UvrB protein plays an essential role in the prokaryotic excision repair system. UvrB protein shows cryptic ATPase activity, DNA binding, helicase-like activity, and incision activity by interacting with UvrA or UvrC proteins. To reveal the structure-function relationship of this multifunctional protein, the domain structure of *Thermus thermophilus* UvrB protein (ttUvrB) was studied by limited proteolysis and denaturation experiments. Proteolytic profiles indicated that ttUvrB consists of four domains: the N domain (residues 2–105), M domain (106–455), C1 domain (456–590), and C2 domain (591–665). The properties of the proteolytic fragments indicated the involvement of the respective domains in the functions of the protein as follows: the N and C1 domains are necessary for ATPase activity, the C1 domain is indispensable for DNA binding, and the N and/or M domains are involved in UvrA binding. The structural stability of the C1 and C2 domains was higher than that of the N and M domains, which supports the proposed domain nature of ttUvrB. Based on these results and the crystal structure of PerA helicase (Subramanya, H. S., Bird, L. E., Brannigan, J. A., and Wigley, D. B. (1996) *Nature* 384, 370–383), the domain organization of ttUvrB was proposed.

DNA reacts easily with various chemical compounds and with some physical agents like UV radiation in the environment. 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 (1). Nucleotide excision repair is one of the most important excision repair systems such as direct repair, base excision repair, nucleotide excision repair, mismatch repair, and recombinational repair (11) and *T. thermophilus* UvrB protein makes 3' and 5' incisions, respectively. After the incision event, UvrD helicase, DNA polymerase I, and DNA ligase complete the repair by excision of the damaged nucleotide, repair synthesis, and DNA ligation, respectively.

Years of study using *E. coli* indicate that UvrB protein is a multifunctional protein; it shows cryptic ATPase activity, DNA binding, helicase-like activity, and incision activity by interacting with UvrA or UvrC proteins (2). Analyses of many mutants with amino acid substitutions and deletions have suggested the functional regions of the *E. coli* UvrB that are involved in binding of UvrA, UvrC, and DNA (3–6); however, the detailed relationship between its structure and function is still uncertain. In particular, there is only limited information available about the structural properties of UvrB protein.

Generally, large proteins contain multiple domains (7). In the case of a multifunctional protein, the respective functions are often associated with individual domains. UvrB protein is a relatively large multifunctional protein (about 80 kDa), so it is probable that UvrB protein is a multidomain protein. Protein domains can often be dissected by limited proteolysis if they are loosely assembled and the connecting segments are exposed to the environment. The domain structure can also be elucidated by denaturation experiments if the structural stabilities of the domains are different from each other.

We have studied several repair systems including nucleotide excision repair by using an extremely thermophilic bacterium, *T. thermophilus* HB8 (8–11). *T. thermophilus* is a Gram-negative eubacterium that can grow at temperatures over 75 °C (12). In general, proteins isolated from *T. thermophilus* are heat-stable and suitable for physicochemical examination, including x-ray crystallographic analysis. *T. thermophilus* *uwra* (11) and *uwrb* (9) genes have already been cloned and sequenced. Since the amino acid sequences of *T. thermophilus* *uwra* (ttUvrA) and *uwrb* (ttUvrB) 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, ttUvrB differs from *E. coli* in terms of its ability to hydrolyze ATP in the absence of *uwra* protein and DNA (9). Additionally, ttUvrB is stable from 5 to 80 °C at pH 7.5 and between pH 6 and pH 11 at 25 °C (9). These features of ttUvrB are useful for elucidating not only its structural properties but also its structure-function relationship using the methods described above.

In this study we conducted limited proteolysis and denaturation experiments of ttUvrB to reveal the organization of its structural domains. The results indicate that ttUvrB consists of at least four domains. Furthermore, three proteolytic frag-
ments were purified, and their activities were assayed. Based on the results, we propose a model of the domain arrangement of UvrB protein and discuss the relationship between its structure and function.

**EXPERIMENTAL PROCEDURES**

**Materials—**ttUvrB was prepared as described previously (9), and its concentration was determined using an ε value of 60,000 M⁻¹ cm⁻¹ at an absorbance maximum of around 277 nm. Enzymes and reagents were purchased as follows: trypsin from bovine pancreas (type IX), soybean trypsin inhibitor (type I-S), pyruvate kinase (type II), ATP, ADP, and AMP from Sigma Chemical Co., MgCl₂, 5'-adenyllylimidodiphosphate (AMP-PNP) from Boehringer Mannheim; and poly(dT) from Pharmacia Biotech Inc. The concentrations of adenine nucleotides were determined from an ε value of 1.54 × 10⁴ M⁻¹ cm⁻¹. The concentration of poly(dT) was determined from an ε value of 8.52 × 10⁴ M⁻¹ cm⁻¹ (13). Thermolysin was a gift from Dr. Keiichi Fukuyama (Osaka University). All the other chemicals and reagents were purchased from commercial sources.

**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. To examine the effect of nucleotides on ttUvrB digestion, ttUvrB (5 μM) was treated in 50 mM Tris-HCl and 40 mM MgCl₂ with trypsin at a protein to protease molar ratio of 200:1 in the presence of 20 mM ATP, AMP-PNP, or ADP. Proteolysis was stopped by adding trypsin inhibitor in a 2-fold excess to trypsin. ttUvrB (13 μM) was digested with thermolysin in 50 mM Tris-HCl (pH 7.5) and 10 mM CaCl₂ at a protein to protease molar ratio of 200:1 for various times at 60°C. The reaction was stopped by adding diisopropylfluorophosphate in a 2-fold excess to thermolysin. The digests were separated by SDS-polyacrylamide gel electrophoresis (PAGE) on 10% gels (14), followed by staining with Coomassie Brilliant Blue R-250.

**Protein Sequence Analysis—**The proteolytic fragments were separated by SDS-PAGE and electroblotted onto a polyvinylidene difluoride membrane (Immobilon-Ptrans membrane transfer membrane, Millipore) (15). Bands stained with Coomassie Brilliant Blue were excised from the membrane, and their N-terminal amino acid sequences were determined with a gas-phase protein sequencer (Applied Biosystems, model 473A). The N-terminal amino acid sequence of the small fragment (Tc fragment) generated by tryptic digestion was determined as follows: ttUvrB (13 μM) was treated with trypsin in 50 mM Tris-HCl (pH 7.5) at a protein to protease molar ratio of 200:1 for 1 min at 37°C; the reaction products were digested with thermolysin. The reaction mixture for the binding assay, 5 mM ATP, and 10 mM MgCl₂. ATP was reacted with 0.2 or 0.5 μM ttUvrB or its fragments were measured in a 5-mm cell with an excitation wavelength of 295 nm light at 25°C after incubation at 25°C for 1 h in 50 mM Tris-HCl (pH 7.5), 100 mM KCl, 1 mM EDTA, 1 mM DTE, and the indicated concentration of GdnHCl. No further change in the ellipticity values was observed after incubation at various concentrations of GdnHCl for 24 h.

**Fluorescence Measurements—**Fluorescence measurements were carried out with a Hitachi spectrofluorometer, model F-4500. The spectra of 5 μM ttUvrB or its fragments were measured in a 5-mm cell with an excitation wavelength of 295 nm light at 25°C after incubation at 25°C for 1 h in 50 mM Tris-HCl (pH 7.5), 100 mM KCl, 1 mM EDTA, and 1 mM DTE, and the indicated concentration of GdnHCl.

**RESULTS**

**Limited Proteolysis of ttUvrB—**To reveal the organization of the structural domains, we analyzed the limited proteolysis process of ttUvrB (76 kDa) (9) with two endoproteases of different substrate specificities, trypsin and thermolysin. 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 Fig. 1A. The fragments produced by trypsin digestion were observed as relatively discrete bands on the gels. Just after the addition of trypsin, a 66-kDa band appeared (lanes 3 and 6). As the band of intact ttUvrB disappeared, bands of two products with masses of 56 and 51 kDa occurred (lanes 4 and 5). After a 4-h incubation, the 66-kDa band had almost disappeared and a 40-kDa band had increased (lane 8). The major tryptic digest fragments with masses of 66, 56, 51, and 40 kDa were designated as T-1, T-2, T-3, and T-4, respectively.

A time course of ttUvrB digestion by thermolysin is shown in Fig. 1B. At a very early time point, a 66-kDa band appeared as in the trypsin digest (lane 3). After incubation for 5 min, the intact band had disappeared and 57- and 50-kDa bands appeared (lane 4). As the reaction proceeded, the 66- and 57-kDa bands disappeared and instead a 36-kDa band appeared (lanes 5-7). The four major products with masses of 66, 57, 50, and 36 kDa were designated as Th-1, Th-2, Th-3, and Th-4, respectively. The array of digest fragments produced by thermolysin was similar to those obtained with trypsin in terms of number and fragment sizes. These results indicate that cleavages by both trypsin and thermolysin occurred at some specific sites of ttUvrB.

To identify the cleavage sites, the N-terminal amino acid sequences of each band were determined. Both the T-1 and T-3 bands started at Thr⁵⁰⁰, 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. Under mild conditions, ttUvrB to trypsin ratio of 200:1, for 1 min), only T-1 and small fragments were yielded (see Fig. 5A). The digest obtained under such mild conditions was directly subjected to sequencing and was shown to contain two peptides: one began from Thr⁵⁰⁰ and the other from Ala⁵⁸⁸. Residue 588 is arginine, so we conclude that cleavage under these conditions occurred between Arg⁵⁸⁸ and Ala⁵⁸⁹. The fragment starting from Ala⁵⁸⁹ was designated as Tc fragment.
The domain structure of the T. thermophilus UvrB protein was studied using limited proteolysis. The N-terminal amino acid residues were determined using a protein sequencer. C-terminal amino acid residues of the fragments were estimated from the mass of the proteolytic fragments detected by SDS-PAGE and from the specificity of the proteases. The structural domains suggested by limited proteolysis results are shown at the bottom. Roman numbers (I–VI) and the letter W represent the positions of helicase motifs (I–VI) and tryptophan residues, respectively.

As shown in Fig. 2, the cleavages by trypsin and thermolysin occurred at sites very close to each other, although the substrate specificities of these proteases differ. In particular, it should be emphasized that thermolysin, a relatively nonspecific protease, produced only a limited number of fragments. The discreteness and restricted distribution of the cleavage sites suggest that they are located in the exposed regions between the domains. Based on the time-dependent changes in the proteolytic products, the initial cleavage by the proteases is considered to have been at around residue 590. The next cleavages are thought to have occurred at around residues 105 and 455 with similar lability. The presence of three preferential cleavage sites suggests that ttUvrB is organized into at least four domains: the N terminus to residue 105, residues 106–455, residues 456–590, and residue 591 to the C terminus (the indicated residue numbers represent the vicinity of that residue). In the following text, these domains are referred to as...

Fig. 1. Limited proteolysis of ttUvrB. A, 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. B, ttUvrB (13 μM) was treated with thermolysin at a protein to protease molar ratio of 200:1. Lane 1, molecular mass markers; lane 2, undigested ttUvrB; lanes 3–7, digests incubated for the time indicated above each lane. C, ttUvrB (5 μM) was treated with trypsin at a protein to protease molar ratio of 200:1 in the presence of nucleotides. The digestions were stopped after 0.5, 5, 15, 30, and 120 min. Panel 1, 20 mM ATP and 40 mM MgCl₂; panel 2, 20 mM AMP-PNP and 40 mM MgCl₂; panel 3, 20 mM ADP and 40 mM MgCl₂; panel 4, 40 mM MgCl₂. T and Th refer to the bands derived from trypsin and thermolysin digestions, respectively.

Fig. 2. Schematic representation of the fragmentation of ttUvrB. The N-terminal amino acid residues were determined using a protein sequencer. C-terminal amino acid residues of the fragments were estimated from the mass of the proteolytic fragments detected by SDS-PAGE and from the specificity of the proteases. The structural domains suggested by limited proteolysis results are shown at the bottom. Roman numbers (I–VI) and the letter W represent the positions of helicase motifs (I–VI) and tryptophan residues, respectively.

Band showed the presence of two fragments starting from Thr² and Asp¹⁰⁹. A comparison of the T-2 fragment and the T-3 and T-4 fragments suggested that the putative fragments starting from Thr² and Asp¹⁰⁹ in the T-2 band were longer than the T-3 and T-4 fragments, respectively.

Similarly, N-terminal sequences of the bands obtained by thermolysin digestion were determined. Both the Th-1 and Th-3 fragments started from Thr², indicating cleavages in the C-terminal region. Both the Th-2 and Th-4 fragments started from Leu¹⁰⁴ or Ile¹⁰⁶. These cleavages were consistent with the substrate specificity of thermolysin, which predominantly hydrolyzes peptide bonds involving the amino groups of the hydrophobic amino acids Leu, Ile, Met, Phe, Trp, and Val (19).

The locations of these fragments in ttUvrB are summarized in Fig. 2. The C-terminal amino acid residues of the fragments were estimated from their apparent mass assessed by their relative mobility on the gels and from the substrate specificity of each protease. As shown in Fig. 2, the cleavages by trypsin and thermolysin occurred at sites very close to each other, although the substrate specificities of these proteases differ. In particular, it should be emphasized that thermolysin, a relatively nonspecific protease, produced only a limited number of fragments. The discreteness and restricted distribution of the cleavage sites suggest that they are located in the exposed regions between the domains. Based on the time-dependent changes in the proteolytic products, the initial cleavage by the proteases is considered to have been at around residue 590. The next cleavages are thought to have occurred at around residues 105 and 455 with similar lability. The presence of three preferential cleavage sites suggests that ttUvrB is organized into at least four domains: the N terminus to residue 105, residues 106–455, residues 456–590, and residue 591 to the C terminus (the indicated residue numbers represent the vicinity of that residue). In the following text, these domains are referred to as...
the N, M, C1, and C2 domains, respectively.

In the presence of ATP, AMP-PNP, or ADP, the proteolysis pattern of ttUvrB with trypsin was essentially unaltered compared with that in their absence (data not shown); however, the rate of the cleavage reaction was affected (Fig. 1C). Addition of ATP significantly lowered the rate of cleavage of the intact protein. AMP-PNP, an unhydrolyzable analogue of ATP, had little effect on the cleavage rate. ADP had no effect on the cleavage rate. These results suggest that binding of ATP or its analogue to ttUvrB causes certain conformational changes that alter the lability of the C1-C2 interdomain linker to trypsin.

Addition of poly(dT) or poly(dA)poly(dT) had no effect on the digestion pattern and had little, if any, effect on the cleavage rate (data not shown).

**Domain Structure of T. thermophilus UvrB Protein**

**FIG. 3.** Effect of GdnHCl concentration on CD spectra of ttUvrB. A, change in the mean residue ellipticity at 222 nm. ttUvrB (2 μM) was denatured at the indicated GdnHCl concentrations. B, change in the mean residue ellipticity at 295 nm. ttUvrB (10 μM) was denatured at the indicated GdnHCl concentrations.

**FIG. 4.** Effect of GdnHCl concentration on fluorescence of ttUvrB. ttUvrB (5 μM) was denatured at the indicated GdnHCl concentrations. The spectra were measured in a 5-mm cell at an excitation wavelength of 295 nm at 25 °C. A, change of wavelength at the maximum intensity. B, change of fluorescence intensity at 340 nm.

To further examine ttUvrB denaturation by GdnHCl, fluorescence measurements were carried out. As shown in Fig. 2, ttUvrB contains three tryptophan residues at 263, 545, and 622 (9). The emission spectrum in the absence of the denaturant was shown in Fig. 6B. As shown in Fig. 4A, the emission maximum shifted toward higher wavelengths as the denaturant concentration increased. The denaturation curve of the wavelength at the maximum intensity consisted of a small change and a large transition. Similar changes were also observed in the denaturation curve of fluorescence intensity at 340 nm (Fig.
The midpoint of the large transition was at 4.0 M GdnHCl, at which concentration large transitions were observed in the CD measurements. A small change at 0 to 1.0 M GdnHCl was also observed in the near UV CD spectra. These results support the notion that a local structural change surrounding the aromatic residues, including the tryptophan residues, occurred at 0 to 1.0 M GdnHCl. At around 2.5 M GdnHCl, at which the first transition was observed in the far UV CD spectra, little change was observed in the fluorescence spectra. This discrepancy suggests that the region without tryptophan residues have lower stabilities than the region containing tryptophan residues.

The unfolding pathway was shown to be at least a three-state process as follows: native, intermediate (at 3.0 M GdnHCl), and unfolded states. If the domains of ttUvrB have distinct stabilities, the protein might have folded and unfolded domains in the intermediate state; however, it is also possible that the intermediate state is a molten globule state. It is generally agreed that a molten globule state has a high content of secondary structure and diminished tertiary contacts compared with the native state (20). Our CD and fluorescence spectroscopy results suggest that at around 3.0 M GdnHCl ttUvrB largely retained its tertiary structure around the tryptophan residues, indicating that the intermediate state corresponds to a protein with some domains unfolded. To verify this notion, a fluorescent dye, 1-anilinonaphthalene-8-sulfonate (ANS), was used. The fluorescence of ANS is enhanced upon specific binding to a hydrophobic surface, such as the surfaces exposed in a molten globule state (21). In the case of ttUvrB, no increase in ANS fluorescence was observed during the denaturation process (data not shown). These observations strongly suggest that the intermediate state of ttUvrB at around 3.0 M is not a molten globule state. Therefore, we think that GdnHCl denaturation of ttUvrB is a stepwise unfolding of individual domains. In other words, these results indicate that ttUvrB consists of domains of at least two different stabilities.

**Purification of Tryptic Fragments**—To examine the functions of the four domains identified by limited proteolysis, we tried to purify the tryptic fragments. The digestion of ttUvrB with trypsin under mild conditions (ttUvrB to trypsin molar ratio of 400:1, for 3 min) produced only two fragments, the T-1 and Tc fragments. These fragments were separated by ion exchange chromatography using a MonoQ HR5/5 column. As shown in Fig. 5A, peak 1 contained only the T-1 fragment. Peak 2 was shown to contain only the Tc fragment by other chromatographic techniques.

To examine the structural properties of the purified T-1, T-3, and Tc fragments, the CD and fluorescence spectra were measured (Fig. 6). The o-helix content of ttUvrB, T-1, T-3, and Tc fragments was estimated to be about 47, 38, 31, and 51%, respectively, using the method of Chen et al. (22). The purified T-1, T-3, and Tc fragments contained two, one, and one tryptophan residues, respectively. All three fragments showed fluorescence emission spectra excited at 295 nm, although their intensities varied. The wavelengths of the maximum emissions for the T-3 and Tc fragments were longer than that of ttUvrB, suggesting that the tryptophan residues within these fragments are in more hydrophilic environments than those within ttUvrB. The results of fluorescence experiment including the denaturation experiment of ttUvrB and its fragments (see Fig. 4B) suggest that each fragment retains its tertiary structure, supporting the hypothesis that the limited proteolysis occurred within the interdomain linkers.

**DNA Binding**—The ability of ttUvrB and the purified fragments to bind DNA was assayed by native PAGE (17). Adding poly(dT) to ttUvrB caused the band of the protein to smear upwards (Fig. 7). 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 m) 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 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

![FIG. 5. Isolation of tryptic fragments. A, isolation of T-1 and Tc fragments. ttUvrB (10 μM) was treated with trypsin at a protein to protease molar ratio of 400:1 for 3 min. The digest was separated on a MonoQ column. Peaks 1 and 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. ttUvrB (10 μM) was treated with trypsin at a protein to protease molar ratio of 50:1 for 60 min. The digest was separated on a MonoQ column. 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. Peak 1 contained the T-3 fragment, whereas peak 2 contained the T-1 and T-2 fragments.](image-url)
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 the C1 domain, it appears that the C1 domain is essential for ssDNA binding.

ATPase Activity—It has been demonstrated that ttUvrB has ATPase activity in the absence of ttUvrA (9). 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 (Fig. 8 and Table I). 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 the C1 domain is indispensable for ATPase activity.

Since the ATPase activity of intact ttUvrB is known to be stimulated by ssDNA (9), the activity of the fragments with poly(dT) was determined (Fig. 8 and Table I). 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 UvrA—Interaction between ttUvrB or its fragments and ttUvrA was examined by native PAGE in the presence of ATP and MgCl$_2$ (Fig. 9). 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 ttUvrB-ttUvrA complex can not be 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 conditions used for the assay.

### Table I

|        | None | Poly(dT) |
|--------|------|----------|
| $K_m$  | mM   | mM       |
| $k_{cat}$ | s$^{-1}$ | s$^{-1}$ |
| ttUvrA | 2.3  | 7.6      |
| T-1    | 4.5  | 7.8      |
| T-3    | ND*$a$ | ND       |

$a$ ND, not detectable.
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. These results suggest that ttUvrA binds to the N domain and/or M domain of ttUvrB. 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 Fig. 10A, 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 were mixed, the ATPase activity was higher than the calculated value but 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. These results suggest that the C2 domain is involved in the stimulation of the ATPase activity of ttUvrA in the 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 to ttUvrA reduced the total ATPase activity (Fig. 10B).

**Denaturation of Fragments**—Denaturation of the T-1 and T-3 fragments by GdnHCl was carried out to examine the stability of the domains. The denaturation curve of the T-1 fragment detected by fluorescence was cooperative, and the midpoint of the transition was at 4.0 M GdnHCl, which was close to that of ttUvrB (Fig. 11A). The denaturation curve in the far UV CD was biphasic, and the midpoints of the two cooperative transitions were 2.5 and 4.0 M GdnHCl, very close to those of the intact protein (Fig. 11B). No change at 5.0 to 7.0 M GdnHCl for T-1 fragment suggests that the C2 domain is denatured at 5.0 to 7.0 M GdnHCl in the intact protein.

In the denaturation curve of fluorescence intensity, the T-3 fragment showed less cooperative transition with a midpoint of about 3.0 M GdnHCl, which was lower than that of the intact protein (Fig. 11A). This result suggests that most of the N and M domains are denatured with the transition centered at 2.5 M GdnHCl. The observation that the fluorescence intensity of the T-3 fragment decreased at 2.5 to 3.5 M GdnHCl may imply that only the part of the M domain containing the tryptophan residue is unfolded in the gentle slope from 3.0 to 4.0 M GdnHCl.

From the results of the fluorescence and CD measurements, we concluded that the C1 and C2 domains are more stable in the presence of the denaturant than the other domains. This supports the notion that ttUvrB consists of at least two domains, as suggested by denaturation experiments using the intact protein. It should be mentioned here that the part of the M domain containing the tryptophan residue may be more stable than the rest part of the M domain. This raises the possibility that the M domain, the largest of all the domains in ttUvrB, consists of two subdomains, with different stabilities.

**DISCUSSION**

Limited proteolysis and denaturation with GdnHCl were shown to be useful for verification of the domain structure of ttUvrB. Scanning calorimetry and heat denaturation, which
are other methods for dissecting the domains, could not be applied because heat denaturation of ttUvrB was irreversible (9). Domain organization can also be predicted from the pattern of amino acid sequence conservation. Comparison of previously known UvrB proteins reveals that the region around the putative linker between the C1 and C2 domains has little sequence similarity and many insertions or gaps. This seems to be consistent with the observation that the C1–C2 linker was most labile to cleavage by proteases and supports the hypothesis that the C2 region forms a separate domain.

Sequence analysis of UvrB proteins also reveals several motifs as follows: a region homologous to transcription-repair coupling factor (TRCF), which interacts with UvrA protein (23); an array of seven conserved regions of so-called “helicase motifs” (24); and two α-helical regions with a high probability score for formation of a coiled-coil motif (4) (Fig. 12A). Based on this information, some deletion and point mutants of E. coli UvrB have been made (3–6). Although these studies can provide valuable information about the residues or regions involved in UvrB function, they may miss the possible involvement of other regions. In this regard, limited proteolysis and denaturation experiments of intact ttUvrB have no bias in principle, toward particular portions of the molecule.

As mentioned above, UvrB protein has a helicase motif and also has a latent helicase activity (25). From the distribution of the motifs in many helicase sequences, it has been suggested that the proximal motifs, I–IV, and distal motifs, V and VI, may belong to distinct domains and that inserted and additional sequences may comprise distinct domains (26). Recently the crystal structure of PcrA DNA helicase was determined (27). According to amino acid sequence analysis, the PcrA and UvrB proteins belong to the helicase superfamilies I and II, respectively, and these superfamilies share a similar pattern of seven conserved sequence motifs (26). Therefore, the structure of the PcrA helicase can help us infer the domain organization of UvrB protein and its structure-function relationship from the results obtained in this study.

The PcrA helicase has two domains, each of which has two subdomains (27), as schematically illustrated in Fig. 12B (panel 1). The seven conserved motifs are assembled around the deep cleft running between the two domains. The relationship between the linear sequence and domain organization of PcrA helicase is shown in Fig. 12A. Subdomain 1A is composed of two segments containing motifs I and Ia and motifs II and III, respectively, which are separated by subdomain 1B. Similarly, the two segments of subdomain 2A, one of which contains motifs V and VI, are separated by subdomain 2B. Domains 1 and 2 are connected by a linker region containing motif IV.

Compared with the PcrA helicase, ttUvrB, as well as its counterparts in other species (9, 28–33), has a longer insertion (about 250 residues) between motifs Ia and II (Fig. 12A). The presence of the inserted sequence suggests that ttUvrB has a larger corresponding subdomain than subdomain 1B of PcrA helicase. This inserted sequence contains a region homologous to TRCF, suggesting that this region forms an additional subdomain. In contrast, ttUvrB has a much shorter insertion (15 residues) between motifs IV and V, suggesting that ttUvrB lacks a subdomain corresponding to subdomain 2B of PcrA helicase. In addition, the extension sequence following motif VI is longer in ttUvrB (about 130 residues) than in PcrA helicase (about 70 residues). Based on this structure and the results obtained in this study, the putative domain organization of ttUvrB is illustrated in Fig. 12B.

This putative domain organization is consistent with that inferred from the limited proteolysis results of this study. The three scissile sites observed for ttUvrB match the connecting regions between the putative domains or subdomains; the linker region between the N and M domains (N–M linker) corresponds to the region between subdomains 1A and 1B, the M–C1 linker corresponds to the region between domains 1 and 2, and the C1–C2 linker corresponds to the region between domain 2 and the additional C-terminal domain. The Tc fragment, which corresponds to the C2 domain, was in a compactly folded, not unstructured state, and its helical content was estimated to be about 51%. These results suggest that the C2 region forms a compact domain by itself. As a whole, the domain organization of ttUvrB can be drawn simply as Fig. 12B (panel 2). According to this model, one significant difference between PcrA helicase and ttUvrB is that subdomain 2B of PcrA helicase appears to be lost in ttUvrB. If the arrangement of four subdomains like those in PcrA helicase is required for helicase activity, alternative organizations can be assumed for ttUvrB, as shown in panels 3 and 4 of Fig. 12B.

Differences in structural stability indicated that ttUvrB can be partitioned into two domains, the N and M domains and C1 and C2 domains. This result agrees well with the proposed structure of ttUvrB. It should be noted here that the T-3 fragment, which corresponds to the N and M domains, appears to
contain the portion that is resistant to denaturation. This observation may imply that a certain subdomain of the M domain is more stable than the rest of the domain.

The properties of the tryptic fragments are summarized in Table II. Although motifs I and II, which are responsible for the binding and hydrolysis of ATP (27), are contained in the N and M domains, our results suggest that the C1 domain is required for the ATPase activity of ttUvrB. In the PcrA structure, bound ADP is located in a pocket at the base of the cleft between subdomains 1A and 2A, which corresponds to the C1 domain (27). Therefore, the putative arrangement of the domains in ttUvrB can explain the requirement for the C1 domain for ATPase activity of ttUvrB. However, this does not necessarily mean that the C1 domain is directly involved in catalytic action of ATP hydrolysis, since in the PcrA helicase motifs V and VI in subdomain 2A contribute to the sides of the nucleotide-binding site. Furthermore, ATP was shown to induce certain conformational changes in ttUvrB, at least around the C1 domain, as ATP made the C1–C2 linker more resistant to proteolysis. This result also supports the involvement of the C1 domain in the binding and hydrolysis of ATP. In this respect, ATP hydrolysis is likely to contribute to such a conformational change because AMP-PNP and ADP had little or no effect on the sensitivity of ttUvrB to proteases.

The ability to bind to DNA also required the C1 domain, containing motifs V and VI, in addition to the N and M domains. The cleft between domains 1 and 2 is likely to be involved in the recognition of the DNA substrate, although the structure of a helicase-DNA complex has not been resolved. It is possible that the loss of the C1 domain disrupts the cleft, leading to an inability to bind to DNA. The C1 domain may be involved in the recognition of DNA since its isoelectric point, calculated from the amino acid sequence, is higher than those of the other domains. This higher basicity is also observed for the corresponding region of UvrB proteins of other species. It has been suggested that ionic interactions play an important role in the formation of UvrB-DNA complexes (6). Alternatively, the loss of ATPase activity may be involved in the inactivation of DNA binding. Site-directed mutagenesis experiments of E. coli UvrB have suggested that motifs V and VI are involved in the induction of ATPase upon DNA binding rather than being involved in DNA binding itself (5).

Our results indicate that the UvrA-binding site is located in the N and/or M domains. The region homologous to TRCF spans residues 112–246 of ttUvrB, indicating that the M domain is involved in UvrA binding. In E. coli, a deletion mutant composed of only the region homologous to TRCF was shown to bind UvrA protein (6). This result suggests that the region homologous to TRCF may form a subdomain by itself. Sequence analysis has suggested that the insertion of such an optional domain is also preferred in the region between motifs Ia and II in other helicases (24).

**TABLE II**

| The activities of ttUvrB and its fragments |
|------------------------------------------|
| ttUvrB | T-1 | T-3 |
| DNA binding | ++ | + | – |
| Interaction with ttUvrA | + | + | + |
| ATPase | + | + | – |
| ATPase stimulation by DNA | +++ | ++ | – |
| by ttUvrA | +++ | + | (Inhibition) |
| Domains | N + M + C1 + C2 | N + M + C1 | N + M |

**FIG. 12. Domain organization of ttUvrB.** A, relationship of primary sequences to putative tertiary structures. Information about the structure of PcrA helicase was obtained from the literature. Roman numerals (I–VI) represent the helicase motifs. TRCF and C/C represent the regions homologous to TRCF and α-helical regions with a high probability score for the formation of a coiled-coil motif, respectively. Labels to brackets (e.g. N) represent the domains suggested by this study. B, schematic model of the domain organization. Panel 1, domain organization of PcrA helicase based on its crystal structure. N and C represent its N and C termini, respectively. Panels 2–4, putative domain organizations of ttUvrB. The patterns for each (sub)domain correspond to those in A.
The binding of ttUvrA to ttUvrB or its fragments required the presence of ATP. In addition, stimulation of ATPase activity was observed when ttUvrA was added to the intact protein and the T-1 fragment but not to T-3 fragment. The T-3 fragment lacked the ability to hydrolyze ATP, suggesting that ATP hydrolysis by ttUvrA is necessary to form UvrA-UvrB complexes. In the case of the T-3 fragment (the N and M domains), total ATPase activity was reduced upon binding to ttUvrA. This is consistent with the observation that in E. coli binding of UvrB to UvrA suppressed the ATPase activity of UvrA (34, 35). This also leads to the notion that the ATPase activity of ttUvrB was stimulated upon binding to ttUvrA since the total activity increased.

The C2 domain was shown to be involved in the stimulation of ATPase activity by DNA or by formation of ttUvrA-UvrB complex. This stimulation is thought to occur via certain conformational changes of ttUvrB upon DNA or UvrA binding, so rearrangements of the C2 domain may affect this process. Addition of Tc fragments (C2 domain) to T1 fragments (N-C1 domain) failed to restore the DNA-dependent ATPase activity to the level of wild-type ttUvrB (date not shown). Although the far UV CD spectrum of reconstituted ttUvrB was almost the same as that of wild-type ttUvrB, the fluorescence spectrum was not. These results indicate the importance of the relative positioning of the domains.

In addition to ATPase activity, formation of UvrA-UvrB complexes activates the helicase activity of the UvrB protein (25, 36). If the arrangement of the four subdomains is essential for helicase activity, a domain of the UvrA protein may fill the position that the subdomain 2B occupies in PcrA helicase. In this respect, a vacant space may be necessary for UvrB function. Such positioning may also be advantageous for the interaction of the UvrA-binding region with the neighboring M domain (Fig. 12B, panel 2). It should be noted, however, that the presence of all four subdomains of ttUvrB may not be essential to helicase activity since PriA helicase is active although it lacks domain 1B (37).

Previously it was reported that UvrB proteins of many organisms including E. coli contain two α-helical regions with a high probability score for formation of a coiled-coil motif (4). A coiled-coil structure is formed by two or more parallel or anti-parallel α-helices. Many coiled-coils lie in a region responsible for oligomerization (38) and, in some cases, for interaction with DNA and RNA (39). Analysis of the ttUvrB amino acid sequence using a computer program, COILS (40), revealed that ttUvrB also contains two regions predicted to form a coiled-coil, residues 249–262 in the M domain and residues 635–651 in the C2 domain (Fig. 12A). Thus, the M and C2 domains may interact with each other through a coiled-coil formation. If this is the case, the C2 domain would be situated in the position occupied by the subdomain 2B in PcrA helicase (Fig. 12B, panel 4). It seems unlikely that a coiled-coil participates directly in DNA binding by ttUvrB because the T-1 fragment, lacking the C2 domain, was able to bind DNA. It has also been proposed, for the E. coli UvrABC system, that the C-terminal coiled-coil of the UvrB protein interacts with the homologous region, which is also predicted to be a coiled-coil, of the UvrC protein (4).

In contrast to E. coli UvrB, ttUvrB can hydrolyze ATP in isolation. This difference implies that the relative positions of the domains of ttUvrB are similar to those of E. coli UvrB in the UvrA-B complex rather than in isolation. In this regard, E. coli UvrB*, which is generated by proteolysis of UvrB protein and which lacks a C-terminal part (28), is of interest. It had been proposed that this cleavage occurs between residues Lys630 and Ala631, based on the apparent molecular mass of isolated UvrB* and its sequence homology with a cleavage site of the Ada protein by OmpT protease (28); however, the actual OmpT cleavage site, which was determined experimentally, was in the region encompassing residues 607–610 (4). This region is relatively close to the C1–C2 linker region of ttUvrB. Therefore, it seems that the domain structure of the C-terminal region is common to both UvrB proteins. Interestingly, E. coli UvrB* shows DNA-dependent ATPase activity in the absence of UvrA protein and is capable of interacting with UvrA protein in a manner that appears to be similar to that of the UvrB protein (34, 41). These observations support our suggestion that the C-terminal domain is involved in the activation of ATPase activity by DNA. The loss of the C-terminal domain would alter the relative positions of the domains in E. coli UvrB, leading to a structure that probably is similar to that of intact ttUvrB or the T-1 fragment. The structural stability of the C-terminal domain may be associated with the difference in activation between E. coli and T. thermophilus UvrB proteins. It has already been reported that when the ttUvrB was over-expressed in E. coli its C terminus was not processed (9); however, it is unclear whether ttUvrB is processed in T. thermophilus cells.

In summary, limited proteolysis and denaturation experiments indicated that ttUvrB consists of four domains. Characterization of proteolytic fragments revealed the involvement of the domains in the various activities of ttUvrB. Based on these results and the crystal structure of PcrA helicase, a spatial arrangement of the domains of ttUvrB was proposed. The insertion, deletion, and addition of subdomains to the basic helicase structure may enable UvrB proteins to function in an excision repair system in which many proteins are involved.

REFERENCES

1. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) DNA Repair and Mutagenesis, American Society of Microbiology Press, Washington, D. C.
2. Sancar, A. (1996) Annu. Rev. Biochem. 65, 43–81
3. Seeley, T. W., and Grossman, L. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 6577–6581
4. Moolenaar, G. F., Franken, K. L. M. C., Dijkstra, D. M., Thomas-Oates, J. E., Visser, R., van de Putte, P., and Gosen, N. (1995) J. Biol. Chem. 270, 30508–30515
5. Moolenaar, G. F., Visser, R., Ortiz-Bayssse, M., Gosen, N., and van de Putte, P. (1994) J. Mol. Biol. 240, 294–307
6. Hsu, D. S., Kim, S.-T., Sun, Q., and Sancar, A. (1995) J. Biol. Chem. 270, 8319–8327
7. Doullitt, R. E. (1995) Annu. Rev. Biochem. 64, 297–314
8. Kato, R., and Kuramitsu, S. (1993) J. Biochem. (Tokyo) 114, 926–929
9. Kato, R., Yamamoto, N., Kito, K., and Kuramitsu, S. (1996) J. Biol. Chem. 271, 9612–9618
10. Takamatsu, S., Kato, R., and Kuramitsu, S. (1996) Nucleic Acids Res. 24, 640–647
11. Yamamoto, N., Kato, R., and Kuramitsu, S. (1996) Gene (Amst.) 171, 103–106
12. Oshima, T., and Inahori, K. (1974) Int. J. Syst. Bacteriol. 24, 102–112
13. Lee, J. W., and Cox, M. M. (1990) Biochemistry 29, 7666–7676
14. Laemmli, U. K., and Favre, M. (1973) J. Mol. Biol. 80, 575–599
15. Matsubara, P. (1987) J. Biol. Chem. 262, 10035–10038
16. Kuramitsu, S., Hirumi, K., Hayashi, H., Morino, Y., and Kagaayamiya, H. (1990) Biochem. Biophys. Res. Commun. 171, 4869–4874
17. Devi, B. J. (1984) Annu. Rev. Biochem. 53, 401–427
18. Pugh, B. P., and Cox, M. M. (1988) J. Biol. Chem. 263, 76–83
19. Matsubara, H., Sasaki, R. M., Singer, A., and Jukes, T. H. (1966) Arch. Biochem. Biophys. 115, 324–331
20. Kowajima, K. (1989) Protein Struct. Funct. Genet. 6, 87–103
21. Pithan, O. B. (1992) in Protein Folding (Creighton, T. E., ed.) pp. 243–300, W. H. Freeman & Co., New York
22. Chen, Y. H., Yan, J. T., and Martinez, H. M. (1972) Biochemistry 11, 4120–4131
23. Selby, C. P., and Sancar, A. (1993) Science 260, 53–58
24. Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, M. V. (1989) Nucleic Acids Res. 17, 7479–7490
25. Oh, E. Y., and Grossman, L. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 3638–3642
26. Gorbalenya, A. E., and Koonin, E. V. (1993) Curr. Opin. Struct. Biol. 3, 419–429
27. Subramanya, H. S., Bird, L. E., Brannigan, J. A., and Wigley, D. B. (1996) Nature 384, 379–383
28. Arikan, E., Kulkarni, S. M., Thomas, D. C., and Sancar, A. (1986) Nucleic Acids Res. 14, 2637–2650
29. Black, C. G., Fyle, J. A., and Davies, J. K. (1995) J. Bacteriol. 177, 1952–1958
30. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F.,
Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995) *Science* **269**, 496–512

31. Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J. F., Dougherty, B. A., Bott, K. F., Hu, P.-C., Luzier, T. S., Peterson, S. N., Smith, H. O., Hutchison, C. A., III, and Venter, J. C. (1995) *Science* **270**, 397–403

32. Shiota, S., and Nakayama, H. (1988) *Mol. & Gen. Genet.* **213**, 21–29

33. Sicard, N., Oreglia, J., and Estevenon, A. M. (1992) *J. Bacteriol.* **174**, 2412–2415

34. Caron, P. R., and Grossman, L. (1988) *Nucleic Acids Res.* **16**, 10891–10902

35. Oh, E. Y., Claassen, L., Thiagalingam, S., Mazur, S., and Grossman, L. (1989) *Nucleic Acids Res.* **17**, 4145–4159

36. Grossman, L., and Thiagalingam, S. (1993) *J. Biol. Chem.* **268**, 16871–16874

37. Zavitz, K. H., and Marians, K. J. (1992) *J. Biol. Chem.* **267**, 6935–6940

38. Harbury, P. B., Kim, P. S., and Alber, T. (1994) *Nature* **371**, 80–83

39. Biou, V., Yaremchuk, A., Tukalo, M., and Cusack, S. (1994) *Science* **263**, 1404–1410

40. Lupas, A., Van Dyke, M., and Stock, J. (1991) *Science* **252**, 1162–1164

41. Caron, P. R., and Grossman, L. (1988) *Nucleic Acids Res.* **16**, 10903–10912