

# Structural insights into dynamics of RecU–HJ complex formation elucidates key role of NTR and stalk region toward formation of reactive state

Sagar Khavnekar<sup>1</sup>, Sarath Chandra Dantu<sup>2</sup>, Svetlana Sedelnikova<sup>3</sup>, Sylvia Ayora<sup>4</sup>, John Rafferty<sup>3,\*</sup> and Avinash Kale<sup>1,\*</sup>

<sup>1</sup>UM-DAE Centre for Excellence in Basic Science, University of Mumbai, Vidhyanagari Campus, Mumbai 400098, India, <sup>2</sup>Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Powai, IIT Bombay, Mumbai 400076, India, <sup>3</sup>The Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Western Bank, Sheffield S10 2TN, UK and <sup>4</sup>Department of Microbial Biotechnology, Centro Nacional de Biotecnología, CNB-CSIC, 28049 Madrid, Spain

Received December 17, 2015; Revised November 04, 2016; Editorial Decision November 07, 2016; Accepted November 09, 2016

## ABSTRACT

Holliday junction (HJ) resolving enzyme RecU is involved in DNA repair and recombination. We have determined the crystal structure of inactive mutant (D88N) of RecU from *Bacillus subtilis* in complex with a 12 base palindromic DNA fragment at a resolution of 3.2 Å. This structure shows the stalk region and the essential N-terminal region (NTR) previously unseen in our DNA unbound structure. The flexible nature of the NTR in solution was confirmed using SAXS. Thermofluor studies performed to assess the stability of RecU in complex with the arms of an HJ indicate that it confers stability. Further, we performed molecular dynamics (MD) simulations of wild type and an NTR deletion variant of RecU, with and without HJ. The NTR is observed to be highly flexible in simulations of the unbound RecU, in agreement with SAXS observations. These simulations revealed domain dynamics of RecU and their role in the formation of complex with HJ. The MD simulations also elucidate key roles of the NTR, stalk region, and breathing motion of RecU in the formation of the reactive state.

## INTRODUCTION

Cells have evolved efficient mechanisms to promote genetic diversity, ensure proper chromosome segregation and restore genome integrity via homologous recombination (1,2). This is an ubiquitous damage control process by which cells repair DNA and facilitate the restart of the stalled replication (1,3–8). This process may involve the formation of a four-way DNA crossover known as a Holliday junction (HJ) from the pairing of homologous duplex DNA.

DNA recombination has been extensively studied in Gram negative bacteria such as *Escherichia coli* in terms of the genetics and the molecular machinery (9). *In vivo* the HJ is formed by the regression of a stalled fork by the action of *E. coli* RecG and may be translocated through thousands of base pairs by the RuvAB translocase complex (10). Branch migration enables the positioning at the junction of HJ intermediates of target sequences for subsequent cleavage by the RuvC HJ resolving enzyme (HJR), allowing for the separation of paired duplex DNA and normal chromosome segregation at cell division (11).

Our understanding of how the various components interact to establish and complete the homologous recombination reaction is less clear in Gram-positive bacteria, as there are marked differences in some components of the process. RecU HJ resolving enzymes are present in all Firmicutes, but they are not present in Gram-positive bacteria with high dG+dC content. These bacteria have RuvC homologues. RecU is involved in DNA repair and recombination (3–8,10). Biochemical data have shown the interaction between the RecU enzyme and the RuvAB complex (12), which is in agreement with the genetics data which show that *recU* and *ruvAB* share common suppressors (6). However, there are other genetic studies that suggest that RecU may participate together with RecG in some resolution reactions (10).

Structural characterization studies of HJ resolvases have classified them into two principal folds (12), with the exception of *E. coli* RusA (13) and bacteriophage T4 endonuclease VII (14). The first group resembles restriction endonucleases and includes archaeal Hjc and Hje resolvases (15,16), as well as bacteriophage T7 endonuclease I (17) and the second group, resembles HIV integrase and includes *E.*

\*To whom correspondence should be addressed. Tel: +91 9167452231; Fax: +91 22 26524982; Email: avinash.kale@cbs.ac.in  
Correspondence may also be addressed to John Rafferty. Tel: +44 114 2222809; Fax: +44 114 2222800; Email: j.rafferty@sheffield.ac.uk

*coli* RuvC (18) and *Schizosaccharomyces pombe* Ydc2 (19). RecU belongs to the first group but has an additional central stalk region not seen in other family members (20).

We have previously reported the structure of RecU from *Bacillus subtilis* (20), which revealed the overall shape and details of the likely active sites in the dimer. RecU has a mushroom-like appearance with a cap (residues 34–55 and 90–199) and stalk sections (residues 56–89) (20,21). However, this structure lacked information on the N-terminal 1–33 residues (NTR) and the stalk region, which emerges from the centre of the dimer interface. The missing ‘mushroom stalk’ could be clearly seen in the structure of a homologous RecU from *Bacillus stearothermophilus* (PDB entry 1Y1O) but the NTR region remained elusive (22). Both regions are essential for RecU activity. Biochemical data showed that the stalk region of the RecU resolvase is essential for HJ recognition and distortion (4). It is also involved in interaction with RecA (23). Furthermore, once HJ is bound, RecU fails to modulate RecA activities, suggesting that both interactions are exclusive (4). Binding of  $\Delta$ NTR-RecU (deleted residues 1–32) to HJ is very unstable, and that is why this mutant does not cleave the HJ properly (24).

In this work, we report the first crystal structure of catalytically inactive RecU (RecU<sub>D88N</sub>) bound to a 12 base pair palindromic dsDNA fragment, with well-defined NTR and stalk regions. The crystal structure of this protein–dsDNA complex provides insights into the binding of the DNA to RecU and the organization of its phosphate backbone relative to the active site and the stalk region. We also present the solution structure analysis of unbound RecU using SAXS, which confirms the conformational variability of the NTR region elusive from the previously reported X-ray structures. A thermofluor-based assay was carried out to test the induced stabilization of the purified protein (RecU<sub>D88N</sub>) against HJs of varying lengths. To gain insights into the conformational properties of RecU and its complex, we have used Molecular Dynamics (MD) simulations of wild type RecU (RecU<sub>WT</sub>), wild type RecU in complex with HJ (RecU<sub>HJ</sub>),  $\Delta$ NTR mutant (residues 34–199) RecU (RecU <sub>$\Delta$ NTR</sub>) and  $\Delta$ NTR mutant RecU in complex with HJ (RecU <sub>$\Delta$ NTR-HJ</sub>). These simulations allowed us to propose how the conformational dynamics of mushroom domain with respect to the stalk region can be used to bind to the HJ. Further, NTR region might have a role in the stabilization of the reactive state during the reaction mechanism.

## MATERIALS AND METHODS

### Protein synthesis and crystallization

Catalytically inactive mutant of RecU<sub>D88N</sub>, which binds but doesn't cleave HJ, was purified as described previously (4). The 12 base palindromic DNA fragment (ACG-CAATTGCGT) was chemically synthesized (Yorkshire Bioscience, York, UK). The synthesized DNA was dissolved to make a stock solution of 1 mM in autoclaved water and was heated at 95°C for 5 min in a beaker of water where it was left to anneal overnight in an insulated flask. The purified protein was concentrated to 10 mg/ml (as estimated by Bradford assay) and RecU–DNA complex was formed by mixing the protein and the DNA to 1:1.25 molar ratio and incubating for 5 h at room temperature prior to

crystallization. Crystals were obtained in sitting drop trials after two weeks at 17°C in a condition composed of 0.1 M BICINE pH 9, 10% PEG 6000.

### Data collection and refinement

Diffraction data for the RecU–DNA complex crystals were measured with synchrotron radiation at the Diamond Light Source, Oxford, United Kingdom on I02 beamline. Before data collection, crystals were transferred into a cryo-protecting solution composed of 0.1 M BICINE pH 9, 10% PEG 6000, 30% ethylene glycol and were further incubated for 30 min at room temperature, before flash cooling in liquid nitrogen. The data were processed with XDS (25) and merged using Scala (26) as implemented in the CCP4 software suite (27). The relevant data statistics are provided in Table 1.

The crystal diffracted to 3.2 Å resolution and belonged to space group R32. Error in the coordinates (0.9 Å) was calculated over all the four molecules of the RecU present in the ASU. Solvent content of 56% was calculated using the program Matthews (28,29). An initial set of phases were obtained by the molecular replacement method using the program Phaser (30) within the PHENIX software suite (31) using the structure of *B. subtilis* RecU (20) (PDB id: 1ZP7) as a search model. Several iterations of manual building using the program COOT (32) were alternated with cartesian and real space refinement using the program Phenix. Refine (33,34). Secondary structure restraints were imposed during refinements owing to low resolution data. Atomic B-factors were refined and were treated as isotropic. NTR atoms were refined for partial occupancies. A model of the flexible NTR region was validated by calculating composite omit maps (35). To further support the model of the NTR, we calculated crystal contacts made by the NTR within the crystals using NCONT in the CCP4 software suite. The model quality was validated for all structures using COOT and MolProbity (36). The statistics of the refined structure are given in Table 1. Structure was analyzed using COOT and PyMol (37). PyMol, VMD (38) and Ligplot+ (39,40) were used to visualize and interpret the interactions and produce figures. The coordinates of this structure has been deposited in Protein Data Bank (PDB) under the accession id: 5FDK.

### SAXS measurements

SAXS data were acquired at the ID14-3 Bio-SAXS beamline at ESRF, Grenoble, France. The typical beam flux was  $2 \times 10^{12}$  photons/s and the size of beam at the detector was  $140 \times 80 \mu\text{m}^2$ . X-ray wavelength was selected using the monochromator to be 0.1 nm. For SAXS measurements, a CCD detector with a pixel size of  $63 \mu\text{m}/\text{pixel}$  and a total size of  $1344 \times 1024$  pixels was used to measure the scattered radiation. The detector was placed behind a vacuum path and the camera length was 1.7 m. The data from SAXS measurements were processed and analyzed using ATSAS (41). The data were scaled averaged and merged using PRIMUS (42). Analysis of the data was carried out with GNOM (43). Ensemble optimized modeling (EOM) was performed for the flexible N-terminal region (33 residues) with EOM 2.0 (44,45). For the EOM, the structure of a

**Table 1.** Data statistics for processing and refinement of diffraction data are shown

Data statistics	
Wavelength (Å)	0.9
Resolution range (Å)	48.7–3.2 (3.3–3.2)
Space group	R 3 2:H
Unit cell	144.4 144.4 310.3 90 90 120
Total reflections	226725 (22642)
Unique reflections	20730 (2042)
Multiplicity	10.9 (11.1)
Completeness (%)	99.8 (99.2)
Mean $I/\sigma(I)$	19.9 (1.6)
Wilson $B$ -factor	118.3
$R$ -merge	0.09081 (1.596)
$R$ -meas	0.0954
<b>Refinement statistics</b>	
$R$ -work	0.2561 (0.3267)
$R$ -free	0.3286 (0.3908)
RMS (bonds)	0.012
RMS (angles)	2.10
Ramachandran favored (%)	84
Ramachandran generously allowed (%)	14.4
Ramachandran outliers (%)	1.6
Clash score	29.4
Molprobability	2.8

Data in parentheses correspond to the highest resolution shell.

(1)  $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$ , where  $I$  is the observed intensity and  $\langle I \rangle$  the average intensity of reflections.

(2)  $R$ -factor =  $\sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$ .

Data in parentheses correspond to the highest resolution shell. The table was generated using the Phenix table utility.

RecU dimer without the 33-residue NTR was given as an input and was treated as the core rigid structure.  $C_{\alpha}$  atoms of these 33 residues were modeled based on the SAXS data with P2 symmetry restrictions.

### Thermofluor assay

All the DNA constructs were ordered from Eurofins, Bangalore, India. Different forms of duplexes and HJs were prepared by combining different mixtures of oligonucleotides (Supplementary Table ST1 and Figure S1) at 100°C in autoclaved nuclease free water to ensure melting of base pairing and then allowing a mixture to anneal overnight in a thermo-jacketed flask prefilled with boiling water. The formation of higher oligomeric species (duplex, or four-way junctions) for these constructs were qualitatively checked using a 2% agarose gel (Supplementary Figure S2). Different DNA constructs and the purified RecU<sub>D88N</sub> were mixed in 1:1 molar ratio. The concentration of RecU<sub>D88N</sub> was 5 µg per reaction well in a purification buffer (100 mM MES pH 6.5). Thermofluor assays were carried out using Sypro orange dye (Invitrogen, Bangalore, India) with a BIO-RAD CFX96 Touch RT PCR machine with a linear temperature gradient from 25°C to 100°C using 1°C step increments. The data, averaged over four experiments, were analyzed using Origin (OriginLab, Northampton, MA, USA) where the curves were smoothed using five points averaging and melting points ( $T_m$ ) were obtained by taking the first derivative of the respective curves. To analyze the stabilization of RecU on binding to HJ we calculated the conformational entropy for control (RecU<sub>D88N</sub>) and RecU<sub>D88N</sub>+HJ24 as follows.

$$\Delta H - T\Delta S = -RT \ln K_{\text{eq}}$$

$$\ln K_{\text{eq}} = \left( \frac{-\Delta H}{R} \right) \frac{1}{T} + \frac{\Delta S}{R}$$

$$\ln K_{\text{eq}} = \frac{[\text{unfolded}]}{[\text{folded}]}$$

$$\Delta H = -R \times \text{slope of } \ln K_{\text{eq}} \text{ versus } 1/T \text{ graph}$$

$$\text{Conformational entropy} = \Delta S = \frac{\Delta H}{T_m}$$

### Molecular dynamics simulations

**Modeling of the starting structure.** To create the starting structure of RecU with the HJ for the molecular dynamics (MD) simulations we initially created the RecU dimer. Of the four polypeptide chains in the asymmetric unit (A–D), chain D was modeled completely and we used it as the template to create the RecU dimer.  $\text{Mg}^{2+}$  ions were added based on their locations in the structure of *B. stearothermophilus* RecU (PDB entry: 2FCO) (21). All the side chains were included in the model. An HJ model was created in COOT using the two DNA duplexes seen in the protein–DNA structure presented here. A copy of the two duplexes present in the X-ray structure was rotated by 90° around the two fold symmetry axis of the RecU dimer. The four duplexes (two originals and two copies) were now treated as the four arms of an HJ and the duplex copies were then rotated around their helical axes to match the 5' and 3' ends of the strands in each adjacent arms. The ends were ligated at the point of crossover. The N-terminal region was used as a guide to dock the two rotated arms (see Supplementary Figure S3 for more details). Finally, any steric clashes were removed with rigid body refinement. A starting structure for the  $\Delta$ NTR



variant simulations was created by deleting residues 1–33 of the NTR region from each monomer.

### Simulation setup

Using RecU dimer and the modeled HJ, we created the starting structures for the MD simulations of wild type RecU (RecU<sub>WT</sub>), wild type RecU in complex with HJ (RecU<sub>HJ</sub>),  $\Delta$ NTR variant RecU (RecU <sub>$\Delta$ NTR</sub>), and  $\Delta$ NTR variant RecU in complex with HJ (RecU <sub>$\Delta$ NTR-HJ</sub>). All our simulations were performed using the GROMACS package version 4.6.4 (46–48). Amber14sb (49) force field for the protein and parmbsc0 forcefield (50) for the DNA were used in all our simulations. TIP3P water model (51) was used and ions were described using parameters from Joung et al. (52). Each protein system was placed in a dodecahedron box with a minimum distance of 1 nm to the box walls. Every simulation box was solvated and the system was neutralized by adding 150 mM NaCl ions. Periodic boundary conditions were applied in all the directions. Each system was energy minimized using a steepest descent algorithm until the largest force acting on the system was smaller than 1000 kJ/(mol nm). The final structure from energy minimization was subjected to temperature equilibration using the Berendsen thermostat (53) to 298 K using a tau-t of 0.1 ps in 100 ps, followed by isotropic pressure equilibration to 1 atm using the Berendsen barostat (54) in 1 ns using a tau-p of 1 ps. During both the equilibration stages, main chain atoms were position restrained using a force constant of 1000 kJ/(mol nm<sup>2</sup>) in all directions. The final pressure equilibrated structure was used for production run simulation during which the temperature was maintained at 298 K using a velocity re-scaling thermostat (55) (tau-t: 1 ps) and Parrinello-Rahman barostat (56) (tau-p: 2 ps). Parallel LINCS (57,58) algorithm was used to constrain all bonds and electrostatics were treated using a Particle mesh Ewald scheme (59,60). Cut-off of 1 nm was used for electrostatics and a short range van der Waals cut-off of 1 nm was used. For each system, three independent 250 ns long production run simulations were carried out. Snapshots were saved at every 40 ps of the respective trajectories. Analysis tools within GROMACS and in-house Python code was used to analyze MD trajectories. Pymol and VMD were used to visualize the trajectories and render the images. All our simulation runs were carried out using the PARAM Yuva-II supercomputer at the National Param Supercomputing Facility (NPSF) at CDAC, Pune, India.

## RESULTS

### Structure of the RecU–DNA complex

Currently there are no structural data available of RecU in complex with HJ. Our attempts to crystallize RecU in complex with HJ (variable lengths/sequences/forms) were unsuccessful. Therefore to gain insights into RecU DNA interactions, we crystallized an inactive mutant RecU<sub>D88N</sub> in complex with a palindromic 12 bp DNA. The asymmetric unit (ASU) of the crystal structure presented here consists of four RecU monomers (two dimers) and two dsDNA molecules (Figure 1). The two protein dimers are constituted by chains A plus C and B plus D and the two DNA

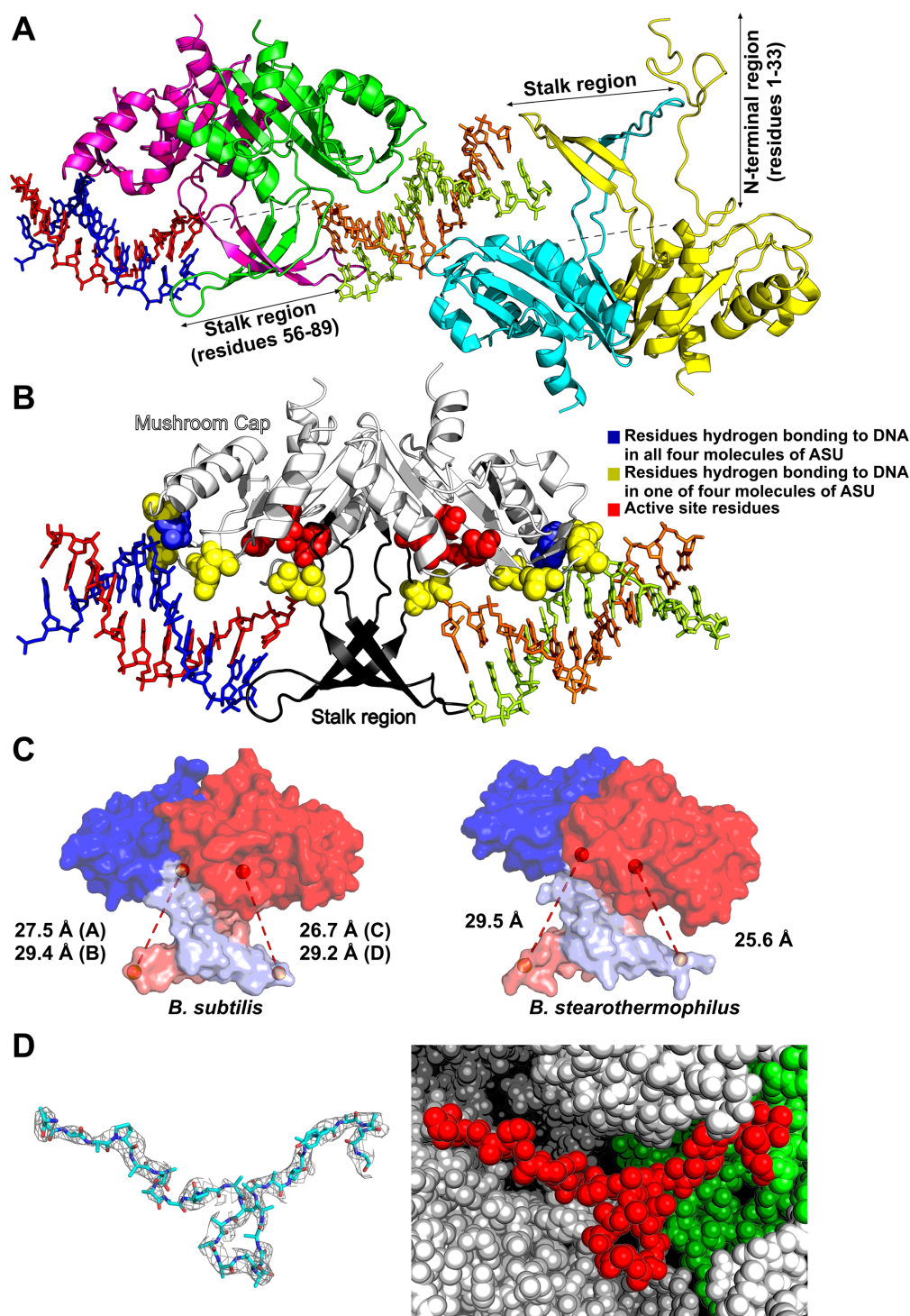
duplexes are constituted by chains E plus F and G plus H. Each RecU monomer displays an  $\alpha/\beta$  architecture and the structure of the protein is identical to the previous structure of unbound wildtype RecU (21) with no indications at this resolution of any kind of large conformational change. The backbone Root Mean Square Deviation (RMSD) between the previously reported structure (PDB ID: 1ZP7) and both the A/C and B/D dimers reported here is 0.95 Å. The complete stalk region of RecU is visible in both dimers within the ASU of our complex structure. The 33 residue long NTR could be completely modeled for chain D. For chains A and C, the NTR could be modeled from residue 23 and 26 onwards, respectively, whereas for chain B we could not model any of the NTR residues. The electron density and crystal environment for this 33 amino acid long region in chain D is shown in Figure 1D. Density for this region in chain-D is also seen in the composite omit map (see Supplementary Figure S4). NTR modeled in chain D shows a total of 76 crystal contacts as calculated using a cut-off distance of 5 Å.

### RecU–DNA interactions

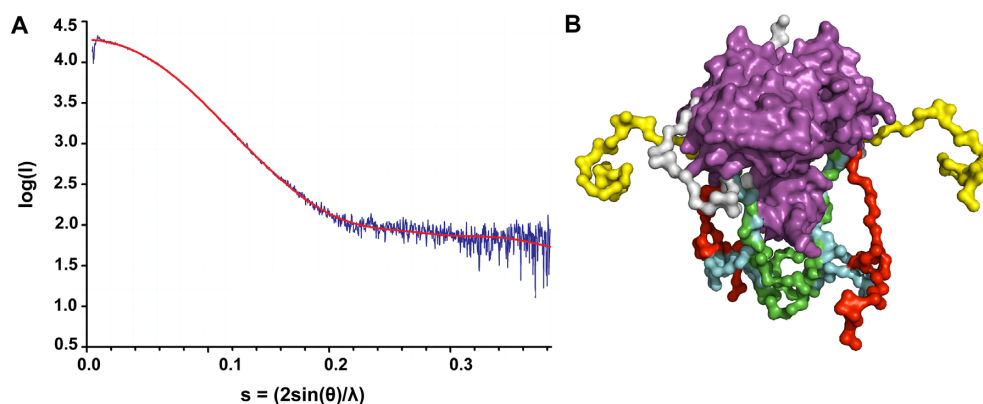
The E/F and G/H DNA duplexes flank the stalk region of the A/C dimer (see Figure 1). Duplexes fit along the concave surface of the mushroom cap of each monomer, which also contains the active site pockets. The E/F DNA duplex is sandwiched between the A/C and B/D dimers. The G/H DNA duplex is sandwiched between A/C and a symmetry related ( $X+1/3$ ,  $Y+2/3$ ,  $Z+2/3$ ) B/D dimer.

Analysis of our complex structure has revealed interactions between the protein and the DNA (Figure 1B). DNA is observed to be bound through a combination of hydrogen bonds (Supplementary table ST2) and long range interactions. A total of 14 protein–DNA hydrogen bonds can be proposed in the existing structure. This number is likely to be a lower estimate of the total number because many side chains could not be modeled fully. However, even at the low resolution, two hydrogen bonds are observed in all four monomers of the ASU: between the N atom of K165 and O1P of the nucleotide backbone; and between the OG atom of S166 and O2P of the nucleotide backbone. The number of hydrogen bonds with DNA per chain A, B, C and D are 5, 4, 2 and 3, respectively.

Apart from these hydrogen bonding interactions, long range interactions were also calculated using PDBSum (61). We could observe a total of 174 long range interactions (cut-off distance of 5.0 Å) between protein and the DNA. Of these, chain A has 49, chain B has 35, chain C has 30 and chain D has 60 non-bonded contacts with DNA. Though the density for the side chains of F81 and Y68 is not seen, they are well positioned to stabilize the approaching DNA by stacking against the nucleobases at the point of HJ crossover. The phosphate atom of the nearest base of the DNA is at a distance of  $\sim 11$  Å from the C $\alpha$  atom of mutated active site residue N88D. Further, in the current study, we could not identify the position of Mg<sup>2+</sup> atoms in the active site.



**Figure 1.** Crystal structure of inactive mutant of RecU (RecU<sub>D88N</sub>)-DNA complex. (A) Crystallographic asymmetric unit (ASU) in crystal structure of RecU-DNA complex consists of four RecU monomers forming two dimers bound to two DNA duplexes. Cartoon representation of protein Chains A, B, C and D are shown in green, cyan, pink and yellow respectively. Chain E, F, G and H of DNA are shown in orange, limon, blue and red respectively. (B) Interactions of RecU with DNA duplexes: Cap and stalk regions are shown in white and black, respectively. Hydrogen bonding and active site residues are shown as van der Waals spheres, while the DNA is shown as stick. Residues which form hydrogen bonds in all four monomers of RecU in ASU are shown in blue. Residues that form hydrogen bonds in only some of monomers are shown in yellow. Active site residues are shown in Red. (C) Distance between C $\alpha$  atoms of residues R71 of one monomer and D88 of its dimerising partner (binding pocket) for both AC and BD dimer are shown for the structure presented here and for the structure of *B. stearothermophilus*. (D) Electron density ( $2F_o - F_c$  map) is shown at sigma level of 0.7 for the 33 residue long NTR region of chain D. the sphere representation is shown for the NTR coloured red. The ASU is coloured green while symmetry mates forming the crystal contacts are coloured white.



**Figure 2.** SAXS analysis of DNA free RecU (D88N). (A) Fit of Ensemble Optimized Modeling (shown in red) of SAXS data to the experimental SAXS data (shown in blue). The  $\chi^2$ -value of the fit is 1.89. (B) Ensemble of conformations of NTR generated from EOM analysis are shown in different colors. The core structure of the RecU enzyme is shown in magenta. Completely extended conformation of NTR are shown in yellow and red, partially collapsed conformation in green and cyan, and completely collapsed conformation of NTR is shown in white.

### Flexible stalk

The stalk consists of a  $\beta$ -hairpin of class 12:12 in each monomer. It comprises a terminal loop (residues 66–77) and two anti-parallel beta sheets (residues 61–65 and 78–81). The  $\beta$ -hairpins from the monomers cross over and extend toward the DNA duplex bound to the body of the other monomer in the dimer (Figure 1A). Hence, the DNA binding pocket consists of a mushroom cap of one monomer and a stalk of a dimerizing partner. Turn region of the  $\beta$ -hairpin consist of residues K70 and R71. This turn region is positioned symmetrically with respect to the mushroom cap such that the distance measured between C $\alpha$  atoms of residues R71 of one monomer and N88 of its dimerising partner (binding pocket) is the same for both pockets of the dimer. This distance is 26 Å for the AC dimer and 29 Å for the BD dimer.

We compared this arrangement of stalk region with that of the previously reported DNA-unbound structure of *B. stearothermophilus* RecU. Its sequence has 57.6% identity and 73.3% similarity with that of *B. subtilis* RecU as observed from global alignment performed using Needle<sup>51</sup>. Further, the stalk region of RecU in both these organisms is completely conserved. In contrast to our structure where we observe a symmetric arrangement, distances between C $\alpha$  atoms of corresponding residues R72 of one monomer and D89 of its dimerising partner were observed to be 29.5 Å and 25.6 Å, indicating an asymmetric arrangement of stalk region with respect to the mushroom cap (Figure 1C).

### SAXS

The poor electron densities of the NTR regions in the X-ray data imply that this region might be conformationally plastic. So, we performed SAXS experiments on the RecU<sub>D88N</sub> in solution. EOM modeled data could be fitted onto the experimental SAXS data with a  $\chi^2$  value of 1.89 (Figure 2A). C $\alpha$  models were generated for 33 residues of NTR. Models with the NTR folded into the protein core structure were rejected. Figure 2B shows multiple possible conformations of the NTR as generated from EOM

### Thermofluor assay

Thermofluor assays were used to assess the stability of RecU in complex with HJ substrates of various arm lengths. Thermofluor data (Supplementary Figure S5) was analysed to calculate the  $T_m$  values for RecU<sub>D88N</sub> and its complexes with duplex (same as in the structure presented here), HJs with oligomer lengths 9, 10, 11 and 12 bp (Figure 3A). Difference in  $T_m$  value due to the binding of duplex is 0.75°C with respect unbound RecU<sub>D88N</sub>. For junctions of lengths 9, 10, 11 and 12 bp the change is –0.5, 6.7, 8.2 and 10.1°C, respectively. Further, the unfolding fraction as a function of the temperature (Figure 3B) was derived from thermofluor data for the RecU<sub>D88N</sub> and RecU<sub>D88N</sub>-HJ complex (oligomeric length of 12 bp). Based on this we calculated the configurational entropy ( $\Delta S$ ) (Figure 3C). It is seen that on interaction of RecU<sub>D88N</sub> with HJ the  $\Delta S$  is reduced indicating the stabilization of RecU in the conformational space.

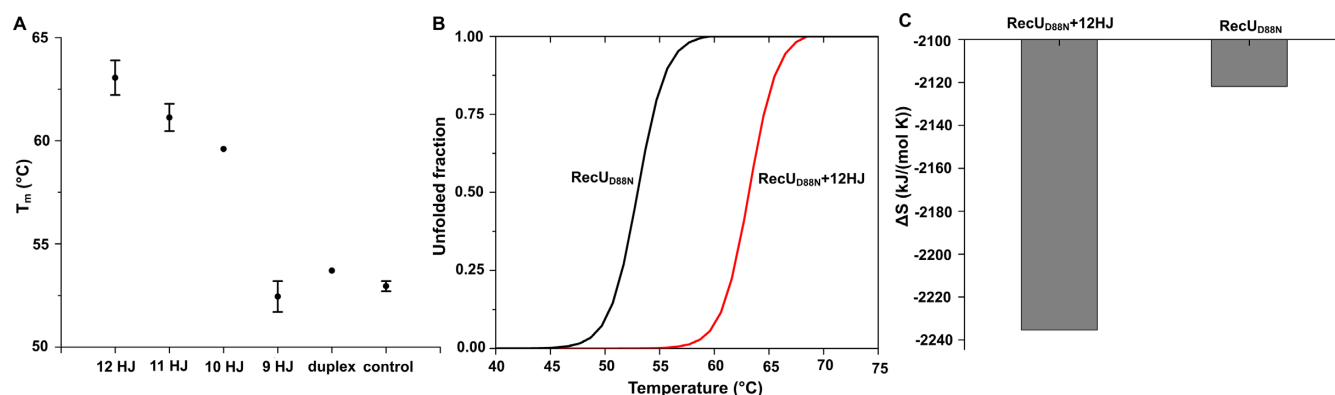
### Molecular dynamics

Our experimental data from X-ray and SAXS provided a picture of a conformationally flexible NTR region and the Thermofluor assays showed that on binding to HJ, stability of RecU<sub>D88N</sub> is increased. To obtain further insights into the solution dynamics of RecU in unbound and HJ bound complex, we carried out MD simulations for RecU<sub>WT</sub>, RecU<sub>HJ</sub>, RecU $\Delta$ NTR and RecU $\Delta$ NTR-HJ. For each protein system three independent simulations, each 250 ns long, with different starting velocities were performed.

### Protein flexibility

To understand the flexibility of the protein and the effect of binding of HJ on RecU, we calculated the root mean square fluctuation (RMSF) on a per residue basis (Supplementary Figure S6) using only the C $\alpha$  atoms. This analysis shows that the stalk and NTR regions are more flexible as compared to the rest of the protein in all the four systems, with the latter being much more flexible. The average RMSF and the standard error values calculated over the entire stalk re-





**Figure 3.** Thermofluor studies. (A) Thermofluor experiments were carried out to study the induced stabilization of RecU on its binding to different lengths of HJ and DNA duplex. Holliday junctions formed by the oligomers of lengths 9 bp, 10 bp, 11 bp, and 12 bp were used. DNA duplex used is same as that of the crystal structure presented here. RecU in the absence of any DNA was used as control. Thermofluor experiments for each of these systems were carried out in quadruplets.  $T_m$  was calculated from first derivatives of fluorescence intensities with respect to temperature. The scatter plot gives a clear picture of stabilization of RecU on HJ binding. (B) Unfolded fractions as a function of temperature for RecU<sub>D88N</sub> and its complex with HJ of an arm length of 12 bp. (C) Conformational entropy calculated from the profiles in (B).

gion residues (NTR region was excluded from this calculation because of its high flexibility) in RecU<sub>WT</sub> ( $0.12 \pm 0.009$  nm) and RecU<sub>HJ</sub> ( $0.07 \pm 0.003$  nm), suggest that on binding to HJ, stalk region becomes less flexible.

#### Comparison of RecU<sub>WT</sub> MD simulation ensemble with SAXS data

To check whether the structures sampled in the MD simulations are identical to the solution SAXS profile, we clustered structures from the RecU<sub>WT</sub> ensemble and clustered the structures based on the backbone RMSD of the NTR region. We used the Jarvis–Patrick algorithm (62) implemented in the g\_cluster tool of GROMACS with a backbone RMSD cut-off of 0.2 nm and a minimum number of 10 neighbors required to form a cluster. Simulated SAXS profiles were generated using the WAXSIS server (63) and representative structures of the top 10 populated clusters. Supplementary Table ST3 summarizes the results of the simulated SAXS profile calculation. Radius of gyration ( $R_g$ ) values from the RecU<sub>WT</sub> simulation (25 Å) are in good agreement with the experimentally calculated value of 24.3 Å. The average chi value for the fit between experimental data and simulated SAXS profiles is 0.4 on the log scale.

#### Principal component analysis (PCA)

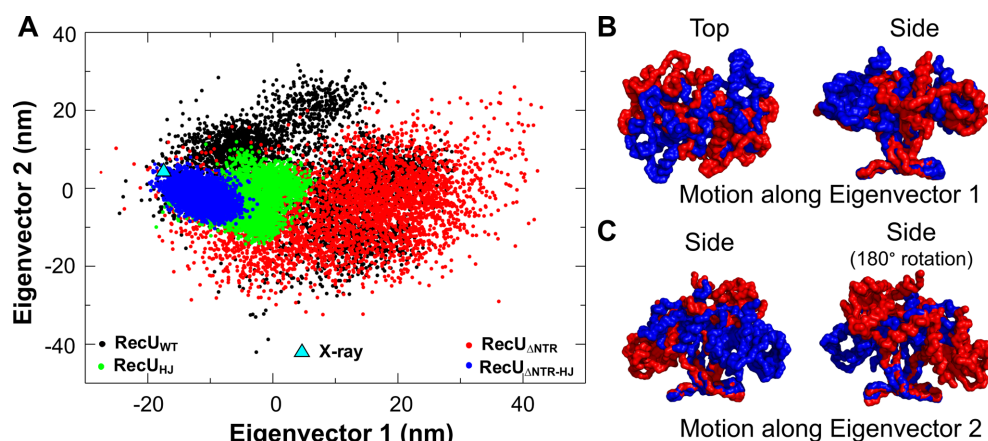
To analyze domain dynamics of RecU, using the g\_covar tool from GROMACS package, we extracted the eigenvectors from the three concatenated trajectories of RecU<sub>WT</sub> using only the backbone atoms. RecU<sub>HJ</sub>, RecU<sub>ΔNTR</sub>, and RecU<sub>ΔNTR-HJ</sub> simulations were projected onto the eigenvectors obtained from the RecU<sub>WT</sub> simulations (Figure 4A). The first five eigenvectors account for 95% of the protein motion, with the first and second eigenvectors (Ev1 and Ev2) accounting for 43% and 41%, respectively. The third and fourth eigenvectors account for 8% and 3% of the protein motion, and were not analyzed further. Ev1 captured the rotation of the mushroom cap domain with respect to the stalk region (Figure 4B and Movie M1). Ev2 described

the rocking motion of the mushroom cap relative to the stalk (Figure 4C and Movie M2). This results in an alternate opening and closing of the two active sites. In the open state, the maximum distance between the Cα atoms of residue R70 (tip of the stalk region) and D88 (active site) is ~32 Å, while in the closed state the minimum distance is ~18 Å (Supplementary Figure S7–S9). RecU<sub>HJ</sub> and RecU<sub>ΔNTR-HJ</sub> show restriction of domain motion along both Ev1 and Ev2. On closer inspection, projection data shows conformational selection of domain motion of RecU on its interaction with HJ, i.e. conformations which are sampled by both RecU<sub>HJ</sub> and RecU<sub>ΔNTR-HJ</sub> are scarcely sampled by RecU<sub>WT</sub> and RecU<sub>ΔNTR</sub> in the absence of HJ.

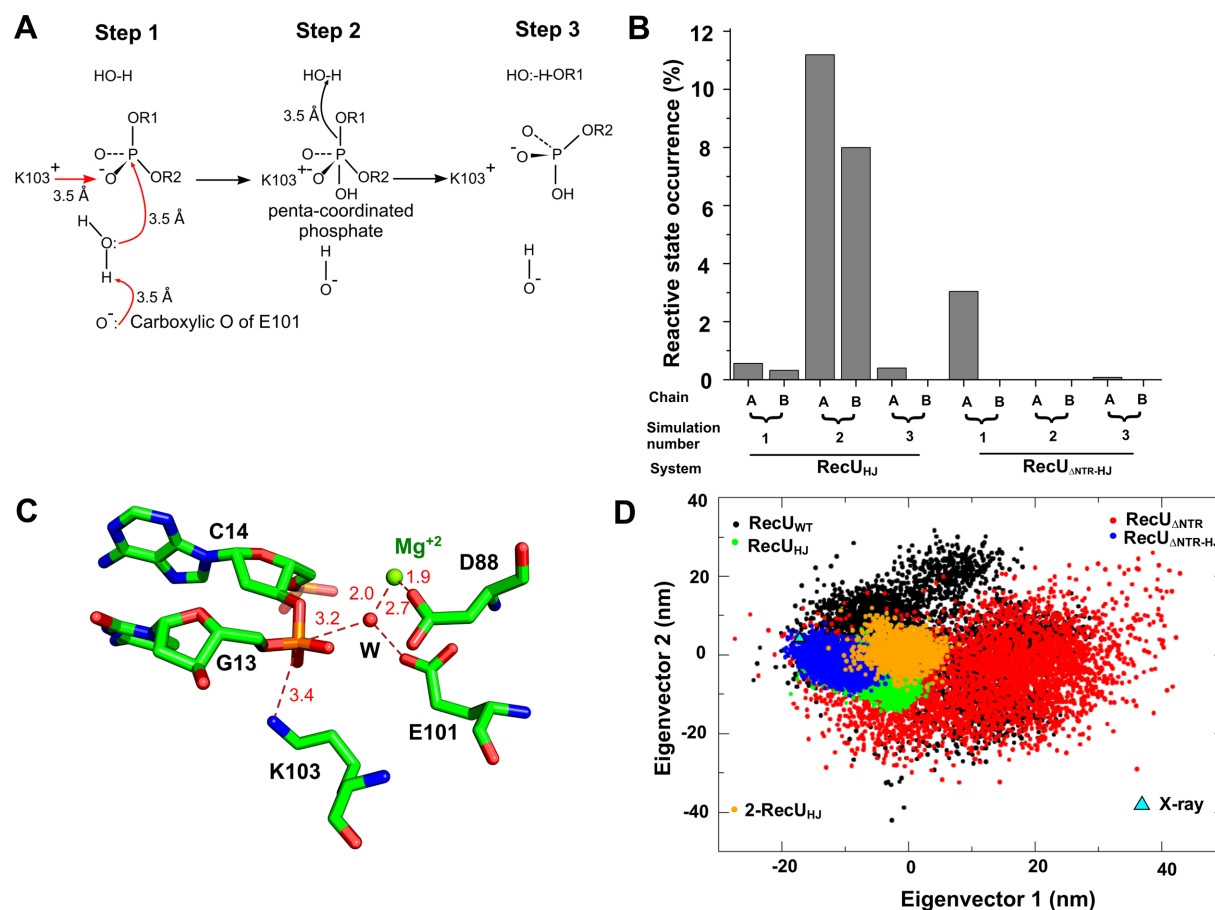
#### Active site

A possible mechanism of phosphodiester bond cleavage has been proposed earlier (64) (see Figure 5A) and is as follows: Step 1: E101 deprotonates one of the coordinating waters of an Mg<sup>2+</sup> ion. Step 2: This water then carries out a nucleophilic attack onto the phosphodiester bond, resulting in a pentavalent phosphate. This pentavalent phosphate is stabilized by K103. Step 3: A second water would act as a general acid via proton donation to the leaving anion which leads to the breakage of the phosphodiester bond.

To check whether in our simulations of RecU<sub>HJ</sub> and RecU<sub>ΔNTR-HJ</sub>, the active site of the protein adopts the reactive conformation, we counted the number of instances when the reactive conformation is sampled. We defined reactive conformation as a state when a water molecule is coordinating with Mg<sup>2+</sup> and the same water molecule is within 3.3 Å of E101 (carboxylate oxygens) and within 3.3 Å of phosphate in the DNA. Figure 5B shows the percentage of simulation time that the reactive conformation is sampled in each chain for both RecU<sub>HJ</sub> and RecU<sub>ΔNTR-HJ</sub>. This percentage ranges from 0.3% to 11% in RecU<sub>HJ</sub> simulations and 0.08–3% in RecU<sub>ΔNTR-HJ</sub> simulations. Figure 5C shows molecular details of one such occurrence of the reactive state. It must be noted that the statistics of reactive state occurrence have not converged across the simulations



**Figure 4.** Principal Component analysis. (A) Projection of wild type RecU (RecU<sub>WT</sub>: shown in black),  $\Delta$ NTR mutant RecU (RecU <sub>$\Delta$ NTR</sub>: shown in red) and their complexes with HJ i.e. (RecU<sub>HJ</sub>: shown in green), (RecU <sub>$\Delta$ NTR-HJ</sub>: shown in blue) along the first two eigenvectors. Sampling along both these eigenvectors is restricted in the presence of HJ. (B) Motion along eigenvector 1 showing rotation of mushroom cap with respect to the stalk region. Two extreme structures along the direction of the eigenvector are shown in red and blue respectively in a surface representation of a backbone trace. (C) Motion along eigenvector 2 showing rocking of mushroom cap with respect to the stalk region.



**Figure 5.** Mechanism of cleavage of phosphodiester bond at the RecU active site. (A) Plausible mechanism of DNA cleavage by RecU based on the mechanism reported earlier (64).  $Mg^{+2}$  will cause polarization of water molecules around it. Carboxylic oxygen of E101 (one that is not coordinated with  $Mg^{+2}$ ) can deprotonate a water coordinating with  $Mg^{+2}$ . The deprotonated water will then attack the phosphate giving rise to penta-coordinated phosphate intermediate that is stabilized by Lewis acid K103. Another  $Mg^{+2}$  coordinated water can then act as general acid by protonating the leaving anion. (B) Percentage occurrence of step 1 of the above mechanism observed in the simulation of RecU in complex with holliday junction (HJ) (RecU<sub>HJ</sub>) and  $\Delta$ NTR mutant in complex with HJ (RecU <sub>$\Delta$ NTR-HJ</sub>). (C) Atomic model of step 1 of proposed mechanism as observed in simulation (one such snapshot). DNA and active site residues are shown in stick representation. Water and  $Mg^{+2}$  are shown as spheres and are coloured in red and green respectively. (D) Projection of simulation-2 of RecU<sub>HJ</sub> (for which the occurrence of step 1 of the proposed mechanism is highest) along eigenvector 1 and 2. RecU<sub>WT</sub>: shown in black; RecU <sub>$\Delta$ NTR</sub>: shown in red; RecU<sub>HJ</sub>: shown in green; RecU <sub>$\Delta$ NTR-HJ</sub>: shown in blue; simulation 2 of RecU<sub>HJ</sub>: shown in orange; projection of X-ray structure is shown as a triangle.



of RecU<sub>HJ</sub> and RecU<sub>ΔNTR-HJ</sub> and are require further investigation via longer simulations. Since the second simulation of the RecU<sub>HJ</sub> complex sampled the most number of reactive state conformations (Figure 5B), we checked the projection of this particular simulation along Ev1 and Ev2. Simulations of RecU<sub>ΔNTR-HJ</sub> do not sample the same conformational space sampled by the second simulation of RecU<sub>HJ</sub> complex, suggesting that this simulation might have accessed the conformational space required for the reactive state formation. No such exclusive sampling was observed along Ev3 and Ev4 indicating the importance of domain dynamics observed along Ev1 and Ev2 in the functioning of RecU.

To identify the motion of RecU that results in the formation of step 1 of the active site, we performed Functional Mode Analysis (FMA) (65) on the second simulation of RecU<sub>HJ</sub>. FMA can be used to isolate the collective motions of a protein that describe an observable of interest, in this case, the formation of the reactive state in the active site. We defined reactive state as per that of Figure 5A, i.e. if there is a water (within 3.3 Å. of Mg<sup>2+</sup> and (OE1 or OE2 of E101)) and (within 3.3 Å. of phosphate atom from the DNA nucleotide). We used the first 25 eigenvectors from the PCA as they capture ~99% motion of the protein. FMA was performed independently for chain A and B, as the occurrence of step 1 is independent in each chain with correlation coefficient of 0.035. It is observed that the formation of step 1 results from the breathing motion of the mushroom cap.

## DISCUSSION

HJ resolvases are specific nucleases which nick the four arms of an HJ at the end of the recombination process. RecU is structurally different from the other HJ resolvases in the family, because of the presence of a central stalk region (20). Though previous biochemical reports elucidated the role of this stalk region (4) and a 33-residue NTR (24) in the activity of RecU, the structural information on a RecU–HJ complex and its dynamics has been limited. Even though a number of resolvase–HJ complex structures are known (11), no HJ complexed structures have been solved from RecU.

Here, we present the first structure of a RecU complex with a bound 12 base palindromic dsDNA fragment, a visible NTR and stalk regions determined at 3.2 Å resolution. Although the low resolution limits the information on the molecular details of interface interactions, we are still clearly able to determine the positioning of the dsDNA. This structure with DNA allowed us to model a bound HJ, was used to carry out MD simulations and revolutionize our understanding of the conformational dynamics of RecU–HJ complex formation. Based on the crystal structure of the RecU–DNA complex, the binding of HJ to RecU could be driven by hydrogen bonds as well as non-bonding interactions. Two hydrogen bonds between K165 and S166 of RecU and the phosphate oxygens of DNA are observed in all four chains. Further, we also observed that F81 is in a position to form a flat hydrophobic interface for binding of HJ, similar to that seen in RuvC (66). Mutational analysis show that a RecU–F81A mutant poorly recognizes and distorts HJs (4).

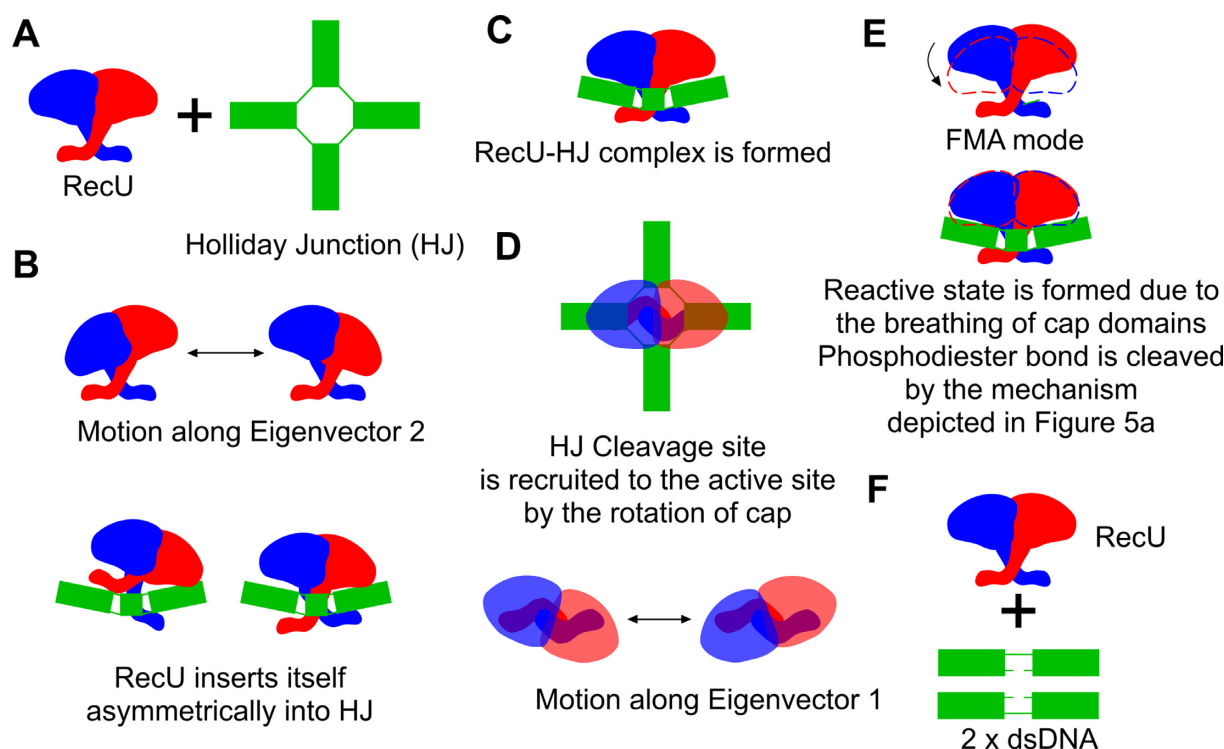
The current crystal structure gives the molecular detail of the stalk region which was unobserved in the electron density maps of the previously reported crystal structure of *B. subtilis* RecU (20). The structure of a RecU homologue from *B. stearothermophilus* (22) shows a stalk region structure stabilized by crystal contacts, whereas in the crystal structure described here, it is stabilized by its interaction with DNA. Further, the arrangement of the main body or mushroom cap region of the enzyme with respect to the stalk region is asymmetric in the structure of *B. stearothermophilus* RecU, unlike the symmetric arrangement in the crystal structure of RecU–DNA complex presented here. This supports the notion that the mushroom cap and stalk region have conformational flexibility. RMSF analysis of our MD simulations confirm the flexibility of the stalk region and when RecU is bound to HJ, there is a decrease in flexibility.

Restricted sampling along Ev1 and Ev2 is observed when RecU is bound to HJ. This is in agreement with the thermofluor studies on complexes of RecU with HJ of varying arm lengths which showed effective stabilization of RecU on binding to HJ. Furthermore, the conformational entropy for RecU alone is more than that of its complex with HJ, indicating restrictive sampling of conformational subspace of RecU upon with HJ. Our results are in agreement with the observations from earlier biochemical studies (4) that point mutation in the stalk region results in decreased RecU cleavage activity, as the stalk region in coordination with mushroom cap is required for the domain dynamics.

Similar to this stalk region, the 33 residue long NTR has been shown to be important for the binding of RecU to the HJ (24). It has been unobserved in previous X-ray structures of RecU (20,64) owing to its high flexibility as also noted in our SAXS analysis. In the crystal structure reported here, this region could be modeled completely in one of four chains (chain D) and could be partially modeled in two other chains. In chain D, the presence of the NTR in the electron density is the result of crystal contacts and we confirmed this using composite omit maps and NCONT analysis.

Conformations of this NTR calculated based on the EOM indicates that this region can exist in multiple conformations in solution. It should also be noted that the conformation of NTR as observed in the chain D of crystal structure is not seen in the NTR ensemble generated by EOM. MD simulations with the NTR region (RecU<sub>WT</sub> and RecU<sub>HJ</sub>) were started from the chain-D NTR conformation. Specifically, the conformations of RecU<sub>WT</sub> simulations can be compared with the RecU<sub>D88N</sub> SAXS data and the agreement between them implies that the MD simulations started from X-ray structure have moved into the conformational space in solution.

Earlier studies have further reported that the NTR deletion mutants have low activity (24). MD simulations of RecU<sub>ΔNTR-HJ</sub> complex allowed us to understand this effect of deletion of the NTR on the activity of RecU by checking the occurrence rate of a reactive conformation in the active site. A RecU<sub>ΔNTR-HJ</sub> complex samples the reactive state of the active site only 3% of the time in contrast to 20% of the time in a RecU<sub>HJ</sub> complex ensemble. The deleterious effect of a ΔNTR mutant is also evident from the projection



**Figure 6.** Plausible mechanism of binding and insertion of RecU in HJ. (A) RecU and a Holliday junction (HJ). (B) RecU in the unbound form exhibits rocking motion as described by the Ev2. With this domain motion RecU inserts itself in the HJ. (C) RecU-HJ complex is formed. (D) RecU then searches for the cleavage site by the rotation of the cap domain (motion observed along eigenvector Ev1). (E) Reactive state in the active site is then formed due to the breathing motion of mushroom cap. This breathing motion was observed in the functional mode analysis on the occurrence of the reactive state. Once reactive state is formed, phosphodiester bond cleavage occurs. (F) Cleaved HJ then diffuses from the RecU.

of MD trajectories along Ev1 and Ev2. The RecU<sub>ΔNTR</sub>-HJ complex scarcely samples in the region sampled by a simulation (simulation-2) RecU<sub>HJ</sub>, which the highest occurrence of reactive states in the active site. This is in agreement with experimental observations that a  $\Delta$ NTR mutant of RecU shows less activity as compared to RecU<sub>WT</sub> (24).

Interestingly, the conformational sampling seen in simulation-2 of RecU<sub>HJ</sub> is exclusive and the same conformational space is scarcely sampled by simulation-1 and simulation-3 of RecU<sub>HJ</sub>, suggesting conformational selection in the formation of reactive state (i.e. step 1 in the proposed mechanism). Formation of the reactive state in each active site is independent of the second active site in agreement with earlier reports on RuvC supporting sequential mode of cleavage (67). As MD simulations do not support bond breaking or making, the effect of the formation of a reactive state in one active site, on the cleavage in the other active site could not be studied. As deduced from FMA, the breathing motion of RecU mushroom cap facilitates in the formation of the reactive state. Though this motion is symmetric, reactive state formation in one active site is independent of the other.

Based on asymmetric and symmetric X-ray structures of mushroom cap with respect to the stalk region from *B. stearothermophilus* and *B. subtilis* respectively, SAXS analysis, thermofluor data and extensive MD simulations we propose the following mechanism of binding of RecU and its action on HJ (Figure 6). The rocking motion of the mushroom cap with respect to the stalk region plays a role in the

insertion of a RecU dimer into the HJ (Figure 6). Then the rotation of the mushroom cap with respect to the stalk region aid RecU to search and locate the cleavage site. Further, NTR might play a role in the stabilization of the reactive state. Phosphodiester bond is cleaved by the mechanism proposed earlier (64) involving nucleophilic attack of water onto the backbone phosphate of DNA. A high resolution structure of RecU-HJ complex and reaction intermediate along with the hybrid quantum mechanics/molecular mechanics calculations might allow us to get more insights into dynamics of reactive state formation and the complete reaction mechanism.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We acknowledge the European Synchrotron Radiation Facility (ESRF) and Diamond Light Source (DLS) for provision of synchrotron radiation facilities and we thank the MX beamline scientists for assistance in beamline usage. We acknowledge the PARAM Yuva II supercomputing facility and NPSF people at CDAC.

## FUNDING

This work was funded in part by a BBSRC UK grant BB/E017576/1. Department of Biotechnology Ramalinga

Swamy fellowship contingency grant of Prof. Ashutosh Kumar, Indian Institute of Technology, Bombay (to S.C.D.). Authors would also like to thank The National PARAM Supercomputing Facility (NPSF), C-DAC for providing us access with Param Yuva-II super computer for carrying out molecular dynamics simulations and for waiving off all the costs incurred towards the computational time.

*Conflict of interest statement.* None declared.

## REFERENCES

- West, S.C. (2003) Molecular views of recombination proteins and their control. *Nat. Rev. Mol. Cell Biol.*, **4**, 435–445.
- Lloyd, K.B. and Low, R.G. (1996) Homologous recombination. In: Neidhardt, F.C. and Curtiss, R. III (eds). *Escherichia coli and Salmonella: Cellular and Molecular Biology*. ASM Press, Washington, D.C., pp. 2236–2255.
- Ayora, S., Carrasco, B., Doncel-Perez, E., Doncel, E., Lurz, R. and Alonso, J.C. (2004) *Bacillus subtilis* RecU protein cleaves Holliday junctions and anneals single-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 452–457.
- Cañas, C., Carrasco, B., García-Tirado, E., Rafferty, J.B., Alonso, J.C. and Ayora, S. (2011) The stalk region of the RecU resolvase is essential for Holliday junction recognition and distortion. *J. Mol. Biol.*, **410**, 39–49.
- Carrasco, B., Ayora, S., Lurz, R. and Alonso, J.C. (2005) *Bacillus subtilis* RecU Holliday-junction resolvase modulates RecA activities. *Nucleic Acids Res.*, **33**, 3942–3952.
- Carrasco, B., Cozar, M.C., Lurz, R., Alonso, J.C. and Ayora, S. (2004) Genetic recombination in *Bacillus subtilis* 168: contribution of Holliday junction processing functions in chromosome segregation. *J. Bacteriol.*, **186**, 5557–5566.
- Rigden, D.J., Setlow, P., Bagyan, I., Stein, R.A. and Jedrzejewski, M.J. (2009) PrfA protein of *Bacillus* species: Prediction and demonstration of endonuclease activity on DNA. *Protein Sci.*, **11**, 2370–2381.
- Pedersen, L.B. and Setlow, P. (2000) Penicillin-binding protein-related factor A is required for proper chromosome segregation in *Bacillus subtilis*. *J. Bacteriol.*, **182**, 1650–1658.
- Bell, J.C. and Kowalczykowski, S.C. (2016) Mechanics and single-molecule interrogation of DNA recombination. *Annu. Rev. Biochem.*, **85**, 193–226.
- Sanchez, H. (2005) The RuvAB branch migration translocase and RecU Holliday junction resolvase are required for double-stranded DNA break repair in *Bacillus subtilis*. *Genetics*, **171**, 873–883.
- Wyatt, H.D.M. and West, S.C. (2014) Holliday junction resolvases. *Cold Spring Harb. Perspect. Biol.*, **6**, a023192.
- Aravind, L., Makarova, K.S. and Koonin, E. V. (2000) SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res.*, **28**, 3417–3432.
- 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.
- Raaijmakers, H., Vix, O., Törö, 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.
- 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. U.S.A.*, **98**, 5509–5514.
- Middleton, C.L., Parker, J.L., Richard, D.J., White, M.F. and Bond, C.S. (2004) Substrate recognition and catalysis by the Holliday junction resolving enzyme Hje. *Nucleic Acids Res.*, **32**, 5442–5451.
- Hadden, J.M., Convery, M.A., Déclais, A.C., Lilley, D.M. and Phillips, S.E. (2001) Crystal structure of the Holliday junction resolving enzyme T7 endonuclease I. *Nat. Struct. Biol.*, **8**, 62–67.
- Ariyoshi, M., Vassilyev, 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.
- Ceschini, S., Keeley, A., McAlister, M.S.B., 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.
- McGregor, N., Ayora, S., Sedelnikova, S., Carrasco, B., Alonso, J.C., Thaw, P. and Rafferty, J. (2005) The structure of *Bacillus subtilis* RecU Holliday junction resolvase and its role in substrate selection and sequence-specific cleavage. *Structure*, **13**, 1341–1351.
- Kelly, S.J., Li, J., Setlow, P. and Jedrzejewski, M.J. (2007) Structure, flexibility, and mechanism of the *Bacillus stearothermophilus* RecU holliday junction resolvase. *Proteins Struct. Funct. Bioinforma.*, **68**, 961–971.
- Osiptuk, J., Li, H., Moy, S., Collart, F. and Joachimiak, A. (2014) X-ray crystal structure of penicillin-binding protein-related factor A from *Bacillus stearothermophilus*.
- Cañas, C., Carrasco, B., Ayora, S. and Alonso, J.C. (2008) The RecU Holliday junction resolvase acts at early stages of homologous recombination. *Nucleic Acids Res.*, **36**, 5242–5249.
- Carrasco, B., Cañas, C., Sharples, G.J., Alonso, J.C. and Ayora, S. (2009) The N-terminal region of the RecU Holliday junction resolvase is essential for homologous recombination. *J. Mol. Biol.*, **390**, 1–9.
- Kabsch, W. (2010) XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 125–132.
- Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **62**, 72–82.
- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A. et al. (2011) Overview of the CCP 4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **67**, 235–242.
- Matthews, B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.*, **33**, 491–497.
- Kantardjiev, K.A. and Rupp, B. (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.*, **12**, 1865–1871.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W. et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
- Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 486–501.
- Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. (2012) toward automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.*, **68**, 352–367.
- Headd, J.J., Echols, N., Afonine, P.V., Grosse-Kunstleve, R.W., Chen, V.B., Moriarty, N.W., Richardson, D.C., Richardson, J.S. and Adams, P.D. (2012) Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **68**, 381–390.
- Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Adams, P.D., Read, R.J., Zwart, P.H., Hung, L.-W., IUCr, W., G.-K.R. et al. (2008) Iterative-build OMIT maps: map improvement by iterative model building and refinement without model bias. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **64**, 515–524.
- Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
- Schrödinger, L. (2010) The PyMOL Molecular Graphics System, Version ~1.3r1.
- Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
- Wallace, A.C., Laskowski, R.A. and Thornton, J.M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.*, **8**, 127–134.



40. Laskowski, R.A. and Swindells, M.B. (2011) LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.*, **51**, 2778–2786.
41. Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D.T., Konarev, P.V. and Svergun, D.I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **45**, 342–350.
42. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. and Svergun, D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **36**, 1277–1282.
43. Svergun, D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.*, **25**, 495–503.
44. Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. and Svergun, D.I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.*, **129**, 5656–5664.
45. Tria, G., Mertens, H.D.T., Kachala, M. and Svergun, D.I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCr*, **2**, 207–217.
46. Hess, B., Kutzner, C., Van Der Spoel, D. and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, **4**, 435–447.
47. Berendsen, H.J.C., van der Spoel, D. and van Drunen, R. (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.*, **91**, 43–56.
48. Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M.R., Smith, J.C., Kasson, P.M., van der Spoel, D. *et al.* (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, **29**, 845–854.
49. Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E. and Simmerling, C. (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.*, **11**, 3696–3713.
50. Pérez, A., Marchán, I., Svozil, D., Sponer, J., Cheatham, T.E., Laughton, C.A. and Orozco, M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J.*, **92**, 3817–3829.
51. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926.
52. Joing, I.S. and Cheatham, T.E. (2009) Molecular dynamics simulations of the dynamic and energetic properties of alkali and Halide ions using water-model-specific ion parameters. *J. Phys. Chem. B*, **113**, 13279–13290.
53. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**, 3684–3690.
54. Berendsen, H.J.C. (1991) Transport Properties Computed by Linear Response through Weak Coupling to a Bath. In: Meyer, M and Pontikis, V (eds). *Computer Simulation in Materials Science*. NATO ASI Series, Springer Netherlands, Dordrecht, Vol. **205**, pp. 139–155.
55. Bussi, G., Donadio, D. and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.*, **126**, 1–7.
56. Parrinello, M. and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, **52**, 7182–7190.
57. Hess, B., Bekker, H., Berendsen, H.J.C. and Fraaije, J.G.E.M. (1997) LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.*, **18**, 1463–1472.
58. Hess, B. (2008) P-LINCS: a parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.*, **4**, 116–122.
59. Darden, T., York, D. and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.
60. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. and Pedersen, L.G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577.
61. Laskowski, R.A. (2001) PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res.*, **29**, 221–222.
62. Jarvis, R.A. and Patrick, E.A. (1973) Clustering using a similarity measure based on shared near neighbors. *IEEE Trans. Comput.*, **100**, 1025–1034.
63. Knight, C.J. and Hub, J.S. (2015) WAXSiS: a web server for the calculation of SAXS/WAXS curves based on explicit-solvent molecular dynamics. *Nucleic Acids Res.*, **43**, W225–W230.
64. Kelly, S.J., Li, J., Setlow, P. and Jedrzejewski, M.J. (2007) Structure, flexibility, and mechanism of the *Bacillus stearothermophilus* RecU Holliday junction resolvase. *Proteins*, **68**, 961–971.
65. Hub, J.S. and de Groot, B.L. (2009) Detection of functional modes in protein dynamics. *PLoS Comput. Biol.*, **5**, e1000480.
66. Górecka, K.M., Komorowska, W., Nowotny, M., Górecka, K.M., Komorowska, W. and Nowotny, M. (2013) Crystal structure of RuvC resolvase in complex with Holliday junction substrate. *Nucleic Acids Res.*, **41**, 9945–9955.
67. Chen, L., Shi, K., Yin, Z. and Aihara, H. (2013) Structural asymmetry in the *Thermus thermophilus* RuvC dimer suggests a basis for sequential strand cleavages during Holliday junction resolution. *Nucleic Acids Res.*, **41**, 648–656.