Rapid Communication J. Biochem. 126, 986-990 (1999)

Crystal Structure of *Thermus thermophilus* HB8 UvrB Protein, a Key Enzyme of Nucleotide Excision Repair[1]

Noriko Nakagawa,*,[2] Mitsuaki Sugahara,*,† Ryoji Masui,* Ryuichi Kato,* Keiichi Fukuyama,* and Seiki Kuramitsu*,†,‡,[3]

Department of Biology, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043; †Harima Institute/SPring-8, The Institute of Physical and Chemical Research (RIKEN), Sayo-gun, Hyogo 679-5148; and ‡Genomic Sciences Center (GSC), RIKEN, Wako, Saitama 351-0198

Received October 18, 1999; accepted October 21, 1999

In the nucleotide excision repair system, UvrB plays a central role in damage recognition and DNA incision by interacting with UvrA and UvrC. We have determined the crystal structure of *Thermus thermophilus* HB8 UvrB at 1.9 Å resolution. UvrB comprises four domains, two of which have an α/β structure resembling the core domains of DNA and RNA helicases. Additionally, UvrB has an α-helical domain and a domain consisting of antiparallel β-sheets (β-domain). The sequence similarity suggests that the β-domain interacts with UvrA. Based on the distribution of the conserved regions and the structure of the PcrA-DNA complex, a model for the UvrB-DNA complex is proposed.

Key words: crystal structure, DNA repair enzyme, nucleotide excision repair, *Thermus thermophilus* HB8, UvrB.

DNA is easily damaged by UV radiation and various chemical agents. Such damage may cause mutagenesis or cell death. Living organisms have DNA repair systems through which genetic information is preserved (*1*). Among various DNA repair systems, nucleotide excision repair is one of the most important repair systems because it repairs a wide range of DNA damage (*2*). Extensive studies on this repair system in *Escherichia coli* (*3*) revealed that the UvrA, UvrB, and UvrC proteins are involved in the recognition and excision of a damaged segment of DNA, and that UvrD helicase, DNA polymerase I, and DNA ligase are involved in completion of the DNA repair.

The $UvrA_2B$ complex recognizes a lesion in DNA as the first step of this excision. After dissociation of UvrA from the $UvrA_2B$-DNA complex through ATP hydrolysis, a stable UvrB-DNA complex is formed, in which the DNA is bent and locally unwound. The binding of UvrC to the UvrB-DNA complex causes incisions at the fourth or fifth phosphodiester bond 3′ from a lesion and subsequently at the eighth bond 5′ from the lesion (*1-3*). Thus, UvrB plays a central role in the recognition and excision of a lesion in DNA, but little is known about the molecular mechanism.

To reveal the structure-function relationship of this multifunctional protein, we have studied the UvrB from an extremely thermophilic bacterium, *Thermus thermophilus* HB8 (*4-6*). The results of limited proteolysis and denaturation experiments on *T. thermophilus* HB8 UvrB (ttUvrB) suggested that it consists of four structural domains (*5*). Moreover, we have already succeeded in crystallizing ttUvrB (*6*). We report here the crystal structure of ttUvrB at 1.9 Å resolution. The implications of its domain organization for the catalytic mechanism are discussed and a model of a complex with DNA is presented.

In addition to the purification procedures described previously (*4*), ttUvrB was further purified by high performance liquid chromatography on a MonoQ HR5/5 column (Pharmacia). Crystallization of ttUvrB was carried out as described previously (*6*). Intensity data were collected by means of synchrotron radiation at the Photon Factory (Nat1, Pt1, Pt2, and Hg3) and SPring-8 (Nat2), as well as Ni-filtered Cu Kα radiation and R-AXIS IV (U1, Hg1, and Hg2). All the data were processed and reduced using DENZO and SCALEPACK (*7*). The structure of ttUvrB was solved by multiple isomorphous replacement with anomalous scattering (MIRAS). The initial heavy atom positions of the mercury derivative (Hg1) were determined from an isomorphous difference Patterson map. Subsequent difference and cross-Fourier maps allowed the location of the heavy atom sites in the other derivatives using XtalView (*8*). The heavy atom parameters were refined and the multiple isomorphous replacement phases were calcu-

[1] This work was supported in part by Grants-in-Aid for Scientific Research on Priority Areas (Nos. 08280104, 10129219, and 11169223) and for Scientific Research (No. 10780385) from the Ministry of Education, Science, Sports and Culture of Japan, and by the Sakabe Project of TARA (Tsukuba Advanced Research Alliance) of the University of Tsukuba. The diffraction data were collected at BL6A (and BL18B) at Photon Factory (No. 98G141) and at BL41XU at SPring-8 (No. 1998A0096-NL).

[2] Recipient of Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists.

[3] To whom correspondence should be addressed. Tel: +81-6-6850-5433, FAX: +81-6-6850-5442, E-mail: kuramitu@bio.sci.osaka-u.ac.jp

Abbreviations: ttUvrB, *Thermus thermophilus* HB8 UvrB; TRCF, transcription-repair coupling factor; ssDNA, single-stranded DNA; dsDNA, double-stranded DNA; MIRAS, multiple isomorphous replacement with anomalous scattering.

lated using MLPHARE in CCP4 (*9*). Due to the difficulty in maintaining the isomorphism of the frozen crystals, the analysis was initiated with the data collected at 293 K. The electron density map showed a clear solvent boundary. After phase improvement by solvent flattening and histogram matching with DM (*9*), the polyalanine model was built using program O (*10*). When the model for about 50% of the residues had been built, the phases were extended to high resolution by combining phase information derived from the polyalanine model and MIRAS based on the data collected at 100 K. The model was built with the aid of the amino acid sequence, and refined using rigid body refinement and simulated annealing at 2.2 Å resolution (CNS 0.9) (*11*). Parts of the model, which were initially difficult to trace, were fitted successively to a $2F_{obs} - F_{calc}$ map (Fig. 1A). Water molecules were added in the refinement at 1.9 Å resolution with the respective temperature factors. The final model consists of 552 residues, three molecules of β-octyl glycoside, one sulfate ion (Fig. 1B), and 335 ordered water molecules. The final phasing and refinement statistics are given in Table I. The coordinates have been deposited at the Protein Data Bank under ID code 1D2M.

The overall structure of ttUvrB consists of four domains (Fig. 2). Two of them (colored blue and red), which contain all of the conserved helicase sequence motifs, comprise a central parallel β-sheet flanked by α-helices. Since these α/β domains are similar to domains 1A and 2A of PcrA (*12*), Rep (*13*), and HCV (*14*) helicases, they are referred to as helicase domains 1A (blue) and 2A (red), respectively. In addition, ttUvrB has an α-helical domain (α-domain), colored green in Fig. 2, and a domain consisting of antiparallel β-sheets (β-domain), colored cyan. Helicase domains 1A and 2A are separated by a large cleft, whereas helicase domain 1A and the α-domain are in extensive contact with each other (Fig. 2A). The β-domain protrudes from the rest of the molecule and its contact with the α-domain is limited (Fig. 2B).

The amino acid sequence of the β-domain (residues 154–242) is homologous to residues 86–172 of transcription-repair coupling factor (TRCF) from *T. thermophilus* HB8 (unpublished data). Since TRCF is known to interact with UvrA (*15*), it is suggested that the β-domain is a UvrA-binding domain. This notion is supported by the observation that a truncated *E. coli* UvrB, which corresponds to the

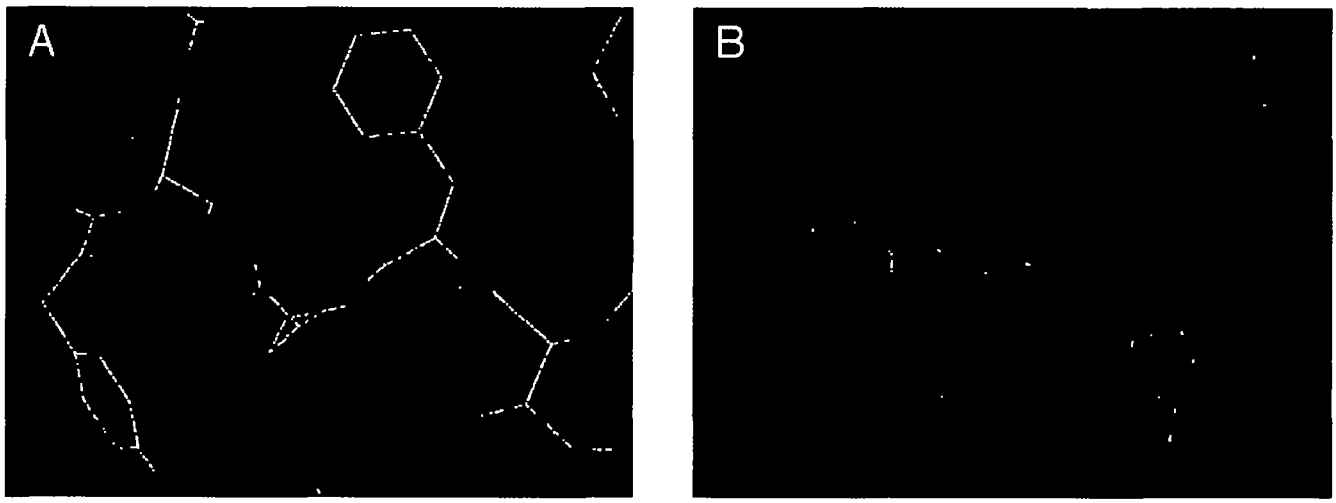

Fig. 1. **Representative regions of the electron density.** (A) The β sheet region in the β-domain (residues 199–203). (B) Binding site of a sulfate ion between helicase domains 1A and 2A. The sulfur atom is colored green. In both panels, the electron density results from a $2F_{obs} - F_{calc}$ map calculated using the refined model.

TABLE I. **Data collection and refinement statistics.**

Phasing statistics	Nat1	Pt1[a]	U1[b]	Hg1[c]	Hg2[d]	Nat2	Pt2[a]	Hg3[c]
Temperature (K)	293	293	293	293	293	100	100	100
Wavelength (Å)	1.00	0.89	1.54	1.54	1.54	0.71	1.00	1.00
Resolution (Å)	2.9	3.2	3.2	4.0	3.9	1.9	2.7	3.1
Completeness (%)	97.6	93.5	99.9	93.3	92.5	93.5	96.1	93.4
R_{merge} (%)	6.2	5.3	6.5	11.1	10.0	6.3	6.2	8.5
No. of sites		3	3	2	1		3	2
R_{Cullis} (%) (centric/acentric)		0.89/0.93	0.86/0.91	0.95/0.95	0.96/0.98		0.68/0.83	0.93/0.96
Phasing power (centric/acentric)		0.49/0.70	0.49/0.74	0.36/0.51	0.24/0.35		1.11/0.99	0.35/0.39

Refinement	Resolution (Å)	Unique reflections	*R*-factor (%)	R_{free}[e] (%)	R.m.s.d bonds (Å)	R.m.s.d angles (deg)
Nat2	30.0–1.9	82,730	23.4	25.3	0.006	1.3

[a] K_2PtCl_4 derivatives (Pt1 and Pt2) were prepared by soaking in an about 2 mM K_2PtCl_4 solution for 14 and 6 h, respectively. [b] A $(CH_3COO)_2UO_2$ derivative (U1) was prepared by soaking in an about 2 mM $(CH_3COO)_2UO_2$ solution for 14 h. [c] Thimerosal derivatives (Hg1 and Hg3) were obtained by cocrystallization with 240 and 400 μM thimerasal, respectively. [d] A $HgCl_2$ derivative (Hg2) was obtained by cocrystallization with 650 μM $HgCl_2$. [e] R_{free} was monitored with 10% of the reflection data excluded from the refinement.

regions of the β-domain, can bind to UvrA (*16*). The β-domain is highly flexible (temperature factor, 50-70 Å^2) compared with the other domains. This suggests fluctuation of the β-domain in the molecule. Furthermore, the electron density for the side chains was not clearly interpretable in the regions formed by residues 159-165 and 223-235, and the region encompassing residues 213-222 was disordered. The formation of the UvrA-UvrB complex may fix these regions.

After DNA damage recognition by the UvrAB complex, UvrB interacts with UvrC to incise the DNA. It has been reported that the C-terminal region of UvrB interacts with UvrC (*17*). In our crystal structure, the C-terminal 82 residues (584-665) were disordered, although SDS-polyacrylamide gel electrophoresis indicated that the apparent molecular weight of ttUvrB in a crystal is the same as that of the intact ttUvrB (data not shown). Previously, we indicated that the fragment composed of the C-terminal 75 residues forms an α-helical structure with an α-helical content of 51% (*5*). As the crystal of ttUvrB has a high solvent content (67%), there is no steric hindrance from the symmetrically related molecule around the C-terminus of the present model. Therefore, the C-terminal region appears to be disordered due to its large motion, which may be restricted by the interaction with UvrC.

UvrB has seven helicase motifs and belongs to helicase superfamily II (*18*). The double-stranded DNA (dsDNA) with a lesion is unwound (*19, 20*) and kinked (*21*) in the UvrAB-DNA complex. However, the UvrAB complex only unwinds a short duplex region formed on annealing of an oligonucleotide to a long single-stranded DNA (ssDNA) (*19*). Compared to the helicases whose structures have been determined, the amino acid sequence of UvrB also has large insertions between helicase motifs Ia and II, and II and III (*5*). However, helicase domains 1A and 2A of ttUvrB are structurally similar to the core domains of

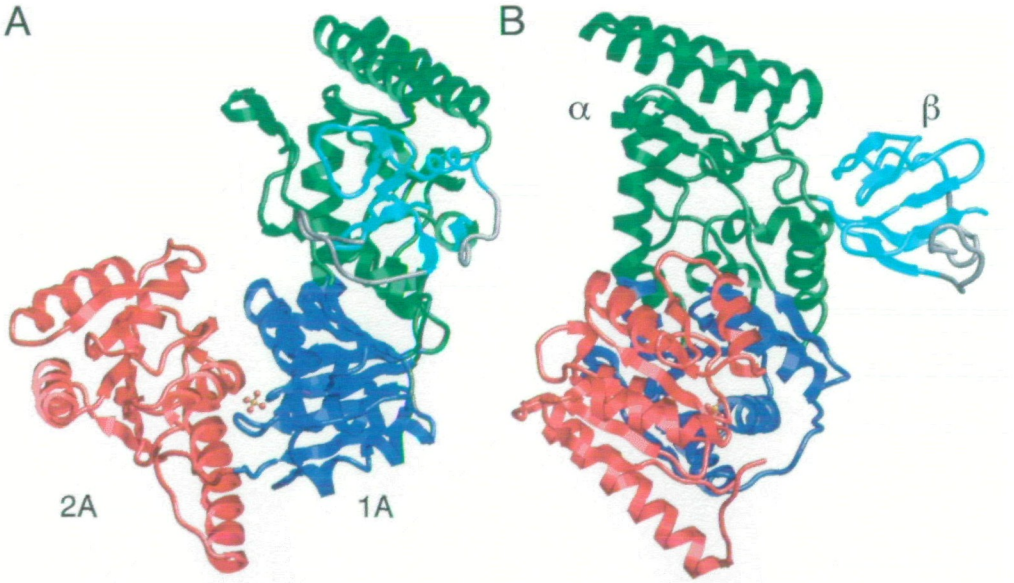

Fig. 2. **Schematic ribbon diagram of the overall structure of ttUvrB.** Panel (B) shows a perpendicular view of Panel (A). Domains 1A, 2A, α, and β are colored blue, red, green and cyan, respectively. The regions colored gray correspond to residues 159-165 and 223-235, where the side chains were disordered. The sulfate ion is represented by the ball and stick model.

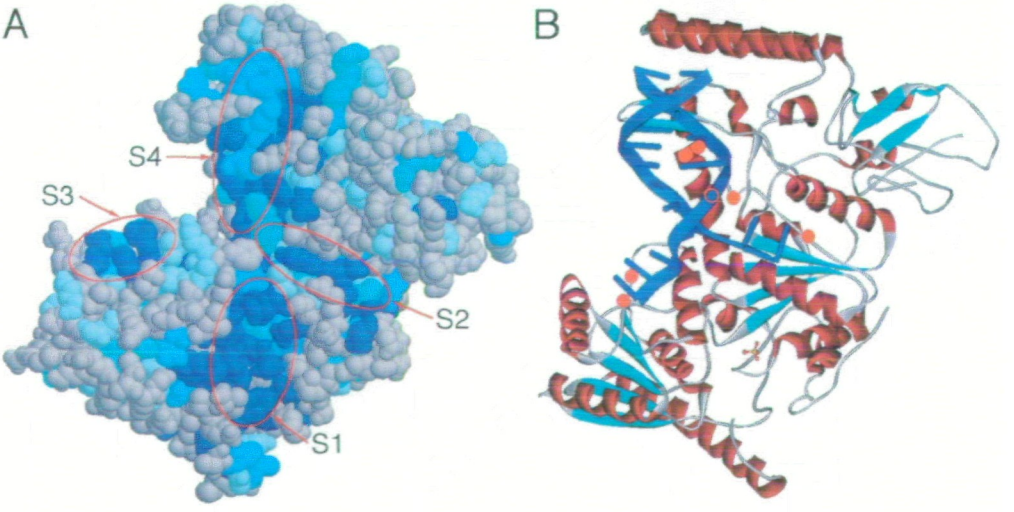

Fig. 3. **Distribution of the conserved residues between species (A), and a putative model of binding of ttUvrB to DNA (B).** (A) The extent of conservation is indicated by the gradient of blue coloring (dark blue, 100%; blue, 90%; cyan, 80%; and light blue, 70%). The four conserved surfaces (S1 to S4) are encircled by red lines. (B) Putative model of binding of ttUvrB to DNA. The model was built based on the distribution of the conserved residues on the surface and the structure of the PcrA-DNA complex (*12*). Secondary structural elements are indicated in cyan for β-strands and red for α-helices. Intervening regions are shown in gray. The positions of the most conserved residues are indicated by the yellow circles. The DNA is colored blue. The sulfate ion is represented by the ball and stick model.

"true" helicases, such as PcrA, Rep, and HCV helicases. Furthermore, Lys42 in helicase motif I, the Walker A-type nucleotide binding motif, interacts with the sulfate ion, whose position is similar to that occupied by the γ-phosphate of the ATP analogue complexed with PcrA helicase (*12*). These results suggest that helicase domains 1A and 2A of ttUvrB are responsible for dsDNA unwinding coupled with ATP hydrolysis. However, ttUvrB alone has no helicase activity (data not shown), like *E. coli* UvrB (*19*), even though ttUvrB has ATPase activity, which is activated by ssDNA (*4*). In this respect, UvrB is not a true helicase, such as PcrA, Rep, and HCV helicases. Association of UvrA may cause a conformational change of UvrB, which is essential for DNA unwinding.

As shown in Fig. 3A, ttUvrB has the four highly conserved surface regions: the cleft between helicase domains 1A and 2A (S1); the "tops" of helicase domains 1A and 2A (S2 and S3, respectively); and the "left" side of the α-domain (S4). Among them, S1 consists of the residues of the helicase motifs, which are involved in ATP hydrolysis and the interaction between helicase domains 1A and 2A.

In the crystal structures of helicase-DNA complexes, the outer surfaces of domains 1A and 2A, corresponding to S2 and S3, are involved in binding to ssDNA (*12-14*). The interfaces between the helicase and ssDNA include not only hydrogen bonds with the DNA backbone, but also stacking interaction of aromatic residues with the base moiety of DNA. The stacking interaction has been suggested to be responsible for several functions: bending of the DNA backbone, such as Trp250 of Rep helicase (*13*); stabilization of the unwound form of DNA, such as Phe626 of PcrA helicase (*12*); and sliding of the DNA, such as Phe64 of PcrA helicase (*12*). Furthermore, the UvrB-DNA complex is stable as to high ionic strength (*22*), and the DNA in this complex is kinked by 130° (*21*). These observations suggest that the interaction between UvrB and DNA involves a hydrophobic interaction and/or intercalation of the aromatic residues into the duplex *via* a stacking interaction. In this respect, it is interesting that the residues present in S2 (residues 63-65 and 81-92) and S3 (residues 448-454 and 469-473) include the aromatic residues (Tyr85, Tyr92, and Tyr469) and histidine (His471). The conservation of these residues suggests these surface regions interact with ssDNA through a stacking interaction. It has been reported that DNA unwinding introduced by the UvrAB complex is localized within 1-3 bp (*19*) or 6 bp (*20*) around a lesion. If the S2 and S3 regions interact with an identical DNA strand, more than 6 bases of the ssDNA region is required. Therefore, it is suggested that the S2 and S3 regions interact with the respective strands of unwound dsDNA.

The S4 region (residues 93-98, 106-108, and 345-353) is located on the left side of the α-domain (Fig. 3A). The electrostatic potential of this surface region (data not shown) is slightly positive due to the presence of Arg347, Arg352, and Lys353. This implies that this region is involved in binding to dsDNA. Based on these structural features of ttUvrB and the crystal structure of the PcrA-DNA complex (*12*), a model for ttUvrB-DNA complex is proposed (Fig. 3B).

This putative model may provide a clue as to the mechanism of the damage recognition by UvrB. Our hypothesis is as follows: When UvrB is loaded onto a damaged site by UvrA, UvrB unwinds the duplex around the lesion through the interaction of the S2 and S3 regions with the respective strands. When the lesion is present in the strand, the stacking interaction of amino acid side chains with the base is weakened or lost. Such alteration of the stacking interaction induces a conformational change of UvrB, which leads to the formation of a stable UvrB-DNA complex.

UvrB is one of the proteins having no helicase activity by themselves in spite of the preservation of the helicase motifs. The structure of ttUvrB represents the first view of such a protein.

We wish to thank Professor Noriyoshi Sakabe of the University of Tsukuba, and Drs. Nobuhisa Watanabe, Mamoru Suzuki, and Noriyuki Igarashi of the National Laboratory for High Energy Physics, and Dr. Masahide Kawamoto of SPring-8 (BL41XU) for their help in the data collection through synchrotron radiation.

REFERENCES

1. Friedberg, E.C., Walker, G.C., and Siede, W. (1995) *DNA Repair and Mutagenesis*, American Society of Microbiology Press, Washington, DC
2. Grossman, L. and Thiagalingam, S. (1993) Nucleotide excision repair, a tracking mechanism in search of damage. *J. Biol. Chem.* **268**, 16871-16874
3. Sancar, A. (1996) DNA excision repair. *Annu. Rev. Biochem.* **65**, 43-81
4. Kato, R., Yamamoto, N., Kito, K., and Kuramitsu, S. (1996) ATPase activity of UvrB protein from *Thermus thermophilus* HB8 and its interaction with DNA. *J. Biol. Chem.* **271**, 9612-9618
5. Nakagawa, N., Masui, R., Kato, R., and Kuramitsu, S. (1997) Domain structure of *Thermus thermophilus* UvrB protein. Similarity in domain structure to a helicase. *J. Biol. Chem.* **272**, 22703-22713
6. Shibata, A., Nakagawa, N., Sugahara, M., Masui, R., Kato, R., Kuramitsu, S., and Fukuyama, K. (1998) Crystallization and preliminary X-ray diffraction studies of a DNA excision repair enzyme, UvrB, from *Thermus thermophilus* HB8. *Acta Crystallogr.* **D55**, 704-705
7. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326
8. McRee, D.E. (1993) *Practical Protein Crystallography*, Academic Press, San Diego
9. Collaborative Computational Project, Number 4. (1994) The CCP4 suite: Program for protein crystallography. *Acta Crystallogr.* **D50**, 760-763
10. Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **21**, 916-924
11. 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., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905-921
12. Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S., and Wigley, D.B. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75-84
13. Korolev, S., Hsieh, J., Gauss, G.H., Lohman, T.M., and Waksman, G. (1997) Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-stranded DNA and ADP. *Cell* **90**, 635-647
14. Kim, J.L., Morgenstern, K.A., Griffith, J.P., Dwyer, M.D., Thomson, J.A., Murcko, M.A., Lin, C., and Caron, P.R. (1998) Hepatitis C virus NS3 RNA helicase domain with a bound

oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure* **6**, 89-100
15. Selby, C.P. and Sancar, A. (1995) Structure and function of transcription-repair coupling factor. I. Structural domains and binding properties. *J. Biol. Chem.* **270**, 4882-4889
16. Hsu, D.S., Kim, S.T., Sun, Q., and Sancar, A. (1995) Structure and function of the UvrB protein. *J. Biol. Chem.* **270**, 8319-8327
17. Moolenaar, G.F., Franken, K.L.M.C., Dijkstra, D.M., Thomas-Oates, J.E., Visse, R., van de Putte, P., and Goosen, N. (1995) The C-terminal region of the UvrB protein of *Escherichia coli* contains an important determinant for UvrC binding to the preincision complex but not the catalytic site for 3′-incision. *J. Biol. Chem.* **270**, 30508-30515
18. Gorbalenya, A.E. and Koonin, E.V. (1993) Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**, 419-429
19. Gordienko, I. and Rupp, W.D. (1997) The limited strand-separating activity of the UvrAB protein complex and its role in the recognition of DNA damage. *EMBO J.* **16**, 889-895
20. Zou, Y. and Van Houten, B. (1999) Strand opening by the UvrA$_2$B complex allows dynamic recognition of DNA damage. *EMBO J.* **18**, 4889-4901
21. Shi, Q., Thresher, R., Sancar, A., and Griffith, J. (1992) Electron microscopic study of (A)BC excinuclease. DNA is sharply bent in the UvrB-DNA complex. *J. Mol. Biol.* **226**, 425-432
22. Orren, D.K. and Sancar, A. (1990) Formation and enzymatic properties of the UvrB-DNA complex. *J. Biol. Chem.* **265**, 15796-15803