Structure and Function of a Nucleotide Excision Repair Enzyme, UvrB

Doctoral Thesis

January 2001

Noriko Nakagawa

Department of Biology, Graduate School of Science

Osaka University

Contents

Abbreviations

bcUvrB	*Bacillus caldotenax* UvrB
CBB	Coomassie Brilliant Blue R-250
CD	circular dichroism
dsDNA	double-stranded DNA
DTE	dithioerythritol
HCV	hepatitis C virus
MIRAS	multiple isomorphous replacement with anomalous scattering
PAGE	polyacrylamide gel electrophoresis
ssDNA	single-stranded DNA
TRCF	transcription-repair coupling factor
ttUvrA	*Thermus thermophilus* UvrA
ttUvrB	*Thermus thermophilus* UvrB

Abstract

UvrA, UvrB and UvrC proteins initiate nucleotide excision repair, in which wide range of DNA damage is removed by excision. Among them, UvrB protein plays a central role in damage recognition and DNA incision by interacting with UvrA and UvrC proteins. Furthermore, UvrB protein has seven conserved regions consisting of helicase motifs. To reveal the structure-function relationship of this multifunctional protein, X-ray crystallographic analysis and functional analysis of the domains were conducted using UvrB from an extremely thermophilic bacterium, *Thermus thermophilus* HB8 (ttUvrB). The results indicated that ttUvrB consists of five domains: 1A, α-, β-, 2A and C2. Among them, the structures of domains 1A and 2A are similar to those of DNA and RNA helicases. The properties of the proteolytic fragments, which lack some of the domains, indicate the involvement of the respective domains in the following functions: domains 1A and 2A are necessary for ATPase activity, domain 2A is indispensable for DNA binding, and the β-domain is involved in UvrA binding. Furthermore, ttUvrB has four surfaces consisting of highly conserved residues. Based on the conserved regions of ttUvrB and the structures of the helicase-DNA complexes, a model for the UvrB-DNA complex is proposed. In this model, the interaction between UvrB protein and ssDNA involves a stacking interaction of aromatic residues with the base moiety of DNA. When a lesion is present on a strand, the stacking interaction can be altered, inducing a conformational change in the UvrB protein. Such change would lead to the formation of the UvrB-DNA preincision complex.

Introduction

DNA reacts readily with various chemical compounds and also with certain physical agents in the environment, such as UV radiation. Such reactions cause alterations in the chemistry or the sequence of DNA, resulting in mutagenesis or even cell death. To avoid these alterations, living organisms possess repair systems such as direct repair, base excision repair, nucleotide excision repair, mismatch repair, and recombinational repair (Friedberg *et al.*, 1995). Nucleotide excision repair is one of the most important repair systems, since this system is involved in the repair of a wide range of DNA damage (Sancar, 1996). Furthermore, this system has been found in all free-living species tested, from the smallest free-living life form *Mycoplasma genitalium* to humans. Nucleotide excision repair mutants of *Escherichia coli* and yeast are sensitive to the mutagenic and lethal effects of UV light and other genotoxic agents. In humans, three diseases are associated with defective nucleotide excision repair: xeroderma pigmentosum, Cockayne's syndrome, and trichothiodystrophy. The patients show hypersensitivity to sunlight and/or progressive neurological degeneration (Friedberg *et al.*, 1995). Although none of the nucleotide excision repair proteins of humans shares significant sequence homology with those of eukaryotes, the overall strategy of the repair system is quite similar in prokaryotes and eukaryotes (Figure 1) (Sancar, 1994).

In *E. coli* this system involves at least six proteins: UvrA, UvrB, and UvrC proteins, UvrD helicase (DNA helicase II), DNA polymerase I, and DNA ligase (Sancar, 1996). UvrA and UvrB proteins form a $UvrA_2B$ complex in an ATP-dependent reaction.

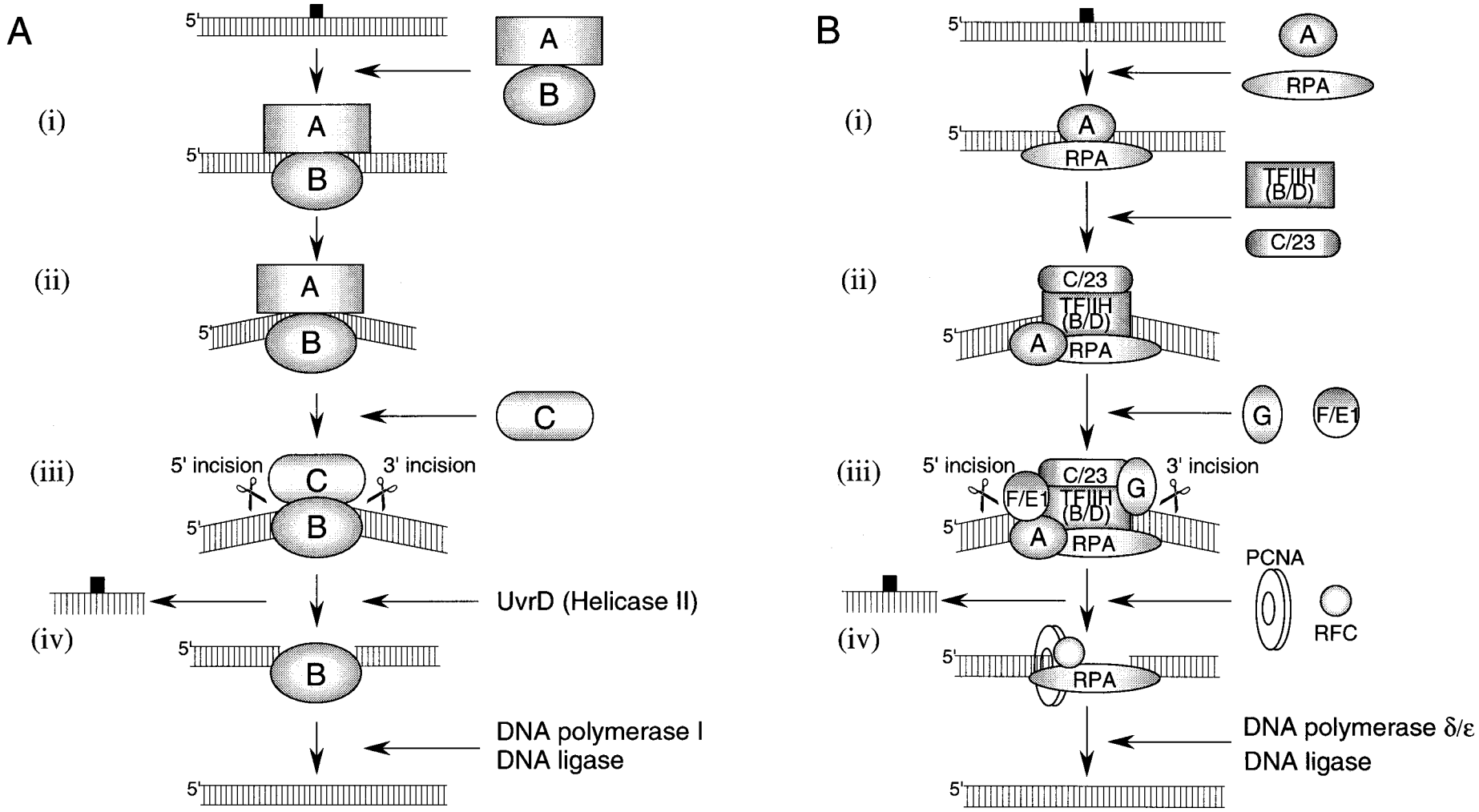

Figure 1. **Molecular mechanisms of nucleotide excision repair.**

The repair proceeds in four steps: (i) Initial damage recognition, (ii) Distortion of the DNA, (iii) Dual incision, (iv) Repair synthesis and ligation. *A*, molecular mechanism in prokaryote. *A*, UvrA protein; *B*, UvrB protein; *C*, UvrC protein. *B*, molecular mechanism in human. *A_G*, XPA through XPG (except XPE) proteins; *C/23*, XPC-HHR23B complex; *F/E1*, XPF-ERCC1 complex.

This complex recognizes a lesion in the DNA and a $UvrA_2B$-DNA complex is formed, in which the DNA is unwound and kinked. The UvrB protein tightly binds to the damaged site, forming a UvrB-DNA preincision complex, and the UvrA protein dissociates from the complex. Then UvrC protein binds to the stable UvrB-DNA complex and make 3' and 5' incisions. Following the incisions, UvrD helicase releases the damage-containing oligonucleotide and UvrC protein, leaving a UvrB-gapped-DNA complex. DNA polymerase I binds to the 3' OH terminus generated at the 5' incision and displaces bound UvrB protein during the course of repair synthesis. The remaining nick after repair synthesis is sealed by DNA ligase.

In the nucleotide excision repair system UvrB protein plays a central role, since it is involved in most of the repair process and interacts with UvrA, UvrC proteins, UvrD helicase, DNA polymerase and DNA. Sequence analysis of UvrB protein reveals several motifs: a region homologous to transcription-repair coupling factor (TRCF) (Selby & Sancar, 1993), an array of seven conserved regions of so-called "helicase motifs" (Gorbalenya *et al.*, 1989), two α–helical regions with a high probability score for formation of a coiled-coil motif, and a region homologous to UvrC protein (Moolenaar *et al.*, 1995). Analyses of many mutants with amino acid substitutions and deletions have suggested the functional regions of *E. coli* UvrB that are involved in binding of UvrA, UvrC, and DNA (Seeley & Grossman, 1989; Moolenaar *et al.*, 1994; Moolenaar *et al.*, 1995; Hsu *et al.*, 1995); however, details of relationship between its structure and function are still uncertain. In particular, there is only limited information available about the structural properties of UvrB protein.

In general, proteins isolated from the extremely thermophilic bacterium, *Thermus*

thermophilus HB8, which can grow at temperatures over 75°C (Oshima & Imahara, 1974), are heat-stable and suitable for physicochemical studies, including X-ray crystallographic analysis (Yokoyama *et al.*, 2000a; Yokoyama *et al.*, 2000b). All the components of the excision repair system from *T. thermophilus* HB8, namely UvrA, UvrB, UvrC proteins, UvrD helicase, DNA polymerase I and DNA ligase have already been purified (Kito, 1997; Kato *et al.*, 1996; Kobayashi, 1999; unpublished data). Since the amino acid sequences of these proteins show homology with those of other prokaryotes including *E. coli*, the mechanism of nucleotide excision repair in *T. thermophilus* is considered to be similar to that of *E. coli*. Nevertheless, *T. thermophilus* UvrB (ttUvrB) differs from *E. coli* in terms of its ability to hydrolyze ATP in the absence of UvrA protein and DNA (Kato *et al.*, 1996). Additionally, ttUvrB is stable from 5°C to 80°C at pH 7.5, and between pH 6 and pH 11 at 25°C (Kato *et al.*, 1996). These features of ttUvrB are useful for elucidating not only its structural properties but also its structure-function relationship.

In this study the three-dimensional structure of ttUvrB has been determined by multiple isomorphous replacement and refined at 1.9 Å resolution. The structure revealed that ttUvrB has four domains. To reveal the structure-function relationship of UvrB protein, the proteolytic fragments of ttUvrB, which lack some of the domains, were purified and their activities were assayed. The results indicated the involvement of the respective domains in the functions of the protein. Furthermore, UvrB protein has four surfaces composed of highly conserved residues. Based on these conserved regions and the structure of the helicase-DNA complex, a model for the UvrB-DNA complex is proposed.

Experimental Procedures

X-ray Crystallographic Analysis

Protein Purification and Crystallization — In addition to the purification procedures described previously (Kato *et al.*, 1996), ttUvrB was further purified by high-performance liquid chromatography using a MonoQ HR5/5 column (Pharmacia). Crystallization of ttUvrB was carried out as described previously (Shibata *et al.* 1999). Thimerosal derivatives (Hg1 and Hg3) were obtained by cocrystallization with 240 and 400 μM thimerasal, respectively. The $HgCl_2$ derivative (Hg2) was obtained by cocrystallization with 650 μM $HgCl_2$. Due to the difficulty in maintaining the isomorphism of the crystal soaked in the reservoir solution, highly concentrated solutions of heavy atoms were made and added directly to the protein droplet containing crystals. K_2PtCl_4 derivatives (Pt1 and Pt2) were prepared by soaking in about 2 mM K_2PtCl_4 solution for 14 and 6 hours, respectively. $(CH_3COO)_2UO_2$ derivative (U1) was prepared by soaking in about 2 mM $(CH_3COO)_2UO_2$ solution for 14 hours. Selenomethionine-labeled ttUvrB was produced by transforming the methionine auxotrophic strain B834(DE3)pLysS with the ttUvrB expression vector pYB1 and growing these cells in a defined medium containing selenomethionine (LeMaster & Richards, 1985). The selenomethionine-labeled protein was purified and crystallized in the same manner as the native protein.

X-ray Diffraction Data Collection — The crystal (Nat1, Pt1, U1 Hg1 and Hg2) was sealed in a glass capillary tube with a small amount of the mother liquor. Diffraction data were collected at room temperature by means of synchrotron radiation at the

Photon Factory, as well as Ni-filtered Cu Kα radiation and R-AXIS IV. In order to stabilizing the crystal against X-ray irradiation the crystal (Nat2, Pt2, Hg3 and Se1) was mounted in the loop and flash-cooled to 100 K using an Oxford Cryosystems Cryostream or Rigaku Cryostat. Data from the derivatives were measured at Photon Factory. Data collection of the native crystal with higher resolution and signal-to-noise ratio was achieved using synchrotron radiation at the SPring-8 (Figure 2). All the data were processed and reduced using DENZO and SCALEPACK (Otwinoski & Minor, 1997).

Structure Determination and Refinement — The structure of ttUvrB was solved by the multiple isomorphous replacement with anomalous scattering (MIRAS). Initial heavy atom positions of the mercury derivative (Hg1) were determined from an isomorphous difference Patterson map. Subsequent difference and cross-Fourier maps located the heavy atom sites in the other derivatives using XtalView (McRee, 1993). The heavy atom parameters were refined and the multiple isomorphous replacement phases were calculated using MLPHARE in CCP4 (Collaborative Computational Project, 1994). Due to difficulty in keeping isomorphism among the frozen crystals, the analysis was initiated with the data collected at room temperature. The electron density map showed a clear solvent boundary. After phase improvement by solvent flattening and histogram matching with DM (Collaborative Computational Project, 1994), the polyalanine model was built using the program O (Jones *et al.,* 1991). When the model for about 50% of the residues was 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

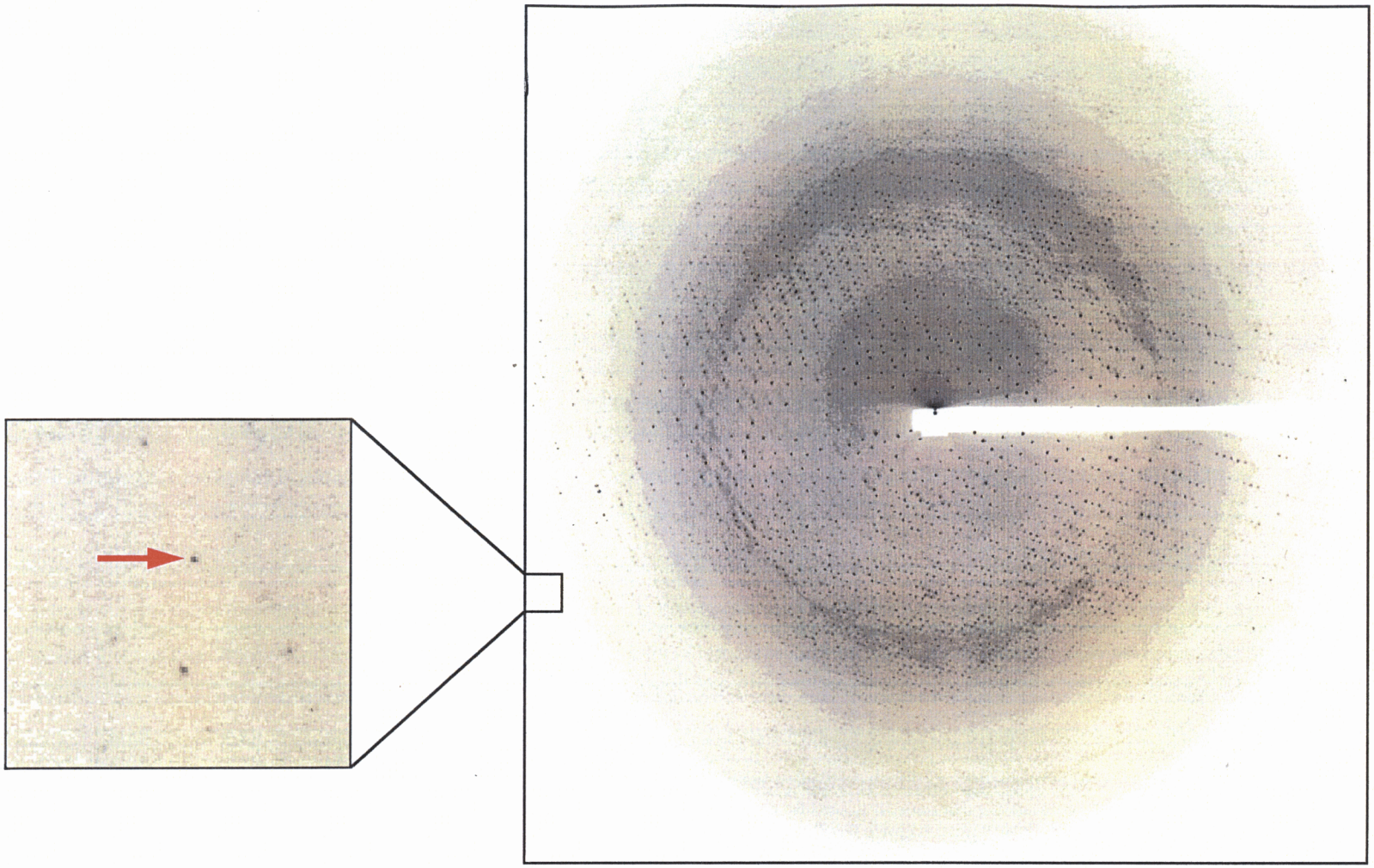

Figure 2. **X-ray diffraction patterns of the native crystal (Nat2).**

A 2° oscillation photograph of ttUvrB crystal was taken at the BL41XU of SPring-8. The left box is an enlargement of the indicated portion of the whole image. The red arrow indicates the diffraction spot at 1.9 Å resolution.

was built with the aid of the amino acid sequence and the selenium coordinates obtained from the difference Fourier map of the selenomethione-labeled crystal. The model was refined using rigid body refinement and simulated annealing at 2.2 Å resolution (CNS 0.9) (Brunger *et al.*, 1998). Parts of the model, which were initially difficult to be traced, were fitted successively to $2F_{obs}$-F_{calc} map (Figure 3A). Water molecules were added in the refinement at 1.9 Å resolution with respective temperature factors. The final model consists of 552 residues, three molecules of β-octyl glycoside, one sulfate ion (Figure 3B), and 335 ordered water molecules. Final phasing and refinement statistics are given in Table 1. The coordinates have been deposited at the Protein Data Bank under ID code 1D2M. Least squares comparison of two structures and calculation of the root-mean-square deviations of the Ca atoms were carried out using LAQKAB in CCP4 (Collaborative Computational Project, 1994).

Functional Analysis

Limited Proteolysis — ttUvrB (15 μM) was treated with trypsin in 50 mM Tris-HCl (pH 7.5) at a protein to protease molar ratio of 100:1 or 20:1 for various times at 37°C. Proteolysis was stopped by adding trypsin inhibitor in a 2-fold excess to trypsin. The digests were separated by SDS-polyacrylamide gel electrophoresis (PAGE) on 10% gels (Laemmli & Favre, 1973), followed by staining with Coomassie Brilliant Blue R-250 (CBB).

Isolation of Tryptic Fragments — ttUvrB (10 μM) was treated with trypsin in 50 mM Tris-HCl (pH 7.5) at a protein to protease molar ratio of 400:1 and 50:1 for 3 min

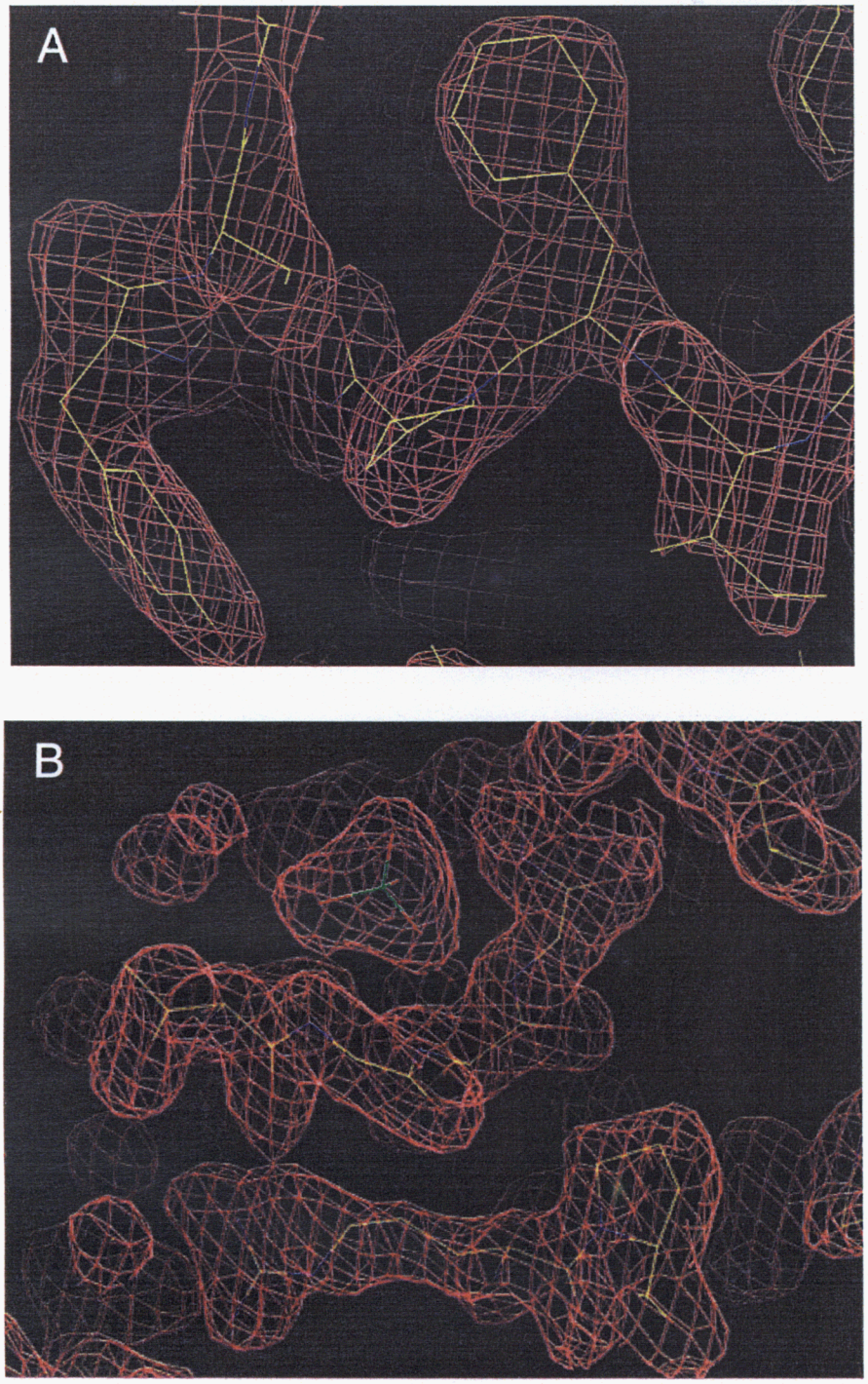

Figure 3. **Representative regions of the electron density.**

A, β sheet region in the β-domain (residues 199-203). *B*, binding site of a sulfate ion between domains 1A and 2A. The sulfur atom is colored in green. In both panels, electron density results from a $2F_{obs}$-F_{calc} map calculated using the refined model. This figure was produced using program O (Jones *et al*., 1991).

Table 1. **Data collection and refinement statistics.**

Phasing statistics	Nat1	Pt1	U1	Hg1	Hg2	Nat2	Pt2	Hg3	Se1
Temperature (K)	293	293	293	293	293	100	100	100	100
Wavelength (Å)	1.00	0.89	1.54	1.54	1.54	0.71	1.00	1.00	0.90
Resolution (Å)	2.9	3.2	3.2	4.0	3.9	1.9	2.7	3.1	3.3
Completeness (%)	97.6	93.5	99.9	93.3	92.5	93.5	96.1	93.4	96.4
R_{merge} (%)[a]	6.2	5.3	6.5	11.1	10.0	6.3	6.2	8.5	8.1
Number 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 (Nat2)	
Resolution (Å)	30.0 – 1.9
Unique reflections	82,730
R-factor (%)	23.4
R_{free} (%)[b]	25.3
R.m.s.d. bonds (Å)	0.006
R.m.s.d. angles (deg)	1.3

[a] $R_{merge} = \Sigma|I_{obs} - \langle I \rangle|/\Sigma I_{obs}$

[b] R_{free} was monitored with 10% of the reflection data excluded from the refinement.

and 60 min at 37°C, respectively. Proteolysis was stopped by adding trypsin inhibitor in a two-fold excess to trypsin. The reaction products were separated by HPLC using a MonoQ HR5/5 column (Pharmacia) eluted with a linear gradient of NaCl (from 0.2 to 0.4 M) in 50 mM Tris-HCl (pH 7.5) at a flow rate of 0.5 ml/min. Fractions were collected and assessed for fragmentation by SDS-PAGE. The tryptic fragments obtained were dialyzed in 50 mM Tris-HCl (pH 7.5), 10 mM 2-mercaptoethanol, 1 mM EDTA, 100 mM KCl and 10% glycerol and stored at 4°C. The concentrations of the tryptic fragments, T-1, T-3, and Tc, were determined using an ε value of 53,000, 41,000 and 8,700 M^{-1} cm^{-1} at an absorbance maximum of around 277 nm, respectively, which were calculated by a previously described procedure (Kuramitsu *et al.*, 1990). Their N-terminal amino acid sequences were determined with a gas-phase protein sequencer (Applied Biosystems, model 473A). Circular dichroism (CD) measurements were carried out with a Jasco spectropolarimeter, model J-720W. The CD spectra of 2 μM ttUvrB or its fragments in 50 mM Tris-HCl (pH 7.5), 100 mM KCl, 1 mM EDTA and 1 mM DTE were measured in a 1 mm cell in the far UV region between 200 and 250 nm at 25°C.

DNA Binding Assay — DNA binding was assayed by native PAGE (Davis, 1964). The assay was carried out using poly(dT) as a substrate. Each reaction mixture for the binding assay, 5 μM ttUvrB or its fragments and 1 mM poly(dT) in 50 mM Tris-HCl (pH 7.5), was incubated at 25°C for 60 min. The reaction mixture was then electrophoresed on a 7.5% polyacrylamide gel under non-denaturing conditions (Davis, 1964), which was followed by staining of the gel with CBB. The concentration of poly(dT) was determined from an ε_{260} value of 8.52×10^3 M^{-1} cm^{-1} (Lee & Cox, 1990).

ATPase Assay — Hydrolysis of ATP was measured at 25°C by an enzyme-coupled spectrophotometric assay (Pugh & Cox, 1988). The change of absorbance at 340 nm was measured with a Hitachi spectrophotometer, model U-3000. ATP was reacted with 0.2 or 0.5 μM ttUvrB or its fragments in 50 mM Tris-HCl (pH 7.5), 100 mM KCl, 10 mM $MgCl_2$, 1 mM dithioerythritol (DTE), 2 mM phosphoenolpyruvate, 0.32 mM NADH, 25 units/ml pyruvate kinase and 25 units/ml lactate dehydrogenase at 25°C. To examine the interaction of ttUvrB or its fragments with single-stranded DNA (ssDNA) or *T. thermophilus* UvrA (ttUvrA), 50 μM poly(dT) or 0.4 μM ttUvrA was added to each reaction mixture containing 0.2 μM ttUvrB. The concentration of ATP was determined from an ε_{259} value of 1.54×10^4 M^{-1} cm^{-1}.

Assay of Interaction with ttUvrA — Interaction of ttUvrB or its fragments with ttUvrA was detected by native PAGE (Davis, 1964). The reaction mixture was composed of the indicated concentrations of ttUvrA, 5 μM ttUvrB or its fragments, 50 mM Tris-HCl (pH 7.5), 5 mM ATP, and 10 mM $MgCl_2$. The reaction mixture was electrophoresed on a 4% polyacrylamide gel containing 5 mM ATP and 10 mM $MgCl_2$ in running buffer containing 5 mM ATP and 10 mM $MgCl_2$. The bands of proteins were visualized with CBB.

Results

Overall Structure of ttUvrB

The overall structure of ttUvrB consists of four domains, which are referred to as 1A, α, β, and 2A (Figure 4). Domains 1A and 2A comprise a central parallel β-sheet flanked by α-helices. The α-domain consists of mainly α-helical polypeptide stretches and one flexible β-hairpin, whereas the β-domain consists of two anti-parallel β-sheets and one α-helix. Domains 1A and 2A are separated by a large cleft, whereas domain 1A and the α-domain are in extensive contacted with each other (Figure 4A). The β-domain protrudes from the rest of the molecule and its contact with the α-domain is limited (Figure 4B). In the crystal structure, the 82 C-terminal residues (584–665) were disordered, although SDS-PAGE indicated that the apparent molecular weight of ttUvrB in the crystal was the same as that of the intact ttUvrB (data not shown).

Recently, the crystal structure of UvrB from *Bacillus caldotenax* (bcUvrB) has been independently determined (PDB code: 1D9X, Theis *et al.*, 1999). The crystal structure of bcUvrB also consists of four domains. Similarly to ttUvrB, the C-terminal 63 residues were disordered, although bcUvrB was crystallized under different conditions using PEG instead of Li_2SO_4 as a precipitant (Theis *et al.*, 1999). The structures of the four domains are quite similar (Figure 5A). The root-mean-square deviations are 0.81 Å for 134 superimposed Cα positions in domain 1A, 1.67 Å for 185 superimposed Cα positions in the α-domain, and 0.98 Å for 174 superimposed Cα positions in domain 2A. Although the β-domains is partially disordered in the

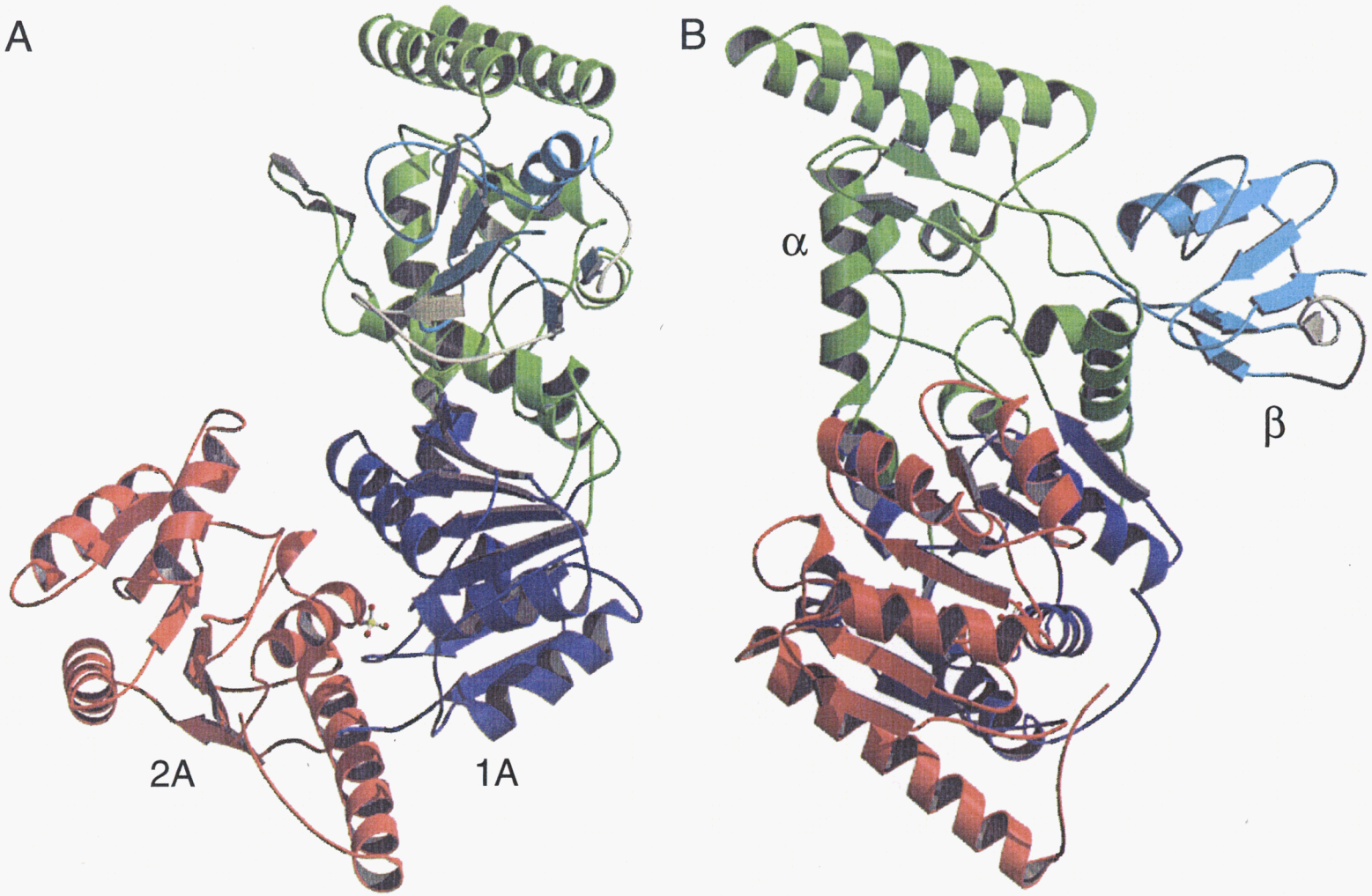

Figure 4. **Schematic ribbon diagram of the overall structure of ttUvrB.**

Panel B shows a perpendicular view of *Panel A*. The 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 the by the ball and stick model. This figure, as well as Figure 5A and 5B, were made using Molscript (Kraulis, 1991) and Raster3D (Merritt and Murphy, 1994).

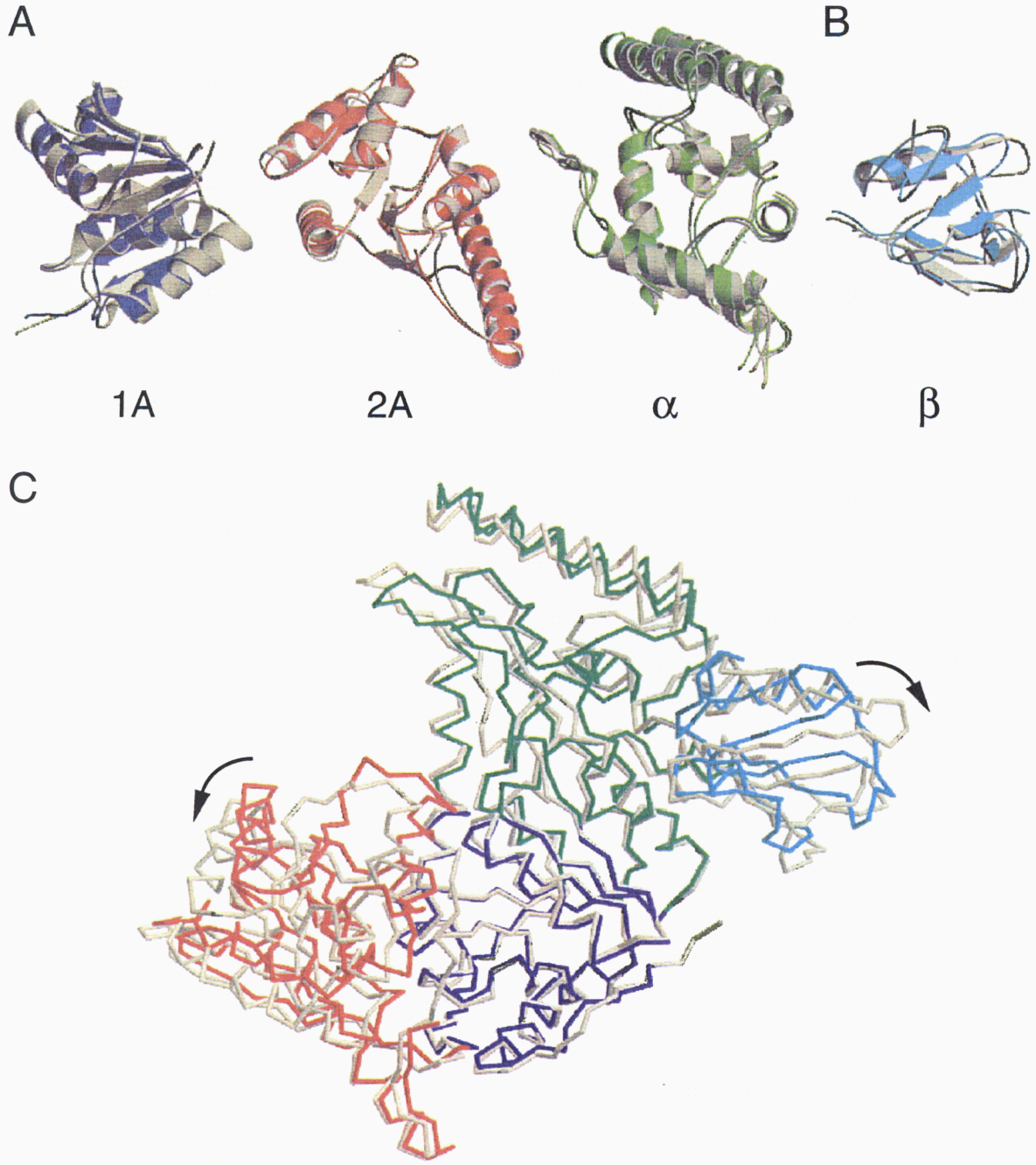

Figure 5. **Superimposed structures of ttUvrB and bcUvrB.**

A, superimposition of each domain (1A, 2A and α) of ttUvrB and bcUvrB from the same view as in Figure 4A. *B*, superimposition of the β-domain from the same view as in Figure 4B. *C*, C_α trace of superimposed structures of ttUvrB and bcUvrB. The 1A domains of each molecule have been superimposed. The color scheme for the domains of ttUvrB is the same as that in Figure 4. The structures of bcUvrB are shown in gray.

structures of both ttUvrB and bcUvrB, the core structures of the β-domain are quite similar (Figure 5B). While the individual domains are superimposed well, there are significant differences in domain orientation. The rigid body movements of domain 2A and the β-domain are most easily illustrated by superposition of the 1A domains of ttUvrB and bcUvrB (Figure 5C). Domains 1A and 2A are connected to each other by a single short linker and separated by a wide cleft. The β-domain also protrudes from the rest of the molecule and the temperature factors of the β-domain are high compared with the other domains. These results suggest that the orientations of domain 1A and the β-domain are variable in solution.

Structural Similarity to Helicase

UvrB protein has seven conserved regions of helicase motifs and belongs to helicase superfamily II (Gorbalenya & Koonin, 1993). Double-stranded DNA (dsDNA) bearing a lesion is unwound (Gordienko & Rupp, 1997; Zou & Van Houten, 1999) and kinked (Shi *et al.*, 1992) in the UvrAB-DNA complex. However, the UvrAB complex unwinds only a short duplex region formed by annealing an oligonucleotide to a long ssDNA (Gordienko & Rupp, 1997). In Comparison with PcrA helicase, Rep helicase, and the hepatitis C virus (HCV) RNA helicase domain whose structures have been determined (Subramanya *et al.*, 1996; Korolev *et al.*, 1997; Yao *et al.*, 1997), the amino acid sequence of ttUvrB also has large insertions between helicase motifs Ia and II and between motifs II and III (Figure 6B). However, domains 1A and 2A of ttUvrB are structurally similar to the core domains of those helicases. As shown in Figure 6A, six helicase motifs (I, Ia, II, III, V and VI) are assembled around the deep

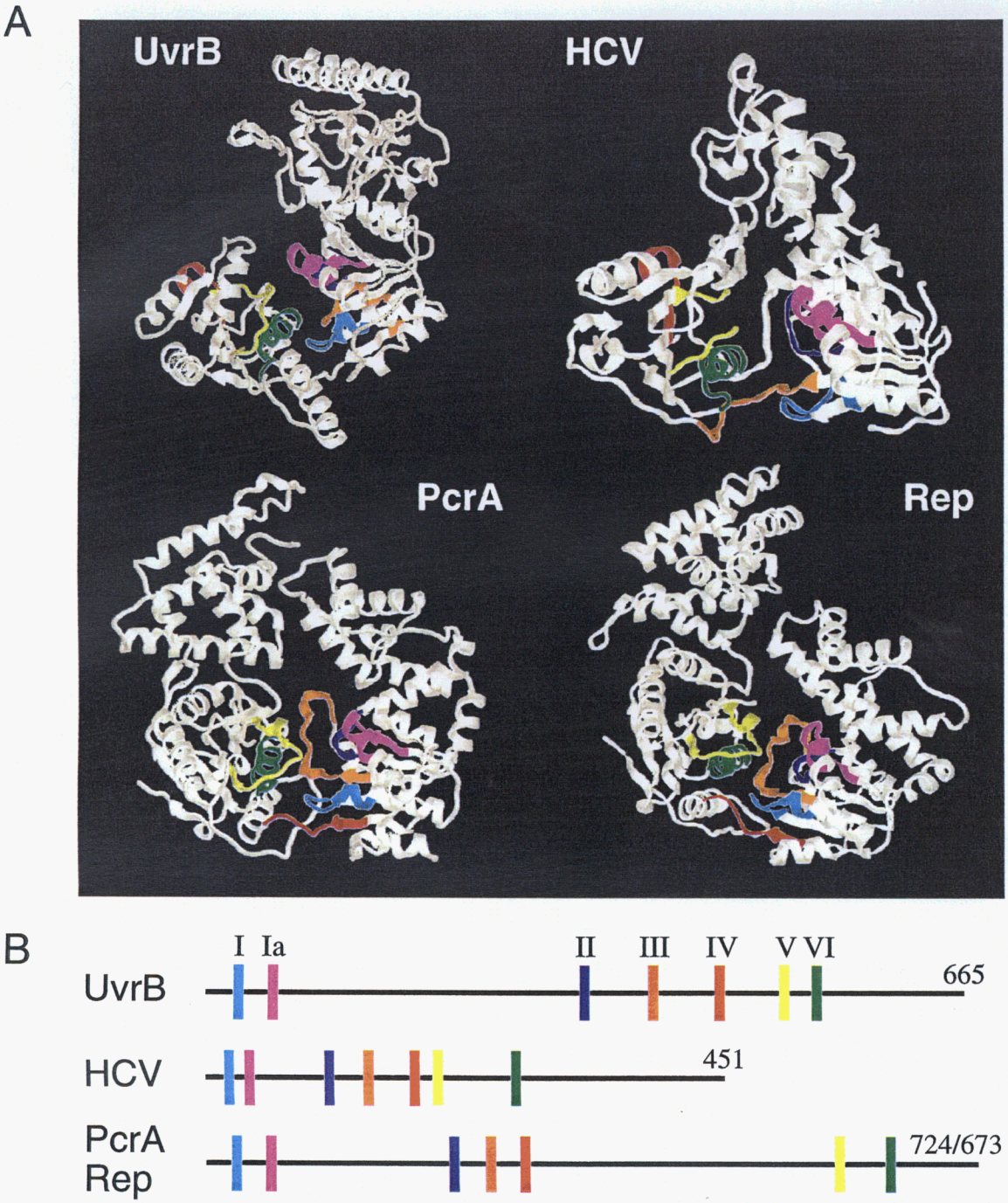

Figure 6. **Positions of the helicase motifs in ttUvrB and helicases.**

A, positions of the helicase motifs within the structures. *B*, relationship between the linear sequence and the position of the helicase motifs. The color scheme of the helicase motifs in *Panel A* is the same as that in *Panel B*. *Roman numerals* (*I–IV*) represent the helicase motifs. *HCV*, hepatitis C virus RNA helicase domain (PDB code: 1HEI); *PcrA*, *Bacillus stearothermophilus* PcrA helicase (PDB code: 1PJR); *Rep*, *E. coli* Rep helicase (PDB code: 1UAA).

cleft running between domains 1A and 2A in the structures of ttUvrB and the helicases, although the distances between the helicase motifs vary. Motifs IV in helicase superfamily I (PcrA and Rep helicases) and helicase superfamily II (ttUvrB and HCV helicase) are located at the linker between domains 1A and 2A and the top of domain 2A, respectively.

The structural similarity to helicases extends beyond the helicase motifs to encompass most of domains 1A and 2A. As shown in Figure 7, the structures of these domains share a core of parallel β-strands, α-helices and helicase motif-containing loops whose connectivity is preserved. Furthermore, in the crystal structure of ttUvrB, Lys42 in helicase motif I interacts with a sulfate ion, whose position is similar to that occupied by the γ-phosphate of the ATP analogue complexed with PcrA helicase (Subramanya *et al.,* 1996). These results suggest that domains 1A and 2A of ttUvrB are responsible for dsDNA unwinding coupled with ATP hydrolysis. In the following description, domains 1A and 2A are referred to as helicase domains 1A and 2A. Apart from the structural similarities to helicase and related ATP-hydrolyzing enzymes, UvrB protein has no other known structural neighbor.

Purification of Tryptic Fragments

The crystal structure of ttUvrB provides detailed insight into the domain architecture of UvrB protein and, through comparison with the helicases, indicates some structural similarity to helicases. Like helicases, UvrB protein shows ATPase activity and DNA binding. Furthermore, in the nucleotide excision repair system, UvrB protein is involved in most of the repair process, such as the recognition of damaged DNA and

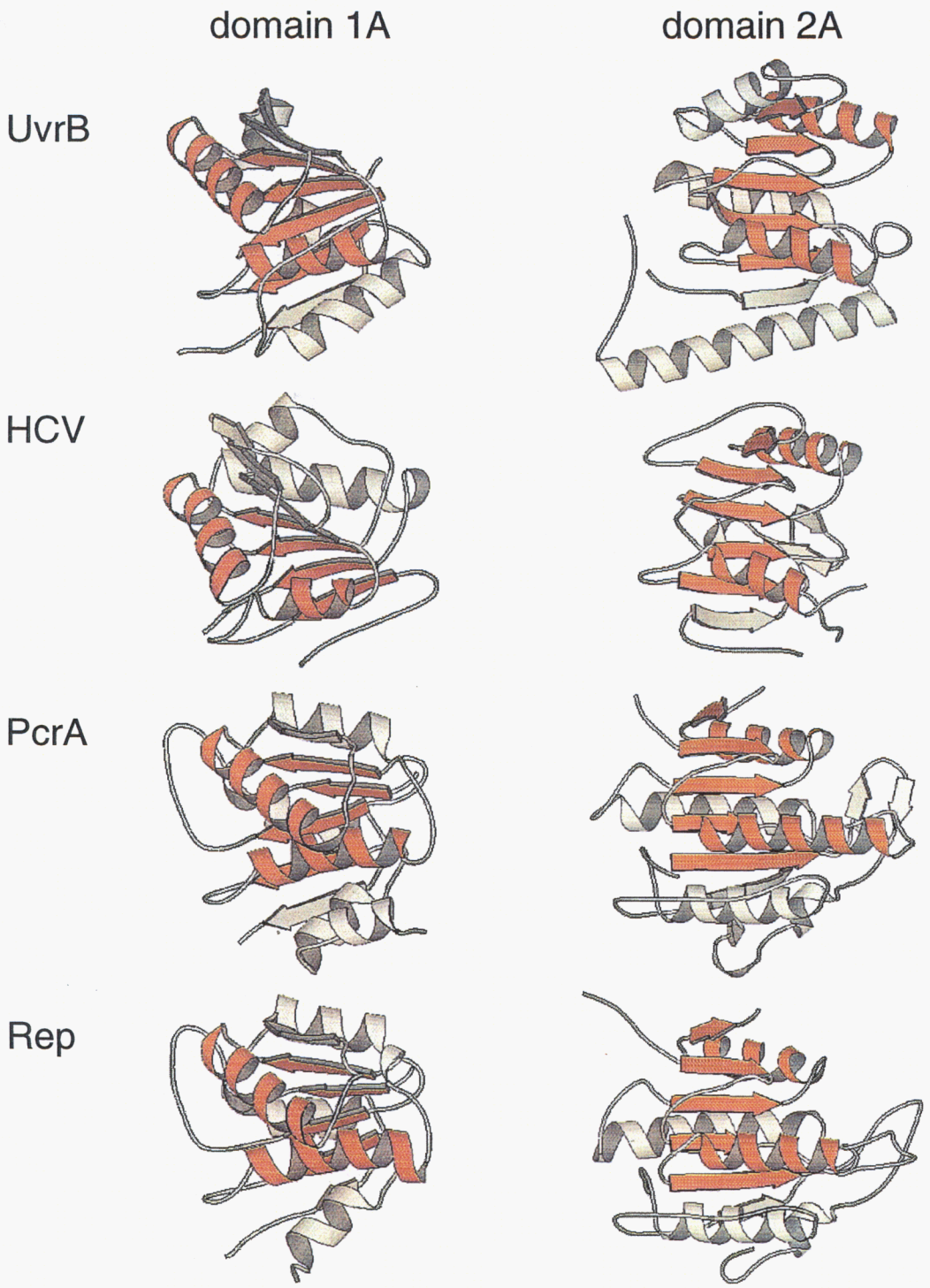

Figure 7. **Common secondary structural elements in helicase domains 1A and 2A of ttUvrB and helicases.**

The common α-helices and β-strands of domains 1A and 2A in all the molecules are shown in orange. This figure was produced using Molscript (Kaulis, 1991).

the incision of DNA by interacting with UvrA and UvrC proteins. To reveal the structure-function relationship of this multifunctional protein, it is necessary to analyze the individual activities in addition to the three-dimensional structure. In the case of a multifunctional protein, the respective functions are often associated with individual domains (Doolittle, 1995). In this regard, the proteolytic fragments of ttUvrB, which lack some of the domains, are useful for elucidating the functions of each domain.

Under mild conditions, endoproteases will cleave a protein preferentially at the sites exposed to the solvent, which are often within interdomain linkers. A time course of digestion by trypsin is shown in Figure 8. The fragments produced by trypsin digestion were observed as relatively discrete bands on the gels. The four major tryptic fragments with masses of 66, 54, 51, and 40 kDa are referred to as T-1, T-2, T-3, and T-4, respectively. Under extremely mild conditions (ttUvrB to trypsin ratio of 200:1, for 1 min), only T-1 and a small fragment (referred to as Tc) were obtained.

To examine the activities of the tryptic fragments, I attempted to purify the tryptic fragments. Digestion of ttUvrB with trypsin at a protein to protease molar ratio of 400:1 for 3 min produced only two fragments, T-1 and Tc. These fragments were separated by ion-exchange chromatography using a MonoQ HR5/5 column. As shown in Figure 9A, peak 1 contained only the T-1 fragment. Peak 2 was shown to contain only the Tc fragment by N-terminal amino acid sequencing and native PAGE (the Tc band being too small to detect by SDS-PAGE: see Figure 11). The digest containing the T-1, T-2, and T-3 fragments (ttUvrB to trypsin molar ratio of 50:1 for 60 min) was also applied to the MonoQ HR5/5 column, but only two peaks were obtained (Figure 9B). Peak 1 contained only the T-3 fragment, whereas peak 2

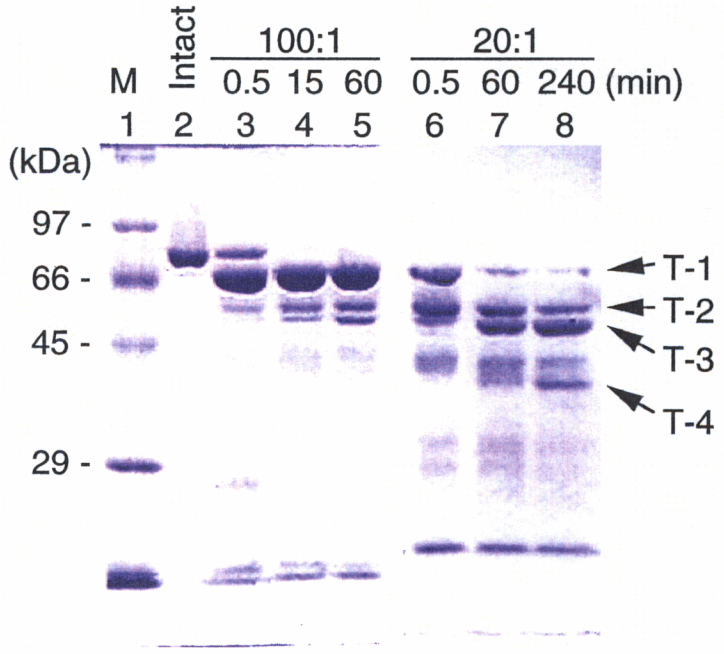

Figure 8. **Limited proteolysis of ttUvrB.**

ttUvrB (5 μM) was treated with trypsin at a protein to protease molar ratio of 100:1 (*lanes 3–5*) or 20:1 (*lanes 6–8*). *Lane 1*, molecular mass markers: rabbit muscle phosphorylase b (97 kDa), bovine serum albumin (66 kDa), egg white albumin (45 kDa) and bovine carbonic anhydrase (29 kDa); *lane 2*, undigested ttUvrB; *lanes 3–8*, digests incubated for the time indicated above each lane.

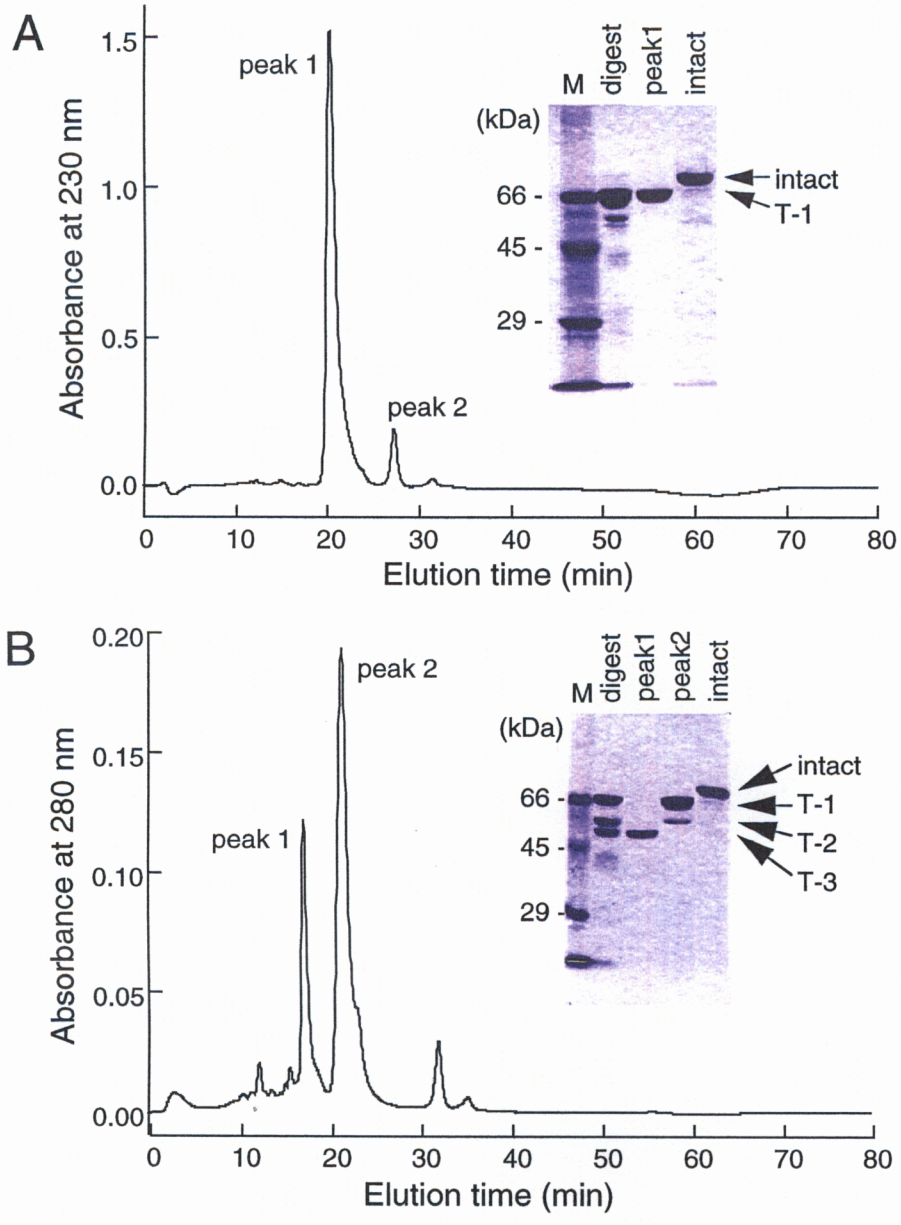

Figure 9. **Isolation of tryptic fragments.**

A, isolation of T-1 and Tc fragments. *Peak 1* and *peak 2* contained the T-1 and Tc fragments, respectively. *Inset*: *M*, molecular mass markers;*digest*, the reaction products before loading onto the MonoQ column; *peak 1*, eluate fraction containing *peak 1*; *intact*, undigested ttUvrB. *B*, isolation of the T-3 fragment. *Peak 1* contained the T-3 fragment, whereas *peak 2* contained the T-1 and T-2 fragments. *Inset*: *M*, molecular mass markers;*digest*, the reaction products before loading onto the MonoQ column; *peak 1*, eluate fraction containing *peak 1*; *peak 2*, eluate fraction containing *peak 2*; *intact*, undigested ttUvrB.

contained both the T-1 and T-2 fragments. The T-1 and T-2 fragments could not be separated by other chromatographic techniques. Therefore, three tryptic fragments, namely T-1, T-3 and Tc, could be purified.

To identify the cleavage sites, the N-terminal amino acid sequences of each fragment were determined. Both the T-1 and T-3 fragments started at Thr2, which is also the N-terminal residue of the intact protein. This indicates that for the T-1 and T-3 fragments the cleavages occurred in the C-terminal region. Based on this result, the molecular mass of the fragments and the crystal structure of ttUvrB, its fragments, T-1 and T-3 are considered to consist of the 1A, 2A, α-, and β-domains and 1A, α- and β-domains, respectively. The digest containing the T-1 and Tc fragments was directly subjected to sequencing, and shown to contain two peptides, one beginning from Thr2 and the other from Ala589. This result indicates that the Tc fragment starts at Ala589. In the crystal structure, the C-terminal 82 residues (584-665) were disordered. Therefore, the Tc fragment corresponds to the C-terminal region, whose structure has not been determined.

To examine the structural properties of the purified T-1, T-3, and Tc fragments, the CD spectra were measured. As shown in Figure 10, the far UV CD spectra of ttUvrB and its fragments had negative double maxima at around 210 and 220 nm, which are characteristic of an α-helical structure. The α-helix content of ttUvrB and its fragments T-1, T-3, and Tc was estimated to be about 47%, 38%, 31%, and 51%, respectively, using the method of Chen *et al.* (Chen *et al.*, 1972). These results indicate that the purified fragments retain their secondary structures. Although the C-terminal region corresponding to the Tc fragment was completely disordered in the

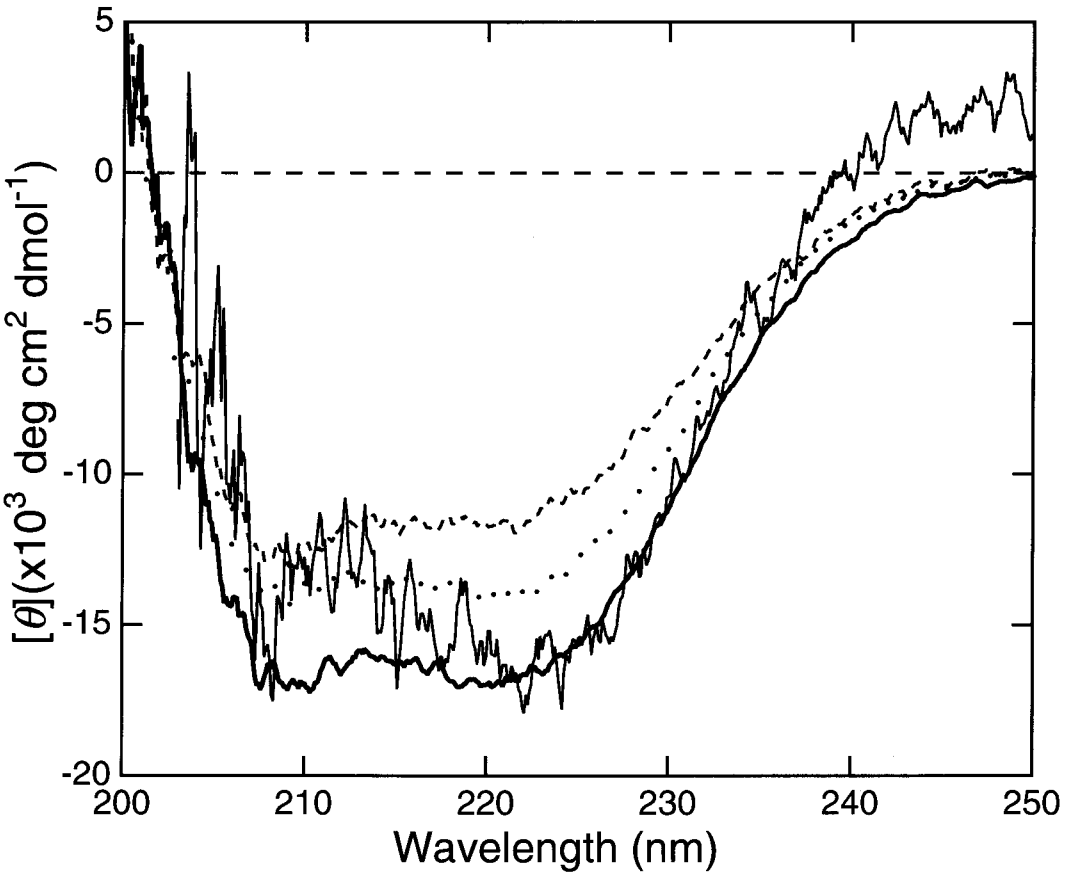

Figure10. **Far UV CD spectra of ttUvrB and its fragments.**

The CD spectra of 2 μM ttUvrB(*thick line*), T-1 (*dotted line*), T-3 (*dashed line*), and Tc fragments (*thin line*) were measured between 200 and 250 nm at 25°C.

crystal structure, the Tc fragment forms an α-helical structure, suggesting that the C-terminal region forms one domain by itself. In the following description, this C-terminal domain is referred to as the C2 domain.

DNA Binding

To reveal the functions of the domains revealed by the crystal structure of ttUvrB, the activities of the purified tryptic fragments were assayed. The ability of ttUvrB and its fragments to bind DNA was assayed by native PAGE (Davis, 1964). Adding poly(dT) to ttUvrB caused the band of the protein to smear upwards (Figure 11). This decrease in mobility indicates that ttUvrB can bind to ssDNA; however, it should be noted that a discrete band could not be obtained despite of the high concentration (1 mM) of poly(dT). This observation indicates that ttUvrB gradually dissociated from the poly(dT) during the electrophoresis. Thus, DNA binding of ttUvrB is considered to be relatively weak.

Adding poly(dT) to the T-1 fragment, which lacks the C2 domain, also caused the band to smear upwards, but less extensively. This observation indicates that the T-1 fragment can bind to ssDNA but with a lower affinity than that of the intact protein. This result also suggests that the C2 domain is involved in, but is not essential for, binding to ssDNA. In contrast, poly(dT) caused no change in the mobility of T-3 or Tc bands, indicating that these fragments are unable to bind to ssDNA. Since the difference between the T-1 and T-3 fragments is the presence of domain 2A, it appears that domain 2A is essential for ssDNA binding.

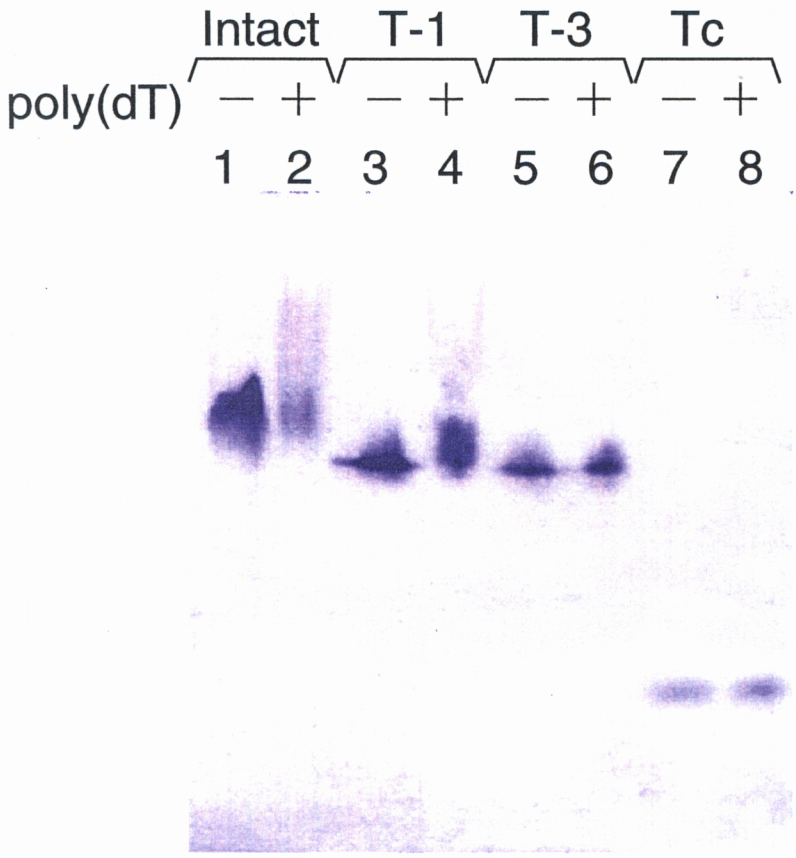

Figure 11. **Binding of ttUvrB and its fragments to ssDNA.**

ttUvrB and its fragments (5 μM) were incubated with 1 mM poly(dT) for 1 h at 25°C. Samples were electrophoresed on a 7.5% polyacrylamide gel and stained with CBB. *Lanes 1, 3, 5,* and *7,* without poly(dT); *lanes 2, 4, 6,* and *8*, with 1 mM poly(dT).

ATPase Activity

It has been demonstrated that ttUvrB has ATPase activity in the absence of ttUvrA (Kato *et al.*, 1996). The K_m and k_{cat} of ttUvrB, T-1, and T-3 fragments were determined at 25°C using an enzyme-coupled spectrophotometric method (Figure 12 and Table 2). The T-1 fragment had almost the same ATPase activity as the intact protein, whereas the T-3 fragment had no activity. These results suggest that domain 2A is indispensable for ATPase activity

Since the ATPase activity of intact ttUvrB is known to be stimulated by ssDNA (Kato, *et al*, 1996), the activity of the fragments with poly(dT) was determined (Figure 12 and Table 2). Although the activity of the T-1 fragment was also stimulated by poly(dT), its k_{cat} in the presence of poly(dT) was lower than that of the intact protein. This observation suggests that the C2 domain is not required for ATPase activity but is concerned with the stimulation of the ATPase activity by ssDNA. The T-3 fragment had no ATPase activity in the presence of ssDNA.

Interaction with ttUvrA

Interaction between ttUvrB or its fragments and ttUvrA was examined by native PAGE in the presence of ATP and $MgCl_2$ (Figure 13). ttUvrA and ttUvrB showed different mobilities on a polyacrylamide gel (lanes 1 and 2). When ttUvrA was added to ttUvrB at a molar ratio of 1:1, one band only was observed close to the position of ttUvrA (lane 4). The band at the original position of ttUvrB had disappeared, so it is probable that ttUvrB was shifted to the upper position by forming a complex with ttUvrA. The positions of ttUvrA and the ttUvrA-ttUvrB complex could not be

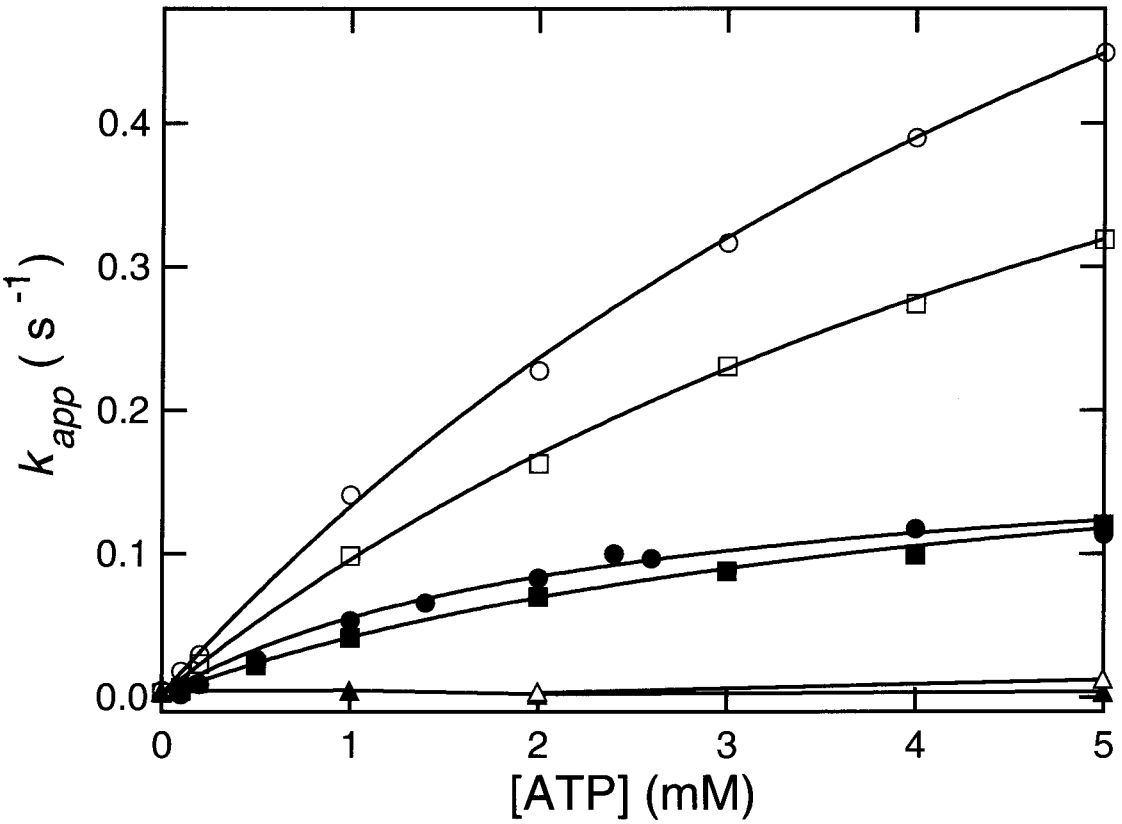

Figure 12. **ATPase activity of ttUvrB and its fragments.**

ATPase activity was assayed as described in Materials and Methods. The ordinate (k_{app}) represents the apparent rate constant. Symbols: *closed circles*, ttUvrB; *open circles*, ttUvrB with 50 μM poly(dT); *closed squares*, T-1 fragment; *open squares*, T-1 fragment with 50 μM poly(dT); *closed triangles*, T-3 fragment; *open triangles*, T-3 fragment with 50 μM poly(dT).

Table 2. **Kinetic parameters for ATPase activity of ttUvrB and its fragments.**

	None		Poly(dT)	
	K_m (mM)	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})
ttUvrB	2.3	0.18	7.6	1.13
T-1	4.5	0.22	7.8	0.78
T-3	ND[a]	ND	ND	ND

[a]ND, not detectable.

The values were calculated from Figure 12.

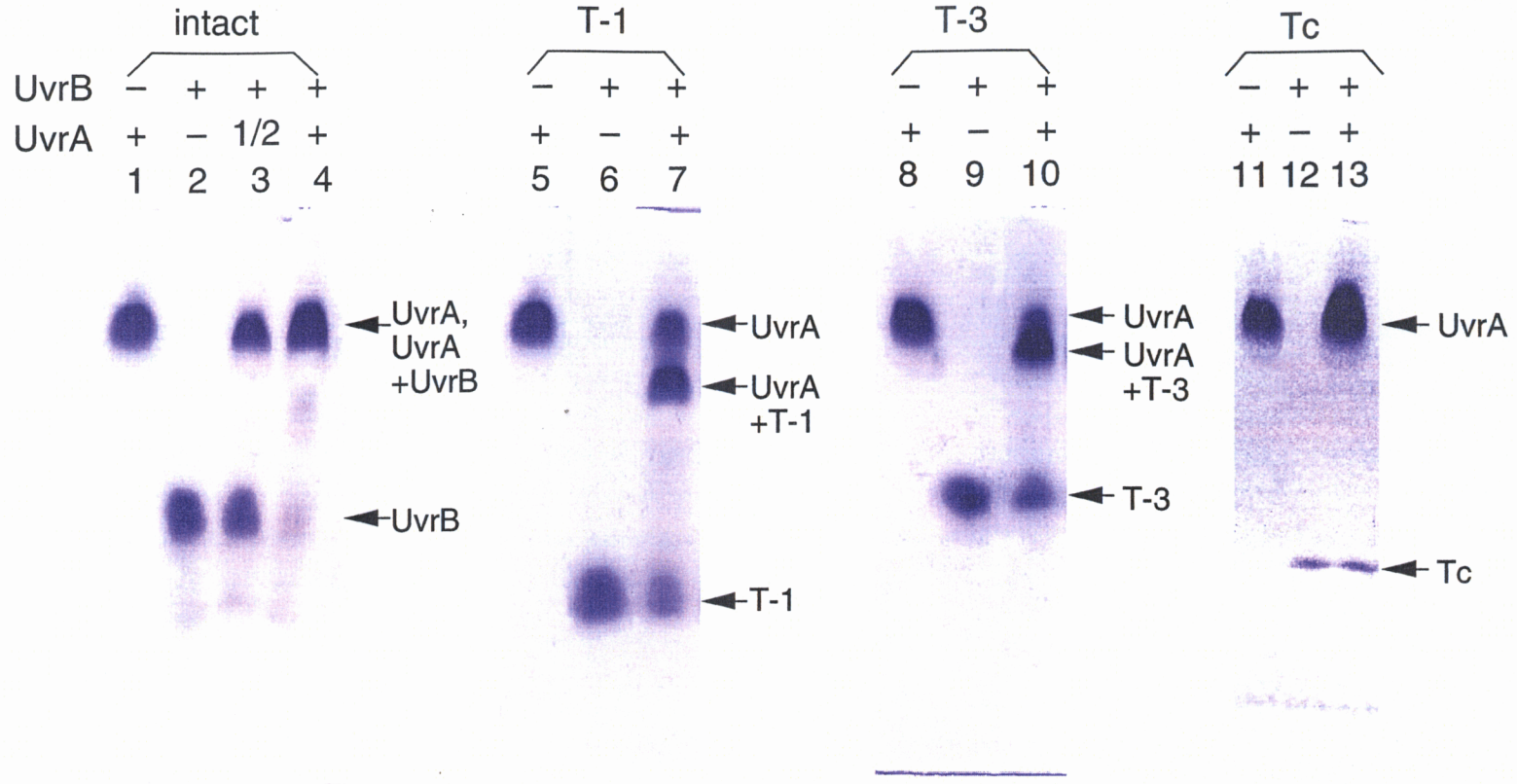

Figure 13. **Interaction between ttUvrA and ttUvrB or its fragments.**

All reaction mixtures contained 5 mM ATP and 10 mM $MgCl_2$. The samples were electrophoresed on a 4% polyacrylamide gel in the presence of 5 mM ATP and 10 mM $MgCl_2$. Proteins were stained with CBB. *Lanes 1, 5*, 8 and *11*, 5 μM ttUvrA; *lane 2*, 5 μM ttUvrB; *lane 3*, 2.5 μM ttUvrA and 5 μM ttUvrB; *lane 4*, 5 μM ttUvrA and 5 μM ttUvrB; *lane 6*, 5 μM T-1 fragment; *lane 7*, 5 μM ttUvrA and 5 μM T-1 fragment; *lane 9*, 5 μM T-3 fragment; *lane 10*, 5 μM ttUvrA and 5 μM T-3 fragment; *lane 12*, 5 μM Tc fragment; *lane 13*, 10 μM ttUvrA and 5 μM Tc fragment.

separated under these electrophoresis conditions.

When ttUvrA was added to the T-1 fragment, the band at the position of the T-1 fragment became weaker, and a new band appeared below the position of the ttUvrA band (lane 7). This result indicates that the T-1 fragment can interact with ttUvrA. However, a weak band was still detected at the original position of the T-1 fragment in the presence of ttUvrA, even though the band of the intact protein disappeared under the same conditions. Therefore, affinity of the T-1 fragment for ttUvrA is considered to be weaker than that of the intact protein. A similar result was observed for the T-3 fragment indicating that it binds to ttUvrA, but with a lower affinity than the intact protein (lane 10). In contrast, the Tc fragment caused no change in the mobility in the presence of ttUvrA, indicating that it is unable to bind to ttUvrA (lane 13). These results suggest that ttUvrA binds to the region within the 1A, α- and β-domains. An interaction between ttUvrA and ttUvrB or its fragments was not detected in the absence of ATP (data not shown).

To further examine the interaction of ttUvrB with ttUvrA, the effect of ttUvrA on the ATPase activity of ttUvrB or its fragments was investigated. As shown in Figure 14A, the ATPase activity of the mixture of ttUvrA and ttUvrB was higher than the sum of the activities of ttUvrA and ttUvrB alone. When ttUvrA and the T-1 fragment, which lacks the C2 domain, were mixed, the ATPase activity was higher than the calculated value (Figure 14A). However the extent of the stimulation was less than that of the mixture of ttUvrA and intact ttUvrB, even though in the absence of ttUvrA the T-1 fragment had the same level of ATPase activity as the intact protein (Figure 12 and Table 2). These results suggest that the C2 domain is involved in the stimulation

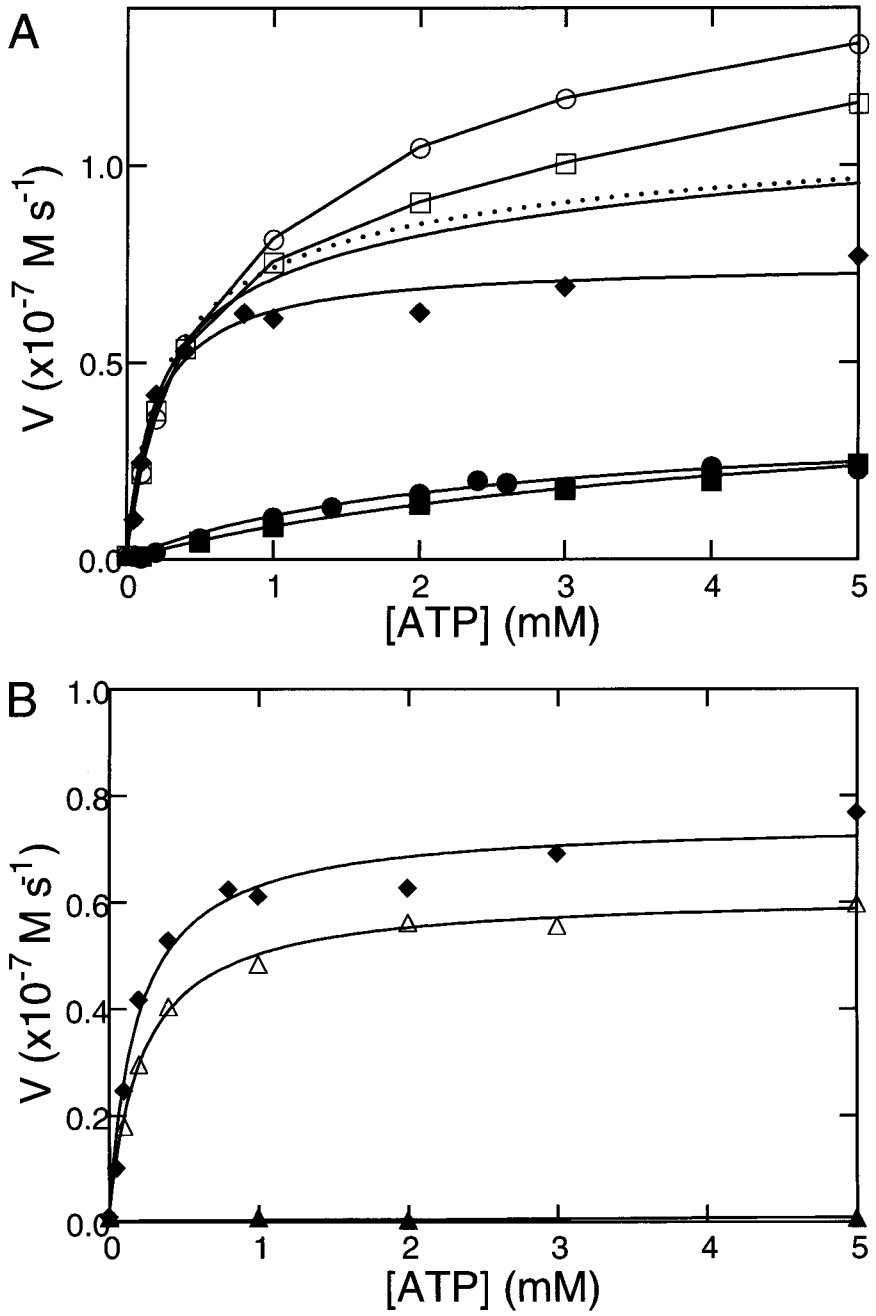

Figure 14. **ATP hydrolysis in the presence of ttUvrA and ttUvrB or its fragments.** The ordinates (*V*) represent the rates of ATP hydrolysis. *A*, *closed circles*, 0.2 μM ttUvrB; *open circles*, 0.2 μM ttUvrB and 0.4 μM ttUvrA; *closed squares*, 0.2 μM T-1 fragment; *open squares*, 0.2 μM T-1 fragment and 0.4 μM ttUvrA; *closed diamonds*, 0.4 μM ttUvrA. The *dotted line* is calculated from the ATPase activities of ttUvrA and ttUvrB and the *dashed line* is calculated from those of ttUvrA and the T-1 fragment. *B*, *closed triangles*, 0.2 μM T-3 fragment; *open triangles*, 0.2 μM T-3 fragment and 0.4 μM ttUvrA; *closed diamonds*, 0.4 μM ttUvrA.

of the ATPase activity of ttUvrA-ttUvrB complex. Involvement of the C2 domain in ATPase activity was also seen in the activation by ssDNA. Addition of the T-3 fragment, which lacks the ATPase activity, to ttUvrA reduced the total ATPase activity (Figure 14B). These results indicate that formation of the UvrAB complex suppresses the ATPase activity of UvrA and stimulate that of ttUvrB.

Discussion

Protein-Protein Interaction

The crystal structure of ttUvrB and the structural properties of the proteolytic fragments revealed that UvrB protein consists of five domains: 1A, α, β, 2A and C2. The structures of the helicase domains 1A and 2A resemble the core structure of helicases and related ATP-hydrolyzing enzymes, whereas the structures of the α- and β-domains have no similarity to other known structures. The C2 domain is considered to form an α-helical structure. The functional properties of ttUvrB and its fragment indicate that ttUvrB hydrolyzes ATP and binds to ssDNA like a helicase, and that these activities require the helicase domains 1A and 2A. In addition to these helicase-related activities, UvrB protein plays multiple roles in the nucleotide excision repair system, involving many proteins. Therefore, the addition of three domains, α, β, and C2 to the helicase domains 1A and 2A may enable UvrB protein to function in the repair system.

During the repair process, UvrB protein forms a complex with UvrA or UvrC proteins. TRCF, which is involved in transcription coupled repair, also interact with UvrA protein (Figure 15). As shown in Figure 16A, the amino acid sequence of residues 154–242 of ttUvrB is homologous to residues 86–172 of TRCF. The crystal structure revealed that the corresponding region forms one domain, the β-domain, suggesting that the β-domain is a UvrA-binding domain. This notion is supported by the observation that the tryptic fragments containing the β-domain can bind to ttUvrA (see Figure 11). In *E. coli*, it was also indicated that the deletion mutant containing

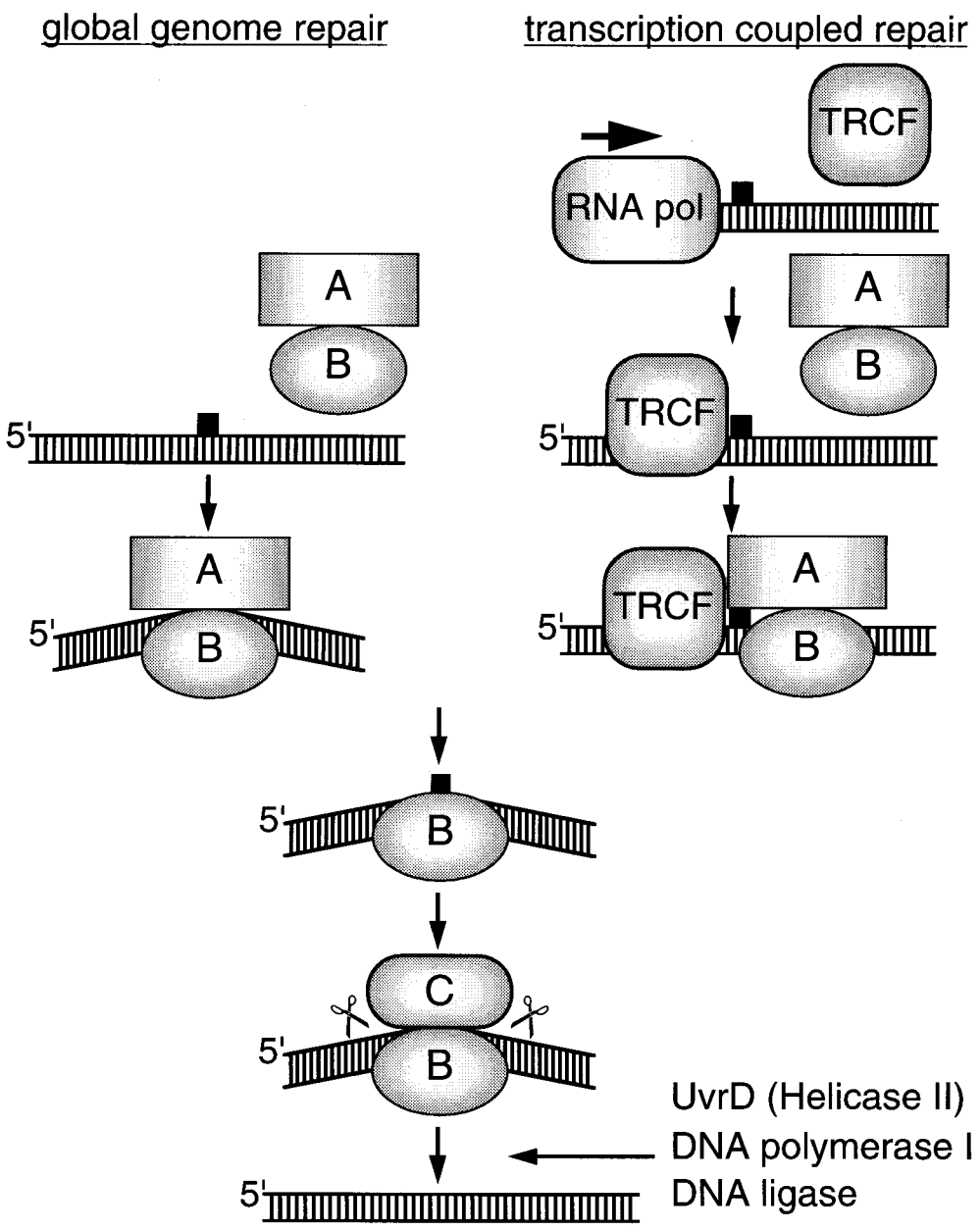

Figure 15. **Molecular mechanisms of nucleotide excision repair.**

Transcription-independent (left) and transcription-dependent (light) forms are shown. In transcription coupled repair, the stalled RNA polymerase complex on a damaged DNA is recognized by TRCF. TRCF releases RNA polymerase and the truncated transcript while simultaneously recruiting the UvrAB complex. TRCF and UvrA protein dissociate leaving a stable UvrB-DNA complex. The subsequent steps are identical in the transcription-independent and transcription-coupled modes. *A*, UvrA protein; *B*, UvrB protein; *C*, UvrC protein; *RNA pol*, RNA polymerase; *TRCF*, transcription-repair coupling factor.

A

```
ttUvrB 154:RNLVVERGKPYPREVLLERLLELGYQRNDIDLSPGRFRAKGEVLEIFPAY
ttTRCF  86:WRLLLEVGRAYPREALLSRLLKLGYARDE------DYRVLGEVVELGEVR
ecUvrB 151:MMLHLTVGMIIDQRAILRRLAELQYARNDQAFQRGTFRVRGEVIDIFPAE
ecTRCF 126:HALVMKKGQRLSRDALRTQLDSAGYRHVDQVMEHGEYATRGALLDLFPMG

ttUvrB      -ETEPIRVELFGDEVERISQVHPVTG-ERLRELPG----FV-----LFPA:242
ttTRCF      -------LEFFGDELERLVVRGEERRRHVLLPKPGKAEGFTSKKVLHFPG:172
ecUvrB      SDDIALRVELFDEEVERLSLFDPLTG-QIVSTIPR----FT-----IYPK:246
ecTRCF      -SELPYRLDFFDDEIDSLRVFDVDSQ-RTLEEVEA----IN-----LLPA:214
```

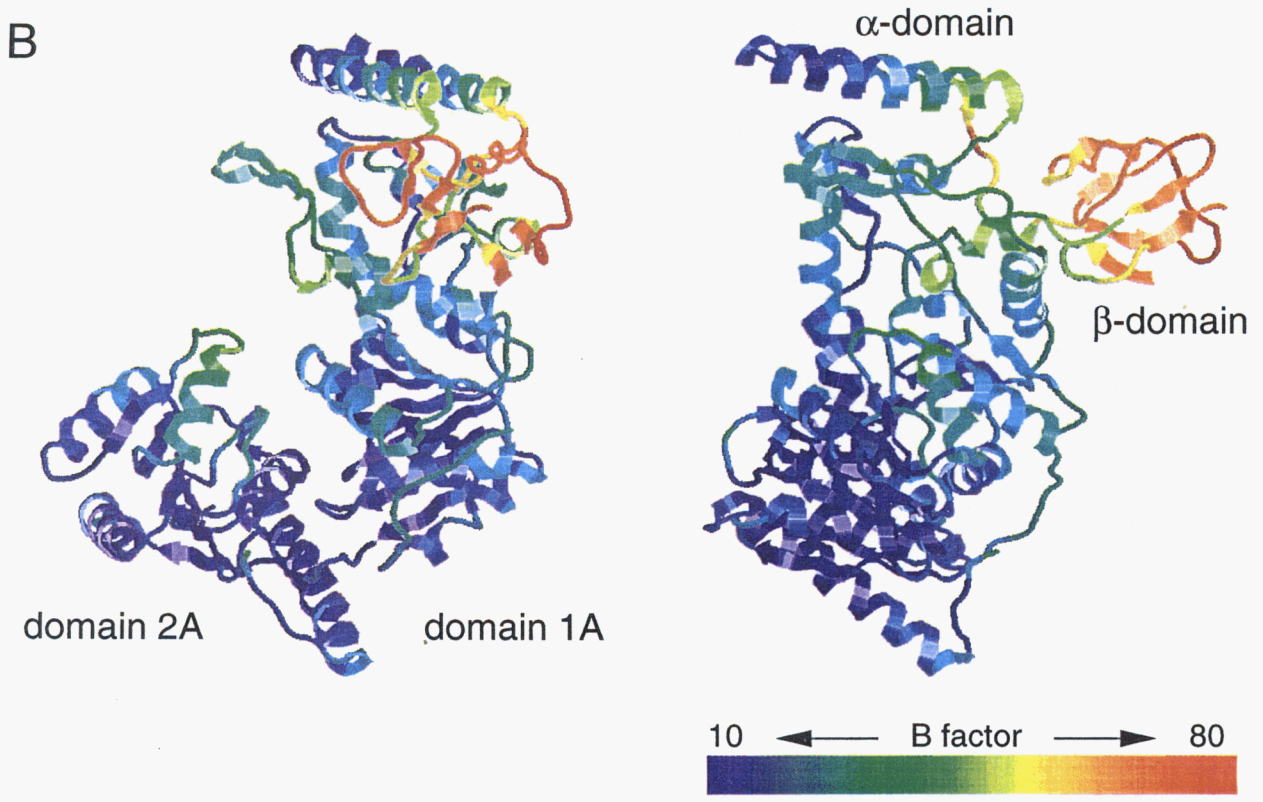

Figure 16. **Features of the β-domain.**

A, sequence similarity between UvrB and TRCF from *T. thermophilus* HB8 and *E. coli*. The identical or homologous residues in more than three sequences are indicated by red letters. Deleted amino acid residues are indicated by bars (–). *ttUvrB*, UvrB from *T. thermophilus* HB8; *ttTRCF*, TRCF from *T. thermophilus* HB8; *ecUvrB*, UvrB from *E. coli*; *ecTRCF*, TRCF from *E. coli*. *B*, distribution of the B-factor values in the structure of ttUvrB. The color scale represents the B-factor value. The view is the same as in Figure 4.

the β-domain interacts with UvrA (Hsu *et al.*, 1995). As shown in Figure 16B, the temperature factors for the β-domain are high compared with those of the other domains. Furthermore, a comparison of ttUvrB and bcUvrB shows that the orientations of the β-domain in ttUvrB and bcUvrB are different (see Figure 5). These results suggest fluctuation of the β-domain in the molecule. This domain motion may change the relative orientation of UvrA and UvrB proteins in the UvrAB complex during the repair process.

The C2 domain was shown to be involved in the stimulation of ATPase activity by ssDNA or by formation of a ttUvrA-ttUvrB complex (see Figure 12 and Figure 14). In the crystal structure the loop that connects domain 2A to the disordered C2 domain is located close to the cleft between helicase domains 1A and 2A. As the ATP-binding site is situated at the bottom of this cleft, the conformational change or rearrangement of the C2 domain upon ssDNA or ttUvrA binding may affect the ATPase activity of ttUvrB. Recently the structure of the 55 C-terminal residues of *E. coli* UvrB has been studied by X-ray crystallography and NMR (Sohi *et al.*, 2000; Alexandrovich *et al.*, 1999). The results indicate that this fragment adopts a helix-loop-helix fold. This is consistent with the observation that the Tc fragment, which corresponds to the C2 domain, forms an α-helical structure (see Figure 10). Furthermore, the fragment of the 55 C-terminal residues was shown to form a dimer in a head-to-head manner (Sohi *et al.*, 2000). Since this region is homologous to the corresponding part of UvrC protein (Moolenaar *et al.*, 1995), it has been proposed that the interaction between UvrB and UvrC proteins is similar to that observed in the dimer of this C-terminal fragment (Sohi *et al.*, 2000). Therefore, the C2 domain,

which contains the 55 C-terminal residues, is considered to be a UvrC-binding domain.

Structural and Functional Similarity to Helicase

UvrB protein has seven conserved regions of helicase motifs and belongs to helicase superfamily II (Gorbalenya & Koonin, 1993). The crystal structure of ttUvrB revealed that the helicase domains 1A and 2A of ttUvrB are similar to those of helicase in terms of not only the positions of the helicase motifs, but also the connectivity of the secondary structural elements. As shown in Figure 17, in all the crystal structures of the helicase-DNA complexes, helicase domain 1A and 2A are involved in binding to ssDNA (Velankar *et al.*, 1999; Korolev *et al.*, 1997; Kim *et al.*, 1997). This is consistent with the observation that domain 2A is indispensable for ssDNA binding (see Figure 11).

In the PcrA-DNA complex, the cleft between helicase domains 1A and 2A was shown to close upon binding to ATP and to open upon hydrolysis of ATP (Velankar *et al.*, 1999). It has been proposed that ATP hydrolysis is coupled to such domain motion, which causes not only translocation along DNA, but also separation of dsDNA into the strands of ssDNA (Velankar *et al.*, 1999). The properties of the proteolytic fragments of ttUvrB indicate that both of helicase domains 1A and 2A are necessary for ATP hydrolysis. Superimposition of the 1A domains of ttUvrB and bcUvrB shows that the two relative orientations of helicase domains 1A and 2A are different. These results suggest that ATP hydrolysis of UvrB protein is also coupled to domain motion. However, UvrB protein alone has no helicase activity, whereas in the UvrAB complex, UvrB protein unwinds a short range of DNA around the lesion and

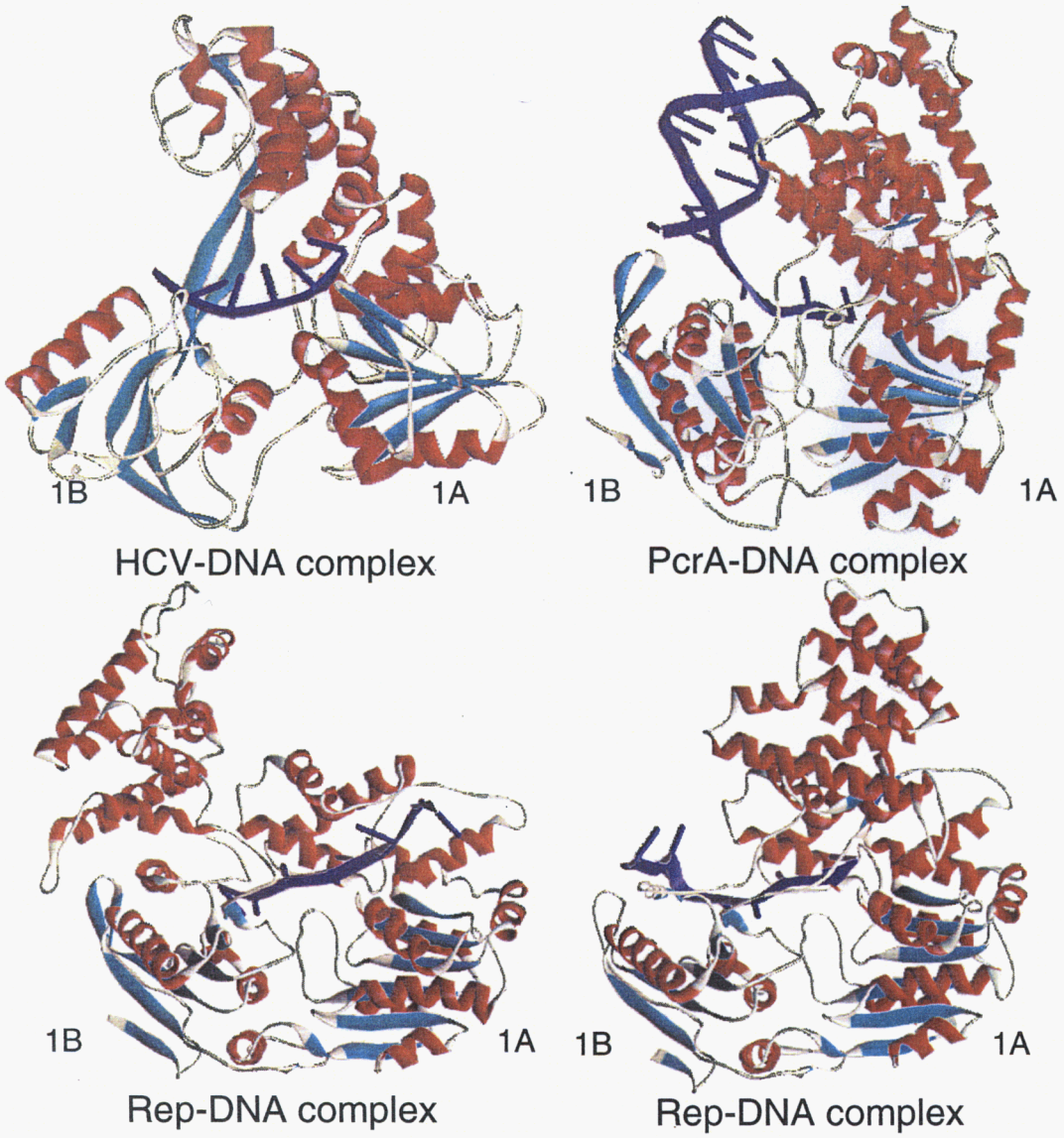

Figure 17. **Crystal structures of the helicase-DNA complexes.**

Secondary structural elements are indicated in cyan for β-strands and red for α-helices. Intervening regions are shown in gray. The DNA is colored blue. *HCV*, hepatitis C virus RNA helicase domain (PDB code: 1A1V); *PcrA*, *B. stearothermophilus* PcrA helicase (PDB code: 3PJR); *Rep*, *E. coli* Rep helicase (PDB code: 1UAA). This figure was produced using WebLab ViewerLite (MSI Inc.).

kinks the DNA (Gordienko & Rupp, 1997; Zou & Van Houten, 1999; Shi *et al.*, 1992). Therefore, in the case of UvrB protein, the domain motion upon ATP hydrolysis may be utilized for unwinding the DNA, but not translocation along the DNA. This would result in low processivity of unwinding activity and recognition of the DNA damage.

Model for Damage Recognition by UvrB

The most remarkable feature of the nucleotide excision repair system is its extremely broad substrate specificity: the system is able to recognize DNA damage of differing size and shape, such as an apurinic/apyrimidinic site, pyrimidine dimer and psoralen adduct (Van Houten & Snowden, 1993). Unlike an enzyme with strict substrate specificity, specific hydrogen bonding and ionic interaction between the protein and DNA damage cannot be the basis for recognition of these lesions. Conformational distortion of DNA structure caused by certain lesions has also been considered a crucial element for damage recognition. However, mismatches and naturally bent DNA are not substrates of this repair system (Selby & Sancar, 1990). It has also been demonstrated that the range of substrate specificity includes base damage, which is not considered distortive (Van Houten & Snowden, 1993). Therefore, different features appear to be involved in DNA damage recognition.

UvrA protein plays an important role in damage recognition as a molecular matchmaker: it brings UvrB protein and damaged DNA together, promotes their association, and then leaves the complex (see Figure 15; Sancar, 1996). Recently it has been reported that UvrB protein alone binds in a damage-dependent manner to

substrates truncated on the 5' side of the damage, leading to a stable UvrB-DNA complex (Moolenaar *et al.*, 2000). This indicates that not only UvrA protein but also UvrB protein has a damage-recognition site within the molecule. The present results provide clues to the mechanism of damage recognition by UvrB protein.

As shown in Figure 18, ttUvrB has 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 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 (Figure 17). 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. It has been suggested that stacking interaction is responsible for several functions: bending of the DNA backbone, such as Trp250 of Rep helicase (Korolev *et al.*, 1997), stabilization of the unwound form of DNA, such as Phe626 of PcrA helicase (Velankar *et al.*, 1999), and sliding of the DNA, such as F64 of PcrA helicase (Velankar *et al.*, 1999). Furthermore, the UvrB-DNA complex is stable to high ionic strength (Orren & Sancar 1990) and the DNA in this complex is kinked by 130° (Shi *et al.*, 1992). These observations suggest that the interaction between UvrB and DNA involves hydrophobic interaction and/or intercalation of the aromatic residues into the duplex via 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 aromatic

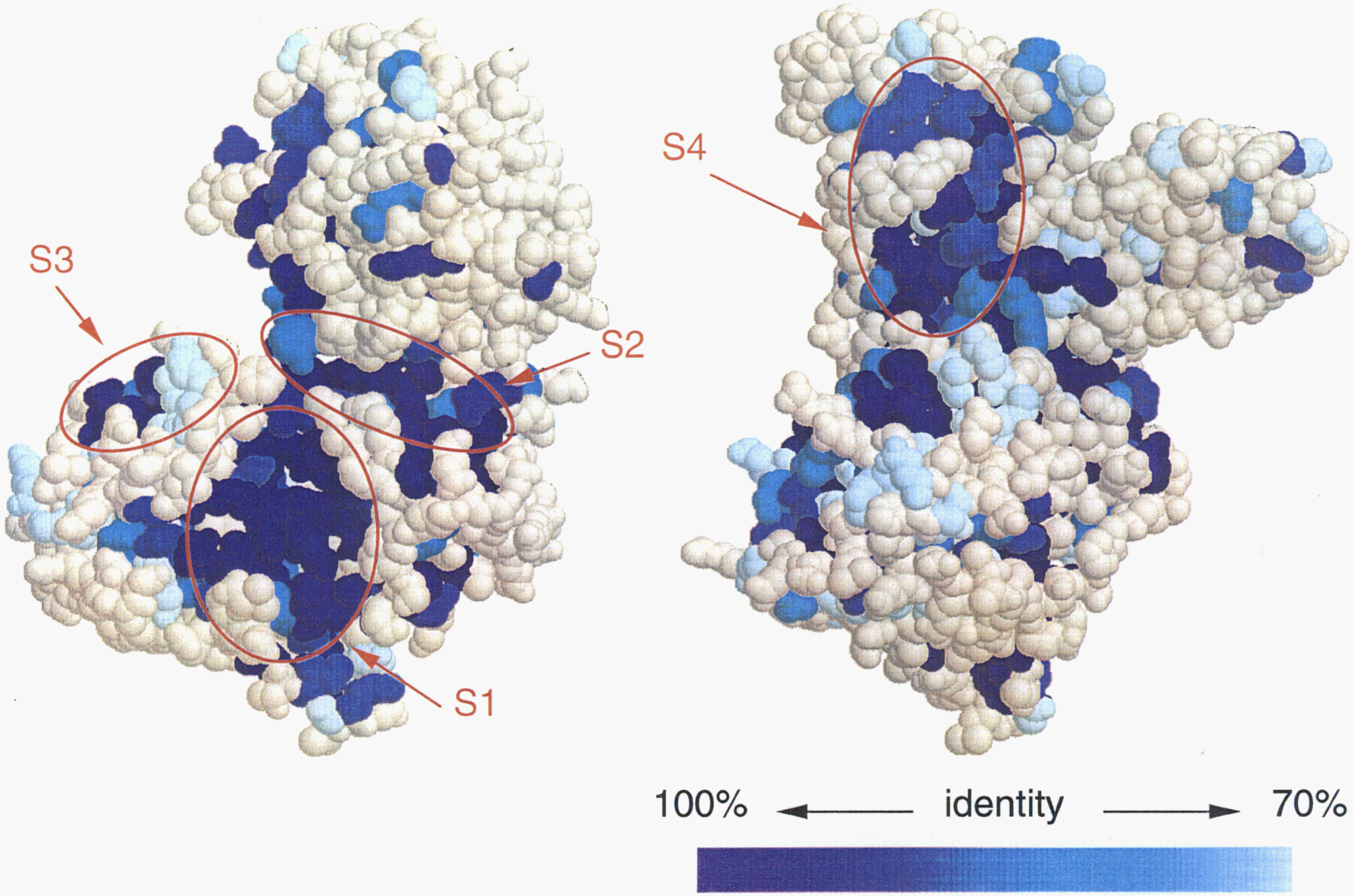

Figure 18. **Distribution of the conserved residues between species.**

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. The view is the same as in Figure 4.

residues (Tyr85, Tyr92, and Tyr469) and histidine (His471). Conservation of these residues suggests that these surface regions interact with ssDNA by stacking. It has been reported that DNA unwinding introduced by the UvrAB complex localizes within 1-3 bp (Gordienko & Rupp, 1997) or 6 bp (Zou & van Houten, 1999) around a lesion. If the S2 and S3 regions interact with an identical DNA strand, more than 6 bases of the ssDNA region are 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 (Figure 18). The electrostatic potential of this surface region is slightly positive due to the presence of Arg347, Arg352, and Lys353. This implies that the region is involved in binding to dsDNA. Based on these structural features of ttUvrB and the crystal structure of the PcrA-DNA complex, which is only the structure of a helicase-dsDNA complex (Velanker, *et al.*, 1999), a model for the ttUvrB-DNA complex is proposed (Figure 19). In this model, the DNA is partially unwound around a lesion. The region of ssDNA interacts with the S2 and S3 regions of helicase domains 1A and 2A, whereas the region of dsDNA binds to the S4 region of the α-domain.

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

Figure 19. **Putative model of binding of ttUvrB to DNA.**

Secondary structural elements are indicated in cyan for β-strands and red for α-helices. Intervening regions are shown in gray. The DNA is colored blue. The positions of the most conserved residues are indicated by the orange circles. This figure was produced using WebLab ViewerLite (MSI Inc.).

change in the UvrB protein, leading to formation of a stable UvrB-DNA complex.

Therefore, UvrB protein may recognize a lesion by checking whether or not the DNA can form a stacking interaction with amino acid side chains of UvrB protein in a normal manner. This hypothesis could explain the observation that the nucleotide excision repair system is able to recognize a wide range of DNA damage, but not mismatches.

Refference

Alexandrovich, A., Sanderson M. R., Moolenaar, G., Gossen, N. and Lane, A. N. (1999) NMR assignments and secondary structure of the UvrC binding domain of UvrB. *FEBS Lett.* **451**, 181-185.

Brünger, 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.

Chen, Y. H., Yang, J. T. and Martinez, H. M. (1972) Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion. *Biochemistry* **11**, 4120-4131.

Collaborative Computational Project, Number 4. (1994) The CCP4 suite: Program for protein crystallography. *Acta Crystallogr.* **D50**, 760-763.

Davis, B. J. (1964) Disc electrophoresis-II. Method and application to human serum proteins. *Ann. N. Y. Acad. Sci.* **121**, 404-427.

Doolittle, R. F. (1995) The multiplicity of domains in proteins. *Annu. Rev. Biochem.* **64**, 287-314.

Friedberg, E. C., Walker, G. C. and Siede, W. (1995) *DNA repair and mutagenesis*, American Society of Microbiology Press, Washington, D. C.

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P. and Blinov, V. M. (1989) Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res.* **17**, 4713-4730.

Gorbalenya, A. E. and Koonin, E. V. (1993) Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**,

419-429.

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.

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.

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.

Theis, K., Chen, P. J., Skorvaga, M., Van Houten, B., and Kisker, C. (1999) Crystal structure of UvrB, a DNA hlicase adapted for nucleotide excision repair. *EMBO J.* **18**, 6899-6907.

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**, 9612-9118.

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.

Kito, K. (1997) Master's Thesis. (Department of Biology, Graduate School of Scinec, Osaka University)

Kobayashi, M. (1999) Master's Thesis. (Department of Biology, Graduate School of Scinec, Osaka University)

Korolev, S., Hsieh, J. and Gauss, G. H. (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.

Kraulis, P. J. (1991) MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures. *J. Appl. Crystallogr.* 24, 946-950.

Kuramitsu, S., Hiromi, K., Hayashi, H., Morino, Y. and Kagamiyama, H. (1990) Pre-steady-state kinetics of *Escherichia coli* aspartate aminotransferase catalyzed reactions and thermodynamic aspects of its substrate specificity. *Biochemistry* **29**, 5469-5476.

Laemmli, U. K. and Favre, M. (1973) Maturation of the head of bacteriophage T4. I. DNA packaging events. *J. Mol. Biol.* **80**, 575-599.

Lee, J. W. and Cox, M. M. (1990) Inhibition of RecA protein promoted ATP hydrolysis. 1. ATPγS and ADP are antagonistic inhibitors. *Biochemistry* **29**, 7666-7676.

LeMaster, D. M. and Richards, F. M. (1985) ^{1}H-^{15}N heteronuclear NMR studies of *Escherichia coli* thioredoxin in samples isotopically labeled by residue type. *Biochemistry* **24**, 7263-7268.

McRee, D. E. (1993) *Practical Protein Crystallography*. Academic Press, Inc., San Diego.

Merritt, E. A., and Murphy, M. E. P. (1994) Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr.* **D50**, 869-873.

Moolenaar, G. F., Visse, R., Ortiz-Buysse, M., Goosen, N. and van de Putte, P. (1994) Helicase motifs V and VI of the *Escherichia coli* UvrB protein of the UvrABC endonuclease are essential for the formation of the preincision complex. *J. Mol. Biol.* **240**, 294-307.

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.

Moolenaar, G. F., Monaco, V., van der Marel G. A., van Boom, J. H., Visse, R. and Goosen, N. (2000) The effect of the DNA flanking the lesion on formation of the UvrB-DNA preincision complex. *J. Biol. Chem.* **275**, 8038-8043.

Orren, D. K. and Sancar, A. (1990) Formation and enzymatic properties of the UvrB-DNA complex. *J. Biol. Chem.* **265**, 15796-15803.

Oshima, T. and Imahori, K. (1974) Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a nonsporulating thermophilic bacterium from a Japanese thermal spa. *Int. J. Syst. Bacteriol.* **24**, 102-112.

Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

Pugh, B. F. and Cox, M. M. (1988) High salt-activation of RecA protein ATPase in the absence of DNA. *J. Biol. Chem.* **263**, 76-83.

Sancar, A. (1994) Mechanisms of DNA excision repair. *Science* **266**, 1954-1956.

Sancar, A. (1996) DNA excision repair. *Annu. Rev. Biochem.* **65**, 43-81.

Seeley, T. W. and Grossman, L. (1989) Mutations in the *Escherichia coli* UvrB ATPase motif compromise excision repair capacity. *Proc. Natl. Acad. Sci. USA* **86**, 6577-6581.

Selby, C. P. and Sancar, A. (1990) Structure and function of the (A)BC excinuclease of *Escherichia coli*. *Mutat. Res.* **236**, 203-211.

Selby, C. P. and Sancar, A. (1993) Molecular mechanism of transcription-repair coupling. *Science* **260**, 53-58.

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.

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.

Sohi, M., Alexandrovich, A., Moolenaar, G., Visse, R., Goosen, N., Venrnede, X., Fontecilla-Camps, J., Champness, J. and Sanderson, M. R. (2000) Crystal structure of *Escherichia coli* UvrB C-terminal domain, and a model for UvrB-

UvrC interaction. *FEBS Lett.* **45**, 161-164.

Subramanya, H. S., Bird, L. E., Brannigan, J. A. and Wigley, D. B. (1996) Crystal structure of a DExx box DNA helicase. *Nature* **384**, 379-383.

Van Houten, B. and Snowden, A. (1993) Mechanism of action of the *Escherichia coli* UvrABC nuclease: clues to the damage recognition problem. *Bioessays*, **15**, 51-59.

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.

Yao, N., Hesson, T., Cable, M., Hong, Z., Kwong, A. D., Le, H. V. and Weber, C. (1997) Structure of the hepatitis C virus RNA helicase domain. *Nature Struct. Biol.* **4**, 463-467.

Yokoyama, S., Hirota, H., Kigawa, T., Yabuki, T., Shirouzu, M., Terada, T., Ito, Y., Matsuo, Y., Kuroda, Y., Nishimura, Y., Kyogoku, Y., Miki, K., Masui, R., and Kuramitsu, S. (2000) Structural Genomics Projects in Japan. *Nature Struct. Biol.* **7**, 943-945.

Yokoyama, S., Matsuo, Y., Hirota, H., Kigawa, T., Shirouzu, M., Kuroda, Y., Kurumizaka, H., Kawaguchi, S., Ito, Y., Shibata, T., Kainosho, M., Nishimura, Y., Inoue, Y., and Kuramitsu, S. (2000) Structural Genomics Projects in Japan. *Prog. Biophys. Molec. Biol.* **73**, 363-376.

Zou, Y. and Van Houten, B. (1999) Strand opening by the $UvrA_2B$ complex allows dynamic recognition of DNA damage. *EMBO J.* **18**, 4889-4901.

Acknowledgments

I would like to express my deep gratitude to Professors Seiki Kuramitsu and Keiichi Fukuyama and Drs. Ryoji Masui and Ryuichi Kato for their continuing guidance and encouragement throughout the course of my study, and for many valuable discussions. I am very grateful to Drs. Mitsuaki Sugahara and Jun Ishijima for their helpful advice on my crystallographic study.

I am sincerely grateful to Professor Noriyoshi Sakabe of the University of Tsukuba, Associate professor Nobuhisa Watanabe of Hokkaido University, Drs. Mamoru Suzuki and Noriyuki Igarashi of the National Laboratory for High Energy Physics, and Dr. Masahide Kawamoto of SPring-8 (BL41XU) for their aid in the data collection with synchrotron radiation. Finally, I thank my colleagues in Professor Kuramitsu's laboratory for their kind help in this work.

List of Publications

Original Papers

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.

Shibata, A., Nakagawa, N., Sugahara, M., Masui, R., Kato, R., Kuramitsu, S. and Fukuyama, K. (1999) Crystallization and Preliminary X-ray Diffraction Studies of a DNA Excision Repair Enzyme, UvrB, from *Thermus thermophilus* HB8. *Acta Cryst. D* **55**, 704–705.

Nakagawa, N., Sugahara, M., Masui, R., Kato, R., Fukuyama, K. and Kuramitsu, S. (1999) Crystal Structure of *Thermus thermophilus* HB8 UvrB Protein, a Key Enzyme of Nucleotide Excision Repair. *J. Biochem.* **126**, 986–990.

Yamagata, A., Masui, R., Kato, R., Nakagawa, N., Ozaki, H., Sawai, H., Kuramitsu, S. and Fukuyama, K. (2000) Interaction of UvrA and UvrB Proteins with a Fluorescent Single-stranded DNA. *J. Biol. Chem.* **275**, 13235–13242.

Kato, R., Kataoka, M., Mikawa, T., Masui, R., Nakagawa. N., Kamikubo, H. and Kuramitsu, S. (2000) Observation of RecA Protein Monomer by Small Angle X-ray Scattering with Synchrotron Radiation. *FEBS Lett.* **482**, 159–162.

Theis, K., Skorvaga, M., Machius, M., Nakagawa, N., Van Houten, B. and Kisker, C. (2000) The Nucleotide Excision Repair Protein UvrB, a Helicase-like Enzyme with a Catch. *Mutation Res.* **460**, 277-300.

Proceedings

Nakagawa, N., Masui, R., Kato, R. and Kuramitsu, S. (1997) Analysis of Domain Organization of *Thermus thermophilus* UvrB Protein Using Limited Proteolysis and Denaturation. *Protein Science* **6** (suppl. 2), 107.

Nakagawa, N., Mikawa, T., Hasegawa, K., Kawaguchi, S., Masui, R., Kato, R. and Kuramitsu, S. (1998) DNA Repair Enzymes from Extremely Thermophilic Bacterium, *Thermus thermophilus* HB8. *Protein Science* **7** (suppl. 1), 136.

総説など

中川紀子, 増井良治, 加藤龍一, 倉光成紀 (1998) 原核生物のヌクレオチド除去修復. 実験医学 **16**, 1069–1075

中川紀子 (2000) 原核生物の紫外線傷害と修復.「シリーズ光が拓く生命科学４. 生物の光傷害とその防御機構」(日本光生物学協会編), 17–35, 共立出版, 東京