THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 274, No. 33, Issue of August 13, pp. 23673–23678, 1999
Printed in U.S.A.

Oligomerization of a MutS Mismatch Repair Protein from *Thermus aquaticus**

(Received for publication, April 26, 1999, and in revised form, June 2, 1999)

Indranil Biswas,[a,b] Changill Ban,[c] Karen G. Fleming,[d,e] Jun Qin,[f,g] Jeffrey W. Lary,[h] David A. Yphantis,[h,i] Wei Yang,[c] and Peggy Hsieh[a,j]

From the [a]Genetics and Biochemistry Branch and [c]Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892, the [d]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, the [f]NHLBI, National Institutes of Health, Bethesda, Maryland 20892, and the [h]Department of Molecular and Cell Biology and National Analytical Ultracentrifuge Facility, University of Connecticut, Storrs, Connecticut 06269, and the [i]Biology Department and Institute of Molecular Biology and Biotechnology, University of Crete, Vasilika Vouton, 71409 Heraklion, Crete, Greece.

The MutS DNA mismatch protein recognizes heteroduplex DNAs containing mispaired or unpaired bases. We have examined the oligomerization of a MutS protein from *Thermus aquaticus* that binds to heteroduplex DNAs at elevated temperatures. Analytical gel filtration, cross-linking of MutS protein with disuccinimidyl suberate, light scattering, and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry establish that the *Taq* protein is largely a dimer in free solution. Analytical equilibrium sedimentation showed that the oligomerization of *Taq* MutS involves a dimer-tetramer equilibrium in which dimer predominates at concentrations below 10 μM. The ΔG^0_{2-4} for the dimer to tetramer transition is approximately -6.9 ± 0.1 kcal/mol of tetramer. Analytical gel filtration of native complexes and gel mobility shift assays of an maltose-binding protein-MutS fusion protein bound to a short, 37-base pair heteroduplex DNA reveal that the protein binds to DNA as a dimer with no change in oligomerization upon DNA binding.

DNA mismatch repair plays an important role in safeguarding the integrity of the genome. This repair pathway recognizes and corrects mispaired bases arising from misincorporation of bases during DNA replication, chemical damage to bases, and the formation of heteroduplex DNA containing mispaired or unpaired bases during homologous recombination (reviewed in Refs. 1 and 2). In addition, mismatch repair safeguards the fidelity of genetic recombination (reviewed in Ref. 3). The demonstration that defects in mismatch repair lead to a significant increase in spontaneous mutation rates and an elevated risk for tumor development highlights the importance of mismatch repair in the maintenance of genome stability (reviewed in Refs. 4 and 5).

The most extensively studied mismatch repair pathway is the DNA adenine methylation-directed mismatch repair pathway of *Escherichia coli*. Genetic and biochemical studies have identified at least 10 gene products required for methyl-directed repair. These include MutS, MutL, MutH, DNA helicase II, single-stranded DNA-binding protein, exonuclease I, exonuclease VII or RecJ exonuclease, DNA polymerase III holoenzyme, and DNA ligase (reviewed in Ref. 2). Repair has been reconstituted *in vitro* from purified components (6, 7).

Recognition of mispaired or unpaired bases is carried out by the MutS protein, a family of related proteins whose members are found in organisms ranging from bacteria to humans. *E. coli* MutS protein recognizes up to four consecutive unpaired bases as well as seven of eight possible mismatches; C:C mismatches are poorly repaired by the methyl-directed mismatch repair pathway (reviewed in Ref. 2). Given this relatively broad substrate specificity of MutS, DNA binding by MutS proteins poses an interesting problem of protein-DNA recognition.

We have identified a MutS homolog from the thermophilic eubacterium *Thermus aquaticus* YT-1 (8). This 90.7-kDa protein binds *in vitro* to heteroduplex DNAs containing mispaired or unpaired bases over a wide temperature range from 4 to 70 °C and has a thermostable ATPase activity. In an effort to elucidate the molecular mechanism of substrate recognition by this protein, we have carried out enzymatic and chemical footprinting of a MutS-heteroduplex DNA complex and identified extensive contacts involving the major and minor grooves of the DNA as well as the phosphate backbone in the immediate vicinity of an unpaired base (9). In addition, photocross-linking and site-directed mutagenesis identified a conserved Phe residue at the amino terminus of the protein that appears to be critical for DNA binding (10).

A critical parameter concerning the mechanism of DNA binding by MutS proteins is protein oligomerization, which has been reported for a number of MutS homologues with varying results. Eukaryotic MutSα and MutSβ proteins are thought to recognize mispaired and unpaired bases as heterodimers of MSH2 and MSH6 or MSH2 and MSH3, respectively (11–19). Co-expression, cross-linking, and co-immunoprecipitation studies indicate a 1:1 molar ratio of MSH2 with MSH3 or MSH6. In the case of hMutSα, gel filtration chromatography and band sedimentation through sucrose density gradients yield a native relative molecular mass of 255 kDa indicative of a heterodimer (11). However, hMutSα can form higher order complexes of unknown stoichiometry with heteroduplex DNAs in gel mobility shift assays (20), suggesting that larger oligo-

* This work was supported in part by National Science Foundation Grants BIR 9218679 and DBI 9603153 (to D. A. Y.) and Grant BIR 9318373 (to the National Analytical Ultracentrifuge Facility). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[b] Present address: National Center for Cell Science, Pune University Campus, Ganeshkhind, Pune 411007 MH, India.

[e] Supported by National Institutes of Health Grants GM 16769 (to K. G. F.) and GM 54160 (to Donald M. Engelman, Yale University, New Haven, CT).

[g] Present address: Dept. of Biochemistry, Baylor College of Medicine, Houston, TX 77030.

[j] To whom correspondence should be addressed: National Institutes of Health, Bldg. 10, Rm. 9D04, 10 Center Dr., MSC 1810, Bethesda, MD 20892-1810. E-mail: hsieh@ncifcrf.gov.

meric species can form under some conditions. *E. coli* MutS protein binds as a homodimer based on electron microscopic images and surface plasmon resonance (21), although MutS-heteroduplex DNA complexes when analyzed by sedimentation (22) exhibit a more complex behavior. Finally, gel filtration chromatography of a thermostable MutS protein from *Thermus thermophilus* HB8 has led to the proposal of a tetrameric structure (23). The 90.7-kDa *Taq* MutS protein migrates anomalously on gel filtration corresponding to a molecular mass of 280 kDa and a Stokes radius of 55.7 Å (9). As part of our efforts to understand the molecular mechanisms underlying substrate recognition by MutS, we use a variety of approaches to determine the oligomerization state of *Taq* MutS protein both in free solution and when bound to a heteroduplex DNA containing an unpaired base.

MATERIALS AND METHODS

MutS Protein—Taq MutS protein was isolated from an *E. coli* BL21(DE3) pLysS (Novagen) overproducing strain and purified to apparent homogeneity on Source 30Q (Amersham Pharmacia Biotech) and Mono Q HR10/10 (Amersham Pharmacia Biotech) anion exchange columns followed by gel filtration on a Sephacryl S-300 Hi-Prep 16/60 column (Amersham Pharmacia Biotech) as described previously (9, 10). The concentration of protein refers to monomers and was determined spectrophotometrically using a molar extinction coefficient of $\epsilon_{280} = 8.1 \times 10^4\ \text{M}^{-1}\ \text{cm}^{-1}$ (10). MutS was also expressed as a fusion protein with *E. coli* maltose-binding protein (MBP)[1] from the plasmid pMALc2 (New England Biolabs). A 220-bp fragment from the amino terminus of the MutS coding region was amplified by polymerase chain reaction using 5′-GGG AAT TCA TGG AAG GCA TGC TGA AGG GCG AGG GCC C-3′ and 5′-CCC AAG CTT TCA GGG GAT CCC CGC CAT GGG GGT GGT GAA GTC C-3′ primer pairs. The amplified polymerase chain reaction product was cloned into pMalc2 at *Hin*dIII and *Eco*RI sites. The resultant recombinant plasmid was further restricted by *Bam*HI and a 2.8-kilobase pair *Bam*HI fragment containing the remainder of the *Taq* MutS coding region from pET3MutS (8) was cloned into it to generate pMBP-MutS. The construct was confirmed by DNA sequencing of the amplified region. MBP-MutS was overproduced in a XL-1Blue host (Stratagene) grown in rich medium (per liter: 10 g of tryptone, 5 g of yeast extract, 5 g of NaCl, 2 g of glucose) with 100 μg/ml ampicillin to an A_{600} of 0.6. isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 0.4 mM, and the cells were incubated for an additional 4 h. Cells were lysed by sonication in 20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA containing Complete protease inhibitor mixture (Roche Molecular Biochemicals). The lysate was cleared by centrifugation and applied to an amylose column (New England Biolabs) and purified according to manufacturer's suggestion. The fusion protein was concentrated in a Centricon-30 (Amicon) and further purified by passage over a MonoQ anion exchange column as described above. The protein concentration was determined by the Bradford assay (Bio-Rad) using wild type MutS protein as a standard.

Chemical Cross-linking—Disuccinimidyl suberate (DSS, Pierce) was added at the indicated concentration to 20 μl of binding buffer (20 mM HEPES, pH 7.8, 5 mM $MgCl_2$, 10% glycerol) and 1 μM MutS protein. Reactions were incubated at room temperature for 30 min and terminated by the addition of 1 μl of 1 M Tris-HCl, pH 7.0, followed by an additional 15 min of incubation. Samples were analyzed on a 6% SDS-polyacrylamide gel after addition of 20 μl of sample loading buffer. For cross-linking with bis(sulfosuccinimidyl)suberate (Pierce) for mass spectroscopy, 0.9 mg/ml MutS was incubated with 0.04 mM of bis(sulfosuccinimidyl)suberate in 20 mM HEPES, pH 7.5, 1 mM EDTA, 1 mM β-mercaptoethanol, and 150 mM NaCl at room temperature for 1 h in a reaction volume of 20 μl.

Laser Light Scattering—Right angle light scattering was performed using a DynaPro-801 molecular sizing instrument (Protein Solutions). MutS protein was analyzed at a concentration of approximately 1 μM in a buffer containing 20 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM dithiothreitol, 300 mM KCl, and 10% glycerol. The scattering of 100 μl of each protein solution was measured seven times with a 5-s interval at room temperature. Data were analyzed using the AutoPro software package. Molecular weight values are reported as the average calculated values, with S.D. values equal to the degree of polydispersity, where a polydispersity index less than 30% of the hydrodynamic radius (R_H) is considered well behaved and unimodal (24).

Mass Spectrometry—A 1.2-m matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer with delayed extraction (Voyager-DE, Perseptive Biosystems, Framingham, MA) was used for measuring free MutS and bis(sulfosuccinimidyl)suberate-cross-linked MutS. The matrix solution was a saturated solution of α-cyano-2,5-dihydroxybenzoic acid in acetonitrile:water (1:1). Aliquots of 1 μl of protein solution were diluted in 10 μl of the matrix solution, and 0.5 μl of the resulting mixture was applied to the sample plate and dried in air prior to analysis. Masses were calibrated externally with bovine serum albumin. Mass accuracy is within ±0.1% .

Sedimentation Equilibrium—Ultracentrifugation experiments were carried out using a Beckman/Coulter XLI ultracentrifuge. Concentration distributions were measured using the Rayleigh interference optics. Care was taken to focus the camera lens on the plane 2/3 of the distance from the entry window to the exit window of the cell, following the procedure of Richards *et al*. (25). Interferometric measurements were usually supplemented by absorbance estimates.

Equilibrium experiments were carried out in ~2.6-mm-high columns using external loading cells (26) to obtain water-water blanks for each speed both before and after running the solutions of interest without disassembly of the ultracentrifuge cells. The program REEDIT9 (or equivalently REEDIT/XLAEDIT)[2] was used to correct all sets of fringe displacements by subtraction of the appropriate blanks and then to edit and select the data ranges to be used for the analyses. Attainment of equilibrium was monitored using one of the MATCH programs (MATCHV7 or MATCHV9)[2] that fit one set of concentration measurements in terms of another set. These programs make it possible to compare data sets, even though there may be offsets in the concentration reference levels or in the radius offsets between two data sets. Experiments were considered to be at equilibrium when the r.m.s. differences between data sets were less than about 0.008 fringe over a period of several hours. Typical equilibrium times for experiments at 20 °C were 17–24 h. The blank-corrected equilibrium distributions were analyzed using WinNONL3[2] or WinNONLR,[2] Windows 95 versions of the NONLIN program of Johnson *et al*. (27). These same programs were used to obtain estimates of $Mw(r)$, the radial weight average molecular weight distribution corresponding to the fitting model. The partial specific volume, $\bar{V}$, of the MutS protein was estimated from the amino acid sequence using the program SEDNTERP (version 1.01)[3] as 0.7446 $\text{cm}^3\ \text{g}^{-1}$ at 20 °C in water. The effects of each concentration of glycerol on the partial specific volume were estimated using the average value of the derivative of $\bar{V}$ with glycerol concentration[4] (estimated as $3.3 \times 10^{-4}\ \text{ml}^2\ \text{g}^{-2}$ by linear regression of the values of ${\phi'}_2^0$ for the several proteins reported by Gekko and Timasheff; Ref. 28). Solvent densities were estimated from composition using SEDNTRP supplemented by data from the International Critical Tables or other standard references.

Analytical Gel Filtration—Fifty pmol of MutS protein was applied on a Superdex 200 PC2.3/30 SMART column (Amersham Pharmacia Biotech) equilibrated with binding buffer. Standards were: thyroglobulin (669 kDa), apoferritin (443 kDa), β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), and albumin (66 kDa) (Sigma). For the analysis of MutS-DNA complexes, 50 pmol of MutS was incubated at 60 °C with 10 pmol of ^{32}P-labeled 37-bp insertion/deletion heteroduplex or homoduplex (9) in 40 μl of binding buffer. The reactions were cooled to room temperature and then applied to a Superdex 200 PC 3.2/30 SMART column (Amersham Pharmacia Biotech) and eluted in binding buffer. Fractions were collected and counted for radioactivity. Peak fractions were also analyzed on native polyacrylamide gels to verify the presence of complexes and on SDS-PAGE to detect the presence of MutS protein.

DNA Binding—DNA binding of MBP-MutS fusion protein was determined by gel mobility shift assays. MBP-MutS protein (1.6 μg) was treated with varying amounts of factor Xa (80, 160, 320, 640, 1000, 1500, 2000, or 4000 ng; Roche Molecular Biochemicals) in 40 μl of volume in binding buffer (20 mM HEPES, pH 7.8, 5 mM magnesium

[1] The abbreviations used are: MBP, maltose-binding protein; DSS, disuccinimidyl suberate; bp, base pair(s); r.m.s., root mean square; PAGE, polyacrylamide gel electrophoresis; ATPγS, adenosine 5′-*O*-(thiotriphosphate); MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight.

[2] These programs, written by J. Lary and D.A. Yphantis, are available on the SPIN6 anonymous FTP site: spin6.mcb.uconn.edu.

[3] This program, written by D. T. Hayes, T. M. Laue, and J. Philo, is available on the RASMB anonymous FTP site: rasmb.bbri.harvard.edu.

[4] This treatment is discussed by James M. Cole on the RASMB E-mail discussion list group.

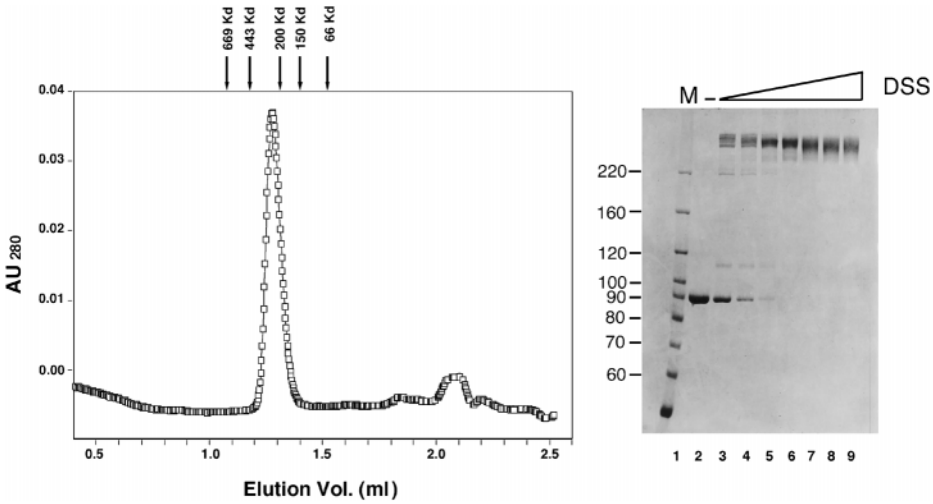

FIG. 1. **Gel filtration chromatography and cross-linking analyses of MutS in free solution.** *Right panel*, MutS in binding buffer containing $MgCl_2$ was applied to a Superdex 200 gel filtration column. *Arrows* indicate the elution positions of standards. *Left panel*, MutS was cross-linked with DSS and analyzed on a 6% SDS-polyacrylamide gel as described under "Materials and Methods." Concentrations of DSS are 12.5, 25, 50, 100, 250, 500, and 1000 μM, in *lanes 2–8*, respectively. *Lane 1*, *M*, molecular size markers.

acetate, 0.1 mM dithiothreitol, 0.1 mM EDTA, and 10% glycerol). After incubation with protease, half of the sample was analyzed by SDS-PAGE to monitor the extent of cleavage. To the remainder, we added 1 pmol of ^{32}P-labeled, 37-bp insertion/deletion heteroduplex. The reaction was incubated for 15 min at 60 °C and electrophoresed on a 4.8% native polyacrylamide gel in TBE containing 10 mM magnesium acetate as described previously (9). To identify the protein species present in complexes bound to DNA, we excised ^{32}P-labeled bands from native gels and electrophoresed them on an 8% SDS-polyacrylamide gel lacking a stacking layer, followed by Coomassie staining. Wild type MutS and intact MBP-MutS fusion protein were also included as positive controls. Molecular weight markers were from Life Technologies, Inc.

RESULTS

Oligomerization of Taq MutS in Free Solution—The oligomerization state of *Taq* MutS protein in free solution under conditions used for heteroduplex DNA binding was analyzed by analytical gel filtration on a Superdex 200 column (Fig. 1). MutS eluted as a 280-kDa protein in the presence of 5 mM $MgCl_2$, corresponding to a Stokes radius of 56 Å similar to previously reported findings obtained in the absence of Mg^{2+} (270 kDa) (8). Since the molecular mass of a *Taq* MutS monomer is 90.6 kDa, these results suggested that MutS protein was present in an oligomeric form in free solution, most likely a dimer or trimer.

Oligomerization of MutS protein was also examined by chemical cross-linking using DSS, a nonspecific cross-linker that targets primary amines. One micromolar MutS protein was cross-linked with varying concentrations of DSS in DNA binding buffer containing 5 mM $MgCl_2$. At low concentrations of DSS, two species predominated representing monomer and a higher order complex migrating near the exclusion limit of a 6% denaturing SDS gel (Fig. 1). In addition, a species migrating at approximately 110 kDa was observed that was probably a result of intrachain cross-links. At higher concentrations of DSS, the monomer was converted entirely to the large, cross-linked species. Similar results were obtained with glutaraldehyde, another nonspecific cross-linker (data not shown).

The finding that MutS protein is an oligomer in solution, as judged by cross-linking and gel filtration chromatography, was further supported by dynamic light-scattering. Experiments were conducted at a protein concentration of approximately 1 μM in a buffer containing 20 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM dithiothreitol, 300 mM KCl, and 10% glycerol at 25 °C. Under these experimental conditions, *Taq* MutS protein was determined to be primarily a dimer with an estimated molecular mass of 187 kDa (see Table I). Significant quantities of monomeric or oligomers larger than dimer were not detected, consistent with the results of cross-linking and gel filtration.

More precise identification of the major *Taq* MutS oligomeric species was obtained by MALDI-TOF mass spectrometry. The mass spectra of the native and glutaraldehyde cross-linked MutS protein in which a distribution of charge state was apparent are shown in Fig. 2 (*A* and *B*, respectively). As noncovalent interactions are too weak to survive the MALDI process, native MutS was detected in the form of a monomer with MALDI-TOF. When MutS was covalently cross-linked with glutaraldehyde prior to MALDI-TOF, a higher molecular mass species of 184.9 kDa was detected, indicating that the cross-linked dimer was the dominant species. Because the mass accuracy of this method is better than 0.1%, these measurements established that the major oligomeric form of MutS is a dimer.

TABLE I
Dynamic light scattering measurements of Taq MutS: estimation of molecular mass and degree of aggregation

D_T	R_H[a]	Molecular mass	Polydispersity
	nm	*kDa*	*%*
336	5.6 ± 1.3	187 ± 47	23

[a] Mean hydrodynamic radius derived from the measured translational diffusion coefficient (D_T) using the Stokes-Einstein equation.

Sedimentation Equilibrium—The analytical ultracentrifuge was used to examine the oligomeric structure of *Taq* MutS protein in solution. Several MutS preparations were examined in 20 mM HEPES, 100 mM NaCl, and 1 mM EDTA, with varying concentrations of glycerol, at pH 7.4 over a range of speeds. Data from the most stable of these preparations are shown here. Fig. 3*A* presents the residuals from a global fit of data from three loading concentrations examined at 9000 and 13,000 rpm. The fitting model was that of a homogeneous, reversibly associating dimer-tetramer equilibrium using the calculated values of σ, the effective reduced molecular weight. The fitting r.m.s. was 0.0138 fringes, and no systematic deviations could be discerned, indicating a good fit to the model. No improvement was seen upon inclusion of other oligomeric species to the fitting model. Convergence for the value of σ also gave no improvement in the fits. Observations of the same solutions at 6000 rpm (data not shown) indicated no detectable contributions from species larger than tetramer. Examination of these solutions at 18,000 rpm (data not shown) showed no evidence of monomer, in agreement with the gel filtration results.

Some marginal improvement was found using fits to the model of a heterogeneous dimer-tetramer. Fig. 3*B* compares the weight-average molecular weight radial distributions,

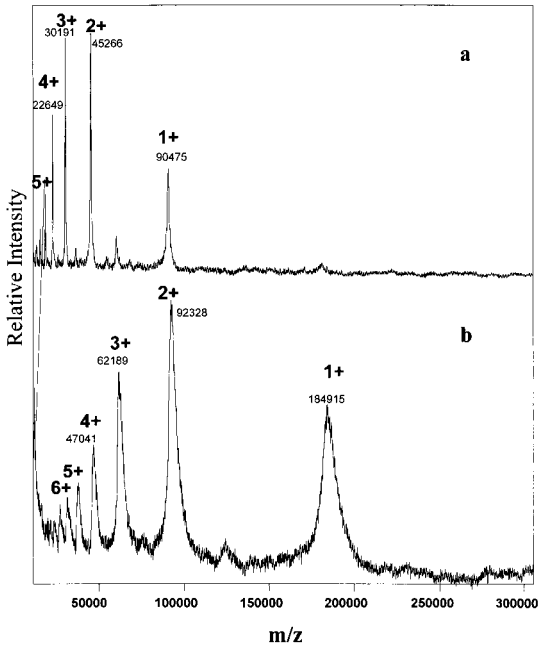

FIG. 2. **MALDI-TOF mass spectra of native *Taq* MutS protein (*a*) and glutaraldehyde cross-linked MutS protein (*b*).** Multiply charged ions with charge states ranging from 1 to 7 were observed.

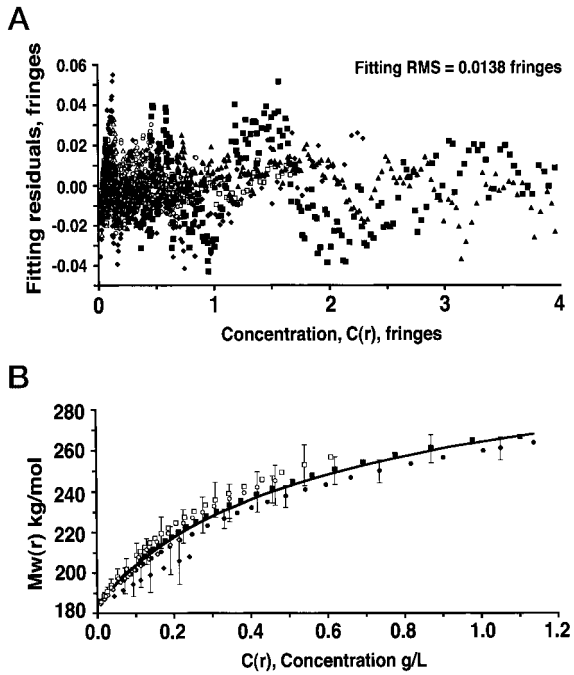

FIG. 3. **Sedimentation equilibrium of MutS protein.** The MutS loading concentrations were 631 μg/ml with 6.6% w/w glycerol (◆, ◇), 312 μg/ml with 3.3% w/w glycerol (●, ○), and 104 μg/ml with 1.1% w/w glycerol (■, □). Solutions were run at 6000, 9000, 13,000, and 18,000 rpm at 20 °C. Only results at 9000 rpm (*filled symbols*) and at 13,000 rpm (*open symbols*) are shown. *A*, the residuals from a least-squares fit of six data sets to the model of a homogeneous self-associating system, using a single, global association constant for all data sets. The r.m.s. residuals are of a magnitude typical for the XLI ultracentrifuge, show no systematic trends, and indicate a reasonable fit of the data to this model. *B*, the weight average molecular weights at each position, $Mw(r)$, estimated from nonlinear least squares fitting as a function of the observation concentrations, $C(r)$. The *solid curve* shows the values found for the fitting model of a homogeneous dimer-tetramer equilibrium, using a *common* equilibrium constant for all data sets. Values of $Mw(r)$ from a fit to the model of a heterogeneous dimer-tetramer equilibrium, with *separate equilibrium constants for each data set*, are shown as discrete points. (About 80% of the data points have been removed for clarity.) The r.m.s. of the residuals for this heterogeneous fit was 0.0132 fringes.

$Mw(r)$, calculated from the least-squares fits for both fitting models (homogeneous and heterogeneous self-association of dimers to form tetramers). Within experimental error, all the $Mw(r)$ data appear to fit a single curve, indicating this preparation to be a homogeneous, reversibly associating system (29–31). However, variance ratio tests show these two fitting models to be different only at the 95% confidence level ($p = 0.05$), indicating the likelihood of some heterogeneity.

The molar equilibrium constants for formation of tetramer from dimer for all data sets and for both models show overlapping ranges, as would be expected from Fig. 3*B*. The average value of the equilibrium constants for formation of tetramer from dimer is 1.3×10^5 liters mol^{-1}. The value of ΔG^0_{2-4}, the Gibbs standard free energy for the formation of tetramer from dimer estimated from these equilibrium constants is -6.9 ± 0.1 kcal mol^{-1} of tetramer, with extreme values of -6.5 to -7.0 kcal mol^{-1} for the values from the individual, separate data sets. Measurements with the absorbance optics at 280 and 260 nm closely mirrored these observations (data not shown).

Analytical Gel Filtration Chromatography of a MutS-Heteroduplex DNA Complex—Determination of the oligomerization state of MutS protein when bound to a heteroduplex DNA was made by analytical gel filtration chromatography. Fifty pmol of *Taq* MutS protein was incubated at 60 °C for 10 min in binding buffer with 10 pmol of ^{32}P-labeled heteroduplex DNA (37 bp) containing an unpaired thymidine residue. The reaction was cooled to room temperature and loaded on a Superdex 200 gel filtration column equilibrated in binding buffer containing Mg^{2+} (Fig. 4*A*). The ^{32}P-labeled heteroduplex bound to MutS eluted at the position corresponding to MutS dimer (see Fig. 1), indicating that MutS binds to DNA as a dimer and that DNA binding does not induce a change in the oligomerization of MutS protein. Fractions corresponding to the bound complex were analyzed by SDS-PAGE to confirm the presence of MutS protein (data not shown) and on native polyacrylamide gels to confirm the presence of MutS-DNA complexes (Fig. 4*B*). The experiment was repeated with a 37-bp homoduplex DNA as a control (Fig. 4*C*). Less than 4% of radiolabeled homoduplex DNAs eluted at the position of MutS dimer reflecting the much lower affinity of *Taq* MutS for homoduplex DNA (8); the remainder eluted at the position of free DNA. Repetition of gel filtration experiments with increasing concentrations of DNA up to a 1:1 molar ratio of DNA heteroduplex:protein gave the same results, namely the elution of bound heteroduplex DNA exclusively at the position of protein dimer. No change in protein oligomerization was observed upon DNA binding.

Heteroduplex DNA Binding by a MBP-MutS Fusion Protein—Supporting evidence for the presence of a dimer of MutS protein bound to a heteroduplex DNA was obtained in gel shift assays using an MBP-MutS fusion protein. The *Taq* MutS protein was expressed in *E. coli* as a fusion with the 42-kDa maltose-binding protein and purified on an amylose affinity column as described under "Materials and Methods." The ATPase activity and heteroduplex binding activity of the MBP-MutS fusion protein was compared with that of the native MutS protein and found to be essentially identical (data not shown). MBP-MutS was incubated with increasing amounts of factor Xa protease, which cleaves the MBP peptide from MutS. The proteolytically treated MutS was incubated with a ^{32}P-labeled 37-bp heteroduplex DNA and the relative mobilities of MutS-DNA complexes were analyzed on native polyacrylamide gels (Fig. 5*A*).

Three distinct protein complexes were observed based on their electrophoretic mobilities. In the absence of factor Xa, only the slowest migrating complex was observed. This species

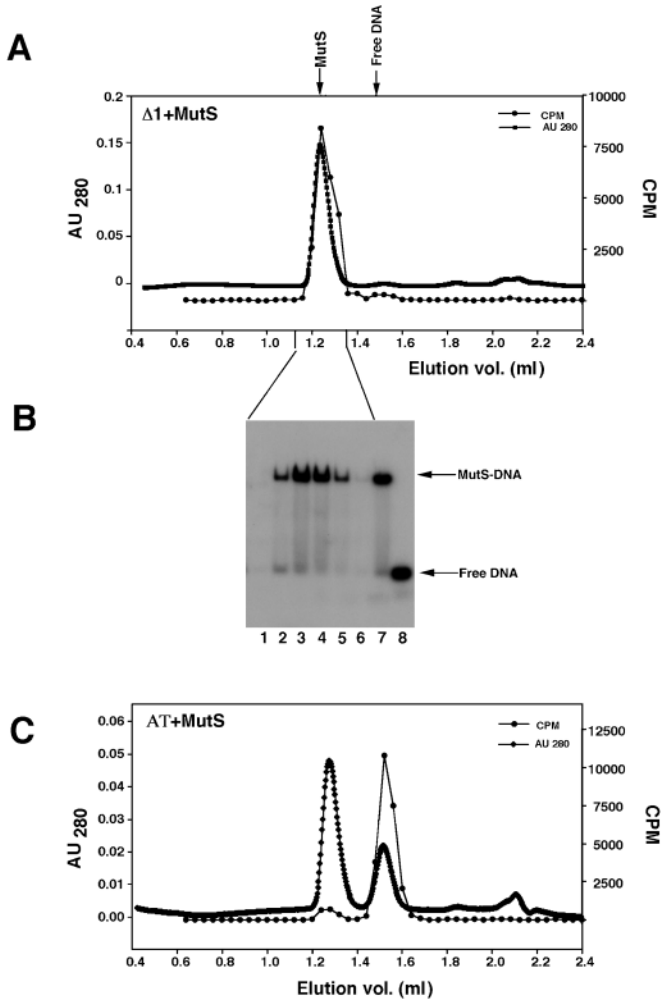

FIG. 4. **Analytical gel filtration of MutS-DNA complexes.** *A*, elution profiles of a MutS-Δ1-heteroduplex DNA from a Superdex S200 column as monitored by absorbance at 280 nm. The presence of heteroduplex DNA in column fractions was determined by scintillation counting (denoted *cpm*). *B*, peak fractions were analyzed on 6% native polyacrylamide gels (*lanes 1–6*) to verify the presence of MutS-Δ1 complexes. As a control, gel mobility shift assays in the presence and absence of MutS protein were included (*lanes 7* and *8*, respectively). *C*, elution profile of MutS-homoduplex DNA complexes (*AT*) from the S200 column.

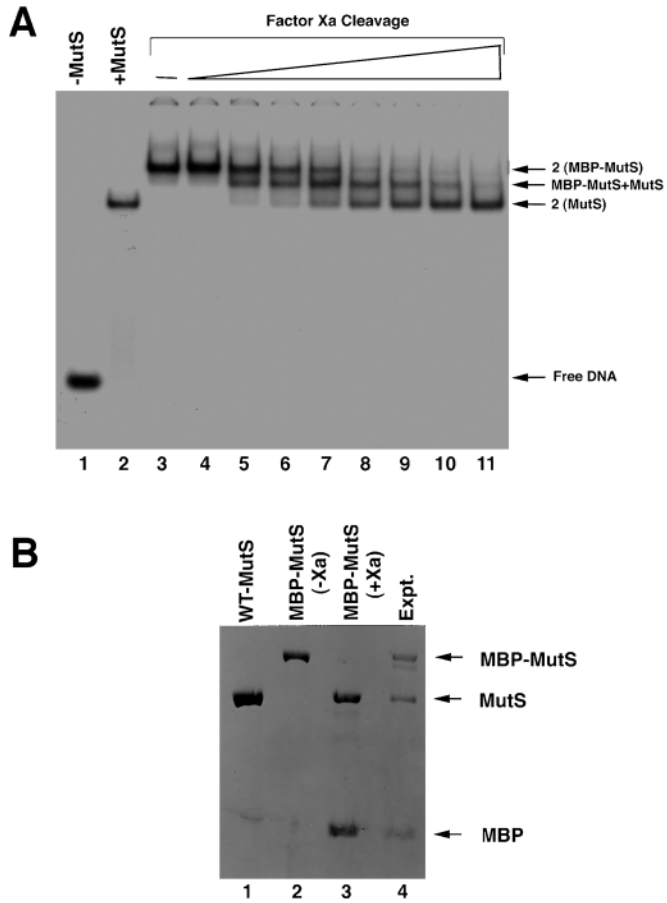

FIG. 5. **DNA binding of a MBP-MutS fusion protein.** *A*, MBP-MutS fusion protein was subjected to proteolysis by increasing amounts of factor Xa. The digested protein was then incubated with ^{32}P-labeled heteroduplex DNA (*Δ1*), and complexes were analyzed on a 6% native polyacrylamide gel as described under "Materials and Methods." *Lane 1*, heteroduplex DNA alone; *lane 2*, heteroduplex DNA with native MutS protein; *lane 3*, heteroduplex DNA incubated with MBP-MutS in the absence of factor Xa pretreatment; *lanes 4–11*, heteroduplex DNA incubated with MBP-MutS protein pretreated with increasing amounts of factor Xa. *B*, SDS-PAGE analysis of the factor Xa-treated MutS-Δ1 DNA complexes. Protein-DNA complexes having intermediate mobility in a gel mobility shift assay after treatment of the MBP-MutS protein with factor Xa were excised from native polyacrylamide gels and subjected to SDS-PAGE analysis in an 8% gel lacking a stacking layer. *Lane 1*, native MutS protein; *lane 2*, MBP-MutS protein; *lane 3*, MBP-MutS protein treated with factor Xa; *lane 4*: the intermediate complex isolated from a gel mobility shift assay as described in *A*, above.

predominated at low concentrations of protease. As the protease concentration was increased, a species with intermediate mobility appeared that was converted to the fastest migrating complex upon more extensive proteolysis. Based on the relative mobilities of complexes obtained after proteolysis, we inferred that the intermediate complex consisted of a dimer of MutS bound to DNA in which one of the MutS subunits contained an intact MBP fusion while the second monomer had had its MBP peptide portion removed. This was confirmed by excising these complexes from the native polyacrylamide gel and subjecting them to SDS-PAGE (Fig. 5*B*). Coomassie staining revealed approximately equimolar amounts of MBP-MutS and MutS protein. Although the gel mobility shift experiment with an MBP-fusion protein cannot distinguish between a dimer or a tetramer of MutS bound to DNA, this experiment, coupled with the result of gel filtration chromatography of complexes shown in Fig. 4, firmly establishes that a dimer of MutS is bound to a heteroduplex DNA.

DISCUSSION

We have examined the oligomerization of *Taq* MutS both in free solution and when bound to a heteroduplex DNA containing an unpaired base. Gel filtration and chemical cross-linking of MutS protein indicated the association of MutS monomers in solution in the absence and presence of Mg^{2+}, with magnesium being required for binding to heteroduplex DNA. Light scattering and MALDI-TOF mass spectrometry established that the predominant form of MutS in free solution at concentrations below 1 μM is a dimer. This was confirmed by sedimentation equilibrium studies that described a homogeneous, reversibly associating system in which the dimer form is in equilibrium with a tetrameric species. The Gibbs standard free energy for the formation of tetramer from dimer was -6.9 ± 0.1 kcal mol^{-1}, consistent with the prevalence of the tetrameric form only at very high protein concentrations ($>10^{-5}$ M). Neither monomeric forms nor oligomers larger than a tetramer were observed, consistent with gel filtration analyses and mass spectrometry. Analytical gel filtration of MutS-heteroduplex DNA complexes and gel mobility shift analysis of MBP-MutS-heteroduplex complexes revealed that *Taq* MutS binds to DNA as a dimer, establishing that there is no change in oligomerization upon binding to short heteroduplex DNAs.

Determining the oligomeric state of MutS protein both in free solution and when bound to DNA is important for understanding the molecular mechanism involved in the repair of mismatches and unpaired bases. Using a variety of approaches, we have determined that *Taq* MutS exists predominantly as a dimer at physiological concentrations and binds DNA as a dimer. Of interest is the fact that *Taq* MutS can oligomerize in solution to form tetramers. The formation of a tetrameric species might explain the heterogeneous behavior of *E. coli* and human MutS protein complexes analyzed by density gradient sedimentation or gel electrophoresis of protein-DNA complexes (20, 22). In addition, the formation of tetrameric MutS complexes may provide a mechanism for the observations that *E. coli* MutS binds to DNA with a stoichiometry of one DNA bound per dimer of protein and that MutS together with MutL translocate along the DNA forming α structures (21). A tetrameric form would provide two DNA binding sites that, when coordinately regulated, would facilitate translocation along the DNA. Finally, it is possible that tetrameric species represent a storage mechanism for the protein. Although the formation of a tetrameric MutS complex is unlikely at physiological concentrations, it is possible that the equilibrium between dimeric and tetrameric forms of MutS might be shifted to favor the larger species as a result of binding to longer DNAs and/or conformational changes arising from the interaction of MutS with other proteins. We note that the nucleotide cofactors ATP, ADP, and ATPγS, while inducing extensive conformational changes in *Taq* MutS, do not affect the oligomerization state of the protein *in vitro*.[5]

Elucidating the structural basis for dimerization and understanding how subunits communicate with each other remain important problems in mismatch repair. Preliminary results indicate that the conserved helix-turn-helix motif at the COOH terminus of MutS proteins serves as a dimerization domain that is essential for both DNA binding and ATP hydrolysis *in vitro*.[6] Recent studies indicate that MutS translocates along the DNA in an ATP-dependent fashion (21, 32). It has also been postulated that ATP triggers a switch between active and inactive DNA binding states of hMutSα (20). These findings suggest that the subunits of MutS act in a coordinated fashion to effect movement along the DNA. Our findings describing the oligomerization of *Taq* MutS protein provide a baseline for future studies aimed at understanding how oligomeric MutS proteins function in mismatch repair.

[5] I. Biswas, unpublished observations.

[6] I. Biswas, G. Obmolova, A. Herr, A. Newman, and P. Hsieh, manuscript in preparation.

Acknowledgments—We are grateful to George Poy for oligonucleotide syntheses and DNA sequencing. We thank Linda Robinson and Arlene Jenkins for their assistance.

REFERENCES

1. Kolodner, R. (1996) *Genes Dev.* **10,** 1433–1442
2. Modrich, P., and Lahue, R. (1996) *Annu. Rev. Biochem.* **65,** 101–133
3. Radman, M. (1989) *Genome* **31,** 68–73
4. Thomas, D. C., Umar, A., and Kunkel, T. A. (1996) *Mutat. Res.* **350,** 201–205
5. Umar, A., and Kunkel, T. A. (1996) *Eur. J. Biochem.* **238,** 297–307
6. Lahue, R. S., Au, K. G., and Modrich, P. (1989) *Science* **245,** 160–164
7. Cooper, D. L., Lahue, R. S., and Modrich, P. (1993) *J. Biol. Chem.* **268,** 11823–11829
8. Biswas, I., and Hsieh, P. (1996) *J. Biol. Chem.* **271,** 5040–5048
9. Biswas, I., and Hsieh, P. (1997) *J. Biol. Chem.* **272,** 13355–13364
10. Malkov, V. A., Biswas, I., Camerini-Otero, R. D., and Hsieh, P. (1997) *J. Biol. Chem.* **272,** 23811–23817
11. Drummond, J. T., Li, G.-M., Longley, M. J., and Modrich, P. (1995) *Science* **268,** 1909–1912
12. Palombo, 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
13. Acharya, S., Wilson, T., Gradia, S., Kane, M. F., Guerrette, S., Marsischky, G. T., Kolodner, R., and Fishel, R. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 13629–13634
14. Alani, E. (1996) *Mol. Cell. Biol.* **16,** 5604–5615
15. Iaccarino, I., Palombo, F., Drummond, J., Totty, N. F., Hsuan, J. J., Modrich, P., and Jiricny, J. (1996) *Curr. Biol.* **6,** 484–486
16. Palombo, F., Iaccarino, I., Nakajima, E., Ikejima, M., Shimada, T., and Jiricny, J. (1996) *Curr. Biol.* **6,** 1181–1184
17. Drummond, J. T., Genschel, J., Wolf, E., and Modrich, P. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 10144–10149
18. Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1997) *Curr. Biol.* **6,** 1185–1187
19. Genschel, J., Littman, S. J., Drummond, J. T., and Modrich, P. (1998) *J. Biol. Chem.* **273,** 19895–19901
20. Gradia, S., Acharya, S., and Fishel, R. (1997) *Cell* **91,** 995–1005
21. Allen, D. W., Makhov, A., Grilley, M., Taylor, J., Thresher, R., Modrich, P., and Griffith, J. D. (1997) *EMBO J.* **16,** 4467–4476
22. Su, S. S., and Modrich, P. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 5057–5061
23. Takamatsu, S., Kato, R., and Kuramitsu, S. (1996) *Nucleic Acids Res.* **24,** 640–647
24. Ferré-D'Amaré, A. R., and Burley, S. K. (1994) *Structure* **2,** 357–359
25. Richards, E. G., Teller, D., Hoagland, V. D., Haschemeyer, R., and Schachman, H. K. (1971) *Anal. Biochem.* **41,** 215–247
26. Ansevin, A. T., Roark, D. E., and Yphantis, D. A. (1970) *Anal. Biochem.* **34,** 237–261
27. Johnson, M. L., Correia, J. J., Halvorson, H. R., and Yphantis, D. A. (1981) *Biophys. J.* **36,** 575–588
28. Gekko, K., and Timasheff, S. N. (1981) *Biochemistry* **20,** 4667–4676
29. Squire, P. G., and Li, C. H. (1961) *J. Am. Chem. Soc.* **83,** 3521–3528
30. Yphantis, D. A. (1964) *Biochemistry* **3,** 297–317
31. Yphantis, D. A., Correia, J. J., Johnson, M. L., and Wu, G. M. (1979) in *Physical Aspects of Protein Interactions* (Catsimpoolas, N., ed) pp. 275–303, Elsevier/North Holland, Amsterdam
32. Blackwell, L. J., Martik, D., Bjornson, K. P., Bjornson, E. S., and Modrich, P. (1998) *J. Biol. Chem.* **273,** 32055–32062