MOLECULAR AND CELLULAR BIOLOGY, May 1997, p. 2436–2447 Vol. 17, No. 5
0270-7306/97/$04.00+0

Genetic and Biochemical Analysis of Msh2p-Msh6p: Role of ATP Hydrolysis and Msh2p-Msh6p Subunit Interactions in Mismatch Base Pair Recognition

ERIC ALANI,[1]* TANYA SOKOLSKY,[1] BARBARA STUDAMIRE,[1] JUAN J. MIRET,[2]† AND ROBERT S. LAHUE[2]†

Section of Genetics and Development, Cornell University, Ithaca, New York 14853-2703,[1] and Department of Biochemistry and Molecular Biology, University of Massachusetts Medical Center, Worcester, Massachusetts 01655[2]

Received 23 September 1996/Returned for modification 25 November 1996/Accepted 25 January 1997

Recent studies have shown that *Saccharomyces cerevisiae* Msh2p and Msh6p form a complex that specifically binds to DNA containing base pair mismatches. In this study, we performed a genetic and biochemical analysis of the Msh2p-Msh6p complex by introducing point mutations in the ATP binding and putative helix-turn-helix domains of *MSH2*. The effects of these mutations were analyzed genetically by measuring mutation frequency and biochemically by measuring the stability, mismatch binding activity, and ATPase activity of msh2p (mutant msh2p)-Msh6p complexes. A mutation in the ATP binding domain of *MSH2* did not affect the mismatch binding specificity of the msh2p-Msh6p complex; however, this mutation conferred a dominant negative phenotype when the mutant gene was overexpressed in a wild-type strain, and the mutant protein displayed biochem ical defects consistent with defects in mismatch repair downstream of mismatch recognition. Helix-turn-helix domain mutant proteins displayed two different properties. One class of mutant proteins was defective in forming complexes with Msh6p and also failed to recognize base pair mismatches. A second class of mutant proteins displayed properties similar to those observed for the ATP binding domain mutant protein. Taken together, these data suggested that the proposed helix-turn-helix domain of Msh2p was unlikely to be involved in mismatch recognition. We propose that the *MSH2* helix-turn-helix domain mediates changes in Msh2p-Msh6p interactions that are induced by ATP hydrolysis; the net result of these changes is a modulation of mismatch recognition.

Mismatch repair is a highly conserved process that results in the removal of base pair mismatches that occur during DNA replication, DNA damage, and genetic recombination (25, 37, 45). Much of our understanding of mismatch repair comes from *Escherichia coli*, where an in vitro repair reaction was developed and individual components were identified and purified to homogeneity (28, 36). This approach led to the identification of three proteins, MutS, MutL, and MutH, that were shown to be important for the initial steps in mismatch repair. These steps are thought to involve MutS binding to the base pair mismatch, followed by ATP-dependent binding of MutL to the MutS-mismatch complex (15). This complex is then thought to activate MutH, an endonuclease that cleaves only the unmethylated strand of hemimethylated GATC sites that are present immediately after passage of a DNA replication fork. The sensitivity of MutH endonuclease to methylation results in incision of the newly replicated strand. Excision of this strand proceeds through the region of DNA containing the base pair mismatch. This is then followed by resynthesis of the excised strand using the parental strand as a template. The mechanistic steps leading to MutH activation are not clear but have been shown to require ATP hydrolysis (4, 57). Interestingly, only MutS displays an ATP hydrolysis activity; this activity results in the loss of mismatch recognition when MutS-dependent mismatch binding reactions are performed in the presence of ATP (15, 53, 54).

As a first step toward understanding the role of ATP in mismatch repair, Haber and Walker (17) showed that MutS bound and hydrolyzed ATP with a K_m of 6 μM and a K_{cat} of 0.26 μM ADP/min/μM MutS. They also showed that mutS-KA622, a mutant protein that contained a mutation in the ATP phosphate binding loop (P-loop) consensus, displayed reduced ATPase activity; the K_m of ATP hydrolysis for mutS-KA622 was 30-fold higher and the K_{cat} was four- to fivefold lower than for the wild-type protein. Interestingly, mutS-KA622 specifically recognized DNA containing base pair mismatches, albeit with reduced affinity. Genetic analysis indicated that the mutS-KA622 mutation displayed a null phenotype and was dominant over wild type when overexpressed. A large number of dominant negative *E. coli mutS* mutations were subsequently identified by Wu and Marinus (61); they also showed that a subset of these mutations mapped to regions containing ATP binding domain consensus sequences. This observation was of interest because the bacterial, yeast, and human MutS homolog protein complexes failed to specifically bind to DNA containing base pair mismatches in the presence of ATP but displayed specific mismatch binding in the presence of nonhydrolyzable ATP analogs (1, 11, 15). Taken together, these data suggest that ATP binding and hydrolysis play a critical role in modulating the mismatch recognition activity of MutS and its homologs.

We are interested in understanding how mismatch repair proteins recognize base pair mismatches and transduce the mismatch recognition signals that result in excision repair steps. We use the yeast *Saccharomyces cerevisiae*, which contains six MutS homologs, Msh1p to Msh6p, as a model system

* Corresponding author. Mailing address: Section of Genetics and Development, Cornell University, 459 Biotechnology Building, Ithaca, NY 14853-2703. Phone: (607) 254-4811. Fax: (607) 255-6249. E-mail: eea3@cornell.edu.

† Present address: Eppley Institute, University of Nebraska Medical Center, Omaha, NE 68198.

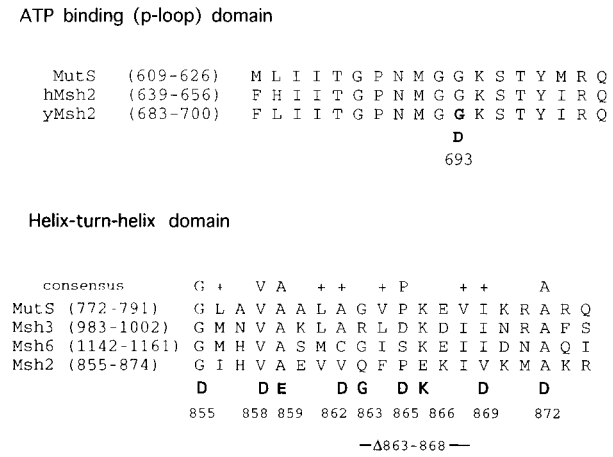

FIG. 1. (A) Alignment of the P-loop motif of purine nucleotide binding proteins found in *E. coli* MutS and human and *S. cerevisiae* Msh2p. (B) Alignment of putative helix-turn-helix domains found in *E. coli* MutS and *S. cerevisiae* Msh2p, Msh3p, and Msh6p. + refers to conserved hydrophobic amino acid residues. Amino acid substitutions in the Msh2p amino acid sequence that were constructed for this study are shown in boldface. In the text, mutations are referred to by the following nomenclature: the first letter after *msh2* refers to the wild-type amino acid sequence and the second letter refers to the amino acid substitution.

with which to study mismatch repair (3, 20, 31, 39, 48, 49, 51). Recent studies are consistent with the idea that Msh2p and Msh6p form a complex that principally recognizes base pair mismatches and single-nucleotide insertions/deletions, and Msh2p and Msh3p form a complex that principally recognizes two- to four-nucleotide insertion/deletion mismatches (1, 18, 22, 24, 31, 43). Two highly conserved domains have been identified in MutS and the yeast and mammalian MutS homologs, an ATP binding domain that has been shown to be involved in ATP hydrolysis in MutS protein and a helix-turn-helix motif (Fig. 1) (12, 14, 16, 20, 29, 47, 48, 51). The helix-turn-helix motif was first identified in a group of prokaryotic DNA binding proteins and was shown to be involved in DNA sequence-specific recognition through both mutational and structural analysis of lambda repressor, Cro, and Cap proteins (see reference 41 for a review). The MutS homologs contain a helix-turn-helix motif that strongly matches the consensus derived from analyzing DNA sequence-specific proteins that have been shown by structural analyses to utilize this motif (Fig. 1) (51). One problem with identifying helix-turn-helix domains in a gene of interest is that the small size of the motif does not allow unambiguous identification of a DNA recognition domain without additional structural, biochemical, and genetic evidence; sequence comparisons made without regard to stereochemical constraints may result in false identification (40, 41).

Previously we showed that the *S. cerevisiae* MutS homologs Msh2p and Msh6p form a complex that specifically recognizes base pair mismatches and one-nucleotide insertions (1). Like the human MSH2-GTBP complex, Msh2p-Msh6p mismatch binding specificity was not observed in the presence of ATP (1, 11, 22, 42). Interestingly, the carboxy-terminal 114 amino acids of Msh2p, which contain the helix-turn-helix domain, were required for both Msh6p interaction and mismatch recognition (1). To examine the role of the ATP binding domain and helix-turn-helix motif in mismatch recognition, we constructed site-specific mutations in both the ATP binding and helix-turn-helix domains of *MSH2* and examined by genetic and biochemical techniques the effects of these mutations on mismatch recognition, modulation of mismatch binding by ATP, and interaction with Msh6p. The *msh2-GD693* ATP binding domain mutation resulted in phenotypes that were similar to those observed for the *mutS-KA622* mutation; mutations in the helix-turn-helix domain of Msh2p resulted in phenotypes and biochemical properties consistent with defects in ATP binding and hydrolysis and/or complex formation with Msh6p. Based on these data, we propose that the helix-turn-helix domain plays a role in modulating mismatch recognition by responding to conformational changes in Msh2p and Msh6p that are induced by ATP hydrolysis.

MATERIALS AND METHODS

Strains. *E. coli* RKY1400 (*thr leuB6 thi thyA trpC1117 hsrK12 hsmK12* str^r *recA13*) was kindly provided by R. Kolodner and was used to amplify and manipulate all plasmids described in this report. *S. cerevisiae* BJ5464 (*MAT*α *ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL*) was obtained from the Yeast Genetic Stock Center and was transformed with pEAE51 (*GAL10-MSH6 TRP1* 2μm) and derivatives of pEAE20 (*GAL10-MSH2 URA3* 2μm) containing wild-type and mutant *msh2* alleles; this strain was used for the overexpression and purification of wild-type and mutant Msh2p-Msh6p complexes (1, 2) (see below). *S. cerevisiae* FY86(*MAT*α *ura3-52 leu2Δ1 his3Δ200*) was used to measure the dominant mutator phenotypes of *msh2* mutant derivatives of plasmids pEAA39 (*MSH2::CA5 ARS-CEN URA3*) and pEAE20 (Table 1) (1, 2). EAY252 (*MAT*α *ura3-52 leu2Δ1 trp1Δ63 msh2Δ::TRP1*) is a *msh2*Δ derivative of FY23 (59) that was used to measure the complementation phenotype of wild-type and *msh2* mutant derivatives of pEAA39 and pEAE20. The *msh2*Δ*::TRP1* allele contains a deletion of amino acids 1 to 898 in the 964-amino-acid *MSH2* open reading frame. The FY strains were kindly provided by K. Dollard and F. Winston (59). Yeast strains were transformed with episomal vectors by using the lithium acetate method described by Gietz and Schiestl (13).

Media, Reagents, and Chemicals. *E. coli* strains were grown in LB broth or on LB agar that was supplemented with ampicillin (100 μg/ml) when required (33). Yeast strains were grown in either YPD or minimal selective media (52). Selective media contained 0.7% yeast nitrogen base, 2% agar, and 0.09% of a dropout mix that lacks the amino acid used for selection; 2% glucose, 2% sucrose, 3% glycerol, 2% lactate, and 2% galactose were included as indicated. Selective medium containing 2% each galactose and sucrose was used in the complementation and dominance tests that involved expression of *GAL10-MSH2* and mutant derivatives expressed in pEAE20. When required, canavanine (Sigma, St. Louis, Mo.) was included in minimal selective media lacking arginine at 60 mg/liter (50).

ATP and AMP-PNP (adenylyl-imidodiphosphate) were from Pharmacia (Uppsala, Sweden) and Boehringer Mannheim (Indianapolis, Ind.), respectively, and BA85 0.45-μm-pore-size nitrocellulose filters were from Schleicher & Schuell (Keene, N.H.). Protein concentrations were determined by the Bradford dye method, using bovine serum albumin (BSA) as a standard (7), and reagents were obtained from Bio-Rad (Richmond, Calif.). Purified antihemagglutinin mouse monoclonal antibody (clone 12CA5) was purchased from Boehringer Mannheim. For column chromatography, PBE94 was purchased from Pharmacia, and single-stranded DNA (ssDNA)-cellulose (catalog no. D8273) was purchased from Sigma; all resins were precycled according to the manufacturer's instructions. Protein standards were myosin (205 kDa), β-galactosidase (116 kDa), phosphorylase *b* (97.4 kDa), serum albumin (66.2 kDa), and ovalbumin (44 kDa) and were obtained from Bio-Rad.

Determination of mutation rates. Colonies of wild-type and *msh2*Δ strains containing *msh2* plasmids described in Table 1 were streaked onto selective minimal plates to form single colonies. Plates contained 2% each galactose and sucrose for experiments involving Msh2p overexpression from *GAL10* 2μm vectors and contained 2% glucose for experiments involving expression of Msh2p from *ARS-CEN* vectors. Strains were grown at 30°C until colonies had formed containing approximately 10^7 cells. Eleven colonies from each strain were resuspended in water and then plated with the appropriate dilution onto arginine dropout plates containing or lacking canavanine (60 mg/liter). The same carbon source was used for the entire experiment. The frequency of canavanine resistance was determined for each colony, and the median frequency was then recorded.

Nucleic acid techniques. All restriction endonucleases, T4 polynucleotide kinase, T4 DNA ligase, T4 DNA polymerase, and Vent polymerase were from New England Biolabs and used according to manufacturer's specifications. PCR was performed for site-directed mutagenesis by the overlap extension method (19). Most reactions involved 12 PCR cycles using a denaturation step of 1 min at 94°C, an annealing step of 1 min at 50°C, and a polymerization step of 2.0 min at 72°C. Reactions were performed in 50 μl with 10 pmol of each primer and 1 μg of yeast pEAE20. Plasmid DNA was isolated by alkaline lysis, and all DNA manipulations were performed as described previously (30). Oligonucleotide synthesis and double-stranded DNA sequencing of the entire subcloned fragments used to make the *msh2* mutant alleles were performed at the Cornell

TABLE 1. Median frequencies of spontaneous mutations in *msh2*Δ (EAY252) and *MSH2* (EAY236) strains bearing the indicated *msh2* alleles on *ARS-CEN* and *GAL10* 2μm plasmids[a]

Allele (class)	Relative median frequency of canavanine resistance			
	*msh2*Δ (*ARS-CEN*)	*msh2*Δ (*GAL10* 2μm)	*MSH2* (*ARS-CEN*)	*MSH2* (*GAL10* 2μm)
Wild type	2.0, 1.7	**1.0**	1.0	**1.0**
Null (III)	54, 54, 32, 84	54, 54, 32	1.3	1.3
GD693 (II)	35, 27	54, 31	1.5	50, 33, 24, 95
GD855 (II)	36	57	1.1	91, 50
VD858 (II)	35	31	1.3	87, 39
AE859 (I)	44, 34	6.7	1.4	5.5
VD862 (I)	41, 20	4.3	1.1	2.3
QG863	NT	1.3	NT	NT
PD865 (I)	8.5, 8.2	1.5, 1.0	1.2	1.8, 1.3, 2.0, 1.9
EK866	NT	1.1, 0.84	NT	NT
VD869 (I)	15, 10	1.0, 1.9	NT	NT
AD872 (I)	45, 37	9.3, 4.1	1.2	4.5
Δ6 (III)	NT	33	NT	4.7
Δ860 (III)	NT	43	NT	1.4
GD693, GD855 (II,II)	NT	100	NT	78, 44
GD693, VD858 (II,II)	NT	33	NT	68
GD693, AE859 (II,I)	NT	35	NT	56
GD693, VD862 (II,I)	NT	47	NT	49
GD693, PD865 (II,I)	NT	29	NT	3.6, 2.6, 3.7
GD693, AD872 (II,I)	NT	35	NT	4.5, 2.8
GD693, Δ6 (II,III)	NT	25	NT	6.7

[a] Canavanine resistance frequencies were determined for *msh2* and *MSH2* strains containing the indicated *msh2* alleles substituted into the pEAA39 (*MSH2 ARS-CEN*) or pEAE20 (*GAL10-MSH2* 2μm) parental plasmid (see Materials and Methods for pEAA and pEAE designations). Strains bearing the YCP50 (*ARS-CEN URA3*) vector were used in the null allele experiments. These strains were grown in glucose or galactose-sucrose medium as described in Materials and Methods. For each strain, the median frequency of canavanine resistance was determined for 11 independent colonies. In some cases, the median canavanine resistance frequencies calculated in independent experiments are shown. All data are presented relative to the median canavanine resistance frequency (values in boldface type) obtained in strains bearing a 2μm vector containing *GAL10-MSH2* (pEAE20). The median frequency of canavanine resistance in *MSH2* and *msh2*Δ strains bearing pEAE20 ranged from 0.5×10^{-6} to 1.0×10^{-6}. See text for designation of mutant classes.

Biotechnology Analytical/Synthesis Facility. Oligonucleotides were 5′ labeled by using [^{32}P]γATP and T4 polynucleotide kinase as described previously (9). Oligonucleotide concentrations and annealing conditions were described previously (2).

***MSH2* and *MSH6* plasmids.** pEN43/pEAE20 (10.3 kb) is a 2μm *URA3* vector that contains a fusion of the *GAL10* promoter to the *MSH2* coding region (2). pEAA39 (9.0 kb) is an *ARS-CEN URA3* vector that contains the *MSH2* gene expressed under its own promoter (1). The *MSH2* gene in both of these plasmids bears the 12CA5 epitope sequence (26) inserted at amino acid 644 in the *MSH2* open reading frame. This epitope tagged variant of *MSH2* is specifically recognized by antibody 12CA5 (1, 2), and both the *ARS-CEN* and 2μm plasmids fully complement the mutator phenotype of a *msh2*Δ strain as measured by the canavanine resistance assay (3). *msh2* mutant derivatives of pEAA39 and pEAE20, shown in Fig. 1, were created by overlapping PCR mutagenesis (1, 19). Two derivatives of pEAE20, pEAE43 and pEAE44, contain deletion mutations in *MSH2* (1). pEAE43 (*msh2-Δ6*) contains a six-amino-acid deletion of *MSH2* from amino acids 863 to 868. pEAE44 (*msh2-Δ860*) contains a frameshift mutation in *MSH2* that results in a deletion of amino acids 860 to 964 in the *MSH2* coding region. pEAE51 (10.8 kb) is a 2μm *TRP1* vector that contains a fusion of the *GAL10* promoter to the *MSH6* coding region (1).

The plasmid number designations for the *msh2* mutant derivatives of pEAA39 are as follows: pEAA51 (*msh2-GD693*), pEAA55 (*msh2-GD855*), pEAA60 (*msh2-VD858*), pEAA56 (*msh2-AE859*), pEAA57 (*msh2-VD862*), pEAA58 (*msh2-PD865*), pEAA61 (*msh2-VD869*), and pEAA59 (*msh2-AD872*). The plasmid designations for the *msh2* mutant derivatives of pEAE20 are as follows: pEAE27 (*msh2-GD693*), pEAE30 (*msh2-GD855*), pEAE31 (*msh2-VD858*), pEAE32 (*msh2-AE859*), pEAE33 (*msh2-VD862*), pEAE34 (*msh2-QG863*), pEAE35 (*msh2-PD865*), pEAE36 (*msh2-EK866*), pEAE37 (*msh2-VD869*), pEAE38 (*msh2-AD872*), pEAE43 (*msh2-Δ6*), pEAE44 (*msh2-Δ860*), pEAE39 (*msh2-GD693*, *GD855*), pEAE40 (*msh2-GD693*, *VD858*), pEAE74 (*msh2-GD693*, *AE859*), pEAE75 (*msh2-GD693*, *VD862*), pEAE41 (*msh2-GD693*, *PD865*), pEAE42 (*msh2-GD693*, *AD872*), and pEAE60 (*msh2-GD693*, *Δ6*).

Biochemical techniques and immunoprecipitation reactions. Overexpression and purification of Msh2p-Msh6p were performed as described previously, using yeast strain BJ5464 transformed with pEAE20 and pEAE51 (1). The purity of protein preparations was monitored by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with 8% polyacrylamide gels (27) and by measuring binding to homoduplex and heteroduplex oligonucleotide substrates (1). msh2p-Msh6p complexes were purified by the same procedure from yeast strains containing the appropriate *MSH2* and *MSH6* plasmids. A summary of the yields obtained during the purification of the wild-type and mutant complexes is presented in Table 2. The approximate molecular weights of Msh2p-Msh6p and mutant derivatives were determined by measuring the time required to elute polypeptides from a Superose 6HR gel filtration column (0.79 cm^2 by 30 cm; void volume = 6.3 ml) controlled by a Waters 650 FPLC (fast protein liquid chromatography) apparatus. The FPLC column was run at room temperature and eluted at a flow rate of 1.0 ml/min in 500 mM NaCl–25 mM Tris (pH 7.5)–1 mM EDTA–10 mM β-mercaptoethanol. The molecular weight standards thyroglobulin, bovine gamma globulin, chicken ovalbumin, equine myoglobin, RNase A, and vitamin B_{12} (Bio-Rad) were run as a mixture in experiments. This analysis was kindly performed by Jinlin Peng at the Cornell Biotechnology Analytical/Synthesis Facility.

Immunoprecipitations from crude extracts derived from galactose-induced BJ5464/pEAE20, BJ5464/pEAE51, or BJ5464/pEAE20,pEAE51 by using antibody 12CA5 were performed as described previously (1, 2).

TABLE 2. Purification of Msh2p-Msh6p and msh2p-Msh6p complexes[a]

Prepn	Protein fraction (mg)			
	I	II	III	IV
Msh2p-Msh6p	35	2.5	0.63	0.23
msh2-GD693p–Msh6p	26	2.0	0.66	0.21
msh2-GD855p–Msh6p	33	2.5	0.74	0.23
msh2-AE859p–Msh6p	23	1.6	0.30	0.068
msh2-VD862p–Msh6p	22	1.3	0.22	0.060
msh2-PD865p–Msh6p	21	1.4	0.17	0.026
msh2-AD872p–Msh6p	19	0.34	0.14	0.015
msh2-Δ6p–Msh6p	28	1.3	0.28	0.057

[a] Msh2p-Msh6p and the indicated msh2p-Msh6p complexes were purified from BJ5464 cells containing *GAL10-MSH6* and *GAL10-MSH2* or *GAL10-MSH6* and *GAL10-msh2* overexpression plasmids as described in Materials and Methods. The starting material for each purification was 0.8 to 1.0 g (wet weight) of cells. Fractions I to IV refer to crude extract, PBE94, ssDNA cellulose, and PBE94 steps, respectively.

DNA binding assays. DNA binding assays were performed as described previously (1, 32). The standard buffer for the DNA binding assay contained 25 mM Tris (pH 7.5), 2.0 mM $MgCl_2$, 0.1 mM dithiothreitol (DTT), 0.01 mM EDTA, and 40 μg of BSA per ml. ATP was added to a final concentration of 1.6 mM as indicated. Binding was performed at 30°C for 15 min in a 60-μl reaction, and each reaction mixture contained 16.7 nM ^{32}P-labeled duplex oligonucleotide, 0 to 1,333 nM unlabeled competitor duplex oligonucleotide, and 0 to 0.26 μg of fraction IV of Msh2p-Msh6p or the indicated mutant derivative. The 37-mer homoduplex and +1 oligonucleotide substrates used in this study are described by Alani (1). Following incubation, samples were analyzed by filter binding to KOH-treated nitrocellulose filters (32), using a Hoefer Scientific Instruments (San Francisco, Calif.) FH225V filtering unit. Filter binding was performed as described by Chi and Kolodner (9).

ATPase assays. The standard ATPase assay was performed in 60-μl reaction mixtures containing 0.27 μg of Msh2p-Msh6p or mutant derivatives, 100 μM [γ-^{32}P]ATP, 25 mM Tris (pH 7.5), 2.0 mM $MgCl_2$, 0.1 mM DTT, 0.01 mM EDTA, and 40 μg of BSA per ml. When specified, homoduplex and +1 oligonucleotide substrates were included at 167 nM. The reaction mixtures were incubated for 15 min at 30°C, and the amount of ATP hydrolyzed was determined by Norit A absorption assays (10). K_m measurements were performed in standard ATPase assay conditions with the exception that the concentration of [γ-^{32}P]ATP was varied from 0.41 to 100 μM (1).

RESULTS

Rationale for site-specific mutagenesis. Biochemical characterization of the *E. coli* MutS, human MSH2-GTBP, and yeast Msh2p-Msh6p complexes revealed that the mismatch binding activity displayed by each of these proteins could be modulated by ATP (1, 11, 15, 22). Protein sequence alignments of bacterial, yeast, and mammalian MutS homologs indicated that these proteins share a highly conserved ATP binding domain (5, 47, 48). The human and yeast *MSH2* ATP binding domains, for example, displayed 81% amino acid identity in a 120-amino-acid region (Fig. 1) (12, 29). Biochemical analysis of the *Salmonella* MutS and *S. cerevisiae* Msh1p and Msh2p-Msh6p proteins indicated that these proteins displayed similar ATPase activities, and genetic analysis of *Salmonella mutS* indicated that the ATP binding domain was important for MutS function (1, 10, 17).

Visual inspection of the prokaryotic and eukaryotic MutS homolog amino acid sequence led to the identification of a highly conserved 20-amino-acid helix-turn-helix domain (Fig. 1); this structure was of interest because such domains had been shown to be important for DNA sequence-specific binding in a variety of prokaryotic DNA binding proteins (16, 20, 41, 47, 51). The helix-turn-helix sequence identified in MutS homologs conformed to most of the stereochemical constraints observed in such motifs with one exception: position nine in the MutS consensus sequence was not occupied by a glycine residue (51). A glycine residue in this position is considered important for the formation of the loop that connects the two alpha helices that contact DNA; however, it is important to note that not all proteins that bind DNA through a helix-turn-helix motif contain a glycine at this position (40, 41, 51).

Based on the homologies described above, we chose to study the function of the ATP binding and helix-turn-helix domains in Msh2p by using site-directed mutagenesis. We chose a loss-of-contact approach where highly conserved amino acids were substituted with residues of opposite charge or hydrophobicity. A similar approach was performed on a variety of proteins that displayed the helix-turn-helix motif, with the finding that such substitutions dramatically affected DNA binding but rarely destabilized the protein or caused disruption of its overall protein structure (56, 60). A summary of the amino acid substitutions made in Msh2p is presented in Fig. 1, Table 1, and Materials and Methods. In addition to the helix-turn-helix mutations, a single substitution was made in the ATP binding domain, a glycine-to-aspartic acid change at amino acid 693 in the P loop. This substitution was chosen because similar substitutions in MutS and in other ATP binding proteins resulted in proteins that displayed decreased ATP hydrolysis and/or binding (17, 55). Two *msh2* deletion alleles were also used in this analysis; the *msh2-Δ6* allele contains a six-amino-acid deletion of *MSH2* from amino acids 863 to 868 with the addition of a single histidine residue at the deletion break point, and the *msh2-Δ860* allele contains a frameshift mutation beginning at amino acid 860 in the *MSH2* coding region (1).

msh2 mutant proteins are expressed at wild-type levels. Before initiating the genetic and biochemical analysis of the Msh2p-Msh6p complex described below, we tested whether mutations in the ATP binding and helix-turn-helix domains of *MSH2* would alter the stability of the Msh2 polypeptide. Mutant *msh2* genes constructed by using site-directed mutagenesis were subcloned into single-copy (*ARS-CEN*) vectors in which the mutant gene was expressed under the native *MSH2* promoter (derivatives of pEAA39) or into high-copy-number (2μm) *GAL10* vectors in which *MSH2* expression was placed under galactose control (derivatives of pEAE20) (Materials and Methods). Previous studies indicated that galactose induction of yeast strains bearing pEAE20 resulted in overexpression of Msh2p to about 0.5% of total protein (2). A subset of wild-type and site-specific mutants were analyzed in crude lysates or by immunoprecipitation to determine whether these mutations affected Msh2p expression and/or stability. All of the mutant alleles tested (*msh2-GD693*, *msh2-GD855*, *msh2-VD858*, *msh2-AE859*, *msh2-VD862*, and *msh2-AD872*) expressed msh2p in crude extracts at levels similar to the wild-type level; furthermore, the overexpressed wild-type and mutant msh2 polypeptides that were immunoprecipitated from these extracts appeared similar in size and integrity (data not shown). These results suggested that the site-specific mutations did not cause destabilization of Msh2p. Previous analysis of the *msh2-Δ6* and *msh2-Δ860* alleles showed that they expressed intact mutant polypeptides at wild-type levels (1). In addition, immunoprecipitation studies of yeast strains expressing *msh2-GD855*, *msh2-AE859*, *msh2-PD865*, and *msh2-AD872* alleles under the native *MSH2* promoter in single-copy vectors revealed that expression of the corresponding polypeptides was also similar to wild-type expression (52a).

Genetic analysis of *msh2* alleles. Previous analysis of proteins containing ATP binding and helix-turn-helix domains indicated that mutations in these domains would disrupt the function of these proteins, resulting in a null phenotype. In some cases, overexpression of these mutant proteins resulted in a dominant negative phenotype (17, 55, 56, 60). The *msh2* mutant alleles present in either single copy or high copy number were transformed into *msh2*Δ (EAY252) and wild-type (EAY236) strains to study whether they could complement the mutator phenotype exhibited by *msh2*Δ mutants and/or confer a dominant negative phenotype in the wild type. The mutator phenotype of *msh2* mutant alleles was assessed by using the canavanine resistance assay (3, 49). In this assay, *msh2*Δ strains exhibited a median canavanine resistance frequency that was approximately 25- to 50-fold higher than that observed in the wild type (Table 1). A summary of the complementation and dominant mutator data is presented in Table 1. Three classes of mutations were identified in this analysis. Class I mutations (*msh2-AE859*, *msh2-VD862*, *msh2-PD865*, *msh2-VD869*, and *msh2-AD872*) poorly complemented the *msh2*Δ phenotype in single copy but showed stronger complementation when expressed in high copy number. These mutations did not display a strong dominant mutator phenotype when expressed in either single copy or high copy number. Class II mutations (*msh2-GD693*, *msh2-GD855*, and *msh2-VD858*) were unable to complement the *msh2*Δ phenotype in single copy or high copy

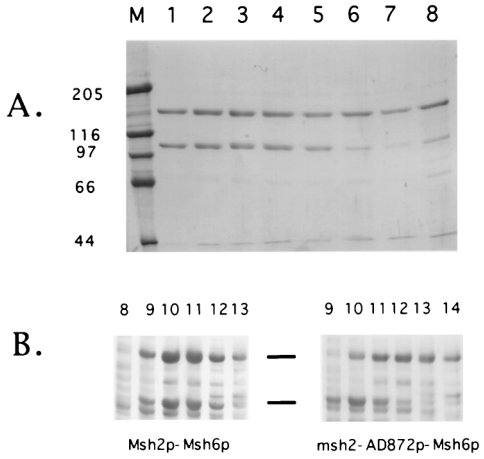

FIG. 2. (A) SDS-PAGE (8% gel) analysis of fraction IV obtained from BJ5464 extracts cooverexpressing Msh6p and each of the following Msh2p derivatives: Msh2p (lane 1), msh2-GD693p (lane 2), msh2-GD855p (lane 3), msh2-AE859p (lane 4), msh2-VD862p (lane 5), msh2-PD865p (lane 6), msh2-AD872p (lane 7), and msh2-Δ6p (lane 8). The purification of Msh2p-Msh6p and each of the msh2p-Msh6p complexes is described in Materials and Methods. Each lane represents approximately 0.9 μg of fraction IV. M, molecular weight standards; molecular masses (in kilodaltons) are provided. (B) SDS-PAGE (8% gel) analysis of fraction I from the purification of Msh2p-Msh6p and msh2-AD872p–Msh6p. The elution profiles of Msh2p, msh2-AD872p (lower bar), and Msh6p (upper bar) are shown.

number. However, they displayed a strong dominant negative phenotype when expressed in high copy number. Class III mutations (*msh2-Δ6* and *msh2-Δ860*) displayed the *msh2*Δ null phenotype. They failed to complement a *msh2*Δ strain and failed to display a strong dominant negative mutator phenotype when expressed in single copy (when tested) or high copy number (1). Two mutations, *msh2-QG863* and *msh2-EK866*, were not characterized in great detail because they showed either no effect or only a weak effect on *MSH2* function.

To examine the genetic relationship between the classes of *msh2* mutations, double-mutant combinations containing class I,II and class II,III mutations were constructed, and the mutant polypeptides were then overexpressed in a wild-type strain (Table 1). A dominant negative phenotype was observed in strains that overexpressed polypeptides bearing two different class II mutations (i.e., *msh2-GD693*, *GD855*). However, overexpression of class II,III and a subset of class I,II mutant proteins did not result in a dominant negative phenotype. Protein expression analysis indicated that the class I,II and class II,III mutant proteins were expressed at levels similar to those observed for strains overexpressing Msh2p (data not shown). These data suggest that class I and class III mutations can disrupt the class II mutant functions that result in a dominant negative phenotype.

Biochemical purification of msh2p-Msh6p complexes. The identification of a helix-turn-helix domain in *MSH2*, coupled with the genetic analysis shown above, suggested that mutations in the *MSH2* helix-turn-helix domain would disrupt mismatch repair by disrupting the mismatch binding specificity of the Msh2p-Msh6p complex. Surprisingly, the biochemical analysis, as described below, suggested that the helix-turn-helix domain was unlikely to be directly involved in mismatch recognition. Based on the genetic analysis described in Table 1, strains bearing *GAL10* 2μm plasmids that contained *MSH6* and representative class I (*msh2-AE859*, *msh2-VD862*, *msh2-PD865*, and *msh2-AD872*), class II (*msh2-GD693* and *msh2-GD855*), and class III (*msh2-Δ6*) mutations were grown in culture and induced for expression of Msh2p and Msh6p. Extracts prepared from these cells overexpressing Msh6p and Msh2p or mutant msh2p were used to purify mutant complexes by using the procedure used to purify Msh2p-Msh6p (1). As shown in Fig. 2A and Table 2, the protein yields and purity of msh2-GD693p–Msh6p and msh2-GD855p–Msh6p complexes were indistinguishable from those obtained from the purification of Msh2p-Msh6p. Copurification of Msh6p and mutant msh2p appeared weaker in strains overexpressing Msh6p and each of the following polypeptides: msh2-PD865p, msh2-AD872p, and msh2-Δ6p (Fig. 2A) (1). For each of these mutant proteins, the msh2 polypeptide eluted in the PBE94 gradient at a lower NaCl concentration than was observed for Msh6p; subsequently, the majority of msh2p that was loaded onto ssDNA-cellulose flowed through the column. As a result, fraction IV for each of these mutant protein preparations contained predominantly Msh6p polypeptide (Fig. 2A). An example of the first PBE94 chromatography step for this class of proteins is shown for the msh2-AD872p in Fig. 2B. It is important to note that the defect in complex formation observed in the purification of msh2-AD872p–Msh6p was also observed in Msh2p-specific immunoprecipitation reactions involving crude extracts that contained overexpressed msh2-AD872p and Msh6p (data not shown).

Previously, we showed that Msh2p-Msh6p eluted as a single complex in Superose 6HR filtration with a molecular mass corresponding to approximately 300 kDa (1). To investigate whether mutant msh2 proteins were defective in interacting with Msh6p, we measured the elution profiles of representative classes of msh2p-Msh6p proteins. As shown in Fig. 3, the msh2-GD693p–Msh6p complex displayed an elution profile that was indistinguishable from that of the Msh2p-Msh6p complex. However, the msh2p-AE859p–Msh6p complex displayed an elution profile that differed from that of the wild-type complex; only about one-third of the material loaded eluted at a molecular mass of 300 kDa. The remaining material eluted at

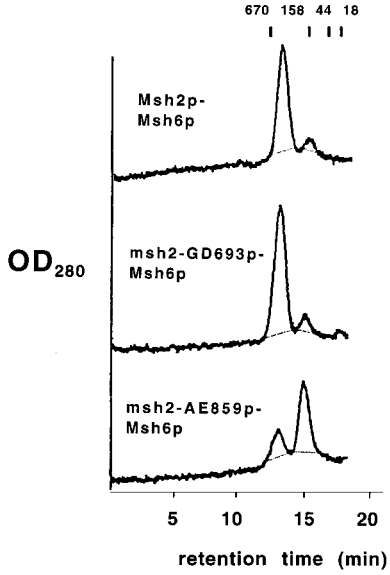

FIG. 3. Gel filtration analysis of Msh2p-Msh6p, msh2-GD693p–Msh6p, and msh2-AE859p-Msh6p. Five to 10 μg of each the indicated preparations was applied to a Superose 6HR gel filtration column (Materials and Methods) and eluted in 500 mM NaCl–25 mM Tris (pH 7.5)–1 mM EDTA–10 mM β-mercaptoethanol. The elution profiles of thyroglobulin (670 kDa), bovine gamma globulin (158 kDa), chicken ovalbumin (44 kDa), and equine myoglobulin (18 kDa) standards are shown. OD_{280}, optical density at 280 nm.

a molecular mass of approximately 100 to 140 kDa. This observation, as well as the chromatographic properties of this complex described below, suggested that the lower-molecular-mass fraction represented monomeric forms of msh2-AE859p and Msh6p polypeptides.

While the stoichiometry of msh2p and Msh6p in the msh2-AE859p–Msh6p and msh2-VD862p–Msh6p preparations appeared indistinguishable from that of the Msh2p-Msh6p complex, four observations suggested that these complexes were unstable: (i) gel filtration analysis revealed that only a fraction of the msh2-AE859p–Msh6p complex eluted at a molecular weight consistent with the purification of an intact Msh2p-Msh6p complex (see Fig. 3 and below); (ii) immunoprecipitation reactions involving Msh2p-Msh6p, msh2-AE859p-Msh6p, and msh2-VD862–Msh6p fraction IV preparations indicated that only a small fraction of Msh6p could be coimmunoprecipitated with msh2-AE859p or msh2-VD862p, while an equal fraction could be immunoprecipitated with Msh2p (reference 1 and data not shown); (iii) as shown in Table 2, the yields of the msh2-AE859p–Msh6p and msh2-VD862p–Msh6p complexes were much lower than those observed for the Msh2p-Msh6p, msh2-GD693p–Msh6p, and msh2-GD855p–Msh6p complexes; (iv) in the first PBE94 gradient, a small proportion of both msh2-AE859p and msh2-VD862p polypeptides eluted earlier in the gradient than the Msh6p, and a small proportion of these msh2 polypeptides flowed through the ssDNA-cellulose loading step.

Mismatch binding specificity of mutant complexes. Previously, we showed in filter binding assays that Msh2p-Msh6p could specifically bind to oligonucleotides containing a single-nucleotide insertion (1). We used this assay, which measures the binding affinity of Msh2p-Msh6p to a ^{32}P-labeled +1 duplex oligonucleotide mismatch substrate in the presence or absence of unlabeled +1 and homoduplex competitors, to assess the mismatch binding specificity of the mutant msh2p-Msh6p complexes (1). Previous analysis showed that MutS, Msh2p-Msh6p, and human MSH2-GTBP all displayed an approximately 5- to 15-fold specificity for oligonucleotide DNA substrates containing base pair mismatches compared to homoduplex oligonucleotide substrates (1, 5, 21, 23). Furthermore, theoretical and experimental considerations have demonstrated that the discrimination between two competitors can be determined by measuring the maximal horizontal separation between the binding curves resulting from such titrations (9). This measurement is best made at the maximum concentration of competitor that still allows accurate determination of substrate binding because the horizontal separation between the binding curves is constant at high degrees of competition. Competition assays shown in Fig. 4 were performed under conditions where 0.26 μg of each type of complex was incubated with stoichiometric amounts of ^{32}P-labeled +1 substrate (1 pmol of Msh2p-Msh6p per pmol of oligonucleotide substrate) plus various amounts of unlabeled competitor. It is important to note that the K_D for binding of Msh2p-Msh6p to the +1 substrate could not be determined because it was not possible to achieve equilibrium binding conditions due to the complex mode of binding of Msh2p-Msh6p (1). For this reason, the binding data are presented relative to the binding of wild-type or mutant complexes in the absence of competitor. The percentage of total binding for the wild-type and mutant complexes to the +1 substrate in the absence of competitor is presented in the legend to Fig. 4.

As shown in Fig. 4A to C, for the Msh2p-Msh6p, msh2-GD693p–Msh6p, and msh2-GD855p–Msh6p complexes, approximately fivefold-higher levels of homoduplex competitor were required to achieve the same degree of competition for the ^{32}P-labeled +1 substrate as was observed with the +1 competitor when high levels of competition were achieved. The overall levels of binding of each of the three complexes to the +1 substrate in the absence of competitor were similar, suggesting that the wild-type and msh2-GD693p–Msh6p and msh2-GD855p–Msh6p mutant complexes showed similar DNA binding affinities (data not shown). Taken together, these studies indicated that the msh2-GD693p–Msh6p and msh2-GD855p–Msh6p complexes were proficient in mismatch recognition.

As shown previously, the msh2-Δ6p–Msh6p preparation, which contained mostly Msh6p protein, did not display mismatch binding specificity (1). These data, combined with the results of Msh2p DNA binding studies performed in the absence of Msh6p, suggested that the specificity of mismatch binding for Msh2p and Msh6p proteins was dependent on the ability of these proteins to form an intact complex (1, 2). Support for this hypothesis came from biochemical analysis of the class I mutant proteins; these proteins were shown by chromatography and by immunoprecipitation analysis to be defective in Msh6p interactions. msh2-AE859p–Msh6p and msh2-VD862p–Msh6p complexes, which displayed moderate subunit interaction defects (Table 2, Fig. 3, and data not shown), also showed weak specificity for mismatch substrates. The msh2-PD865p–Msh6p and msh2-Δ6p–Msh6p preparations, which showed a strong defect in Msh2p-Msh6p interactions, were completely defective for mismatch specificity and showed a slightly higher specificity for homoduplex substrates (reference 1, Fig. 4, and data not shown). Similar conclusions were reached in mismatch binding studies performed with nuclear extracts that had been prepared from cells expressing mutant msh2 proteins in single copy (34, 35). In these studies, *MSH2*-dependent mismatch binding was observed in nuclear extracts prepared from cells expressing either msh2-GD693p or msh2-GD855p but was not observed in extracts prepared from cells expressing either msh2-AD872p or msh2-Δ6p (34a).

msh2-GD693p–Msh6p and msh2-GD855p–Msh6p complexes bind to mismatches in a reaction that is insensitive to ATP. Previous analysis of the Msh2p-Msh6p complex indicated that mismatch substrate specificity could be abolished if binding reactions were performed in the presence of ATP (1). We tested the effects of ATP on the binding properties of msh2-GD693–Msh6p and msh2-GD855p–Msh6p complexes by incubating wild-type and mutant complexes in the presence of ^{32}P-labeled +1 and unlabeled +1 and homoduplex competitor in the presence and absence of 1.6 mM ATP. As shown in Fig. 5, mismatch discrimination was not observed for the Msh2p-Msh6p complex when binding reactions were performed in the presence of ATP. However, mismatch discrimination could still be observed for the msh2-GD693p–Msh6p and msh2-GD855p–Msh6p complexes under the same conditions. This finding was strikingly similar to that observed when the nonhydrolyzable ATP analog AMP-PNP was included in Msh2p-Msh6p binding reactions (1).

Previously we showed that the Msh2p-Msh6p complex displayed an ATPase activity (K_m = 6 μM) that could be inhibited when Msh2p-Msh6p was incubated in the presence of the +1 substrate. To test whether the msh2-GD693p–Msh6p, msh2-GD855p–Msh6p, and msh2-AE859p–Msh6p complexes displayed defects in ATPase activity, ATPase assays were performed at various ATP concentrations in the presence and absence of homoduplex and +1 DNA substrates. As shown in Fig. 5D, the ATPase activities of msh2-GD693p–Msh6p and msh2-GD855p–Msh6p were weaker in the presence of homoduplex and +1 substrates than was observed for the Msh2p-Msh6p. Consistent with this finding was the observation that

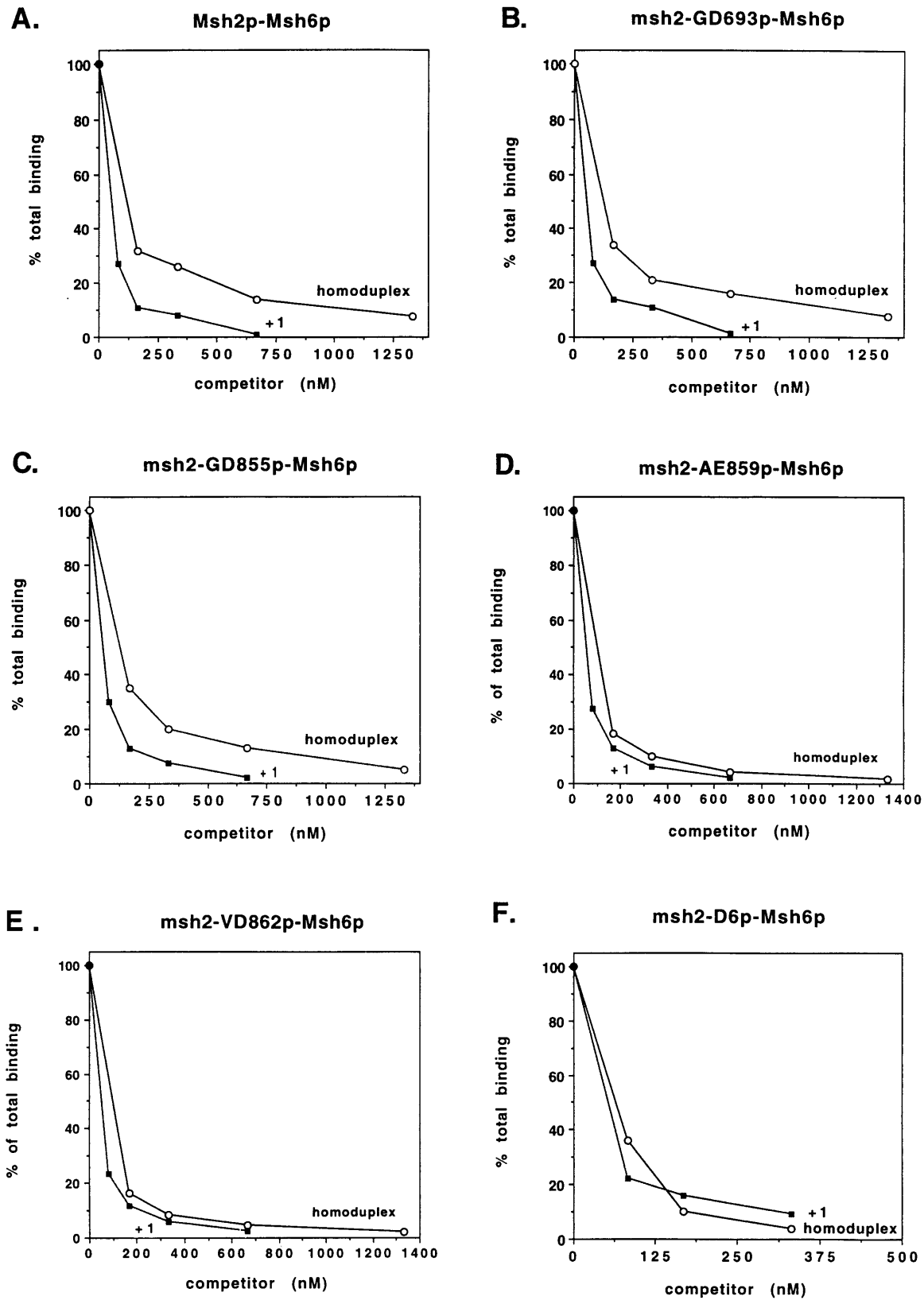

FIG. 4. Mismatch binding assays performed with Msh2p-Msh6p and class II (msh2-GD693p–Msh6p and msh2-GD855p–Msh6p) and class I (msh2-AE859p–Msh6p, msh2-VD862p–Msh6p, and msh2-Δ6p–Msh6p) mutant heterodimers. Binding was performed at 30°C in 60-μl reaction mixtures containing 25 mM Tris (pH 7.5), 2 mM $MgCl_2$, 0.1 mM DTT, 0.01 mM EDTA, 40 μg of BSA per ml, 0.3 μg of Msh2p-Msh6p or msh2p-Msh6p fraction IV, 16.7 nM ^{32}P-labeled +1 substrate, and the indicated amount of unlabeled +1 and homoduplex competitor substrates. After a 15-min incubation, the amount of ^{32}P-labeled +1 substrate that remained bound to Msh2p-Msh6p or msh2p-Msh6p was measured by filter binding. Competitive binding reactions involving the indicated complexes are shown. Data for msh2-Δ6p–Msh6p competition experiments are from reference 1. The percentages of input +1 substrate bound for each of the complexes in the absence of competitor were as follows: Msh2p-Msh6p, 26%; msh2-GD693p–Msh6p, 11%; msh2-GD855p–Msh6p, 14%; msh2-AE859p–Msh6p, 30%; msh2-VD862p–Msh6p, 23%; and msh2-Δ6p–Msh6p, 19%.

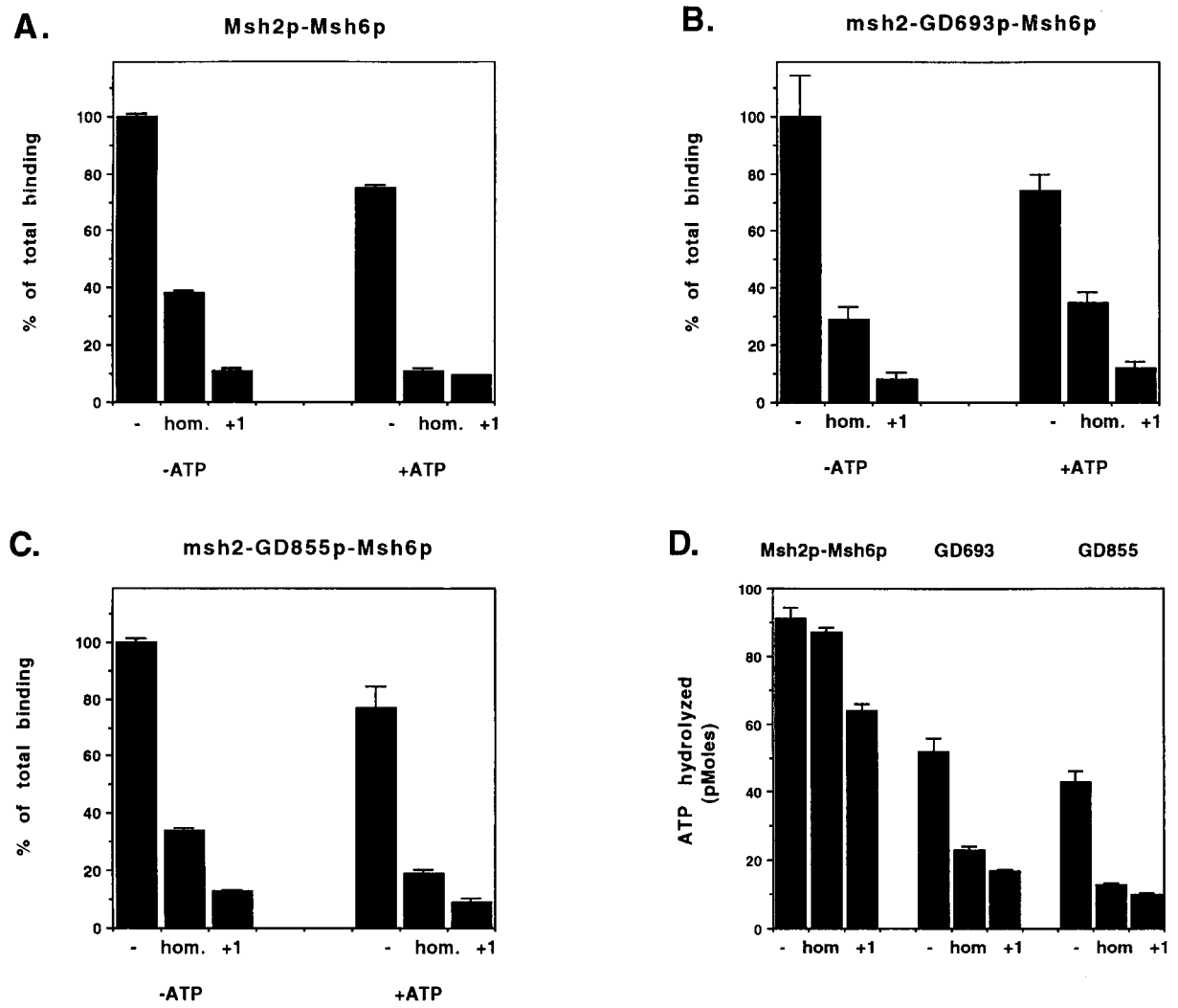

FIG. 5. Addition of ATP to mismatch binding reactions eliminated the mismatch binding specificity of Msh2p-Msh6p (A) but not msh2-GD693p–Msh6p (B) or msh2-GD855p–Msh6p (C) complexes. Binding was performed in duplicate reactions at 30°C in 60 μl-reaction mixtures containing 25 mM Tris (pH 7.5), 2 mM $MgCl_2$, 0.1 mM DTT, 0.01 mM EDTA, 40 μg of BSA per ml, 0.30 μg of the respective fraction IV preparation, 16.7 nM ^{32}P-labeled +1 substrate, and 83 nM unlabeled +1 or homoduplex competitor substrates (hom.); 1.6 mM ATP was included in binding reaction mixtures as indicated. After a 15-min incubation, the amount of ^{32}P-labeled +1 substrate that remained bound to Msh2p-Msh6p was measured by filter binding. Binding data are presented relative to binding observed in the absence of ATP and unlabeled competitor substrate. (D) ATP hydrolysis activity in binding reactions performed with Msh2p-Msh6p, msh2-GD693p–Msh6p and msh2-GD855p–Msh6p. Three-tenths microgram of fraction IV of each preparation was incubated in the presence of 100 μM [γ-^{32}P]ATP and 167 nM the indicated oligonucleotide DNA substrate. The amount of ATP hydrolyzed was determined after a 15-min incubation at 30°C (Materials and Methods). In all panels, the results from duplicates were averaged and the range between the two values is shown. The percentages of input +1 substrate bound for each of the complexes in the absence of competitor and ATP were as follows: Msh2p-Msh6p, 26%; msh2-GD693p–Msh6p, 11%; and msh2-GD855p–Msh6p, 14%.

the K_m for ATP hydrolysis of these mutant complexes was approximately twofold higher than that observed for Msh2p-Msh6p (Table 3). The K_m of the msh2-AE859p-Msh6p complex, which displayed defects in both complex formation and mismatch recognition, was approximately fivefold higher than that measured for Msh2p-Msh6p.

DISCUSSION

We performed site-directed mutagenesis of the *MSH2* ATP binding and helix-turn-helix domains to determine whether these domains play an important role in Msh2p function. Genetic and biochemical analyses revealed two classes of non-null mutant proteins. Class I proteins, which contained mutations that mapped to the helix-turn-helix domain of *MSH2*, were defective in forming complexes with Msh6p and also displayed defects in mismatch recognition. Class II proteins, which contained mutations that mapped to either the ATP binding or helix-turn-helix domain, formed complexes with Msh6p that were proficient in mismatch recognition; however, the mismatch specificity of these complexes was not modulated by ATP. Genetic analysis indicated that overexpression of the class II mutant proteins in a wild-type strain resulted in a dominant negative phenotype.

Msh2p-Msh6p complex formation is required for mismatch recognition. Previous analysis indicated that purified Msh2p or fractions highly enriched for Msh6p did not display the mismatch binding specificity observed for the Msh2p-Msh6p complex. Instead, these fractions displayed DNA binding specificities that did not correlate with specificities predicted by in vivo mismatch repair studies (1, 2, 24, 31, 37, 44). In this report, we showed that fractions containing Msh6p and msh2 proteins

TABLE 3. K_m for ATP hydrolysis of Msh2p-Msh6p and mutant derivatives

Complex	K_m(s) (μM ATP)[a]
Msh2p-Msh6p	5.0, 7.4
msh2-GD693p–Msh6p	13, 13
msh2-AE859p–Msh6p	21, 38
msh2-GD855p–Msh6p	10

[a] Determined as described in Materials and Methods. For the Msh2p-Msh6p, msh2-GD693p–Msh6, and msh2-AE859p–Msh6p complexes, K_m measurements were made on two independently purified protein preparations; the value determined for each preparation is shown.

defective in Msh6p interaction (msh2-AE859p and msh2-VD862p) were also defective in mismatch binding. It is important to note that these fractions contained Msh6p and the mutant msh2 proteins in stoichiometries similar to that observed for the wild-type complex (Fig. 2 and 4). These results support a model in which Msh2p-Msh6p complex formation is required for mismatch recognition rather than a model involving the independent binding of Msh2p and Msh6p to base pair mismatches. Double-mutant analyses involving *msh2* alleles bearing mutations that displayed a dominant negative phenotype (*msh2-GD693*, *msh2-GD855*, and *msh2-VD858*) and those that displayed a strong defect in complex formation (*msh2-Δ6*, *msh2-AD872*, and *msh2-PD865*) (Table 1) provide further support for this model. The dominant negative phenotype exhibited by the *msh2-GD693* mutation could be suppressed by *msh2* mutations that caused strong defects in complex formation, suggesting that the dominant negative phenotype reflected the activity of mutant complexes that could bind to base pair mismatches but could not act in subsequent mismatch repair steps (see below). According to this model, suppression of the dominant negative phenotype could be achieved by disrupting complex formation and thus prevent complex binding to the mismatch. We cannot exclude the possibility that mutations that disrupted Msh2p complex formation can also act to disrupt other Msh2p functions such as mismatch recognition; we do not favor this possibility, as our data suggest that the helix-turn-helix domain in Msh2p does not play a role in mismatch recognition (see below).

The helix-turn-helix domain in Msh2p is unlikely to serve as a DNA recognition domain analogous to that found in λ repressor. The identification of a helix-turn-helix domain in the MutS homologs encouraged us to create site-specific mutations in *MSH2* with the expectation that mutations in the helix-turn-helix domain would result in a mismatch recognition defect. Our data, however, suggested that the helix-turn-helix domain identified in the *mutS* homologs was unlikely to play a direct role in mismatch recognition, as described below.

First, mutations in the Msh2p helix-turn-helix domain that would be expected, based on mutational analysis of DNA binding proteins that utilize a helix-turn-helix domain, to cause a disruption of contacts between polypeptide sequences and the DNA phosphate backbone did not result in defects in mismatch recognition (41, 58, 60). For example, overexpression of msh2-GD855p and msh2-VD858p in wild-type strains resulted in a mutator phenotype, consistent with the idea that the mutant msh2 proteins could interact with Msh6p and compete with Msh2p-Msh6p for binding to DNA containing base pair mismatches. Biochemical analysis supported this idea: the msh2-GD855p–Msh6p complex was proficient in mismatch recognition but displayed an ATP binding and/or hydrolysis defect that resulted in a mutant complex that remained bound to a mismatch even in the presence of ATP.

Second, the *msh2-AE859* and *msh2-VD862* mutations, which caused defects in both Msh2p-Msh6p interactions and mismatch recognition, did not cause a mutator phenotype when overexpressed in wild-type strains but did so in conjunction with the *GD693* mutation. The biochemical and genetic analyses presented here and elsewhere (17, 61) are consistent with the idea that the dominant negative *msh2* mutations are defective in mismatch repair in steps that occur after mismatch recognition. If we assume that the dominant negative mutator phenotypes conferred by *msh2-GD693-AE859* and *msh2-GD693-VD862* mutations were the result of defects following mismatch recognition, then it seems unlikely that the *AE859* and *VD862* mutations directly disrupted contacts between Msh2p-Msh6p and the base pair mismatch, as the double mutant combinations displayed phenotypes consistent with a proficiency in mismatch recognition. An explanation for the dominant negative phenotype observed in *msh2-GD693-AE859* and *msh2-GD693-VD862* mutants but not in *msh2-GD693-PD865* or *msh2-GD693-AD872* mutants was that the msh2-AE859p and msh2-VD862p proteins displayed a less severe Msh6p interaction defect than the msh2-AD872p and msh2-PD865p proteins (Fig. 2). Based on these observations, we hypothesized that overexpression of the msh2-GD693–AE859p or msh2-GD693–VD862p complex resulted in a dominant negative phenotype because a large enough proportion of the mutant proteins could complex with Msh6p to bind to base pair mismatches and interfere with Msh2p-Msh6p binding. Unfortunately, we were unable to purify intact msh2-GD693-AE859p–Msh6p or msh2-GD693-VD862p–Msh6p complexes to directly test this hypothesis in vitro. It is important to note that the foregoing hypothesis can also help to explain why some of the class I *msh2* alleles displayed a weak dominant negative phenotype when overexpressed if a fraction of the mutant proteins could compete with wild-type Msh2p for binding to Msh6p (Table 1).

The points summarized above argued against the idea that the helix-turn-helix domain in Msh2 acts to directly bind to base pair mismatches in a way that is structurally analogous to that observed for helix-turn-helix domains in λ, Cro, and Cap repressor. However, we cannot formally rule out the possibility that the helix-turn-helix domain in the MutS homologs functions as part of a domain that is directly required for mismatch recognition. Structural analysis of DNA binding proteins such as λ repressor and Gal4p indicated that homodimer formation was required for these proteins to display sequence specific recognition; in some cases, mutations that resulted in the loss of dimerization also mapped to the DNA binding domain (8, 41). Based on these findings, it is possible that a mismatch recognition domain could include the Msh2p and Msh6p interaction domains. The isolation of mutant msh2p and msh6p proteins that can form stable complexes with their wild-type counterpart proteins and also display a defect in mismatch recognition could enable us to directly identify a mismatch recognition domain. Alternatively, physical methods, such as UV cross-linking of mismatched DNA to the Msh2p-Msh6p complex could allow us to directly identify amino acid residues required for mismatch recognition.

Model for early steps in mismatch repair. We propose a model in Fig. 6 to explain how the ATP binding and helix-turn-helix domains in Msh2 could function in the Msh2p-Msh6p complex. In this model, the two domains coordinately act in response to ATP binding and/or hydrolysis signals to modulate mismatch recognition of the Msh2p-Msh6p complex. This model, which also proposes that the conformational changes

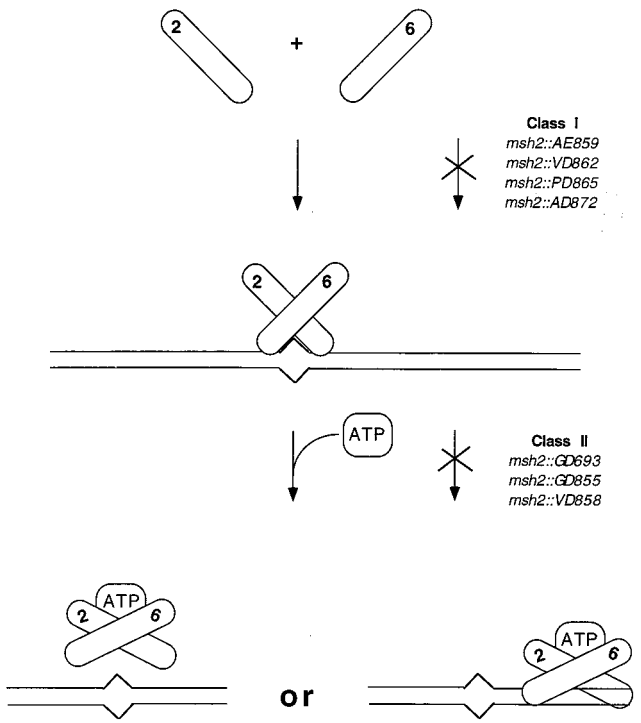

FIG. 6. A model describing Msh2p-Msh6p interactions with DNA containing base pair mismatches. Following mismatch recognition by the Msh2p-Msh6p complex, the ATP hydrolysis activity of Msh2p-Msh6p results in a conformational change within the interaction domain of the two proteins that results in the loss of mismatch binding specificity. This change then allows the Msh2p-Msh6p complex to interact with mismatch repair proteins that participate in subsequent steps in mismatch repair. Mutations in the *MSH2* ATP binding and helix-turn-helix domains that disrupt steps proposed in this model are shown.

that occur during ATP binding and/or hydrolysis are productive only if both Msh2p and Msh6p are present, is based on the following observations.

(i) *msh2* mutations affecting either ATP binding and hydrolysis or Msh2p-Msh6p interactions can be identified within the same region of the Msh2 polypeptide sequence. For example, the *GD855* and *VD858* mutations resulted in phenotypes consistent with ATP binding/hydrolysis defects, while the *AE859*, *VD862*, *PD865*, and *AD872* mutations resulted in defects in Msh6p interaction. The finding that mutations that affected ATP binding and hydrolysis were located within a region that was also required for Msh2p-Msh6p interaction suggested the possibility that ATP binding and hydrolysis could modulate mismatch recognition through conformational changes in the Msh2p-Msh6p interaction domain.

(ii) The K_m for ATP hydrolysis in msh2-AE859p–Msh6p preparations was several fold higher than the K_m observed for either the Msh2p-Msh6p complex or the msh2-GD855p–Msh6p or msh2-GD693p–Msh6p mutant complex. Of these four complexes, only the msh2-AE859p–Msh6p complex displayed a defect in complex stability as well as mismatch recognition. In addition, the mutator phenotype of the *msh2-AE859* allele but not the *msh2-GD693*, *msh2-GD855*, and *msh2-VD858* alleles could be suppressed by overexpression. These findings suggested that the interaction between the two proteins played a critical role in modulating the ATPase activity of the complex. It is important to note that the ATPase activity of the msh2-GD693p–Msh6p complex was not reduced to the extent that would be expected for an amino acid substitution located in a conserved sequence in the P loop (see references 17 and 55 for examples). One hypothesis that will be tested in the future and is consistent with this result is that both Msh2p and Msh6p contain ATPase activities and that the residual ATPase activity displayed in msh2-GD693p–Msh6p and msh2-GD855p–Msh6p reflected the Msh6p ATPase activity.

ATP hydrolysis is required in mismatch repair steps that follow mismatch recognition. Previously we showed that the mismatch binding activity of Msh2p-Msh6p could be eliminated when binding reactions were performed in the presence of ATP (1). However, we have now observed two cases where mismatch binding can occur in the presence of ATP. In the first case, Msh2p-Msh6p was observed to bind to mismatched substrates containing palindromic insertions that are poorly repaired in vivo (38) with specificity similar to that observed for single-base-pair mismatches. However, unlike binding to GT or +1 mismatches, binding to palindromic substrates was not eliminated when reactions were performed in the presence of ATP (1). In the second case, site specific mutations were created in *MSH2* that resulted in the formation of complexes that could bind DNA mismatches with the same specificity as the wild type and could still display mismatch recognition in the presence of ATP. Taken together, these observations indicate that ATP-dependent steps following mismatch recognition play a critical role in deciding whether a base pair mismatch will be repaired. Supporting this idea were experiments performed by Au and coworkers, who showed that ATP was required for MutS-MutL-mismatch DNA interactions and ATP hydrolysis was required to activate MutH endonuclease activity (4). It is important to note that the activation of MutH correlated with the efficiency in which a base pair mismatch was subject to repair. Based on this information and the fact that neither MutL and MutH displayed an intrinsic ATP binding/hydrolysis activity, Au and coworkers proposed that MutS-dependent ATP hydrolysis participated in a signal transduction step that resulted in the activation of MutH endonuclease (4). This proposal as well as the model presented in Fig. 6 provide possible mechanisms to explain how the transduction step could be accomplished. This interaction is analogous to that proposed for the hydrolysis of GTP by trimeric G proteins, where the exchange of GDP for GTP allows for the relay of messages from receptor proteins to effector targets (6). Based on studies presented here, we propose that the conformational changes in the Msh2p-Msh6p interaction domains that are induced by ATP hydrolysis could act to signal interactions with other mismatch repair proteins such as the *S. cerevisiae mutL* homolog Mlh1p-Pms1p complex (46). Further analysis of the effect of Msh2p ATPase activity on downstream mismatch repair factors awaits analysis of purified mismatch repair components in an in vitro mismatch repair system that displays strand-specific repair.

ACKNOWLEDGMENTS

We thank Leon Heppel, Gerry Marsischky, Jinlin Peng, Jeff Roberts, and members of the Alani lab for technical advice and helpful discussions and John Lis for comments on the manuscript. E.A. thanks Richard Kolodner for his interest and support during the early part of this work.

E.A. was supported by National Institutes of Health grant GM53085 and USDA Hatch Grant NYC-186424, T.S. was supported by a National Institutes of Health predoctoral training grant, and B.S. was supported by a State University of New York fellowship.

REFERENCES

1. **Alani, E.** 1996. The *Saccharomyces cerevisiae* Msh2p and Msh6p form a complex that specifically binds to duplex oligonucleotides containing mis-

matched DNA base pairs. Mol. Cell. Biol. **16:**5604–5615.
2. **Alani, E., N. W. Chi, and R. D. Kolodner.** 1995. The *Saccharomyces cerevisiae* Msh2 protein specifically binds to duplex oligonucleotides containing mismatched DNA base pairs and loop insertions. Genes Dev. **9:**234–247.
3. **Alani, E., R. A. G. Reenan, and R. D. Kolodner.** 1994. Interaction between mismatch repair and genetic recombination in *Saccharomyces cerevisiae*. Genetics **137:**19–39.
4. **Au, K. G., K. Welsh, and P. Modrich.** 1992. Initiation of methyl-directed mismatch repair. J. Biol. Chem. **267:**12142–12148.
5. **Biswas, I., and P. Hsieh.** 1996. Identification and characterization of a thermostable MutS homolog from Thermus aquaticus. J. Biol. Chem. **271:** 5040–5048.
6. **Bourne, H. R.** 1995. Trimeric G. Proteins: surprise witness tells a tale. Science **270:**933–934.
7. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
8. **Carey, M., H. Kakidani, J. Leatherwood, F. Mostashari, and M. Ptashne.** 1989. An amino-terminal fragment of GAL4 binds as a dimer. J. Mol. Biol. **209:**423–432.
9. **Chi, N., and R. D. Kolodner.** 1994a. Purification and characterization of Msh1, a yeast mitochondrial protein that binds to DNA mismatches. J. Biol. Chem. **269:**29984–29992.
10. **Chi, N., and R. D. Kolodner.** 1994b. The effect of DNA mismatches on the ATPase activity of Msh1, a protein in yeast mitochondria that recognizes DNA mismatches. J. Biol. Chem. **269:**29993–29997.
11. **Drummond, J. T., G.-M. Li, M. J. Longley, and P. Modrich.** 1995. Mismatch recognition by an hMSH2-GTBP heterodimer and differential repair defects in tumor cells. Science **268:**1909–1912.
12. **Fishel, R., M. K. Lescoe, M. R. S. Rao, N. G. Copeland, N. A. Jenkins, J. Garber, M. Kane, and R. D. Kolodner.** 1993. The human mutator gene homolog Msh2 and its association with hereditary nonpolyposis colon cancer. Cell **75:**1027–1038.
13. **Gietz, R. D., and R. H. Schiestl.** 1991. Applications of high efficiency lithium acetate transformation of intact yeast cells using single-stranded nucleic acids as carrier. Yeast **7:**253–263.
14. **Gorbalenya, A. E., and E. V. Koonin.** 1990. Superfamily of *uvrA*-related binding proteins. Implications for rational classification of recombination repair systems. J. Mol. Biol. **213:**583–591.
15. **Grilley, M., K. M. Welsh, S.-S. Su, and P. Modrich.** 1989. Isolation and characterization of the *Escherichia coli mutL* gene product. J. Biol. Chem. **264:**1000–1004.
16. **Haber, L. T., P. P. Pang, D. I. Sobell, J. A. Mankovich and G. C. Walker.** 1988. Nucleotide sequence of the *Salmonella typhimurium mutS* gene required for mismatch repair: homology of *mutS* and *hexA* of *Streptococcus pneumoniae*. J. Bacteriol. **170:**197–202.
17. **Haber, L. T., and G. C. Walker.** 1991. Altering the conserved nucleotide binding motif in the *Salmonella typhimurium* MutS mismatch repair protein affects both its ATPase and mismatch binding activities. EMBO J. **10:**2707–2715.
18. **Habraken, Y., P. Sung, L. Prakash, and S. Prakash.** 1996. Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. Curr. Biol. **6:**1185–1187.
19. **Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease.** 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene **77:**51–59.
20. **Hollingsworth, N. M., L. Ponte, and C. Halsey.** 1995. *MSH5*, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair. Genes Dev. **9:**1728–1739.
21. **Hughes, M. J., and J. Jiricny.** 1992. The purification of a human mismatch-binding protein and identification of its associated ATPase and helicase activities. J. Biol. Chem. **267:**23876–23882.
22. **Iaccarino, I., F. Palombo, J. Drummond, N. F. Totty, J. J. Hsuan, P. Modrich, and J. Jiricny.** 1996. MSH6, a Saccharomyces cerevisiae protein that binds to mismatches as a heterodimer with MSH2. Curr. Biol. **6:**484–486.
23. **Jiricny, J., S. Su, S. G. Wood, and P. Modrich.** 1988. Mismatch-containing oligonucleotide duplexes bound by the *E. coli mutS*-encoded protein. Nucleic Acids Res. **16:**7843–7853.
24. **Johnson, R. E., G. K. Kovvali, L. Prakash, and S. Prakash.** 1996. Requirement of the yeast *MSH3* and *MSH6* genes for *MSH2* dependent genomic stability. J. Biol. Chem. **271:**7285–7288.
25. **Kolodner, R. D.** 1995. Mismatch repair: mechanisms and relationship to cancer susceptibility. Trends Biochem. **20:**397–401.
26. **Kolodziej, P. A., and R. A. Young.** 1991. Epitope tagging and protein surveillance. Methods Enzymol. **194:**508–519.
27. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
28. **Lahue, R. S., K. G. Au, and P. Modrich.** 1989. DNA mismatch correction in a defined system. Science **245:**160–164.
29. **Leach, F. S., N. C. Nicolaides, N. Papadopoulos, B. Liu, J. Jen, R. Parsons, P. Peltomaki, P. Sistonen, L. A. Aaltonen, M. Nystrom-Lahti, X.-Y. Guan, J. Zhang, P. S. Meltzer, J.-W. Yu, F.-T. Kao, D. J. Chen, K. M. Cerosaletti, R. E. K. Fournier, S. Todd, T. Lewis, R. J. Leach, S. L. Naylor, J. Weissenbach, J.-P. Mecklin, H. Jarvinen, G. M. Petersen, S. R. Hamilton, J. Green, J. Jass, P. Watson, H. T. Lynch, J. M. Trent, A. de la Chapelle, K. W. Kinzler, and B. Vogelstein.** 1993. Mutations of a MutS homolog in hereditary nonpolyposis colorectal cancer. Cell **75:**1215–1225.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Marsischky, G. T., N. Filosi, M. F. Kane, and R. Kolodner.** 1996. Redundancy of *Saccharomyces cerevisiae MSH3* and *MSH6* in *MSH2*-dependent mismatch repair. Genes Dev. **10:**407–420.
32. **McEntee, K., G. Weinstock, and I. R. Lehman.** 1980. recA protein-catalyzed strand assimilation: stimulation by *Escherichia coli* single-stranded DNA-binding protein. Proc. Natl. Acad. Sci. USA **77:**857–861.
33. **Miller, J.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Miret, J. J., M. G. Milla, and R. S. Lahue.** 1993. Characterization of a DNA mismatch-binding activity in yeast extracts. J. Biol. Chem. **268:**3507–3513.

34a. **Miret, J. J., and R. Lahue.** Unpublished data.

35. **Miret, J. J., B. O. Parker, and R. S. Lahue.** 1996. Recognition of DNA insertion/deletion mismatches by an activity in *Saccharomyces cerevisiae*. Nucleic Acids Res. **24:**721–729.
36. **Modrich, P.** 1991. Mechanisms and biological effects of mismatch repair. Annu. Rev. Genet. **25:**229–253.
37. **Modrich, P., and R. S. Lahue.** 1996. Mismatch repair in replication fidelity, genetic recombination and cancer biology. Annu. Rev. Biochem. **65:**101–133.
38. **Nag, D. K., and T. D. Petes.** 1991. Seven-base-pair inverted repeats in DNA form stable hairpin *in vivo* in *Saccharomyces cerevisiae*. Genetics **129:**669–673.
39. **New, L., K. Liu, and G. F. Crouse.** 1993. The yeast gene *MSH3* defines a new class of eukaryotic MutS homologues. Mol. Gen. Genet. **239:**97–108.
40. **Ohlendorf, D. H., W. F. Anderson, and B. W. Mathews.** 1983. Many gene-regulatory proteins appear to have a similar alpha-helical fold that binds DNA and evolved from a common precursor. J. Mol. Evol. **19:**109–114.
41. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53:**293–321.
42. **Palombo, F., P. Gallinari, I. Iaccarino, T. Lettieri, M. Hughes, A. D'Arrigo, O. Truong, J. J. Hsuan, and J. Jiricny.** 1995. GTBP, a 160 kD protein essential for mismatch binding activity in human cells. Science **268:**1912–1914.
43. **Palombo, F., I. Iaccarino, E. Nakajima, M. Ikejima, T. Shimada, and J. Jiricny.** 1996. hMutSβ, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. Curr. Biol. **6:**1181–1184.
44. **Parker, B. O., and M. G. Marinus.** 1992. Repair of DNA heteroduplexess containing small heterologous sequences in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **89:**1730–1734.
45. **Petes, T. D., R. E. Malone, and L. S. Symington.** 1991. Recombination in yeast, pp. 407–521. *In* J. R. Broach, J. R. Pringle, and E. W. Jones (ed.), The molecular and cellular biology of the yeast *Saccharomyces*: genome dynamics, protein synthesis and energetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
46. **Prolla, T. A., Q. Pang, E. Alani, R. D. Kolodner, and R. M. Liskay.** 1994b. Interactions between the Msh2, Mlh1 and Pms1 proteins during the initiation of DNA mismatch repair. Science **265:**1091–1093.
47. **Reenan, R. A. G.** 1991. Identification and characterization of genes involved in mismatch repair in *Saccharomyces cerevisiae*. Ph.D. thesis. Harvard University, Cambridge, Mass.
48. **Reenan, R. A. G., and R. D. Kolodner.** 1992a. Isolation and characterization of two *Saccharomyces cerevisiae* genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins. Genetics **132:**963–973.
49. **Reenan, R. A. G., and R. D. Kolodner.** 1992b. Characterization of insertion mutations in the *Saccharomyces cerevisiae MSH1* and *MSH2* genes: evidence for separate mitochondrial and nuclear functions. Genetics **132:** 975–985.
50. **Rose, M. D., F. Winston, and P. Hieter.** 1990. Methods in yeast genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
51. **Ross-Macdonald, P., and G. S. Roeder.** 1994. Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction. Cell **79:**1069–1080.
52. **Sherman, F., G. Fink, and J. Hicks.** 1983. Methods in yeast genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

52a. **Sokolsky, T., and E. Alani.** Unpublished data.

53. **Su, S., R. S. Lahue, K. G. Au, and P. Modrich.** 1988. Mispair specificity of methyl-directed DNA mismatch correction *in vitro*. J. Biol. Chem. **263:**6829–6835.
54. **Su, S., and P. Modrich.** 1986. *Escherichia coli mutS*-encoded protein binds to mismatched DNA base pairs. Proc. Natl. Acad. Sci. USA **83:**5057–5061.
55. **Sung, P., D. Higgins, L. Prakash, and S. Prakash.** 1988. A mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its ATPase and

DNA helicase activities but not its ability to bind ATP. EMBO J. **7:**3263–3269.

56. **Wang, J., and L. Grossman.** 1993. Mutations in the helix-turn-helix motif of the *Escherichia coli* UvrA protein eliminates its specificity for UV-damaged DNA. J. Biol. Chem. **268:**5323–5331.
57. **Welsh, K. M., A.-L. Lu, S. Clark, and P. Modrich.** 1987. Isolation and characterization of the *Escherichia coli mutH* gene product. J. Biol. Chem. **262:**15624–15629.
58. **Wharton, R. P., and M. Ptashne.** 1986. An alpha-helix determines the DNA-binding specificity of a repressor. Trends Biochem. Sci. **11:**71–73.
59. **Winston, F., C. Dollard and S. L. Ricupero-Hovasse.** 1995. Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. Yeast **11:**53–55.
60. **Wissman, A., R. Baumeister, G. Muller, B. Hecht, V. Helbl, K. Pfleiderer, and W. Hillen.** 1991. Amino acids determining operator binding specificity in the helix-turn-helix motif of Tn10 Tet repressor. EMBO J. **10:** 4145–4152.
61. **Wu, T.-H., and M. G. Marinus.** 1994. Dominant negative mutator mutations in the *mutS* gene of *Escherichia coli*. J. Bacteriol. **176:**5393–5400.