640–647 Nucleic Acids Research, 1996, Vol. 24, No. 4

Mismatch DNA recognition protein from an extremely thermophilic bacterium, *Thermus thermophilus* HB8

Satoko Takamatsu, Ryuichi Kato and Seiki Kuramitsu*

Department of Biology, Faculty of Science, Osaka University, 1-1 Machikaneyama-cho, Toyonaka, Osaka 560, Japan

Received October 31, 1995; Revised and Accepted January 8, 1996

DDBJ accession no. D63810

ABSTRACT

The *mutS* gene, implicated in DNA mismatch repair, was cloned from an extremely thermophilic bacterium, *Thermus thermophilus* HB8. Its nucleotide sequence encoded a 819-amino acid protein with a molecular mass of 91.4 kDa. Its predicted amino acid sequence showed 56 and 39% homology with *Escherichia coli* MutS and human hMsh2 proteins, respectively. The *T.thermophilus mutS* gene complemented the hypermutability of the *E.coli mutS* mutant, suggesting that *T.thermophilus* MutS protein was active in *E.coli* and could interact with *E.coli* MutL and/or MutH proteins. The *T.thermophilus mutS* gene product was overproduced in *E.coli* and then purified to homogeneity. Its molecular mass was estimated to be 91 kDa by SDS–PAGE but ~330 kDa by size-exclusion chromatography, suggesting that *T.thermophilus* MutS protein was a tetramer in its native state. Circular dichroic measurements indicated that this protein had an α-helical content of ~50%, and that it was stable between pH 1.5 and 12 at 25°C and was stable up to 80°C at neutral pH. *Thermus thermophilus* MutS protein hydrolyzed ATP to ADP and Pi, and its activity was maximal at 80°C. The kinetic parameters of the ATPase activity at 65°C were K_m = 130 μM and k_{cat} = 0.11 s^{-1}. *Thermus thermophilus* MutS protein bound specifically with G–T mismatched DNA even at 60°C.

INTRODUCTION

All living organisms have DNA repair systems to counteract DNA damage caused by sunlight, chemical agents and DNA replication errors (1). DNA repair systems involve photoreactivation, base excision repair, nucleotide excision repair, mismatch repair and recombinational repair (1). Mismatched base pairs are produced by either genetic recombination, DNA damage or DNA replication errors. In *Escherichia coli*, they are repaired by the *mutH*, *L* and *S* gene products (2). MutS homologues have been found not only in other bacteria, but also in yeast, *Xenopus*, mouse and human (3–9). Moreover, MutL homologues have also been found in organisms ranging from bacteria to the higher eukaryotes (10–12). These findings suggest that the mismatch repair system is ubiquitous for all living organisms (1,3).

In humans, kindred analysis of hereditary nonpolyposis colorectal cancer (HNPCC) has implicated four genes in the development of the disease (8–11,13). These are *hMSH2*, which is a *MutS* homologue, and *hMLH1*, *hPMS1* and *hPMS2*, which are *MutL* homologues in man. A defect in one of these genes apparently predisposes to tumor formation. Molecular analysis of the mismatch repair system is required in order to understand the mechanism of tumor development. *E.coli* has been extensively used as a model to study the mismatch repair system at the molecular level (2). MutS protein specifically binds to mismatched base pairs in the DNA. MutL protein then interacts with MutS protein to increase the stability of the MutS–mismatch DNA complex. MutH protein, a single-strand endonuclease, binds to the complex and then recognizes and nicks the unmethylated (i.e. newly replicated) strand of DNA at the hemi-methylated GATC site. The newly replicated strand is degraded, resynthesized and then the mismatched base pair is repaired. In the above model, MutS protein recognizes mismatched base pairs in the DNA, although the details of the reaction mechanism by which this recognition process occurs remain to be elucidated.

Of the three proteins involved in the mismatch repair process, MutS protein is the most important because of its initial role in mismatched base pair recognition. Therefore, we decided to analyze MutS protein in detail. In order to understand the structure–function relationships for this protein at the molecular level, enzymatic and physicochemical analysis is required. Proteins from thermophilic bacteria are particularly useful for study because they are stable and easily crystallized. The extremely thermophilic bacterium, *Thermus thermophilus* HB8, is an aerobic, rod-shaped, non-sporulating Gram-negative eubacterium, which can grow at temperatures in excess of 75°C (14). In this report, we describe the cloning, sequencing and overexpression of the *mutS* gene from *T.thermophilus*. We also describe the purification and characterization of MutS protein and demonstrate that it was stable at temperatures ≤80°C, it had ATPase activity, and that it could recognize G–T mismatches. These results help to define the molecular role of MutS protein during substrate recognition.

* To whom correspondence should be addressed

Table 1. Bacterial strains and plasmids used

Strain or plasmid	Genotype	Source or reference
***E.coli* strains**		
DH5α	*supE44* Δ*lacU169* (ϕ*80 lacZ*Δ*M15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	(15)
GM4244[a]	*ara thi* Δ(*gpt-lac*)*5* / F′ [*lacI*^{q} *lacZ proAB*$^{+}$]	(43)
GM4271	Same as GM4244, but *mutS458*::mini Tn*10* Kan^{r}	Marinus
JM109(DE3)	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1* Δ(*lac-pro*) DE3 / F′ [*traD36 proAB*$^{+}$ *laqI*^{q} *lacZ*Δ*M15*]	(15)
Plasmids		
pUC118,119	Ap^{r}; high-copy-number plasmids	(15)
pET3a	Ap^{r}; pBR322 derivative, vector for gene expression	(44)
pLysE	Cm^{r}; pACYC184 derivative, represses the expression of T7 RPase	(44)
pTS3	Ap^{r}; pUC118 derivative, carrying the entire *T.thermophilus mutS* gene	This work
pSS1	Ap^{r}; pET3a derivative, a plasmid for expression of the *T.thermophilus mutS* gene	This work

[a]Another name for this strain is CC106.

MATERIALS AND METHODS

Bacterial strains, media, plasmids and chemicals

All of the *E.coli* strains and plasmids used in this study are listed in Table 1. They were grown in LB medium or Terrific broth at 37°C (15). *Thermus thermophilus* HB8 was grown at 75°C as described elsewhere (16). DNA manipulations were carried out using standard procedures (15). The reagents were purchased as follows: DEAE–cellulose DE52 from Whatman Biochemicals; Phenyl-Toyopearl 650M from Tosoh; Sephacryl S-300 HR from Pharmacia. The DNA oligomers used were synthesized by a Cyclone Plus DNA synthesizer or purchased from Kurabo or Nihon Seifun. All other reagents used in this study were of the highest grade commercially available.

Cloning, sequencing and overexpression of the *T.thermophilus mutS* gene

Two highly conserved regions taken from five MutS homologues (*E.coli* MutS, *Salmonella typhimurium* MutS, *Streptococcus pneumoniae* HexA, *Saccharomyces cerevisiae* Msh2 and human DUC-1) (4) were used as the bases for the designed synthesis of mixed oligonucleotide primers for the polymerase chain reaction (PCR). The left and right mixed oligonucleotide primers were 5′-AC(C/G)GG(C/G)CC(C/G)AACATGGG(C/G)GG(C/GAT)-AA-3′ and 5′-AA(G/A)CG(G/C/A)TG(G/C)GTGATGAA(G/A)-CT(C/T)-3′, respectively. Using *T.thermophilus* HB8 genomic DNA as a template, PCR was carried out under the condition described previously (16). Genomic DNA from *T.thermophilus* was digested with *Pst*I and then subjected to Southern hybridization using the ^{32}P-labelled PCR fragment as a probe. The DNA fragments which hybridized with the probe were purified and ligated into pUC119 to produce a mini gene bank of *T.thermophilus* genomic DNA. *Escherichia coli* DH5α cells transformed with the mini gene bank were screened by colony hybridization. Among the positive signals, the DNA fragment (pTS3) containing the entire *T.thermophilus mutS* gene was obtained. The nucleotide sequence of the *Xba*I–*Bam*HI region of pTS3 was determined on both strands by the dideoxy method (Applied Biosystems, Taq cycle sequencing system) using model 373A ABI automated DNA sequencer.

An *Nde*I restriction site was created at the first ATG codon of the *T.thermophilus mutS* gene by PCR-available site-directed mutagenesis. Using the *Nde*I site, *T.thermophilus mutS* gene was cloned into the plasmid pET3a. The resulting plasmid, named pSS1, was used to transform the *E.coli* strain JM109(DE3) harboring the plasmid pLysE. The transformant was cultivated at 37°C in LB medium containing 50 μg/ml ampicillin and 25 μg/ml chloramphenicol until the density of the culture reached 4 × 10^{8} cells/ml. The cells were then incubated for a further 2 h in the presence of 50 μg/ml IPTG and then harvested by centrifugation and stored at –80°C.

Purification of *T.thermophilus* MutS protein

Frozen cells (100 g) were thawed, suspended in buffer I [50 mM Tris–HCl (pH 7.5), 5 mM EDTA, 5 mM β-mercaptoethanol and 25% (w/v) sucrose] and disrupted by sonication on ice. Brij-58 was added to a final concentration of 0.5% (w/v) and then the cell extract was stirred for 30 min at 4°C. Following this, the cell lysate was incubated at 70°C for 10 min and centrifuged (15 000 *g*) for 40 min at 4°C. A DEAE–cellulose DE52 column (bed volume 500 ml) was equilibrated with buffer II [50 mM Tris–HCl (pH 7.5), 1 mM EDTA, 5 mM β-mercaptoethanol and 10% (v/v) glycerol] before the supernatant was loaded. The column was then washed once with 150 ml of buffer II and eluted with a 3000 ml gradient of 0–500 mM NaCl in buffer II. Ammonium sulfate was added to the MutS fraction to give a final concentration of 15% (w/v) and then it was loaded onto a Phenyl-Toyopearl 650M column (bed volume 60 ml) previously equilibrated with 15% (w/v) ammonium sulfate in buffer II. The column was washed with 50 ml of the same buffer and then eluted with a 600 ml gradient of 15–0% (w/v) ammonium sulfate in buffer II. Fractions containing the protein were loaded onto a Sephacryl S-300 HR column (bed volume 400 ml) previously equilibrated with buffer III [50 mM Tris–HCl (pH 7.5), 100 mM

KCl, 0.1 mM EDTA, 5 mM β-mercaptoethanol and 10% (v/v) glycerol]. The column was eluted with buffer III, the fractions containing MutS protein were concentrated and then stored at 4°C. At each step, the column fractionation pattern was analyzed by SDS–PAGE. The final amount of purified MutS protein was 40 mg and the yield was 0.4 mg protein/g cells.

The N-terminal amino acid sequence of the purified protein was analyzed using an automated protein sequencer (ABI, model 473A). The molar extinction coefficient of *T.thermophilus* MutS protein was calculated to be 51 000 $M^{-1}cm^{-1}$ at an absorption maximum of ~278 nm, using the procedure described previously (17).

ATPase assay

Hydrolysis of [α-^{32}P]ATP by *T.thermophilus* MutS protein was assayed by thin-layer chromatography (TLC) between 4 and 95°C (18). The radioactive counts from ATP and its hydrolysis product, ADP, were quantified using a BAS2000 image analyzer (Fuji photo film).

Mismatch DNA binding assay

Mismatch DNA binding was measured by gel retardation assay. Complementary strands of a 16mer or a 37mer oligonucleotide were synthesized (Fig. 6). Each top strand was labeled at the 5′ end with [γ-^{32}P]ATP using with polynucleotide kinase. The oligonucleotides were annealed at 70°C (16mer) or 95°C (37mer) for 10 min, allowed to cool to 30°C (at 1°C/min), and then placed on ice until required. When the assay was carried out using 37 bp duplex DNA as a substrate, each binding assay reaction mixture (total volume 10 µl), comprising 0.2 pmol oligonucleotide duplex, the assay buffer [50 mM Tris–HCl (pH 7.5), 100 mM KCl, 0.1 mM EDTA, 5 mM β-mercaptoethanol, 10% (v/v) glycerol and 20 mM $MgCl_2$] and *T.thermophilus* MutS protein, was incubated at 35 or 60°C for 30 min. After adding 2 µl 50% (w/v) sucrose, the binding assay reaction mixtures were loaded onto a 6% (w/v) acrylamide gel containing 20 mM magnesium acetate and run at 35°C for 90 min or 60°C for 70 min at 76 V in 1× TAE buffer (15) containing 20 mM magnesium acetate. The gel was then dried and placed in contact with an imaging plate. The bands were analyzed using a BAS2000 image analyzer. When the assay was carried out using 16 bp duplex DNA, concentration of magnesium ion was adjusted to 5 mM in the assay buffer and 1 mM in the running buffer, respectively.

RESULTS

Cloning, sequencing and primary structures of the *T.thermophilus mutS* gene

PCR was used to amplify a part of the *T.thermophilus mutS* gene using primers designed from highly conserved regions of five MutS proteins. The entire length of the *mutS* gene was cloned, as described in Materials and Methods, and its nucleotide sequence was determined. The G+C content of the DNA fragment was 68% (data not shown), which agrees well with that of genomic DNA from *T.thermophilus* HB8 (69%) (14). One open reading frame was found consisting of 2457 nucleotides, which encoded a protein comprising 819 amino acid residues with a molecular mass estimated at 91 379 Da. The amino acid sequence of *T.thermophilus* MutS protein is similar to those of the other MutS homologues (Fig. 1). It showed 56 and 39% homology with *E.coli* MutS (19) and human hMsh2 (8) proteins respectively and common to all MutS homologues, the sequence around Gly596 containing the Walker's A-type nucleotide binding motif (20), was also conserved in *T.thermophilus* MutS protein.

Complementation of the high mutagenicity of an *E.coli mutS* mutant by the *T.thermophilus mutS* gene

A complementation test was carried out to determine whether *T.thermophilus* MutS protein had a similar activity to that of *E.coli*. The spontaneous mutation rates of wild-type and *mutS* *E.coli* strains, containing the appropriate plasmids, were assessed by measuring the frequency of rifampicin resistance mutations in each strain. As summarized in Table 2, the mutability of *mutS* mutant harboring the plasmid carrying the *T.thermophilus mutS* gene was only a quarter of that for the *E.coli mutS* mutant strain, GM4271. Thus, *T.thermophilus* MutS protein would appear to play the same function that of *E.coli*.

Overproduction and purification of *T.thermophilus* MutS protein

In order to analyze the biochemical properties of *T.thermophilus* MutS protein, it was overproduced in *E.coli* and purified. The original clone of *T.thermophilus mutS* gene from pTS3 was subcloned into the expression vector pET3a. As shown in Figure 2, an IPTG-induced band, observed at ~91 kDa, was identical to the molecular mass calculated from the amino acid sequence of *T.thermophilus* MutS protein.

Table 2. Mutation frequencies in *E.coli* strains harboring plasmid

Strain (relevant genotype)	Plasmid	Frequency of Rif^r mutants/10^8 cells	Relative mutation frequency
GM4244 (wild-type)	pUC118	0.6	1
GM4244 (wild-type)	pTS3	1.0	1.7
GM4271 (*mutS*)	pUC118	37.2	62.0
GM4271 (*mutS*)	pTS3	9.9	16.5

The rates of spontaneous mutation were assessed by measuring the frequency of mutation to rifampicin resistance (45). *Escherichia coli* cultures were grown until stationary phase, then, the cells were diluted and plated onto LB medium containing either 50 µg/ml ampicillin alone or combined with 100 µg/ml rifampicin. After incubation at 37°C, the numbers of colonies were counted. The mutation frequencies were calculated as the rate of formation of colonies resistant to rifampicin relative to those resistant to ampicillin.

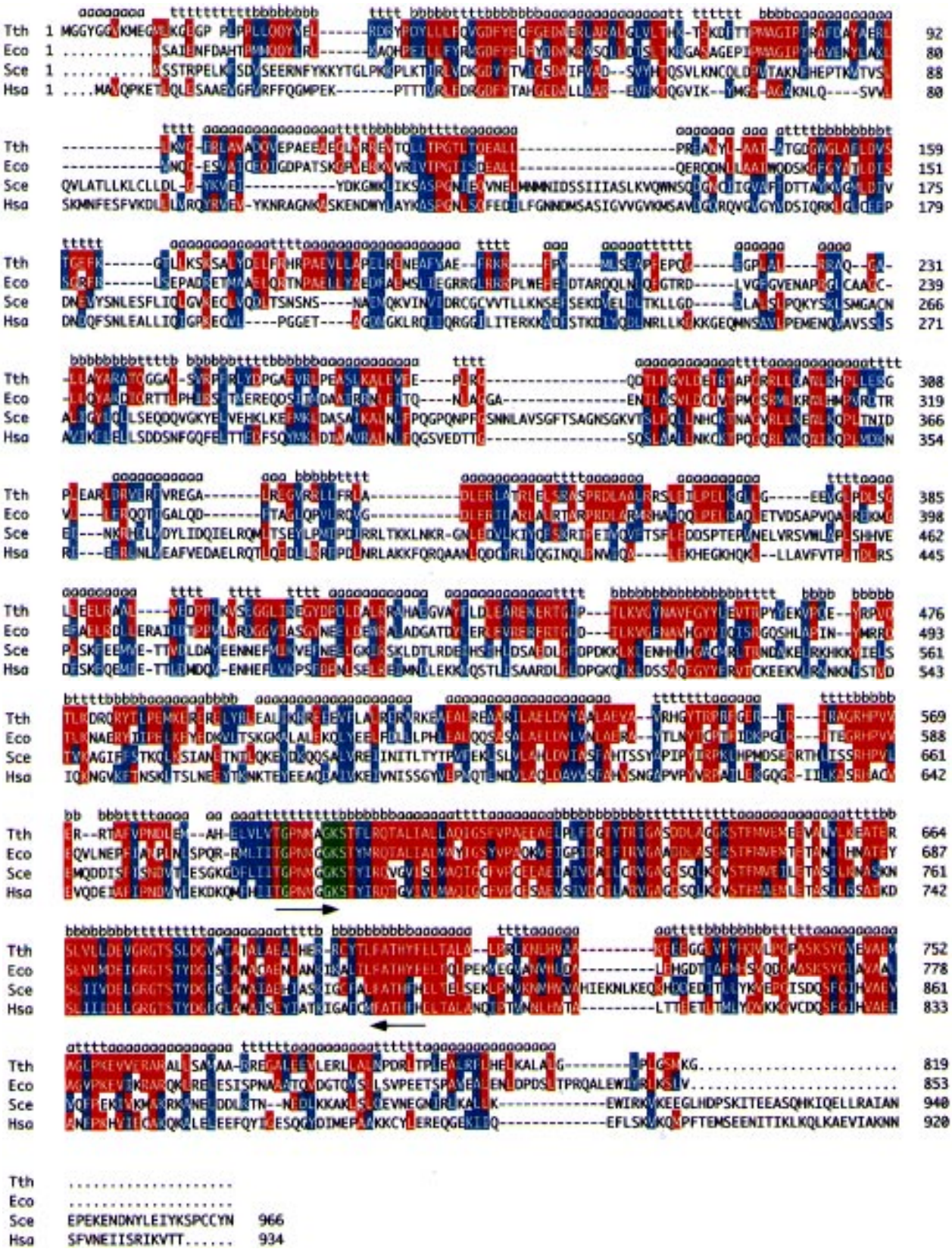

Figure 1. The amino acid sequence and alignment of *T.thermophilus* MutS (Tth), *E.coli* MutS (Eco), *S.cerevisiae* Msh2 (Sce) and *H.sapience* hMsh2 (Hsa) proteins. Identical amino acid residues are indicated white letters on red and homologous residues (R/K, D/E, S/T,N/Q or L/I/V/F/Y/M/W/A) are white letters on blue. The Walker's A-type motif is indicated white letters on green. Deleted amino acid residues in each protein are indicated by bars (–). The regions used for PCR primers are indicated by arrows. Predicted secondary structures of α-helix, β-sheet and turn are indicated by 'a', 'b' and 't', respectively.

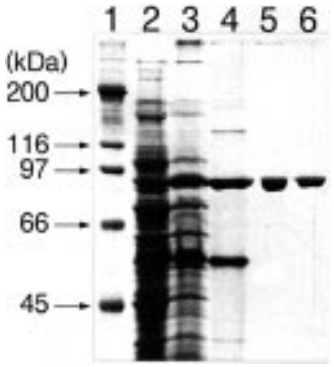

Figure 2. Overproduction and purification of *T.thermophilus* MutS protein were analyzed staining the SDS–PAGE gel [7.5% (w/v) acrylamide] with Coomassie brilliant blue. Lane 1, molecular mass markers (myosin, 200 kDa; β-galactosidase, 116 kDa; phosphorylase B, 97 kDa; bovine serum albumin, 66 kDa; ovalbumin, 45 kDa); lane 2, total cell extract; lane 3, supernatant after heat treatment; lanes 4, 5 and 6, DEAE–cellulose DE52, Phenyl-Toyopearl 650M and Sephacryl S-300 HR chromatography fractions, respectively. About 5–10 μg of protein was loaded into each lane of the gel.

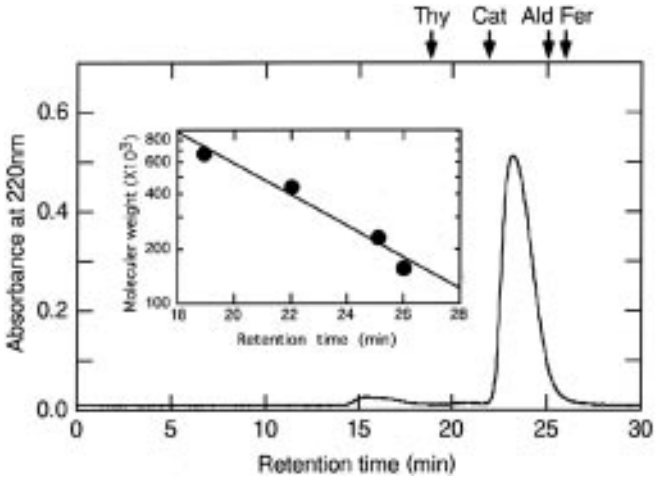

Figure 3. Size-exclusion chromatography of *T.thermophilus* MutS protein. Measurements were performed at room temperature using a buffer solution comprising 100 mM KCl and 50 mM Tris–HCl (pH 7.5), 0.1 mM EDTA, 5 mM β-mercaptoethanol and 10% glycerol. A column used was Pharmacia Superdex 200HR (1 cm × 30 cm). 2.3 μl of 7.85 mg/ml MutS protein was injected to the column and analyzed by absorbance at 220 nm with a flow rate of 0.5 ml/min. The molecular size markers were: Thy, thyroglobulin (669 kDa); Fer, ferritin (440 kDa); Cat, catalase (232 kDa); Ald, aldolase (158 kDa). The inset shows the plots of molecular weight of protein against retention time.

T.thermophilus MutS protein was purified from the *E.coli* cells, which overproduced the protein. After heat treatment, which removed most of the endogenous *E.coli* proteins, the protein was purified to homogeneity by sequential column chromatography on DEAE–cellulose DE52, Phenyl-Toyopearl 650M and Sephacryl S-300 HR (Fig. 2). The N-terminal amino acid sequence was found to be G-G-Y-G-G-V-K-M-E-G-M-L-K-G-E-G-P-G-P-L-. The sequence was identical to that expected after translation from the nucleotide sequence of the *T.thermophilus mutS* gene expect that the N-terminal Met residue was truncated.

Physicochemical characteristics of *T.thermophilus* MutS protein

Molecular size. We attempted to characterize the purified *T.thermophilus* MutS protein using various physicochemical approaches. First, we used size-exclusion chromatography to estimate the molecular size of *T.thermophilus* MutS protein in its native state. Its molecular mass was calculated to be ~330 kDa (Fig. 3), which was 3.6 times larger than its monomer (91.4 kDa). This indicated that *T.thermophilus* MutS protein exists as tetramer in solution.

Secondary structure. The far-UV circular dichroic (CD) spectrum of the recombinant protein was examined to obtain information on the conformation of its polypeptide backbone. It had negative double maxima at ~210 and 220 nm (Fig. 4A), characteristic of an α-helical structure. The helical content of *T.thermophilus* MutS protein was estimated to be ~47% using the method of Chen *et al.* (21). Summers and co-workers (22) found that, for proteins with >40% homology, ≥80% of the side-chain orientations are identical. Since the homologies between the four MutS proteins (*T.thermophilus* MutS, *E.coli* MutS, *S.cerevisiae* Msh2 and human hMsh2), were >~40%, they are expected to have conformations which are almost identical. The method of Chou and Fasman (23) was used to predict the common secondary structures among the homologues (Fig. 1). The α-helical and β-sheet contents of these MutS proteins were estimated to be 49 and 19%, respectively. The predicted α-helical content is well in agreement with the value (47%) estimated from the CD measurement.

Temperature and pH stability. We have used the residue molar ellipticity at 222 nm ($[\theta]_{222}$) to determine the thermal stability of *T.thermophilus* MutS protein. As shown in Figure 4B, the protein was stable between 1 and 80°C, above this temperature, it was denatured irreversibly. The effect of pH on the stability of *T.thermophilus* MutS protein was studied on the basis of its $[\theta]_{222}$. The protein was stable between pH 6 and 12 and between pH 1.5 and 3.5 whereas it aggregated between pH 4 and 6 and its ellipticity could not be measured (Fig. 4C).

Biochemical characteristics of *T.thermophilus* MutS protein

ATPase activity. The Walker's A-type nucleotide binding motif was conserved in MutS proteins (Fig. 1) and ATPase activity has been shown in *S.typhimurium* MutS (24), *E.coli* MutS (25) and *S.cerevisiae* Msh1 (26). In order to find out whether the purified *T.thermophilus* MutS protein also has ATPase activity, we incubated it with [α-^{32}P]ATP and separated out the reaction products using TLC. *T.thermophilus* MutS protein hydrolyzed ATP to ADP and Pi (data not shown) and its activity was maximal at 80°C (Fig. 5). To determine the K_m and k_{cat} values of *T.thermophilus* MutS protein, ATPase activity was measured in the presence of various concentrations of ATP. At 65°C, the values for K_m and k_{cat} were calculated to be 130 μM and 0.11 s^{-1}, respectively.

DNA binding activity. In order to assess the ability of *T.thermophilus* MutS protein to bind to heteroduplex and homoduplex DNAs, we used a gel shift assay to measure MutS binding to oligonucleotide duplexes with and without a G–T mismatch. *Thermus thermophilus* MutS protein specifically bound to 16 bp heteroduplex DNA at 35°C (data not shown). In addition, the binding of *T.thermophilus* MutS protein to DNA was studied at 35°C and at 60°C using a 37 bp duplex DNA as a substrate.

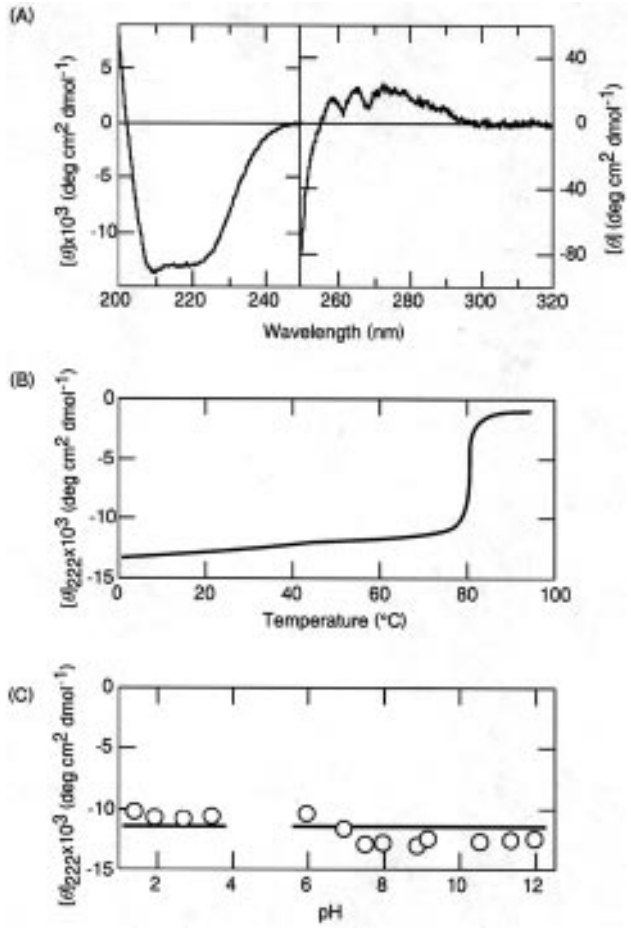

Figure 4. (**A**) Far-UV (left) and near-UV (right) CD spectra of *T.thermophilus* MutS protein. Measurements were performed at 25°C in a buffer solution containing 50 mM Tris–HCl (pH 7.5), 100 mM KCl and 1 mM (200–250 nm) or 5.47 mM (>250 nm) *T.thermophilus* MutS protein. The residue molar ellipticity [θ] is defined as $100.\theta_{obs} \cdot (lc)^{-1}$, where θ_{obs} is the observed molar ellipticity, *l* is the length of the light path in cm and *c* is the residue molar concentration of *T.thermophilus* MutS protein. (**B**) Temperature dependence of the residue molar ellipticity [θ] at 222 nm. Measurement was performed in a solution containing 50 mM Tris–HCl (pH 7.5), 100 mM KCl and 1 μM *T.thermophilus* MutS protein. The heating rate was 1°C/min. (**C**) pH-dependent changes in residue molar ellipticity [θ] at 222 nm. Measurements were performed at 25°C after incubating 1 μM *T.thermophilus* MutS protein with each buffer containing 100 mM KCl at 25°C for 24 h. The buffers used were: 70 mM HCl (pH 1.13), 20 mM glycine–HCl (pH 1.89–3.41), 20 mM Na-acetate (pH 4.13–5.34), 20 mM Na-phosphate (pH 5.95–6.92), 20 mM Tris–HCl (pH 7.49–8.89), 20 mM Na-borate (pH 9.14) and 20 mM Na-carbonate (pH 10.52–11.30).

Thermus thermophilus MutS protein specifically bound to heteroduplex DNA at 35°C (Fig. 7A) and also at 60°C (Fig. 7B). The binding affinity of *T.thermophilus* MutS protein for the G–T containing heteroduplex was four to five times greater rather than that for the G–C containing homoduplex at 35°C (Fig. 7C) and was 8 to 10 times at 60°C (Fig. 7D).

DISCUSSION

MutS protein from *T.thermophilus*

The *mutS* gene from an extremely thermophilic bacterium, *T.thermophilus* HB8, was cloned and its nucleotide sequence determined.

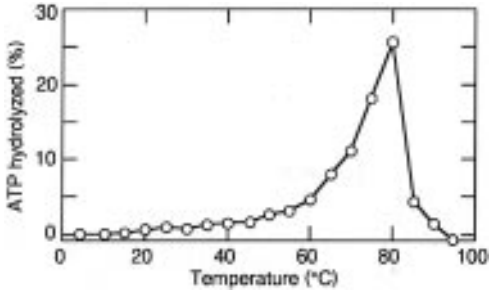

Figure 5. The ATPase activity of *T.thermophilus* MutS protein. The activity was measured as the amount of ATP hydrolysis (%) which occurred after 15 min of incubation at the indicated temperature. The reaction mixture contained 1 μM *T.thermophilus* MutS protein, 50 mM Tris–HCl (pH 7.5), 100 mM KCl, 1 mM DTT, 10 mM $MgCl_2$ and 1 mM ATP.

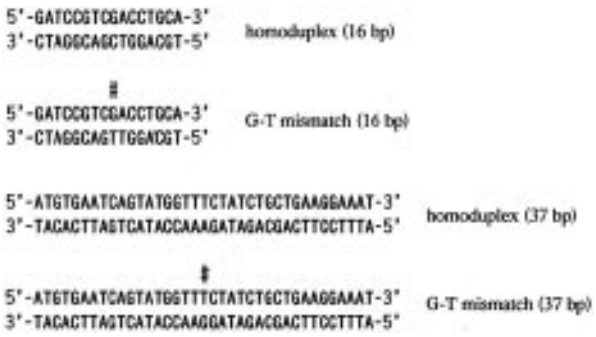

Figure 6. Oligonucleotide duplexes used as substrate for gel shift analysis. G–T mismatch base pairs are indicated by '#'.

It encoded a 91.4 kDa protein which had 56 and 39% homology with *E.coli* MutS and human hMsh2 proteins, respectively. This is the first time that the *mutS* gene has been isolated from a thermophilic bacterium. We found that the *T.thermophilus mutS* gene could complement the high mutability of the *E.coli mutS* mutant. This suggests that there is a mismatch repair system in *T.thermophilus* which is similar to that found in *E.coli* and that the functions of MutS protein from both types of bacteria are the same.

The purified *T.thermophilus* MutS protein, overproduced in *E.coli*, maintained its secondary structure (Fig. 4A) and activities (Figs 5 and 7). We predicted the secondary structures of the MutS homologues (Fig. 1), based on the observation that the three-dimensional structure of homologous proteins are almost identical (22). The α-helical content of the MutS proteins was estimated to be 49%, and this value agreed well with that predicted from the CD measurement (47%). It should be noted that the predicted secondary structure of the N-terminal region for MutS proteins are similar even though the homology of this region is relatively low.

Measurements of its secondary structure and ATPase activity have shown that *T.thermophilus* MutS protein was heat stable at temperatures ≤80°C. This thermostability could be related to its amino acid sequence. The following amino acid residues, known to be chemically unstable at high temperatures, were decreased in number in *T.thermophilus* MutS protein when compared to those of *E.coli* MutS protein: Cys (6 to 2), Asn (27 to 7), Gln (42 to 17) and Met (25 to 14). Whereas the number of Pro residues was increased from 40 in *E.coli* MutS protein to 47 in *T.thermophilus*.

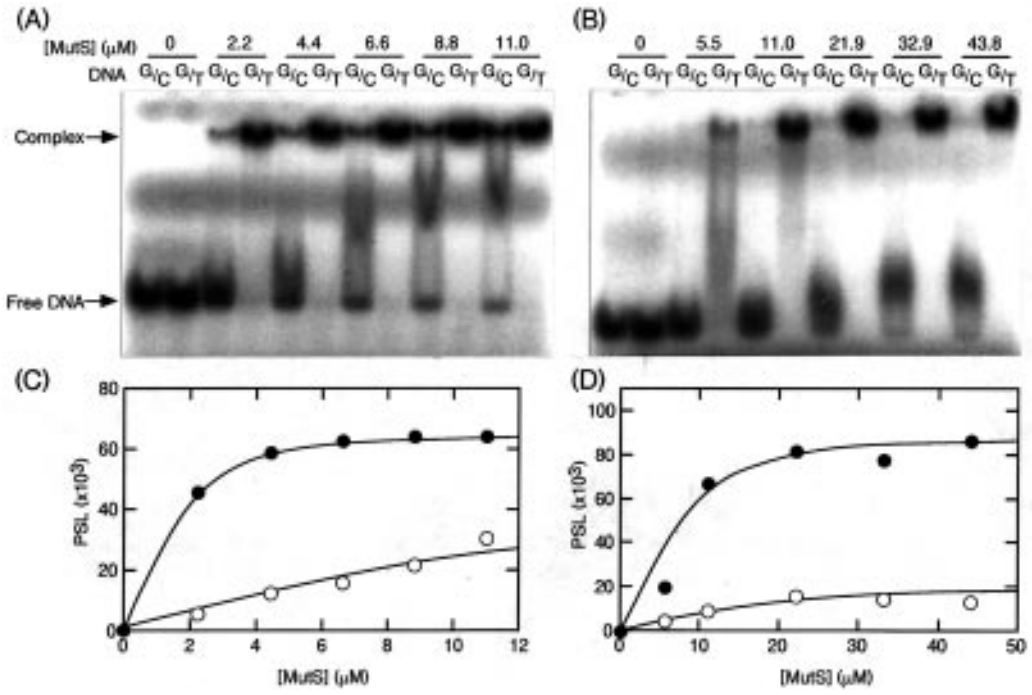

Figure 7. DNA binding of *T.thermophilus* MutS protein. A gel shift analysis of *T.thermophilus* MutS protein binding to homoduplex (complementary duplex) and heteroduplex (mismatched duplex) was carried out. Complementary ($^{G}/_{C}$) or mismatched ($^{G}/_{T}$) oligoduplexes were incubated with the indicated amount of *T.thermophilus* MutS protein at 35°C for 30 min, the samples were then electrophoresed using non-denatured acrylamide gel at 35°C (**A**). The same reactions were carried out at 60°C and the samples were electrophoresed at 60°C (**B**). The radioactivity present in the shifted bands at 35°C (**C**) or at 60°C (**D**) was determined using a BAS2000 image analyzer and plotted. Open circles and closed circles indicate the amount of the protein bound complex to G–C homoduplex and to G–T heteroduplex, respectively.

Pro residues have the highest β-turn potential of all the amino acids and play an important role in peptide folding and globular structure formation. They are considered to decrease the entropy of the protein in its denatured state or increase the conformational enthalpy in its native state, thereby increasing protein stability (27). Similar changes have also been observed in other thermostable proteins (16,28,29).

Protein–protein interaction of MutS protein

Escherichia coli MutS protein interacts with MutL homodimers forming a ternary complex (30). Therefore, the result of the complementation test strongly suggests that *T.thermophilus* MutS protein could be capable of interacting with *E.coli* MutL protein to produce functional repair complexes. Incomplete complementation could be due either to a low level of *T.thermophilus mutS* gene expression in *E.coli*, to weak interactions between inter-species Mut proteins and/or to submaximal activity of *T.thermophilus* MutS protein in *E.coli* cultured at 37°C.

The results of the size-exclusion chromatography strongly suggested that *T.thermophilus* MutS protein exists as a tetramer in its native state. *E.coli* MutS protein is known to exist as both monomer and oligomer (31), however, human hMsh2 protein forms heterodimers with GTBP (G–T binding protein), which is a new MutS homologue in eukaryotes (32,33). On the other hand, hMsh2 protein forms a protein–mismatch DNA complex where it can exist as either the monomer, dimer or multimer (34). For MutL homologues, *E.coli* MutL forms a homodimer (30), and human MutL homologues form heterodimers (35,36). These reports suggest that the oligomerization of the mismatch repair proteins is conserved between prokaryotes and eukaryotes. Therefore, eukaryotic MutS homologues may also be capable of forming tetramers.

Although at neutral pH, *T.thermophilus* MutS protein was a tetramer in the absence of SDS, it was highly aggregated between pH 4 and 6, most probably due to isoelectric precipitation. The isoelectric point of *T.thermophilus* MutS protein, calculated from its amino acid sequence, is 5.6 and therefore, its net charge will approach zero at pH 4–6. The region of *T.thermophilus* MutS protein responsible for protein–protein interaction may be hydrophobic, as reported for many other proteins (37).

Table 3. Kinetic parameters of MutS homologues

	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} M^{-1})	Temp.(°C)	References
T.thermophilus MutS	1.1×10^{-1}	130	8.5×10^{2}	65	This work
S.typhimurium MutS	4×10^{-3}	6	6.7×10^{2}	37	(24)
E.coli MutS	5×10^{-3}			37	(25)
S.cerevisiae Msh1	1×10^{-6}			30	(26)

Biochemical properties

The amino acid sequences of *T.thermophilus* MutS protein and its counterparts in prokaryotes and eukaryotes are similar, and the ATP binding motifs are conserved in all of these proteins. Since it has been shown that *S.typhimurium* MutS (24), *E.coli* MutS (25) and *S.cerevisiae* Msh1 (26) proteins have ATPase activity, it seemed reasonable to expect that *T.thermophilus* MutS protein would also have this activity. Our results show that *T.thermophilus* MutS protein was able to hydrolyze ATP to ADP and Pi. The kinetic parameters (k_{cat} and K_m) and catalytic efficiencies (k_{cat}/K_m) of its ATPase activity are summarized in Table 3. The catalytic efficiencies for many well evolved enzymes are similar over a range of 1×10^6 to 3×10^8 s^{-1} M^{-1} (38), however, the values for MutS proteins were all ~10^3 s^{-1} M^{-1} which are three to five orders of magnitude lower than those of the well evolved enzymes. Since the catalytic efficiencies of MutS proteins were similar at the different physiological temperatures, a low rate constant must be a characteristic of these MutS proteins. This low rate constant suggests the presence of the slow conversion step in the MutS protein–substrate complex. The K_m value of *T.thermophilus* MutS protein at 65°C was more than one order of magnitude higher than that of other prokaryotic MutS protein at 37°C (Table 3). The value of K_m may become relatively high as the temperature rises because the binding step is exothermic ($\Delta H < 0$) reaction.

Thermus thermophilus MutS protein bound efficiently and with high specificity to the G–T mismatched heteroduplex DNA than it did to homoduplex DNA. This mismatch binding activity is similar to those of MutS homologues from other species (31,34,39–42) and suggests that *T.thermophilus* MutS may also play a role in mismatch repair initiation by providing a target for the excision process. The finding of which the non-specific binding of this protein to homoduplex DNA was weaker at 60°C than at 35°C, suggests either that mismatched base pair recognition ability of *T.thermophilus* MutS protein is more efficient at higher temperature or that non-specific DNA–protein complex dissociate more easily. More efficient specific binding at 35°C than 60°C suggests that the MutS–mismatched DNA complex is more unstable at higher temperature. It is interesting to speculate on the nature of the protein–DNA interaction in thermophiles. At higher temperatures, the secondary structure of their DNA alters, therefore, the interaction of proteins with DNA at these temperatures must be stabilized in such a way as to allow their association. The stable complex formed between *T.thermophilus* MutS protein and DNA is a useful model in which to study the molecular mechanism of mismatch recognition.

ACKNOWLEDGEMENT

We are grateful to Dr M. G. Marinus (University of Massachusetts Medical School) for providing the *E.coli* strains GM4244 and GM4271.

REFERENCES

1 Friedberg,E.C., Walker,G.C. and Siede,W. (1995) *DNA Repair and Mutagenesis*. ASM Press, Washington, DC.
2 Modrich,P. (1991) *Annu. Rev. Genet.* **25**, 229–253.
3 Kolodner,R.D. (1995) *Trends Biochem. Sci.* **20**, 397–401.
4 Reenan,R.A.G. and Kolodner,R.D. (1992) *Genetics* **132**, 963–973.
5 Varlet,I., Pallard,C., Radman,M., Moreau,J. and de Wind,N. (1994) *Nucleic Acids Res.* **22**, 5723–5728.
6 Linton,J.P., Yen,J.J., Selby,E., Chen,Z., Chinsky,J.M., Liu,K., Kellems,R.E. and Crouse,G.F. (1989) *Mol. Cel. Biol.* **9**, 3058–3072.
7 Fujii,H. and Shimada,T. (1989) *J. Biol. Chem.* **264**, 10057–10064.
8 Fishel,R., Lescoe,M.K., Rao,M.R.S., Copeland,N.G., Jenkins,N.A., Garber,J., Kane,M. and Kolodner,R. (1993) *Cell* **75**, 1027–1038.
9 Leach,F.S., Nicolaides,N.C., Papadopoulos,N., Liu,B., Jen,J., Parsons,R., Peltomaki,P., Sistonen,P., Aaltonen,L.A., Nystrom-Lahti,M. *et al.* (1993) *Cell* **75**, 1215–1225.
10 Bronner,C.E., Baker,S.M., Morrison,P.T., Warren,G., Smith,L.G., Lescoe,M.K., Kane,M., Earabino,C., Lipford,J., Lindblom,A. *et al.* (1994) *Nature* **368**, 258–261.
11 Papadopoulos,N., Nicolaides,N.C., Wei,Y.-F., Ruben,S.M., Carter,K.C., Rosen,C.A., Haseltine,W.A., Fleischmann,R.D., Fraser,C.M., Adams,M.D. *et al.* (1994) *Science* **263**, 1625–1629.
12 Prolla,T.A., Christie,D.M. and Liskay,R.M. (1994) *Mol. Cell. Biol.* **14**, 407–415.
13 Nicolaides,N.C., Papadopoulos,N., Liu,B., Wei,Y.F., Carter,K.C., Ruben,S.M., Rosen,C.A., Haseltine,W.A., Fleischmann,R.D., Fraser,C.M. *et al.* (1994) *Nature* **371**, 75–80.
14 Oshima,T. and Imahori,K. (1974) *Int. J. Syst. Bacteriol.* **24**, 102–112.
15 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
16 Kato,R. and Kuramitsu,S. (1993) *J. Biochem.* **144**, 926–929.
17 Kuramitsu,S., Hiromi,K., Hayashi,H., Morino,Y. and Kagamiyama,H. (1990) *Biochemistry* **29**, 5469–5476.
18 Weinstock,G. M., McEntee,K. and Lehman,I. R. (1981) *J. Biol. Chem.* **256**, 8856–8858.
19 Schlensog,V. and Bock,A. (1991) *J. Bacteriol.* **173**, 7414–7415.
20 Walker,J.E., Saraste,M., Runswick,M.J. and Gay,N.J. (1982) *EMBO J.* **1**, 945–951.
21 Chen,Y.H., Yang,J.T. and Martinez,H.M. (1972) *Biochemistry* **11**, 4120–4131.
22 Summers,N.L., Carlson,W.D. and Karplus,M. (1987) *J. Mol. Biol.* **196**, 175–198.
23 Chou,P.Y. and Fasman,G.D. (1978) *Adv. Enzymol.* **47**, 45–148.
24 Haber,L.T. and Walker,G.C. (1991) *EMBO J.* **10**, 2707–2715.
25 Au,K.G., Welsh,K. and Modrich,P. (1992) *J. Biol. Chem.* **267**, 12142–12148.
26 Chi,N.-W. and Kolodner,R.D. (1994) *J. Biol. Chem.* **269**, 29993–29997.
27 Watanabe,K., Masuda,T., Ohashi,H., Mihara,H. and Suzuki,Y. (1994) *Eur. J. Biochem.* **226**, 277–283.
28 Argos,P., Rossmann,M.G., Grau,U.M., Zuber,H., Frank,G. and Tratschin,J.D. (1979) *Biochemistry* **18**, 5698–5703.
29 Okamoto,A., Kato,R., Masui,R., Yamagishi,A., Oshima,T. and Kuramitsu,S. (1996) *J. Biochem.* **119**, In press.
30 Grilley,M., Welsh,K.M. Su,S.-S. and Modrich,P. (1989) *J. Biol. Chem.* **264**, 1000–1004.
31 Su,S.S. and Modrich,P. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5057–5061.
32 Drummond,J.T., Li,G.-M., Longley,M.J. and Modrich,P. (1995) *Science* **268**, 1909–1912.
33 Polombo,F., Gallinari,P., Iaccarino,I., Lettieri,T., Hughes,M., D'Arrigo,A., Truong,O., Hsuan,J.J. and Jiricny,J. (1995) *Science* **268**, 1912–1914.
34 Fishel,R., Ewel,A., Lee,S., Lescoe,M.K. and Griffith,J. (1994) *Science* **266**, 1403–1405.
35 Prolla,T.A., Pang,Q., Alani,E., Kolodner,R.D. and Liskay,R.M. (1994) *Science* **265**, 1091–1093.
36 Li,G.-M. and Modrich,P. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 1950–1954.
37 Janin,J. and Chothia,C. (1990) *J. Biol. Chem.* **265**, 16027–16030.
38 Radzicka,A., and Wolfenden,R. (1995) *Science* **267**, 90–93.
39 Jiricny,J., Su,S.-S., Wood,S.G. and Modrich,P. (1988) *Nucleic Acids Res.* **16**, 7843–7853.
40 Fishel,R., Ewel,A. and Lescoe,M.K. (1994) *Cancer Res.* **54**, 5539–5542.
41 Chi,N.-W. and Kolodner,R.D. (1994) *J. Biol. Chem.* **269**, 29984–29992.
42 Alani,E., Chi,N.-W. and Kolodner,R.D. (1995) *Genes Dev.* **9**, 234–247.
43 Wu,T.-H. and Marinus,M.G. (1994) *J. Bacteriol.* **176**, 5393–5400.
44 Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol.* **185**, 60–89.
45 Pang,P.P., Lundberg,A.S. and Walker,G.C. (1985) *J. Bacteriol.* **163**, 1007–1015.