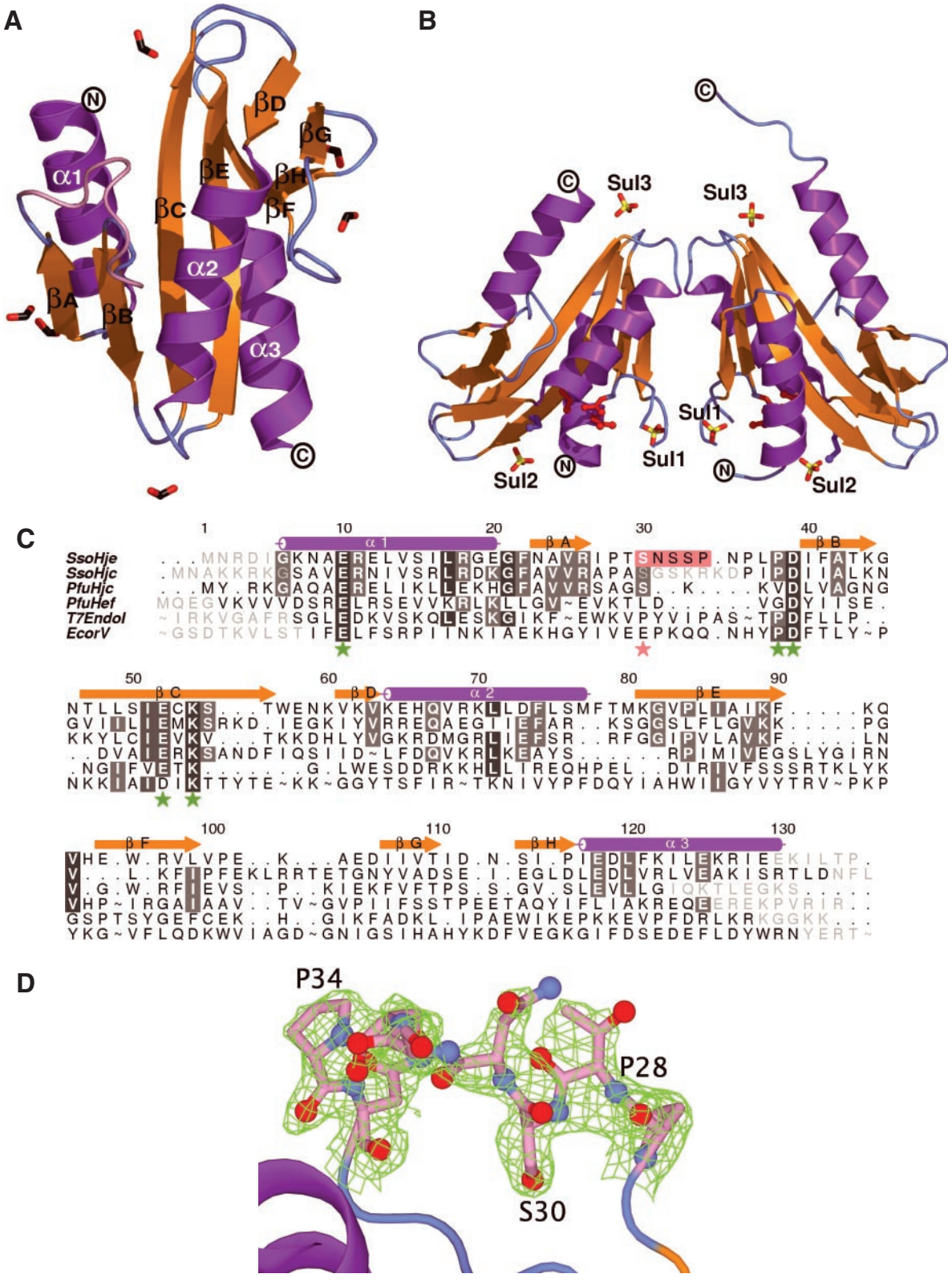

Figure 1. The structure of *Sso*Hje. Ribbon diagram of (**A**) the *Sso*Hje monomer (tetragonal form) and (**B**) the *Sso*Hje dimer (hexagonal form). The colour scheme (strand, gold; helix, purple; loop; cyan) is used in all figures. Formate ions, black/red sticks; sulfate ions, yellow/red sticks. (**C**) Structure-based sequence alignment of nuclease superfamily members. Green asterisks highlight nuclease motif residues; a pink asterisk the conserved serine; pink shading indicates the flexible loop; A tilde indicates sections of EcoRV, Hef or T7eI sequence which have been omitted for clarity. Residues absent from crystal structures are shown in grey font. (**D**) σ_a-weighted electron density (0.9 σ, green mesh) for the flexible loop (pink, residues 28–36) from the tetragonal form of *Sso*Hje. [All molecular graphics prepared with PYMOL (52), sequence alignments prepared using INDONESIA (31) and ALINE (available from the authors)].

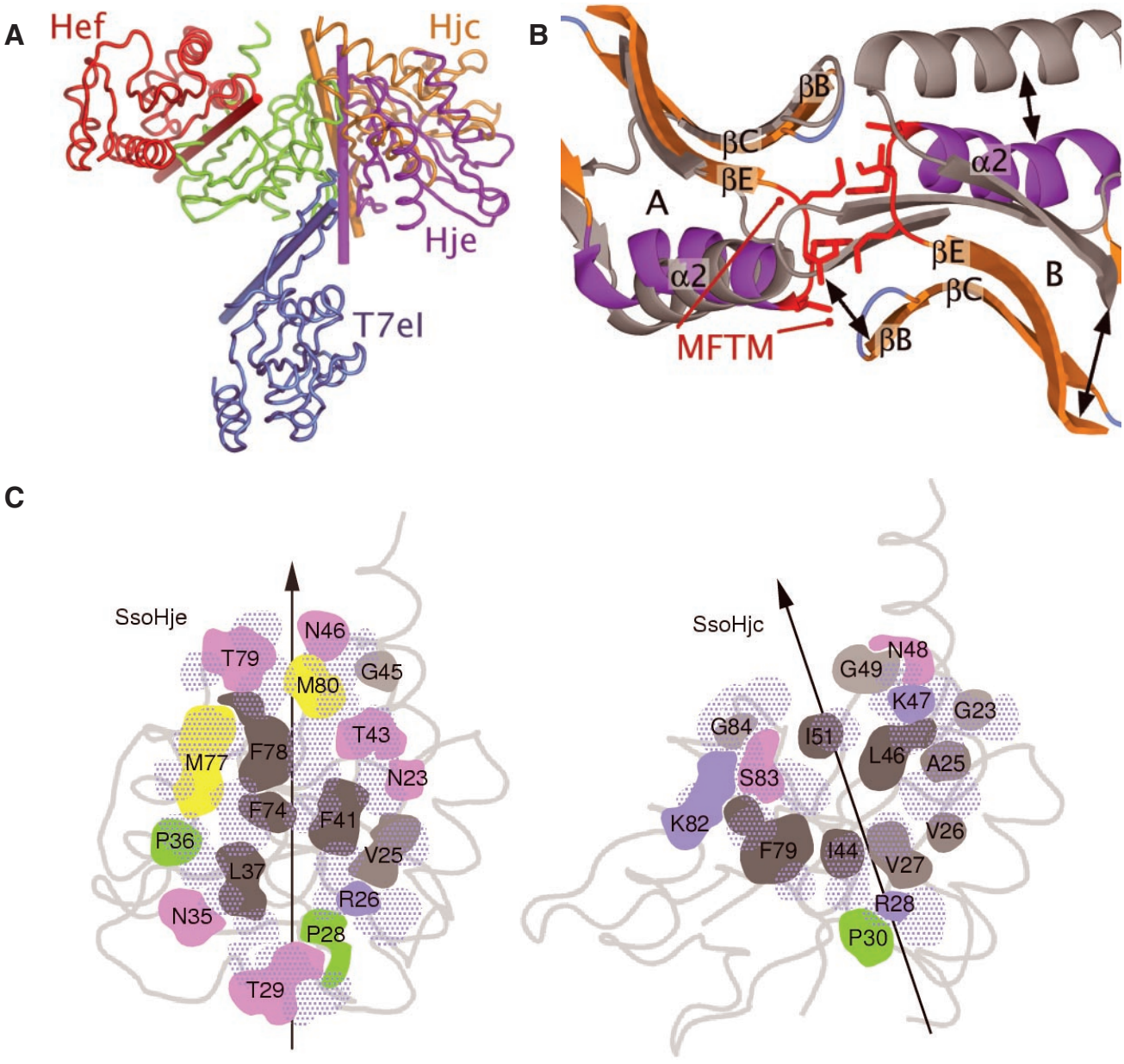

Figure 2. Structural variation at the dimer interface. (**A**) View of the *Sso*Hje (purple), *Sso*Hjc (gold), T7eI (blue) and *Pfu*Hef (red) dimers superposed using the atoms of one subunit (green). Rods indicate the orientations of the molecular dyads. (**B**) *Sso*Hje (colour) superposed on *Sso*Hjc (grey, subunit A; black, subunit B) using the atoms of subunit A. The segment involving residues Met77-Met80 (red) of *Sso*Hje subunit A clashes with part of *Sso*Hjc subunit B. (**C**) Schematic representation of residues contributing to the dimer interface of *Sso*Hje and *Sso*Hjc. Interface residues from the subunit facing the reader are coloured by residue type (grey, hydrophobic; blue, basic; pink, polar; yellow, methionine, green, proline). The common orientation of this subunit is indicated by the grey backbone trace. Residues from the partner subunit are shown in dotted blue and the molecular dyad is marked as a black arrow.

whereas in *Sso*Hje, Met-77 satisfies these interactions and Phe-74 forms sidechain–sidechain interactions with Phe-41'. These differing arrangements are nevertheless accompanied by conservation of surface area buried at the interface, for which *Sso*Hje (740 Å^2) and *Sso*Hjc (758 Å^2) have corresponding values.

Possible viral origin of Hje

S.solfataricus is unique, among the archaea for which genome sequences have been determined, in having two Hjc homologues (Figure 3A). Even the closely related *Sulfolobus tokodaii* genome has only a single Hjc gene, which is very similar (67% identity) to the *Sso*Hjc sequence. The *Sulfolobus islandicus*-infecting rudiviruses *Sirv*1 and *Sirv*2 contain an Hjc homologue (36), raising the possibility of a viral origin for the Hje gene. Of the archaeal Hjc homologues, *Sso*Hje is the most similar in sequence to *Sirv*Hjc, suggesting a close relationship. A similarly altered quaternary structure in the viral Hjc would support this hypothesis. In the absence of a structure of *Sirv*Hjc, sequence analysis (Figure 3B) highlights similarities in residues involved in

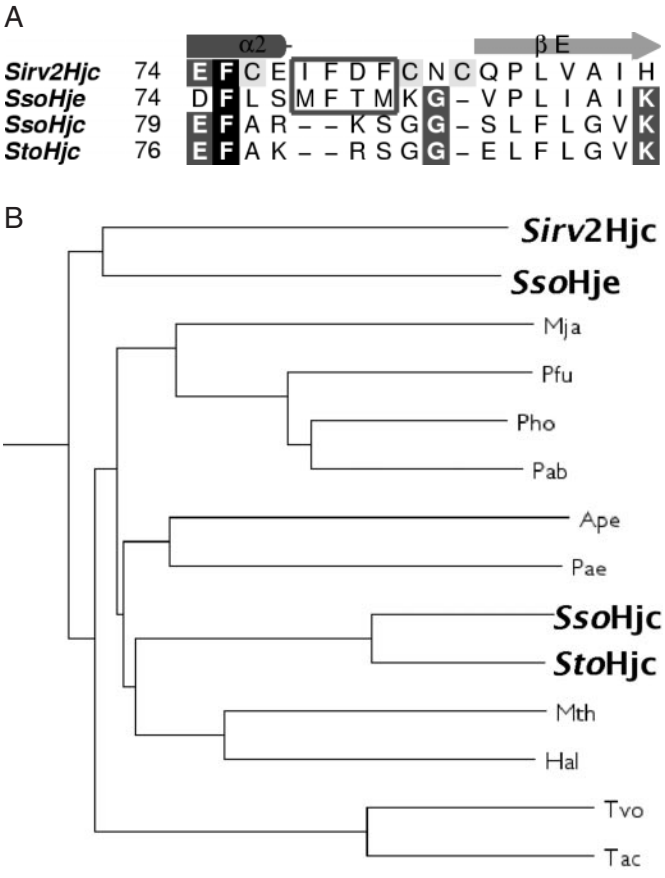

Figure 3. *Sso*Hje is more similar in sequence to *Sirv*2Hjc than to *Sso*Hjc, which is more similar to *S.tokodaii* Hjc. (**A**) Sequence alignment highlighting the region implicated in altering the dimer interface (boxed). Conserved residues are shaded. (**B**) Phylogenetic tree of Hjc sequences, calculated by the neighbour-joining method with midpoint rooting, based on a ClustalX multiple sequence alignment (*MACVECTOR*, Accelrys Inc.). Labels indicate the following species: Sso, *Sulfolobus solfataricus*; Sto, *Sulfolobus tokodaii*, Ape, *Aeropyrum pernix*; Pae, *Pyrobaculum aerophilum*; Afu, *Archaeoglobus fulgidus*; Pfu, *Pyrococcus furiosus*; Pho, *Pyrococcus horikoshii*; Pab, *Pyrococcus abyssi*; Mja, *Methanococcus jannaschii*; Hal, *Halobacterium NRC1*; Tac, *Thermoplasma acidophilum*; Tvo, *Thermoplasma volcanii*; Mth, *Methanothermobacter thermoautotrophicum*; *Sirv*2, *Sulfolobus islandicus*-infecting rudivirus 2.

dimerization, particularly those corresponding to the *Sso*Hje hydrophobic MFTM motif, including conservation of the phenylalanine residue. In addition, a cluster of three cysteine residues on either side of this motif could confer structural differences compared with the canonical Hjc dimer interface.

Wider comparison with other family members

The nuclease superfamily also includes the TIIRE as well as another Holliday junction resolving enzyme, T7eI. A feature of these enzymes is that they all cleave specific DNA molecules (with sequence specificity in the case of the TIIRE and structure specificity in the junction resolving enzymes). TIIRE produce a variety of cutting patterns, leaving blunt ends or 3′ or 5′ overhangs (37). This variety is produced by variation in the arrangement of pairs of nuclease domains by inserted or deleted sequence elements. A comparison of *Sso*Hje, *Sso*Hjc and T7eI indicates that this paradigm extends to the junction resolving enzymes.

While Hjc family members consist of straightforward dimers with no domain swapping, T7eI contains swapped secondary structure elements. In the T7eI structure (13), the loop following strand βA functions differently as the point of secondary structure swapping: instead of forming a flexible loop as in *Sso*Hje, these residues contribute to the β-bridge that links the subunits. It is clear from a superposition of common structural elements in one monomer that the second monomer is displaced greatly from the position of the partner subunit in either *Sso*Hje or *Sso*Hjc (Figure 2A). The *Pyrococcus* protein Hef is a homologue of eukaryotic XPF, the endonuclease involved in nucleotide excision repair (15,38). The nuclease domain of Hef shares significant structural similarity with the Hjc family, including conservation of active site acidic residues involved in binding the catalytic metal ion (20), yet it provides a further example of distinctive dimer formation in the nuclease superfamily (Figure 2A). The significance of Hef dimerization, however, remains unclear given that the protein cleaves only one strand of DNA substrates.

Relative orientation of pairs of active sites

Superposition of Hje onto the crystal structure of EcoRV bound to substrate DNA [1RVB (39); r.m.s.d. 2.0 Å for 58 Cα atoms] places the scissile phosphodiester bond of the DNA in a reasonable position for catalysis in the *Sso*Hje active site. Using this as a point of reference, the distance between the scissile bonds (marked yellow in Figure 4A) in the two active sites of *Sso*Hje is 22 Å compared to 28 Å in *Sso*Hjc. This explains the observation that Hje introduces paired strand scissions near the centre of the four-way junction than Hjc [2–3 bases 3′ of the centre for Hje; 3–4 bases 3′ of the centre for Hjc (3)]. T7eI cleaves DNA 1–2 bases from the junction centre on the continuous strands with minimal disruption of the junction centre (40). This position is located on the outside of a stacked-X junction, requiring that the active sites must face each other, hence the need for the long connecting bridge to wrap the enzyme around the junction. When the cleavage positions of *Sso*Hje and *Sso*Hjc are mapped on an undisrupted junction, they all occur on the face of the junction that has minor groove character at the centre (3), and therefore both *Sso*Hje and *Sso*Hjc must present their pairs of active sites in a coplanar manner.

Sulfate-binding may mimic coordination of DNA backbone

In hexagonal Hje crystals, six sulfate ions are bound to the protein (Figures 1B and 4A). They comprise three pairs related by the protein dimer's molecular dyad and are labelled accordingly (residue numbers 1–3 and corresponding to protein chains A and B). Sulfates A1 and B1 (0.8 Å sulphur atom deviation) and A/B3 (1.5 Å) are related strictly, while sulfates A/B2 (6.0 Å) occupy less similar positions. Sulfates A/B1 are located close to the active site and are coordinated by the sidechain atoms of Arg-11, Arg-18 and Arg-26 that are highly conserved and, in the case of Arg-18 and Arg-26, have been shown by mutational analysis in Hjc (41) to be essential for enzyme activity. Similarly, sulfate B2 is coordinated by the highly conserved Lys-89 and Lys-91, the latter having also been mutated in Hjc with deleterious

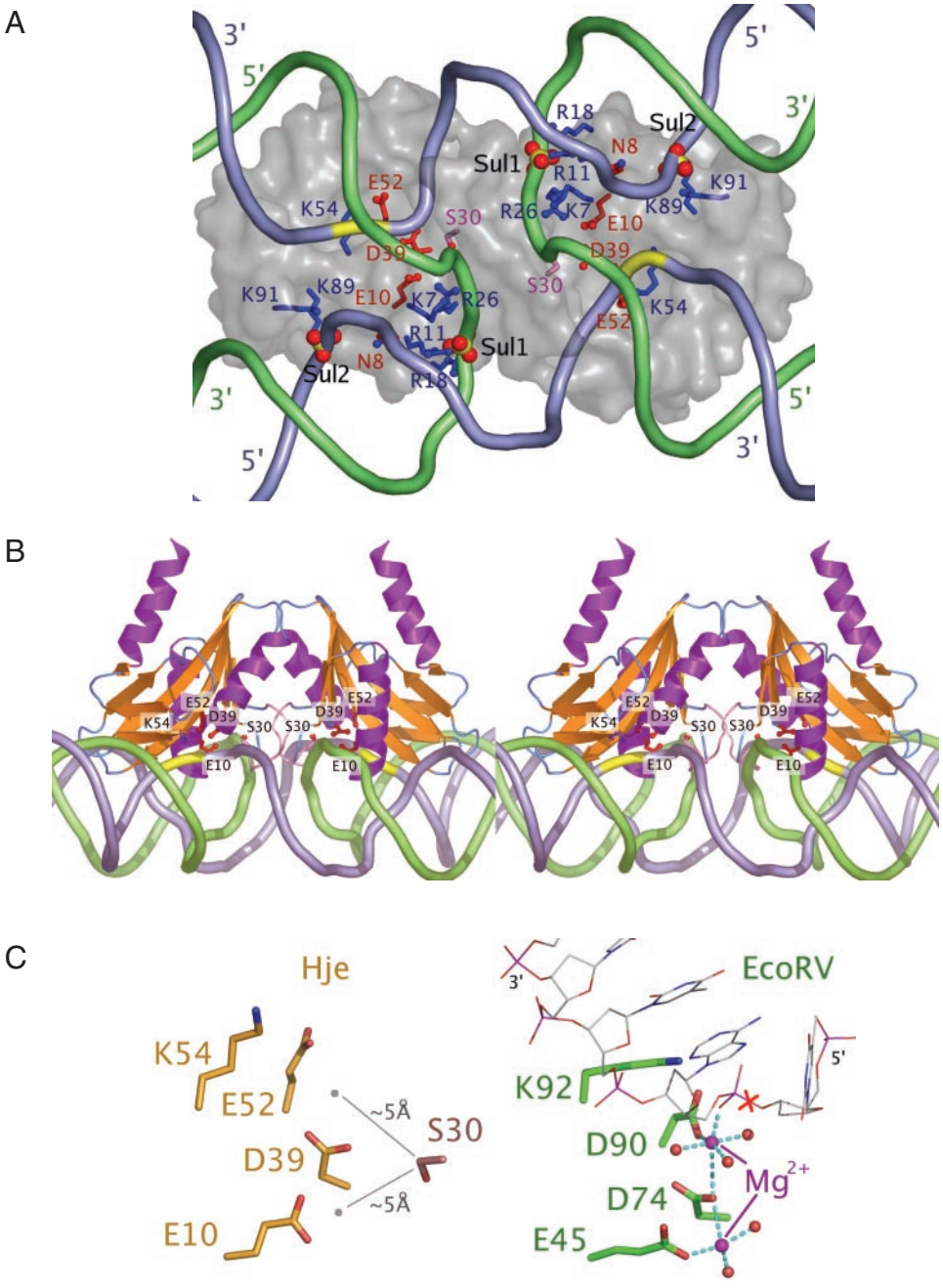

Figure 4. A model of *Sso*Hje bound to a four-way junction. (**A**) View along the molecular dyad; semi-transparent protein surface with selected residues shown as ball-and-stick. The scissile phosphate is highlighted in yellow. (**B**) Stereo view perpendicular to the molecular dyad. (**C**) Active site residues of Hje compared with an EcoRV–DNA substrate complex [1RVB (39)], indicating the proximity of Ser-30 to the catalytic metal centre.

effects (41). We propose that these sulfates mimic the positions of DNA-backbone phosphate moieties in the enzyme:junction complex. In contrast, sulfates A/B3 are located on the surface of the protein diametrically opposite the active site, adjacent to helix α, and have no obvious relationship to the binding of substrate DNA (Figure 1B).

A model for the Hje-junction complex

To extend the insight gained from the *Sso*Hje structure, we exploited the similarity of the active site to the known structures of TIIRE complexed with DNA, along with the observation of sulfate ions bound to residues presumed to bind DNA, to develop a model of the complex of *Sso*Hje with a four-way junction. A non-redundant set of coordinates [1D02, 1DFM, 1EYU, 1FIU, 1IAW, 1KC6, 1M0D and 1RVB; PDB (42)] were superimposed on monomer A of *Sso*Hje using the four catalytic residues of the nuclease motif. The inspection of the resulting superpositions indicated that DNA molecules from EcoRV [1RVB (39)] and NaeI [1IAW (43)] occupied suitable orientations with minimal steric clash between DNA and *Sso*Hje. The EcoRV duplex DNA coordinates

superimposed on active site A were extended with idealized B-DNA. With minor adjustment, the phosphate backbone of the cleaved strand passes through the observed position of sulfate B2 and the uncleaved strand through sulfate B1. This DNA segment was then transformed by the molecular dyad to produce an intermediate model of *Sso*Hje bound to two DNA duplexes. The only significant steric clash in this model, of the uncleaved strand with the loop residues 30–34, is relieved by 'cleaving' this strand and 'ligating' it to the other duplex, thus forming a four-way junction (Figure 4). Further appealing features of this model include the interaction of the N-terminal helix of *Sso*Hje with the major groove of the uncleaved arm of the junction where Lys-7 and Arg-11 [essential for DNA binding (7) and cleavage (41), respectively] can interact with the phosphate backbone and Asn-8 can form hydrogen bonds to the bases.

Substrate recognition and manipulation

The largest structural differences between a native stacked-X junction and the model of the junction bound to *Sso*Hje involve the opening up of the junction centre and the widening of the acute angle between non-stacked junction arms. The electrostatic analysis [GRASP (44)] of the structure of a four-way junction [PDB entry 1NVY (45) with arms extended using idealized B-DNA; data not shown] indicates that the extremum of negative potential is found on the minor groove face at the centre of the junction (although this analysis disregards the charge-screening effect of any counterions). A similar analysis of *Sso*Hje shows the extremum of positive potential at the middle of the molecular dyad on the face that includes both active sites, pointing to simple electrostatic attraction providing a 'first approach' of substrate recognition and the specificity for binding the minor groove face of the junction. The dipoles of all six α-helices (46), which are oriented with their N-termini towards the DNA-binding face, contribute to this positive potential. The skewed arrangement of the stacked-X junction allows the major grooves of both uncleaved arms to interact with the conserved phosphate-binding residues (Lys-7, Arg-11, Arg-18, Lys-89 and Lys-91) and helix α1. It is probably that induced fit of both enzyme and junction contribute to the formation of a catalytically competent complex: distortion of the junction centre would allow the residues of the flexible loop (see below) to insert between the DNA strands at the junction centre. Such distortions are a feature of several Holliday junction resolving enzymes, including Hjc (47), Cce1 (48) and RuvC (47).

Serine 30 is catalytically essential and lies on a flexible loop

The flexible loop located between strands βA and βB in Hjc/Hje is positioned close to the dimer interface in a position where it is obliged to interact with the centre of the DNA junction substrate (Figure 4). The sequence of this segment is variable in length throughout the Hjc family (residues 30–36 in *Sso*Hje, 32–40 in *Sso*Hjc and 29–31 in *Pfu*Hjc) and exhibits characteristics of low complexity, with low scores for globularity [GLOBPLOT (49)]. In the hexagonal form of *Sso*Hje, five residues (30–34) cannot be modelled due to disorder, but these residues are visible in electron density in the tetragonal form. *Sso*Hjc contains the longest example of this loop, which is not observed in electron density, while for *Pfu*Hjc the short three-residue segment is observed (8). Despite considerable variation in the length and sequence of the loop, it contains one of the few residues invariant across the entire Hjc family: a serine (Ser-30 in *Sso*Hje; Ser-32 in *Sso*Hjc).

To assess its importance for catalysis, Ser-30 in *Sso*Hje was mutated to alanine, cysteine and threonine, and the equivalent Ser-32 in *Sso*Hjc was mutated to an alanine. All mutants were tested for the ability to cleave a fixed four-way junction substrate under single turnover conditions at 65°C (Figure 5A and Table 2). The activities of wild-type *Sso*Hjc and *Sso*Hje were measured at 35°C, as their activities at 65°C were too fast to allow accurate determination of reaction rates. *Sso*Hje is ∼30-fold more active than *Sso*Hjc under single turnover conditions with this junction substrate (Figure 5B). This explains the observation that *Sso*Hje is expressed at a very low level in *Sulfolobus* cells, but has an easily detectable activity (50). The activities of the wild-type enzymes at 65°C were extrapolated from the data at 35°C based on the observation that the activity of *Sso*Hje doubles with every 10°C increase in reaction temperature under these conditions (J. L. Parker and M. F. White, unpublished data). The *Sso*Hjc S32A mutant had no detectable catalytic activity at 65°C. As the lower limit of detection of this assay corresponds to a k_{cat} value of $\sim 5 \times 10^{-4}$ min^{-1}, we estimate a decrease in catalytic activity due to this mutation of at least three orders of magnitude. The higher specific activity of *Sso*Hje allowed for accurate determination of the catalytic rates for these mutants. All three mutants, S30A, S30C and S30T, showed a decrease in catalytic rate of 3–4 orders of magnitude. The S30T mutation, which conserves the hydroxyl group of the serine sidechain, has a slightly higher activity than the other two mutants but is still severely compromised. Analysis of the binding affinity of wild-type and S32A *Sso*Hjc shows that the mutation does not affect junction binding, with the mutant binding a four-way junction substrate with a K_D similar to wild-type *Sso*Hjc (Table 2). These data point to an important catalytic role for this conserved serine (in the Hjc family) that is highly sensitive to even conservative substitutions.

Analysis of the precise catalytic mechanism of members of the nuclease superfamily has yielded a number of similar models (37). Common features include the binding of between one and three divalent metal ions by conserved acidic residues resulting in the polarization of a water molecule that is then stabilized by an additional conserved (usually) basic residue. In concert with the nucleophilic attack of this water on the target phosphoryl group, the 3′-O$^-$ leaving group is protonated by another water molecule that is also (usually) coordinated by one of the metal ions. For the Hjc family, we have implicated a serine residue in the catalytic mechanism, and this is, to the best of our knowledge, without precedent in the nuclease superfamily. [Although a serine residue substitutes for one of the typically conserved acidic residues in Cfr10I, the missing metal-binding function is provided by an additional acidic residue (51)]. In order to define possible roles for Ser-30, we have used EcoRV as a model for the active site because, of the well-characterized nuclease family members, it is the most similar in local structure to Hje/Hjc. Ser-30 might play a role in satisfying hydrogen bonds with interrupted base pairs, although if this was the case, one would expect threonine and possibly cysteine mutants to maintain reasonable levels of activity. Given that Ser-30-Oγ is positioned ∼5 Å from the

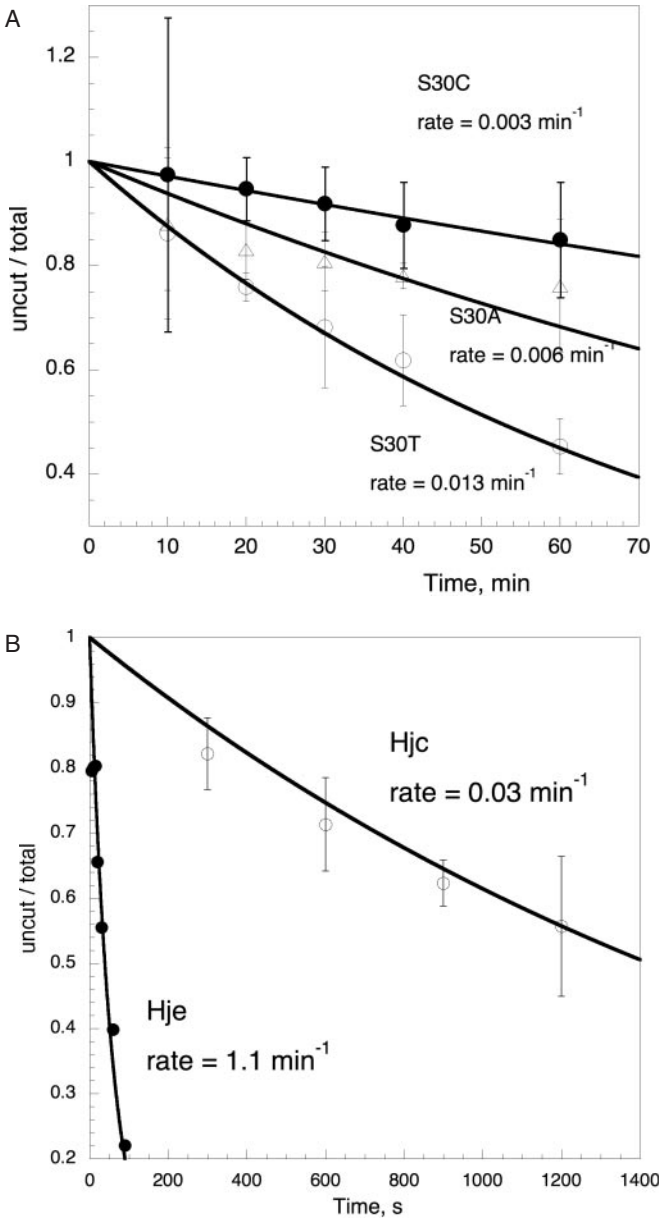

Figure 5. Mutants of serine 30 have significantly reduced activity. (**A**) Single turnover reaction rates for cleavage of a fixed four-way junction by site directed mutants of Hje at 65°C. Junction 1, radioactively 5′-^{32}P-labelled on the b-strand, was cleaved for the indicated times. The reactions were stopped by the addition of EDTA, and the extent of cleavage was determined by the separation of substrate and products by gel electrophoresis and phosphorimaging. Reaction rates are also reported in Table 1. (**B**) Comparison of single turnover reaction rates for the wild-type Hje and Hjc enzymes at 35°C. Experimental conditions were as described for (A). All data points are the mean of triplicate measurements, and standard errors are indicated.

predicted position of the site II catalytic metal ion, a role in the highly defined network of hydrogen bonding that bridges the hydrated metal ions and the substrate would be consistent with the observation that serine cannot be substituted by conservative changes to threonine (steric hindrance of the side-chain methyl group) or cysteine (sulphur being a weaker nucleophile).

Table 2. Kinetic analysis of four-way junction cleavages by *Sso*Hjc and *Sso*Hje mutants

Enzyme	k_{cat} (min^{-1}) 35°C ($\times 10^3$)	k_{cat} (min^{-1}) 65°C ($\times 10^3$)	k_{cat} mutant/ k_{cat} WT	K_D[b](nM)
Hjc WT	27 ± 1	240[a]	1	1
Hjc S32A	—	<0.5[c]	$<2 \times 10^{-3}$	1.2 ± 0.1
Hje WT	1000 ± 50	8800[a]	1	—
Hje S30A	—	5.7 ± 0.7	6×10^{-4}	—
Hje S30C	—	2.8 ± 0.1	3×10^{-4}	—
Hje S30T	—	13 ± 0.3	1.5×10^{-3}	—

[a]Values for k_{cat} for wild-type Hjc and Hje were measured at 35°C, as the rates were too fast at 65°C to allow accurate measurement. These values were estimated based on the observation that the activity of Hje doubles for each 10°C increase in the reaction temperature.
[b]Wild-type Hje does not retard labelled four-way DNA junctions at concentrations below 1 μM protein under electrophoretic gel retardation conditions.
[c]Values of $k_{cat} < 0.5 \times 10^{-3}$ min^{-1} could not be measured. The rates for mutants where no activity could be detected are therefore recorded as '<0.5' above. Actual rates under these conditions may be far $<0.5 \times 10^{-3}$ min^{-1}.

The positioning of this essential catalytic residue on a flexible loop on the junction-binding surface of the protein offers the possibility that this loop acts as a molecular switch during DNA recognition and catalysis. Junction binding, coupled with distortion of the junction centre, could allow the reorientation of the flexible loop containing the catalytic serine, completing the active site and allowing catalysis to proceed. Such a mechanism would ensure that only correctly bound substrate is cleaved, and that nicks occur at both active sites within the lifetime of the complex, thus explaining the exquisite specificity of the Hjc family for four-way DNA junctions, which is achieved without any detectable sequence specificity.

ACKNOWLEDGEMENTS

We thank the staff at the ESRF (Gordon Leonard) and Daresbury Laboratory for help with data collection. This work was funded by the BBSRC. C.S.B. is a BBSRC David Phillips Research Fellow and M.F.W. a Royal Society University Research Fellow.

REFERENCES

1. Lilley,D.M. and White,M.F. (2001) The junction-resolving enzymes. *Nature Rev. Mol. Cell Biol.*, **2**, 433–443.
2. Komori,K., Sakae,S., Shinagawa,H., Morikawa,K. and Ishino,Y. (1999) A Holliday junction resolvase from *Pyrococcus furiosus*: functional similarity to *Escherichia coli* RuvC provides evidence for conserved mechanism of homologous recombination in Bacteria, Eukarya, and Archaea. *Proc. Natl Acad. Sci. USA*, **96**, 8873–8878.
3. Kvaratskhelia,M. and White,M.F. (2000) Two Holliday junction resolving enzymes in *Sulfolobus solfataricus*. *J. Mol. Biol.*, **297**, 923–932.
4. Daiyasu,H., Komori,K., Sakae,S., Ishino,Y. and Toh,H. (2000) Hjc resolvase is a distantly related member of the type II restriction endonuclease family. *Nucleic Acids Res.*, **28**, 4540–4543.
5. Kvaratskhelia,M., Wardleworth,B.N., Norman,D.G. and White,M.F. (2000) A conserved nuclease domain in the archaeal Holliday junction resolving enzyme Hjc. *J. Biol. Chem.*, **275**, 25540–25546.
6. Bond,C.S., Kvaratskhelia,M., Richard,D., White,M.F. and Hunter,W.N. (2001) Structure of Hjc, a Holliday junction resolvase, from *Sulfolobus solfataricus*. *Proc. Natl Acad. Sci. USA*, **98**, 5509–5514.

7. Nishino,T., Komori,K., Ishino,Y. and Morikawa,K. (2001) Dissection of the regional roles of the archaeal Holliday junction resolvase Hjc by structural and mutational analyses. *J. Biol. Chem.*, **276**, 35735–35740.
8. Nishino,T., Komori,K., Tsuchiya,D., Ishino,Y. and Morikawa,K. (2001) Crystal structure of the archaeal Holliday junction resolvase Hjc and implications for DNA recognition. *Structure*, **9**, 197–204.
9. Ariyoshi,M., Vassylyev,D.G., Iwasaki,H., Nakamura,H., Shinagawa,H. and Morikawa,K. (1994) Atomic structure of the RuvC resolvase: a Holliday junction-specific endonuclease from *E.coli*. *Cell*, **78**, 1063–1072.
10. Rafferty,J.B., Bolt,E.L., Muranova,T.A., Sedelnikova,S.E., Leonard,P., Pasquo,A., Baker,P.J., Rice,D.W., Sharples,G.J. and Lloyd,R.G. (2003) The structure of *Escherichia coli* RusA endonuclease reveals a new Holliday junction DNA binding fold. *Structure*, **11**, 1557–1567.
11. Ceschini,S., Keeley,A., McAlister,M.S., Oram,M., Phelan,J., Pearl,L.H., Tsaneva,I.R. and Barrett,T.E. (2001) Crystal structure of the fission yeast mitochondrial Holliday junction resolvase Ydc2. *EMBO J.*, **20**, 6601–6611.
12. Raaijmakers,H., Vix,O., Toro,I., Golz,S., Kemper,B. and Suck,D. (1999) X-ray structure of T4 endonuclease VII: a DNA junction resolvase with a novel fold and unusual domain-swapped dimer architecture. *EMBO J.*, **18**, 1447–1458.
13. Hadden,J.M., Convery,M.A., Declais,A.C., Lilley,D.M. and Phillips,S.E. (2001) Crystal structure of the Holliday junction resolving enzyme T7 endonuclease I. *Nature Struct. Biol.*, **8**, 62–67.
14. Hadden,J.M., Declais,A.C., Phillips,S.E. and Lilley,D.M. (2002) Metal ions bound at the active site of the junction-resolving enzyme T7 endonuclease I. *EMBO J.*, **21**, 3505–3515.
15. de Laat,W.L., Appeldoorn,E., Jaspers,N.G. and Hoeijmakers,J.H. (1998) DNA structural elements required for ERCC1-XPF endonuclease activity. *J. Biol. Chem.*, **273**, 7835–7842.
16. Boddy,M.N., Gaillard,P.H., McDonald,W.H., Shanahan,P., Yates,J.R.,III and Russell,P. (2001) Mus81-Eme1 are essential components of a Holliday junction resolvase. *Cell*, **107**, 537–548.
17. Osman,F., Dixon,J., Doe,C.L. and Whitby,M.C. (2003) Generating crossovers by resolution of nicked Holliday junctions: a role for Mus81-Eme1 in meiosis. *Mol. Cell*, **12**, 761–774.
18. Komori,K., Fujikane,R., Shinagawa,H. and Ishino,Y. (2002) Novel endonuclease in Archaea cleaving DNA with various branched structure. *Genes Genet. Syst.*, **77**, 227–241.
19. Roberts,J.A., Bell,S.D. and White,M.F. (2003) An archaeal XPF repair endonuclease dependent on a heterotrimeric PCNA. *Mol. Microbiol.*, **48**, 361–371.
20. Nishino,T., Komori,K., Ishino,Y. and Morikawa,K. (2003) X-Ray and biochemical anatomy of an archaeal XPF/Rad1/Mus81 family nuclease. Similarity between its endonuclease domain and restriction enzymes. *Structure*, **11**, 445–457.
21. Kvaratskhelia,M. and White,M.F. (2000) An archaeal Holliday junction resolving enzyme from *Sulfolobus solfataricus* exhibits unique properties. *J. Mol. Biol.*, **295**, 193–202.
22. Declais,A.C., Fogg,J.M., Freeman,A.D., Coste,F., Hadden,J.M., Phillips,S.E. and Lilley,D.M. (2003) The complex between a four-way DNA junction and T7 endonuclease I. *EMBO J.*, **22**, 1398–1409.
23. Middleton,C.L., Parker,J.L., Richard,D.J., White,M.F. and Bond,C.S. (2003) Crystallization and preliminary X-ray diffraction studies of Hje, a Holliday junction resolving enzyme from *Sulfolobus solfataricus*. *Acta Crystallogr. D Biol. Crystallogr.*, **59**, 171–173.
24. White,M.F., Giraud-Panis,M.J., Pohler,J.R. and Lilley,D.M. (1997) Recognition and manipulation of branched DNA structure by junction- resolving enzymes. *J. Mol. Biol.*, **269**, 647–664.
25. Otwinowski,Z. and Minor,W. (1996) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
26. CCP4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 760–764.
27. Perrakis,A., Morris,R. and Lamzin,V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.*, **6**, 458–463.
28. Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.
29. Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **26**, 283–291.
30. Lu,X.J., Shakked,Z. and Olson,W.K. (2000) A-form conformational motifs in ligand-bound DNA structures. *J. Mol. Biol.*, **300**, 819–840.
31. Kleywegt,G.J. and Jones,T.A. (1997) Detecting fold motifs and similarities in protein structures. *Methods Enzymol.*, **277**, 525–545.
32. Bond,C.S. (2003) Easy editing of Protein Data Bank formatted files with EMACS. *J. Appl. Cryst.*, **36**, 350–351.
33. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S.*et al.* (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
34. Kvaratskhelia,M., Wardleworth,B.N., Bond,C.S., Fogg,J.M., Lilley,D.M. and White,M.F. (2002) Holliday junction resolution is modulated by archaeal chromatin components *in vitro*. *J. Biol. Chem.*, **277**, 2992–2996.
35. Hayward,S. and Lee,R.A. (2002) Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J. Mol. Graph. Model.*, **21**, 181–183.
36. Birkenbihl,R.P., Neef,K., Prangishvili,D. and Kemper,B. (2001) Holliday junction resolving enzymes of archaeal viruses SIRV1 and SIRV2. *J. Mol. Biol.*, **309**, 1067–1076.
37. Pingoud,A. and Jeltsch,A. (2001) Structure and function of type II restriction endonucleases. *Nucleic Acids Res.*, **29**, 3705–3727.
38. Park,C.H. and Sancar,A. (1994) Formation of a ternary complex by human XPA, ERCC1, and ERCC4(XPF) excision repair proteins. *Proc. Natl Acad. Sci. USA*, **91**, 5017–5021.
39. Kostrewa,D. and Winkler,F.K. (1995) $Mg2^{+}$ binding to the active site of EcoRV endonuclease: a crystallographic study of complexes with substrate and product DNA at 2 A resolution. *Biochemistry*, **34**, 683–696.
40. Declais,A.C., Hadden,J., Phillips,S.E. and Lilley,D.M. (2001) The active site of the junction-resolving enzyme T7 endonuclease I. *J. Mol. Biol.*, **307**, 1145–1158.
41. Komori,K., Sakae,S., Daiyasu,H., Toh,H., Morikawa,K., Shinagawa,H. and Ishino,Y. (2000) Mutational analysis of the *Pyrococcus furiosus* Holliday junction resolvase Hjc revealed functionally important residues for dimer formation, junction DNA binding and cleavage activities. *J. Biol. Chem.*, **275**, 40385–40391.
42. Bernstein,F.C., Koetzle,T.F., Williams,G.J., Meyer,E.E.,Jr, Brice,M.D., Rodgers,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.*, **112**, 535–542.
43. Huai,Q., Colandene,J.D., Topal,M.D. and Ke,H. (2001) Structure of NaeI-DNA complex reveals dual-mode DNA recognition and complete dimer rearrangement. *Nature Struct. Biol.*, **8**, 665–669.
44. Nicholls,A., Bharadwaj,R. and Honig,B. (1993) GRASP: Graphical representation and analysis of surface properties. *Biophys. J.*, **64**, 166–170.
45. Thorpe,J.H., Gale,B.C., Teixeira,S.C. and Cardin,C.J. (2003) Conformational and hydration effects of site-selective sodium, calcium and strontium ion binding to the DNA Holliday junction structure d(TCGGTACCGA)(4). *J. Mol. Biol.*, **327**, 97–109.
46. Hol,W.G., van Duijnen,P.T. and Berendsen,H.J. (1978) The alpha-helix dipole and the properties of proteins. *Nature*, **273**, 443–446.
47. Fogg,J.M., Kvaratskhelia,M., White,M.F. and Lilley,D.M. (2001) Distortion of DNA junctions imposed by the binding of resolving enzymes: a fluorescence study. *J. Mol. Biol.*, **313**, 751–764.
48. Declais,A.C. and Lilley,D.M. (2000) Extensive central disruption of a four-way junction on binding CCE1 resolving enzyme. *J. Mol. Biol.*, **296**, 421–433.
49. Linding,R., Russell,R.B., Neduva,V. and Gibson,T.J. (2003) GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic Acids Res.*, **31**, 3701–3708.
50. Kvaratskhelia,M., Wardleworth,B.N. and White,M.F. (2001) Multiple Holliday junction resolving enzyme activities in the Crenarchaeota and Euryarchaeota. *FEBS Lett.*, **491**, 243–246.
51. Skirgaila,R., Grazulis,S., Bozic,D., Huber,R. and Siksnys,V. (1998) Structure-based redesign of the catalytic/metal binding site of Cfr10I restriction endonuclease reveals importance of spatial rather than sequence conservation of active centre residues. *J. Mol. Biol.*, **279**, 473–481.
52. DeLano,W.L. (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA.