DNA Binding and Protein-Protein Interaction Sites in MutS, a Mismatched DNA Recognition Protein from Thermus thermophilus HB8*

Hidehisa Tachiki, Ryuichi Kato, and Seiki Kuramitsu‡

From the Department of Biology, Graduate School of Science, Osaka University, 1-1 Machikaneyama-cho, Toyonaka, Osaka 560-0043, Japan

The mismatch repair system repairs mismatched base pairs, which are caused by either DNA replication errors, DNA damage, or genetic recombination. Mismatch repair begins with the recognition of mismatched base pairs in DNA by MutS. Protein denaturation and limited proteolysis experiments suggest that Thermus thermophilus MutS can be divided into three structural domains as follows: A (N-terminal domain), B (central domain), and C (C-terminal domain) (Tachiki, H., Kato, R., Masui, R., Hasegawa, K., Itakura, H., Fukuyama, K., and Kuramitsu, S. (1998) Nucleic Acids Res. 26, 4153–4159). To investigate the functions of each domain in detail, truncated genes corresponding to the domains were designed. The gene products were overproduced in Escherichia coli, purified, and assayed for various activities. The MutS-MutS protein interaction site was determined by size-exclusion chromatography to be located in the B domain. The B domain was also found to possess non-specific double-stranded DNA-binding ability. The C domain, which contains a Walker’s A-type nucleotide-binding motif, demonstrated ATPase activity and specific DNA recognition of mismatched base pairs. These ATPase and specific DNA binding activities were found to be dependent upon C domain dimerization.

In living organisms, DNA damage often arises as a result of errors introduced by DNA replication, genetic recombination, and other processes (1). These DNA lesions can result in mutations, genetic diseases, and tumors. To remove such lesions, all organisms have developed DNA repair systems. The mismatch repair (MMR)1 system is one of such repair systems and is conserved significantly throughout all organisms. In Escherichia coli, MutS, MutL, and MutH proteins are included in the MMR system (2). The pathogenic genes of human hereditary nonpolyposis colorectal cancer appear to share a high degree of homology (~30%) with bacterial MutS and MutL (3, 4). Moreover, MutS homologues have also been isolated from plants (5).

These observations suggest that the MMR system is essential for all living organisms from bacteria to eukaryotes. Recently, the three-dimensional structures of E. coli MutL and MutH have been reported, and the relationships between structures and functions have been addressed (6–8).

MutS plays a key role in the early processes of MMR, mediating mismatched base pair recognition. Eukaryotic MutS homologues have been found to bind to mismatched/looped out DNA (9–12). It has also been reported that E. coli MutS binds to mismatched DNA as a dimer, forming an α-shaped loop structure (13). This finding suggests that this protein has two DNA-binding sites, one of which binds to mismatched DNA and the other to homoduplex DNA. There have been many studies exploring the DNA-binding region of MutS. A C-terminal mutant of Salmonella typhimurium MutS was found to have reduced affinity for heteroduplex DNA (14). Furthermore, the C-terminal region of hMSH2, a human MutS homologue, has been shown to be sufficient for binding to mismatches in DNA (15). On the other hand, a phenylalanine residue (Phe-39) in the N-terminal region of Thermus aquaticus MutS was found to be important for heteroduplex DNA binding (16). We have reported that the central region of Thermus thermophilus HB8 MutS binds homoduplex DNA (17). Recently, the N-terminal end and central regions of E. coli MutS were demonstrated to be required for binding of mismatched DNA (18).

Overproduced and purified MutS protein (91.4 kDa) from T. thermophilus HB8 possesses ATPase activity and binds specifically to mismatched DNA (19). T. thermophilus HB8 is an aerobic, rod-shaped, and non-sporulating Gram-negative eubacterium, which can grow at temperatures over 80 °C (20). Generally, the proteins from T. thermophilus are suitable material for structural and functional analyses because they are stable under a wide range of conditions (19, 21–26). T. thermophilus MutS protein is stable from pH 1.5 to 12 at 25 °C and at a neutral pH up to 80 °C (19). Results from protein denaturation and limited proteolysis experiments suggest that T. thermophilus MutS can be divided into three domains as follows: A (N terminus to residue 274), B (residues 275–570), and C (residues 571 to C terminus) (17). We have previously shown that the B domain interacts with double-stranded DNA (dsDNA); however, the assay system used in that study employed a mixture of all three domains. The functions of the other domains, A and C, have not been analyzed. To investigate these domains further, a more refined assay system was required. Truncated genes corresponding to the domains were designed based on the results of proteolytic analysis. Three truncated proteins corresponding to the A–C domains were overproduced in E. coli and purified. To analyze the functions of the individual domains, we assayed for MutS-MutS protein interaction, ATPase activity, and DNA-binding ability. The DNA binding

* This work was supported in part by Grants-in-aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan 08280104, 09780547, 10179209, and 12780511 (to R. K.).
† To whom correspondence should be addressed: Dept. of Biology, Graduate School of Science, Osaka University, 1-1 Machikaneyama-cho, Toyonaka, Osaka 560-0043, Japan. Tel.: 81-6-6850-5433, Fax: 81-6-6850-5442; E-mail: kuramitu@bio.sci.osaka-u.ac.jp.
‡ To whom correspondence should be addressed: Dept. of Biology, Graduate School of Science, Osaka University, 1-1 Machikaneyama-cho, Toyonaka, Osaka 560-0043, Japan.
activity of the B domain was confirmed in this study. The results also suggest that the MutS-MutS protein interaction site is located in the B domain, whereas the mismatched DNA-binding site, as well as the ATPase activity, is located in the C domain, and the ATPase site is also in the C domain. On the basis of these results, the relationships between the functions and the structures of MutS are discussed.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Media, Plasmids, and Chemicals—E. coli strains DH5α (Promega Biotech) and BL21(DE3) carrying either pET16b or pSS1 plasmids were cultured in Terrific Broth or LB broth at 37 °C (27, 28). DNA purification was performed using standard procedures (27). The enzymes and reagents used were purchased from the following sources: restriction enzymes from Toyobo, New England Biolabs, and Nippon Gene; Taq DNA polymerase from PerkinElmer Life Sciences; [α-32P]ATP and [γ-32P]ATP were from ICN; plastic-backed polyethyleneimine-cellulose sheets (MN-Polygram CEL200PEUl) were from Machery and Nagel; glutathione-Sepharose 4B was from Amersham Pharmacia Biotech; and His-Bind resin was from Novagen. Factor Xa was kindly provided by Dr. Kato (National Cardiovascular Center, Japan). Plasmid pSS1, derived from a pET3a (Novagen), contains the complete T. thermophilus mutS gene under control of a T7 promoter (19). Plasmid pET16b (Novagen), which is derived from pET3a, was used for cloning of 5′-terminal target proteins used for overproduction of the N-terminal and central domains of MutS proteins. Plasmid pGEXX-1 (Amersham Pharmacia Biotech) was used to express N-terminal glutathione S-transferase (GST)-fused C-terminal MutS. All the other chemicals and reagents used were obtained from commercial sources.

Construction of Truncated mutS Genes—DNA fragments expressing each domain of T. thermophilus MutS (Fig. 1) were generated by the polymerase chain reaction using pSS1 as a template. The following pairs of primers were used for amplification of each domain: 5′-AGAAGCCGATATGGGGGGGATGGGCGAGTTAAAG-G-3′ and 5′-CTGGGTTGAGTTAATGAAACCTGGAGG-3′ (A domain), 5′-CTCGAGCTCTTGAGACTGCTGGACATGGACCCCTCTTGGGCTTC-CTCTGGGCGC-TG-3′ and 5′-AGCCTAGCTGACACTCCCTCCCTTCGATGACGACGCGGCTATCCAGCGGCTGCCTCCGTTTAAGATATCATATGAGGGCGGGCCACC-3′ and 5′-GGGGAATGCTTACCCTCTTCTATGACCCAGG-G-3′ (C domain). All of the forward primers contained NdeI restriction sites at the first ATG codon (underlined). After subcloning into vectors, the nucleotide sequences of the amplified DNA fragments were confirmed by the dyeoxy method (Applied Bio-systems). Taq cycle sequencing was performed with 373S ABI automated DNA sequencer. By using the NdeI restriction sites, the DNA fragments corresponding to the A and B domains were each ligated with pET16b, and the resulting plasmids were named pSSA and pSSB, respectively. The DNA fragment corresponding to the C domain was inserted into pGEXX-1, and the resulting plasmid was named pSSC2. The His-tagged sequences and the GST-fused region were designed so that the purifying amino acids derived from the temperature elution fraction containing the C domain protein were collected and loaded onto a SuperQ-Toyopearl (Toyo) column (bed volume of 10 ml) equilibrated with buffer I. The B domain protein was designated as MutS-B.

To obtain purified the C domain protein, which is composed of the C-terminal domain, E. coli cells carrying pSSC2 were induced by IPTG and harvested. Frozen cells (20 g) were suspended in 20 ml of PBS buffer (10 mM NaHPO4, 1.8 mM KH2PO4, 140 mM NaCl, 2.7 mM KCl, pH 7.3). Lysozyme was added to cell suspension to a final concentration of 0.1 mg/ml and the mixture was incubated for 15 min at 37 °C. Following this, 180 ml of PBS buffer was added to the solution, and the cells were disrupted by sonication on ice. Triton X-100 was added to a final concentration of 1% (v/v), and the cell extract was stirred for 30 min at 4 °C, and then was centrifuged (39,000 × g) for 30 min at 4 °C. After centrifugation, most of the C domain protein (about 70–80%) was found in the supernatant, however, a considerable amount of protein remained in the supernatant. The supernatant was filtered through a 0.22-μm membrane (Millex-GV, Millipore) and was loaded onto a glutathione-Sepharose 4B column (bed volume 2 ml) that was equilibrated with PBS buffer. The column was washed with 100 ml of the same buffer, and the GST-fused C domain was then eluted with 10 ml of glutathione elution buffer (50 mM Tris-HCl, 10 mM glutathione, pH 8.0) at room temperature. The elution fraction containing the C domain protein were collected and loaded onto a SuperQ-Toyopearl (Toyo) column (bed volume of 10 ml) equilibrated with buffer I. The column was washed with 20 ml of the same buffer, and the protein was eluted with a 100-ml gradient of 0–1 M NaCl in buffer I. Fractions containing the C domain protein were collected and loaded onto a SuperQ S-300HR column. The protein fractions were collected and dialyzed three times against 10-fold volumes of buffer I. The purified C domain protein had a tendency to precipitate at concentrations higher than 10 μM, its concentration was kept below 10 μM.

The amount of the purified C domain protein was 10 mg, and the yield was 0.5 mg of protein per 1 g wet cells. The B domain protein was designated as MutS-B.

To remove the His tag or the GST region from the purified fused proteins, 100 μl of 10 μM fusion protein was incubated with 5 μl of 0.36 mg/ml factor Xa at 37 °C overnight. The factor Xa was removed from the reaction mixture by the following procedure. The reaction solution was loaded onto a Superdex 200HR column (Amersham Pharmacia Biotech) that had been equilibrated with buffer I. The column was then eluted with the same buffer at room temperature and at a flow rate of 0.5 ml/min. Prior to loading onto the Sephacryl S-300HR column, 1 ml of the fusion proteins was cleaved by factor Xa. The protein fractions were collected and characterized by the molar extinction coefficients of each purified truncated protein were calculated to be 23,400 (MutS-A), 51,000 (MutS-B), and 19,500 M−1 cm−1 (GST-C), using a procedure described previously (28). CD measure-
ments were carried out in a sample solution containing 50 mM Tris-HCl, 100 mM KCl, 0.1 mM dithioerythritol, pH 7.5, and 1 μM of each protein using a 0.1-cm cell. The residual molar ellipticity [θ] was defined as 100 θ_mar (lc)⁻¹, where θ_mar is the observed molar ellipticity; l is the length of the light path in centimeters, and c is the residual molar concentration of each protein. CD measurements were performed using a Jasco spectropolarimeter, model J-720W.

Size-exclusion Chromatography—In order to analyze the relationship between the apparent molecular mass and the protein concentration of full-length MutS, 50 μL of the protein sample at various concentrations was injected onto a Superdex 200 HR column (inner diameter, 1 × 30 cm, Amersham Pharmacia Biotech) that had been equilibrated with buffer comprising of 50 mM Tris-HCl, 100 mM KCl, 5 mM β-mercaptoethanol, 0.1 mM EDTA, pH 7.5. The measurements were carried out at room temperature at a flow rate of 0.5 mL/min, and the absorption was detected at 280 nm. The following proteins were used as molecular size markers: thyroglobulin (669 kDa), apoferritin (443 kDa), β-lactamase (200 kDa), alcohol dehydrogenase (150 kDa), albumin (66 kDa), carbonic anhydrase (31 kDa), and cytochrome c (12.8 kDa).

ATPase Assay—The ATPase activity of each protein was analyzed using a TLC method (19). The reaction mixture contained 1 μM of each protein, 1 mM [γ-32P]ATP, 50 mM MgCl₂, 100 mM KCl, 1 mM DTT, and 50 mM Tris-HCl, pH 7.5. Reactions were carried out at 25 °C for 10 min and stopped by addition of 0.5% SDS. The samples were spotted on PEI cellulose sheets and eluted with 0.75 M potassium phosphate, pH 3.4. The sheets were placed in contact with an imaging plate, and the resultant hydrolyzed products were analyzed using by a BAS2000 image analyzer (Fuji photo film).

Gel Retardation Assay—Gel retardation assays using non-denaturing polyacrylamide gel electrophoresis were carried out as described previously (19). A 37-mer oligonucleotide TAKA1, 5'-ATGTGAATCAG-TATGTTTCATCA-GAGGAAT-3', was synthesized, and its 5'-region of MutS contributes to the stability of the full-length MutS (data not shown). This suggests that the N-terminal structural domains.

From the intensities of the resonance unit before and after the annealing of the second oligonucleotide, it was estimated that about 90% of the second oligonucleotide was annealed to the first oligonucleotide. After immobilization of the oligonucleotide, 30 μL of a solution containing 0, 0.2, 0.4, 0.7, 1.0, or 2.0 μM of each protein was applied at a flow rate of 20 μL/min at 25 °C over flow cells in sensor chips, either without oligonucleotide, with TAKA1, with TAKA1-TAKA2, or with TAKA1-TAKA4. Overlay plots were drawn from the data obtained between 10 and 90 s after injection. The data were averaged and corrected by the blank controls, obtained with eluent buffer A and the sensor chip without DNA. Data calculation was carried out using the BLAevaluation software (Amersham Pharmacia Biotech) with algorithms for numerical integration (30). Apparent rate constants (k_on) and dissociation rate constants (k_off) were calculated from the slope and the intercept of these plots, respectively. The value of dissociation constant, K_d, was calculated from those of k_on and k_off.

RESULTS

Construction of Truncated Proteins Corresponding to Each Domain—As we have already shown, MutS protein can be divided into at least three domains based on the results of limited proteolysis and the denaturation (17). In order to analyze the functions of each domain in more detail, truncated proteins that correspond to the proteolytic fragments were designed (Fig. 1). Each truncated gene was inserted into an expression vector for overexpression in E. coli as described under "Experimental Procedures." These proteins were designated MutS-A (A domain), MutS-B (B domain), and GST-C (C domain), respectively (Fig. 1). MutS-A and MutS-B contain a short His repeat on their N termini, and GST-C was constructed as a fusion protein with GST on the N terminus. MutS-A, MutS-B, and GST-C were purified as described under "Experimental Procedures" (Fig. 2). GST-C could be cleaved by factor Xa protease into two products, GST-N and MutS-C, by the sequence-specific endoprotease, factor Xa (arrow).
In order to elucidate the oligomerization state of T. thermophilus MutS in more detail, size-exclusion chromatography experiments at various protein concentrations were carried out. As shown in Fig. 4A, the apparent molecular mass of T. thermophilus MutS at lower concentrations (1 or 5 μM MutS) was estimated to be 220 kDa, which corresponds to a dimer. As the protein concentration increased, the apparent molecular mass increased to about 350 kDa. This suggests that the minimal unit of T. thermophilus MutS is a dimer. The highly oligomerized state of MutS at higher protein concentrations may have some physiological functions that need to be elucidated.

In order to determine where the MutS-MutS interaction site is located, the molecular sizes of the truncated proteins in solution were measured using size-exclusion chromatography. The apparent molecular mass of MutS-A was calculated to be 30 kDa (Fig. 4B, open circle), which is almost identical to the size predicted by the amino acid sequence (32.6 kDa). The MutS-B eluted at a volume corresponding to apparent molecular size of 65 kDa (Fig. 4B, square) which is two times larger than the predicted molecular mass (35.8 kDa). The GST-C, which is the GST-fused C-terminal region of MutS, eluted at a volume corresponding to 110 kDa (Fig. 4B, diamond) which is two times larger than the predicted molecular size (54.6 kDa). Since GST alone forms a dimer (33), we measured the molecular masses of GST-N and MutS-C, which were derived from cleavage of GST-C by factor Xa. The molecular mass of GST-N was calculated to be 60 kDa (Fig. 4B, inverted triangle), which is two times larger than the predicted size (26.3 kDa), and that of MutS-C was calculated to be 25 kDa (Fig. 4B, triangle), which is almost identical to the predicted size (28.3 kDa).
conclusion, MutS-A and MutS-C exist as monomers, whereas MutS-B exists as a dimer. This result suggests that the MutS-MutS interaction site is located within the B domain.

ATPase Activity of Truncated Proteins—MutS contains a Walker’s A-type nucleotide-binding motif (34) and shows weak ATPase activity, the activity of each truncated domain contains the motif. In order to determine which domain is responsible for ATPase activity, the activity of each truncated domain was measured using the TLC method. The ATPase activity of \( T. \) thermophilus MutS was maximal at 80 °C (19); however, the assay was carried out at 25 °C because MutS-B and GST-C were not heat-stable. As shown in Fig. 5, GST-C showed ATPase activity, but MutS-A and MutS-B did not. Interestingly, MutS-C, a product of factor Xa cleavage of GST-C, showed no ATPase activity. Since it has not been reported that GST shows any ATPase activity, the loss of the ATPase activity of MutS-C may be related to the lack of dimer formation without the GST fusion.

DNA Binding Activity of the Truncated Proteins—\( T. \) thermophilus MutS binds with higher affinity to G/T mismatched heteroduplex DNA than to homoduplex DNA (19). To clarify whether or not the truncated proteins bind to DNA or not, the binding activity of each domain to a \( \beta\)-labeled 37-base pair dsDNA with or without a G/T mismatched base pair was examined by a gel retardation assay. Although the full-length MutS bound with higher affinity to mismatched DNA than to homoduplex DNA, MutS-A could not bind to either DNA, whereas MutS-B could bind to both (Fig. 6, upper panel). In contrast, GST-C bound with higher affinity to heteroduplex DNA, in a similar manner to the full-length MutS (Fig. 6, lower panel). These results suggest that there are two DNA-binding sites in MutS, a nonspecific site in the B domain and a specific site in the C domain. Interestingly, GST-N and MutS-C, which are the cleaved products from GST-C, could not bind to either mismatched or complementary DNA (Fig. 6, lower panel). This observation suggests that the specific binding to mismatched DNA by GST-C is related to dimer formation.

The binding specificity of \( T. \) thermophilus MutS for mismatched DNA increases with temperature (19). Small differences between homo- and heteroduplex DNA-binding abilities were observed by the gel retardation experiments shown in Fig. 6; however, the experiments could not be carried out at higher temperatures because of the protein temperature instability. To confirm the DNA-binding specificity with more detail, SPR measurements were carried out. Single-stranded and double-stranded oligonucleotides were immobilized on the surface of the sensor chip, and the sample solution containing full-length MutS or each truncated protein was applied over the chip. As shown in Fig. 7A, increments of resonance units corresponding to binding of full-length MutS to dsDNAs were observed. When the full-length MutS flowed over the chip, higher resonance units were observed for heteroduplex DNA than for homoduplex DNA. This suggests that the amount of MutS bound to heteroduplex DNA was greater than the amount of MutS bound to homoduplex DNA. On the other hand, effective interaction of full-length MutS with single-stranded DNA (ssDNA) was not observed. Next, MutS-B or GST-C was applied over the DNA-immobilized sensor chip, and the changes in the resonance units were observed. As shown in Fig. 7B, MutS-B could interact with both dsDNAs but not ssDNA. When MutS-B was applied onto the sensor chip, the resonance units increased immediately and reached plateaus at similar values for both homo- and heteroduplex DNA. Due to fast association and dissociation of MutS-B and dsDNAs, the rate constants for association and dissociation could not be determined. The sensorgrams in Fig. 7C showed that GST-C interacts with the dsDNA with the mismatch specifically but not with ssDNA. No interaction was observed between GST alone and any DNA (data not shown). Thus, the specific interaction between GST-C and heteroduplex DNA is purely due to the C-terminal region of MutS and not from GST. This result indicates that the C-terminal region has the mismatch-specific dsDNA binding activity.

In order to estimate the kinetic parameters, the SPR measurements were carried out at various concentrations of each protein. Dissociation constant (\( K_d \)) were calculated to be 0.52 (homoduplex) and 0.45 \( \mu \)M (heteroduplex) for full-length MutS, and 4.2 (homoduplex) and 3.2 \( \mu \)M (heteroduplex) for GST-C,
MutS proteins for substrate DNA. Each of the proteins was applied
been reported that the activity requires the Walker’s nucleo-
logues, possesses weak ATPase activity (14, 19, 35–37). It has
had immobilized ssDNA, homoduplex DNA, or G/T-mismatched hetero-
dNA interactions. In our study, the gel retardation and SPR
experiments indicate that the B domain interacts with DNA
nonspecifically. We have already shown that the B domain of
Thermus thermophilus MutS, as well as the other MutS homo-
logs, possesses weak ATPase activity (14). These findings suggest that the C-terminal
region is important not only for ATPase activity but also for
mismatched DNA recognition and protein-DNA interactions.

Other candidate DNA-binding sites have been reported. A
phenylalanine residue (Phe-39) in the N-terminal region of
Thermus aquaticus MutS was affinity-labeled with a single base, looped
duplex DNA (16). Moreover, N-terminal and central regions of
Escherichia coli MutS are required for binding to mismatched DNA
(18). These findings suggest that not only the C-terminal region
of MutS but also the other regions are involved in mismatched
DNA interactions. In our study, the gel retardation and SPR
experiments indicate that the B domain interacts with DNA
nonspecifically. We have already shown that the B domain of
Thermus thermophilus MutS binds to homoduplex DNA (17). These
results suggest that the B domain is also involved in the DNA
binding. The nonspecific DNA binding ability of the B domain
may help increase the total binding affinity of MutS to DNA.

The results of DNA-binding experiments show that Thermus
thermophilus MutS and MutS-C bind specifically to mismatched
DNA. In the case of both proteins, the values of the
Kd for DNA with and without mismatches were significantly different. In
our study, the SPR experiments were performed at 25 °C be-
cause of mechanical limitations and protein instability; how-
ever, the growing temperature of Thermus thermophilus is about
75 °C (20). As previously shown, the ability of Thermus thermophilus
MutS to specifically bind mismatched DNA decreases as the
temperature decreases (19). The difference of Kd with and
without a mismatch may be greater at higher temperatures.

There were some inconsistent observations concerning the
oligomerization state of MutS. Previously, we have shown by
size-exclusion chromatography that the apparent molecular
mass of Thermus thermophilus MutS in solution is about 330 kDa
which corresponds to a trimer or tetramer (19). A similar size,
280 kDa, has also been reported by the same method using
Thermus aquaticus MutS (16, 31). However, studies using cross-linking,
light scattering, mass spectrometry, and equilibrium sedimen-
tation measurements concluded that the protein is a dimer (31). In a study on Escherichia coli MutS using size-exclusion chromatography,
it was reported that the protein forms dimers or oligomers (32, 38). We conclude from this study that Thermus thermophilus MutS exhibits concentration-dependent oligomeriza-
tion, forming dimers at low concentrations and higher order
oligomers at higher concentrations (Fig. 4A). Similar observa-
tions have been reported from studies on Thermus aquaticus MutS using equilibrium sedimentation analysis (31) and on Escherichia coli MutS using size-exclusion chromatography (32). Recently, we
observed that Thermus thermophilus MutS forms a dimer in solution
using small angle x-ray scattering measurements. From these
observations, we conclude that bacterial MutS forms a dimer in
solution. The results of size-exclusion chromatography suggest that the MutS-MutS protein interaction site is located in the B
domain. The observation of MutS-B aggregation at high protein
concentrations further suggests that the B domains have such
interaction sites responsible for the oligomerization of MutS.

According to these results, the B domain has two different

respectively. These results indicate that the binding specificity
of MutS between homo- and heteroduplex DNA is rather small.
We found that the C-terminal region of MutS (GST-C) could
interact with heteroduplex DNA; however, the affinity was
about 1 order lower than that of full-length MutS. This differ-
ence may be caused by a nonspecific DNA binding activity of
the B domain.

DISCUSSION

Based on protein denaturation and limited proteolysis stud-
ies, we proposed a structural organization of Thermus thermophilus
MutS into three structural domains as follows: A (N terminus
to residue 274), B (residues 275–570), and C (residue 571 to C
terminus) (17). Truncated proteins corresponding to the do-
main were constructed and purified. Each purified protein
maintained its secondary structure and demonstrated charac-
teristic activities. Our results show that the three structural
domains of Thermus thermophilus MutS correspond to functional
domains. Purified MutS-A was stable up to 80 °C like the
full-length Thermus thermophilus MutS; however, the other two trunc-
cated proteins were not heat-stable. This suggests that the A
domain may contribute to increase the protein stability of Thermus thermophilus MutS. The B and C domains may be stabilized by
protein-protein and/or domain-domain interaction.

Thermus thermophilus MutS, as well as the other MutS homo-
logs, possesses weak ATPase activity (14, 19, 35–37). It has
been reported that the activity requires the Walker’s nucleo-
tide-binding motif which is located on the C-terminal region of
the protein (14, 15, 18). A C-terminal mutant of Salmonella typhi-
murium MutS containing a modified Walker’s A-type motif
bound to heteroduplex DNA with reduced affinity (14). The
C-terminal region of the human MutS homologue, which cor-
responds to the C domain of Thermus thermophilus MutS, was
required for both ATPase activity and binding to mismatched
heteroduplex DNA (15). In addition, our results show that the
C domain of Thermus thermophilus MutS interacts specifically with
mismatched DNA. These findings suggest that the C-terminal
region is important not only for ATPase activity but also for
mismatched DNA recognition and protein-DNA interactions.

FIG. 7. SPR sensorgrams of the full-length and the truncated
MutS proteins for substrate DNA. Each of the proteins was applied
to the BIAcore 1000 for 90 s at a flow rate of 20 μl/min at 25 °C, and
then the sensor chip was washed with a buffer of 50 mM Tris-HCl, 100
mM KCl, 1 mM MgCl2, 1 mM DTT, 0.1 mM EDTA, pH 7.5. The sensor chip
had immobilized ssDNA, homoduplex DNA, or G/T-mismatched hetero-
duplex DNA. The panels show the sensorgram of Thermus thermophilus MutS
(A), MutS-B (B), and GST-C (C). The concentrations of the proteins were
1 μM (A and C) and 0.2 μM (B).

2 R. Kato, M. Kataoka, and S. Kuramitsu, manuscript in preparation.
activities, nonspecific DNA-binding and MutS-MutS protein interaction.

Among the truncated MutS proteins, the mismatch-specific DNA-binding ability and the ATPase activity were observed only in the GST-fused C domain. These activities were lost when GST-C was cleaved by factor Xa. Size-exclusion chromatography experiments showed that the GST-fused C domain exists as a dimer, but the C domain digested by factor Xa exists as a monomer. Since GST forms dimers (33), the dimerization of GST-C may be mediated by its GST domain. Thus, it is possible that although the C domain dimers possess DNA binding and ATPase activities, monomers do not. These results suggest that the minimal functional unit of MutS is a dimer. It has been demonstrated that E. coli MutS (13) and T. aquaticus MutS (31) bind to DNA as a dimer. It has been reported that E. coli recognized by the C domain. Electron microscopy revealed that dsDNA after which the mismatched region in the dsDNA is ated by the B domain of the protein. The B domain binds is the minimal functional unit. Dimerization of MutS is medi-

the initial step of MMR. In solution, MutS forms a dimer which

possess DNA-binding ability and the ATPase activity were observed

ing and ATPase activities, nonspecific DNA-binding and MutS-MutS protein interaction.

REFERENCES

1. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) DNA Repair and Mu-
tagenesis, American Society for Microbiology, Washington, D. C.
2. Modrich, P. (1991) Annu. Rev. Genet. 25, 229–253
3. Kolodner, R. (1996) Genes Dev. 10, 1433–1442
4. Fishel, R., and Wilson, T. (1997) Curr. Opin. Genet. & Dev. 7, 105–113
5. Kevin, M. C., and John, B. H. Culligan, K. M., and Hays, J. B. (1997) Plant

Physiol. 115, 833–839
6. Ban, C., and Yang, W. (1998) EMBO J. 17, 1526–1534
7. Ban, C., and Yang, W. (1998) Cell 95, 541–552
8. Ban, C., Junop, M., and Yang, W. (1999) Cell 97, 85–97
9. Drummond, J. T., Li, G. M., Longley, M. J., and Modrich, P. (1995) Science 268, 1909–1912
10. Palombo, F., Gallinari, P., Iacarino, I., Lettieri, T., Hughes, M., D’Arrigo, A.,

Truong, Q., Husain, J. J., and Jiricny, J. (1996) Science 268, 1912–1914
11. Palombo, F., Iacarino, I., Nakajima, E., Ikejima, M., Shimada, T., and Jiricny,

J. (1996) Curr. Biol. 6, 1181–1184
12. Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1996) Curr. Biol. 6, 1185–1187
13. Allen, D. J., Makrov, A., Grilley, M., Taylor, J., Thresher, R., Modrich, P., and

Griffith, J. D. (1997) EMBO J. 16, 4467–4476
14. Haber, L. T., and Walker, G. C. (1991) EMBO J. 10, 2797–2715
15. Whitehouse, A., Deeble, J., Taylor, G. R., Guilhou, P. J., Phillips, S. E.,

Meredith, D. M., and Markham, A. F. (1997) Biochem. Biophys. Res. Com-
mun. 232, 10–13
16. Malkov, V. A., Biewas, I., Camerini-Otero, R. D., and Hsieh, P. (1997) J. Biol.

Chem. 272, 23811–23817
17. Tachiki, H., Kato, R., Masui, R., Hasegawa, K., Itakura, H., Fukuyma, K.,

and Kuramitsu, S. (1998) Nucleic Acids Res. 26, 4153–4159
18. Wu, T. H., and Marinus, M. G. (1999) J. Biol. Chem. 274, 3948–3962
19. Takamatsu, S., Kato, R., and Kuramitsu, S. (1996) Nucleic Acids Res. 24, 640–647
20. Osima, T., and Imahori, K. (1974) Int. J. Syst. Bacteriol. 24, 102–112
21. Kato, R., Yamamoto, N., Kito, R., and Kuramitsu, S. (1996) J. Biol. Chem.

9612–9618
22. Okamoto, A., Kato, R., Masui, R., Yamagishi, A., Osima, T., and Kuramitsu,

S. (1996) J. Biochem. (Tokyo) 119, 135–144
23. Yamamoto, N., Kato, R., and Kuramitsu, S. (1996) Gene (Amst.) 171, 103–106
24. Kato, R., Hasegawa, K., Hidaka, Y., Kuramitsu, S., and Hoshino, T. (1997) J.

Bacteriol. 179, 6499–6503
25. Mikawa, T., Kato, R., Sugahara, M., and Kuramitsu, S. (1998) Nucleic Acids

Res. 26, 903–910
26. Kato, R., and Kuramitsu, S. (1998) Eur. J. Biochem. 259, 592–601
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A

Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring

Harbor, NY
28. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990)

Methods Enzymol. 185, 60–89
29. Kuramitsu, S., Hiromi, T., Hayashi, H., Morino, Y., and Kagamiyama, H.

(1990) Biochemistry 29, 5469–5476
30. Rovert, K., and Anders, P. (1997) J. Immunol. Methods 200, 121–133
31. Biewas, I., Ban, C., Fleming, R. G., Qin, J., Luary, J. W., Yphantis, D. A., Yang,

W., and Hsieh, P. (1999) J. Biol. Chem. 274, 23673–23678
32. Bjornson, K. P., Allen, D. J., and Modrich P. (2000) Biochemistry 39, 3176–3183
33. Kaplan, W., Huser, P., Klump, H., Erhardt, J., Sluis-Cremer, N., and D irr, H.

(1997) Protein Sci. 6, 399–406
34. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) EMBO J. 1, 945–951
35. Au, K. G., Walsh, K., and Modrich, P. (1992) J. Biol. Chem. 267, 12142–12148
36. Chi, N. W., and Kolodner, R. D. (1994) J. Biol. Chem. 269, 29993–29997
37. Gradia, S., Chakravarty, S., and Fishel, R. (1997) Cell 91, 995–1005
38. Su, S. S., and Modrich, P. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 5057–5061
39. Blackwell, I. J., Martik, D., Bjornson, K. P., Bjornson, E. S., and Modrich, P.

(1998) J. Biol. Chem. 273, 32055–32062
40. Gradia, S., Subramanian, D., Wilson, T., Acharya, S., Makrov, A., Griffith, J.,

and Fishel, R. (1999) Mol. Cell 3, 255–261
41. Iacarino, I., Marra, G., Palombo, F., and Jiricny, J. (1998) EMBO J. 17, 2677–2686

DNA Binding and Protein-Protein Interaction Sites in MutS

308–309

40709

2001