Deletion Mutation Analysis of the mutS Gene in Escherichia coli*

Te-Hui Wu and Martin G. Marinus‡

From the Department of Pharmacology and Molecular Toxicology, University of Massachusetts Medical School, Worcester, Massachusetts 01655

The MutS protein is part of the dam-directed MutHLS mismatch repair pathway in Escherichia coli. We have constructed deletion derivatives in the mutS gene, which retain the P-loop coding region for ATP binding. The mutant proteins were assayed for ATP hydrolysis, heteroduplex DNA binding, heterodimer MutS formation, and the ability to interact with MutL. Dimerization was assayed by expressing His6-tagged wild-type and non-tagged deletion mutant proteins in the same cell and isolating the His6-tagged protein followed by MutS immunoblotting after SDS-polyacrylamide gel electrophoresis. MutS-MutL interaction was measured using the same technique except that the MutS protein carried the His6 tag. Our results indicate that DNA binding ability resides in the N-terminal end of MutS, and dimerization and MutL interactions are located in the C-terminal end. Given the extensive amino acid homology in the MutS family, our results with E. coli should be applicable to MutS homologues in other prokaryotes and eukaryotes.

The dimeric MutS protein is part of the Escherichia coli MutSLH DNA repair system that corrects mismatches arising in DNA as biosynthetic errors (1). The mutator phenotype of strains with an inactivated mutS gene is consistent with this idea (2). In addition MutS also participates in homologous recombination, transcription-coupled repair, very short patch repair, resolution of directly repeated DNA sequences, and sensitivity to cis-platin and alkylating agents (3–6). MutS binds specifically to mismatched base pairs and insertion/deletion mispairs of up to four nucleotides as well as adducted base mismatches (7, 8). The DNA-bound MutS protein binds ATP and in association with dimeric MutL can form α-helical structures with concomitant hydrolysis of ATP, suggesting movement of the protein complex on DNA away from the mismatch (9). An alternative model suggests that ADP stabilizes DNA binding and ATP increases dissociation (10). After addition of MutH, the MutSLH complex cleaves DNA 5’ to hemimethylated -GATC- sequences in the unmethylated strand. Subsequent excision repair from this nick can occur in either direction by one of several exonucleases, and directionality is imparted through interaction of the MutLS complex with helicase II, the uvrD gene product (11). Resynthesis of the gapped DNA is accomplished by DNA polymerase III holoenzyme and subsequent ligation by DNA ligase.

The basic elements of the scheme outlined above have been conserved in eukaryotes (3–5). This includes MutS and MutL homologues, each of which shows conservation with the E. coli protein at the amino acid level. Unlike bacteria, however, multiple MutS (MutS homologue) and MutL (MutL homologue, postmeiotic segregation) homologs are present as heterodimers in eukaryotes. Mutation in these genes results in microsatellite sequence instability and may dispose individuals to hereditary or sporadic cancer, especially colon carcinoma. The mechanism for strand discrimination during the repair process in eukaryotes is unknown.

At present, there is no structural information for any MutS homologue. The amino acids that are important for specific and nonspecific contact with DNA are not known. Similarly the residues important for dimerization and contacting MutL are unknown. To identify which regions of MutS are essential for these functions, we have made specific deletions in the mutS gene and have examined the properties of the truncated proteins. The results obtained with the E. coli protein should be applicable to other members of the MutS family given their overall amino acid conservation (4).

EXPERIMENTAL PROCEDURES

Bacterial Strains and Culture Conditions—Strain E. coli K-12 GM5864 is GM4271 lysogenized with phage ADE3. GM4271 is a mutant of E. coli which carries a single plasmid pMQ372 which is a pET3d (Novagen) derivative that contains the full coding sequence of a defective mutS gene. Plasmid pMQ372 is derived from pMQ372 but contains the N-terminal His6 tag of pET15b fused to the mutS coding sequence. Plasmids pMQ382 and pMQ395 are derivatives of pACYC184 and carry His6-tagged mutS and mutL, respectively (12).

Deletions in the E. coli mutS gene were made in pMQ372 at naturally occurring restriction enzyme recognition sites, with appropriate endonucleases (New England Biolabs), to produce derivatives with promoter-proximal, promoter-distal, or midgene mutations (see Fig. 1). The fusion joint for each deletion end point is shown in Table I. The method of constructing the deletion at the SupI site resulted in the addition of 13 additional amino acids to the mutant protein. Mutations from the pMQ372 series were moved into pMQ395 by fragment swapping to generate proteins with an N-terminal His6 tag. The His6 tag wild-type and deleted mutant proteins were prepared as recommended by the manufacturer (Novagen) in strain GM5864. Although the specific conditions for cultivating cells varied for each construct, in general GM5864 transformed with a pMQ395 mutant derivative was grown at 37 °C to an A600 of 0.8, shifted to room temperature, and isopropyl-1-thio-β-D-galactopyranoside (IPTG)1 was added to 50 μM final concentration. Incubation was continued for 2–3 h at room temperature, and the cells were harvested and lysed in a French pressure cell (Amino) in

* This work was supported by Grant GM33233 from the National Institute of General Medical Sciences. 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.

‡ To whom correspondence should be addressed: Dept. of Pharmacology and Molecular Toxicology, University of Massachusetts Medical School, 55 Lake Ave., Worcester, MA 01655. Tel.: 508-856-3330; Fax: 508-856-5080; Internet: Martin.Marinus@ummed.edu.

1 The abbreviations used are: IPTG, isopropyl-1-thio-β-D-galactopyranoside; bp, base pair; PAGE, polyacrylamide gel electrophoresis.
lysis buffer (20 mM Tris-HCl, pH 7.9, 5 mM imidazole, 500 mM NaCl) followed by a brief sonication to reduce viscosity. The lysate was centrifuged in a Beckman Ti70 rotor at 39,000 rpm for 30 min at 4 °C, followed by filtration through a 0.45-µm syringe filter (Acrodisc). His6-tag protein was bound to and eluted from nickel-affinity resin (HisBind, Novagen) as recommended by the manufacturer. At least 80% of the wild-type and 10–30% of the mutant MutS were recovered as soluble protein. All proteins were at least 95% pure as judged by Coomassie Brilliant Blue staining of polyacrylamide gels and by Western blotting (e.g. see Fig. 4). Protein concentration was assayed using the Bradford reagent (Bio-Rad). ATPase activity was assayed by the method of Weinstock et al. (13), and percent [α-32P]ATP hydrolysis was measured by scanning polyethyleneimine chromatography plates using a Molecular Dynamics PhosphorImager equipped with ImageQuant software.

Heteroduplex DNA Construction—A 154-bp DNA oligonucleotide with a centrally located G-T mismatch at bp 76 was constructed as a Molecular Dynamics PhosphorImager equipped with ImageQuant manufacturer’s (Tropix) instructions. MutS and MutL (BAbCO) using chemiluminescence according to the method of Weinstock et al. (13), and percent [α-32P]ATP hydrolysis was measured by scanning polyethyleneimine chromatography plates using a Molecular Dynamics PhosphorImager equipped with ImageQuant software.

**RESULTS**

**Construction of Deletion Mutations**—The amino acids important for MutS dimerization, for binding to heteroduplex DNA, and for interaction with MutL are not known. We have constructed deletion derivatives in the *E. coli* mutS gene to determine whether these functions can be localized to specific regions of the MutS protein. Because ATP binding and/or hydrolysis is necessary for heteroduplex DNA-induced MutS conformational change (15) as well as movement of MutS along the heteroduplex DNA (9) and for MutS-MutL interaction (16), we constructed deletion mutations in the mutS gene that retained the coding sequence of the P-loop motif for nucleoside triphosphate binding (between the HpaI and SgrI sites in Fig. 1). All mutant proteins derived from the deletion constructs had similar low ATPase activity (about 0.34 µM ATP hydrolyzed/µM MutS/min) as the wild type.

**Mutator Phenotype of Cells with Plasmid-borne mutS Deletions**—Plasmids containing the wild-type and mutant mutS deletion alleles were introduced into wild-type and *mutS E. coli* strains. The presence of the plasmids in the wild-type strain did not significantly alter the mutation frequency to rifampicin resistance (Table II) or the frequency of Lac+ rever- sion (Table II) as measured by papillation (less than 1% of colonies showing one or more papillae). Similarly the plasmids did not significantly alter the mutation frequency of the *mutS* strain to rifampicin resistance (Table II) or Lac+ papillation (>90% of cells showing papillation). The deletion mutations, therefore, inactivate the protein for genetic complementation and do not show dominance in the wild type.

**Heteroduplex DNA Binding of Mutant Proteins**—MutS protein binds specifically to base mismatches and insertion/deletions in DNA (7, 8). The wild-type and deletion mutant proteins were tested for their ability to bind to a 154-bp heteroduplex containing a G-T mismatch at position 76 by monitoring bandshifts in DNA sequences of fusion joints using the restriction enzymes shown in Fig. 1. The H after the plasmid name indicates an N-terminal His6-tag.

| Plasmid | Sequence | Codons |
|---------|----------|--------|
| pΔ1–311H | AGC.CAT.ATG.CCA.GTG | 312 start |
| pΔ1–260 | AGATATACC.ATG.GAA | 261 start |
| pΔ26–260H | CAT.CCC.GAG.GAA.CGT | 26–260 deleted |
| pΔ26–260 | CAT.CCC.GAG.GAA.CGT | 26–260 deleted |
| pΔ261–556H | ATC.ACC.ATG.AAC.CTG | 26–260 deleted |
| pΔ261–556 | GCC.AAT.TTG.ATC.GTG.GAC.CGG.AGA. | 26–260 deleted |
| pΔ560–853H | GTC.TTG.ACG.TCC.TGC.GCC.TGC.TAA | 680–853 deleted |
| pΔ560–853 | GTC.TTG.ACG.TCC.TGC.GCC.TGC.TAA | 13 additional codons |

**TABLE I**

**DNA sequences of fusion joints**

The plasmid structures are shown in Fig. 1. The underlined sequences are the fusion joints using the restriction enzymes shown in Fig. 1. The H after the plasmid name indicates an N-terminal His6-tag.
A similar result was obtained if both His-6-tagged and non-tagged wild-type proteins were present in cell extracts (Fig. 3B, pMQ372). We assume the lower molecular weight band in these two figures is because of protein degradation. The distribution of wild-type and mutant proteins in eluates from the affinity column was similar if not identical for pD261–556 and pD1–260 encoded mutant proteins, indicating efficient heterodimer formation in vivo. Deletion of the N-terminal region of MutS encompassed by these two mutations (Fig. 1) therefore does not affect dimerization. For the pD680–853 and pD261–556 encoded proteins, however, the distributions of wild-type and mutant proteins were not the same (Fig. 3), indicating less efficient heterodimer formation. The basis for the altered distribution of the mutant proteins is unknown. These mutant proteins are as stable and soluble as those made from pD261–556 and pD1–260, suggesting that lack of heterodimer formation is not because of these causes. None of the mutant non-His6-tagged MutS proteins bind to the affinity column (data not shown). The results suggest that amino acid residues in the C-terminal end of the protein are responsible for efficient dimerization.

**E. coli MutS Deletion Mutants**

**Fig. 1.** Structure and properties of mutS deletion mutants. Plasmids pMQ372 and pMQ395 have the wild-type mutS sequence without (open rectangle) and with (gray box) an N-terminal His6-tag DNA, respectively. Pertinent restriction enzyme recognition sequences are shown. The DNA encoding the P-loop domain for ATP binding is between the HpaI and SspI sites. The sequence of fusion joints and extent of the deletions are given in Table I. The gray box at the 3′-end of pD680–853H and pD680–853 represents DNA encoding 13 additional amino acids. The ATPase activity of the proteins is given as µM ADP produced/µM MutS. MW, calculated molecular weight; G-T, ability to bind to G-T mismatched substrate DNA; S-S, extent of dimerization; S-L, extent of MutS-MutL interaction; ND, not determined.

**Table II**

| Plasmid        | Mutation frequency (× 10^-8) in strain | GM4244 (wild) | GM4271 (mutS) |
|----------------|----------------------------------------|---------------|---------------|
| pET3a         | 2.0                                    | 20            |               |
| pMQ395        | 1.4                                    | 5             |               |
| pMQ372        | 3.1                                    | 3             |               |
| pΔ1–311H      | 0.2                                    | 110           |               |
| pΔ1–311       | 0.3                                    | 37            |               |
| pΔ261–260H    | 0.6                                    | 126           |               |
| pΔ261–260     | 2.0                                    | 45            |               |
| pΔ261–556H    | 0.9                                    | 52            |               |
| pΔ261–556     | 2.7                                    | 83            |               |
| pD680–853H    | 0.6                                    | 76            |               |
| pD680–853     | 0.4                                    | 56            |               |

(Fig. 3B, vector). A similar result was obtained if both His6-tagged and non-tagged wild-type proteins were present in cell extracts (Fig. 3B, pMQ372). We assume the lower molecular weight band in these two figures is because of protein degradation. The distribution of wild-type and mutant proteins in eluates from the affinity column was similar if not identical for pΔ261–556 and pΔ1–260 encoded mutant proteins, indicating efficient heterodimer formation in vivo. Deletion of the N-terminal region of MutS encompassed by these two mutations (Fig. 1) therefore does not affect dimerization. For the pD680–853 and pΔ261–556 encoded proteins, however, the distributions of wild-type and mutant proteins were not the same (Fig. 3), indicating less efficient heterodimer formation. The basis for the altered distribution of the mutant proteins is unknown. None of the mutant non-His6-tagged MutS proteins bind to the affinity column (data not shown). The results suggest that amino acid residues in the C-terminal end of the protein are responsible for efficient dimerization.

**Fig. 2.** MutS protein binding to a 154-bp G-T heteroduplex substrate. A, construction of the 154-bp DNA fragment substrate. B, the three lanes for each mutant protein indicate (from left to right) that fractions eluting at 30, 55, and 80 mM imidazole from the affinity column were tested. The 0 lane represents heteroduplex DNA only. The figure shows that 50-fold excess of unlabeled competitor DNA was used. nt, nucleotide.

**MutS-MutL Interaction**—MutS protein bound to heteroduplex DNA is thought to recruit MutL protein into a complex (16). We have been unable to detect MutS-MutL complexes by co-immunoprecipitation.2 We have used, instead, a variation of the technique to detect dimerization in which His6-tagged MutL was co-expressed in vivo with non-His6-tagged MutS. Affinity purification of MutL followed by SDS-PAGE and immunoblotting with MutS antibody should allow for detection of MutS-MutL complexes (Fig. 4A). The control experiments showed that a band at the expected Mw was detected in cells expressing MutS (Fig. 4B, pMQ372) but was not detected in its

---

2 T-H. Wu and M. G. Marinus, unpublished data.
absence (Fig. 4B, vector). Reduced amounts of mutant MutS were detected in cells containing the C-terminal and midgene mutS deletions (Fig. 4B, pΔ680–853 and pΔ261–566). One mutation deleting the N-terminal end produced almost as much cross-reacting material as wild-type (Fig. 4B, pΔ1–260) whereas another resulted in a vastly excess of it (Fig. 4B, p/Δ26–260). We assume the latter result indicates a higher affinity of the mutant protein for MutL than wild type or binding of more MutS molecules to MutL than normal. The results above suggest that the MutS-MutL interaction region is at the C-terminal end of MutS, perhaps overlapping with the P-loop domain of the ATP binding site (Fig. 1).

**DISCUSSION**

The mutant proteins we have used all have approximately the same ATPase specific activity as the wild-type MutS protein, suggesting that gross structural alterations are not present. This is supported by the observation that the mutant proteins are able to dimerize and to interact with MutL unless specifically defective in one or both of these processes.

The mutS deletion mutations produce little, if any, residual mismatch repair in vivo, and they do not impart a dominant phenotype in a wild-type strain (Table II). The latter finding is surprising for those mutant proteins, such as pΔ1–311, that fail to bind DNA but still interact with downstream proteins. It should be noted that IPTG was not added to the cultures to induce the phage T7 RNA polymerase because the wild-type mutS plasmids complemented the mutator phenotype in its absence (Table II) and because the plasmid instability increased in its presence. Perhaps the ratio of mutant protein to wild type is low in vivo in uninduced cultures because of a lower than expected level of expression and/or poor solubility of the mutant proteins, thereby precluding sequestration of sufficient wild-type and mutant monomers into inactive dimers.

The results from Fig. 1 suggest that the N-terminal end of MutS is important for heteroduplex DNA binding. This is in agreement with the observation of Malkov et al. (17) that Phe-39 of the Thermus aquaticus (Phe-36 in E. coli) MutS cross-links to mismatched DNA. We also find, however, that specific DNA binding occurs at a low level in the mutant deleted for residues 26–260 but is not detectable if residues 261–556 are deleted (Fig. 2). In Thermus thermophilus, a proteolytic fragment containing residues 275–570 of MutS is able to bind double-stranded DNA (18), confirming the importance of this region for DNA binding.

The ability of MutS to dimerize appears to reside within the region containing residues 557–853 (Fig. 3). This agrees with the finding by Alani (19) that in yeast the C-terminal 114 amino acids of Msh2 are important for interaction with Msh6.
Taken together these findings suggest that the same region of MutS homologs is used for homo- and heterodimerization. Alani et al. (15) have also described a mutation (A859E) in yeast Msh2 located in the putative helix-turn-helix domain, which affected Msh2-Msh6 heterodimer formation. The equivalent residue in E. coli MutS is Ala-776 that is within the predicted interaction region (residues 557–853). Because none of the mutations we tested were completely defective in dimerization, it is highly probable that additional residues are involved.

The C-terminal region of MutS also seems to be important for MutS-MutL interaction, and the results in Fig. 4 are the only available mutational data addressing this question. Given that MutS-MutL interaction was detected in the presence of MutS, but not in its absence, and that the extent of the interaction could be varied by mutation (Fig. 4), we believe the assay detected bonafide interaction. However, given the high NaCl concentration (500 mM) used in these experiments, it is probable that the MutS-MutL complexes detected in these experiments are not bound to heteroduplex DNA. This suggests that a fraction of the total MutS protein in vivo may be loosely bound to MutL in the absence of DNA. Even when MutS is bound to DNA in vitro, its interaction with MutL is not readily detectable (20). This might explain why the MutS-MutL complexes are detected in column eluates only at low concentrations of imidazole (Fig. 4). None of the mutant proteins tested was totally deficient in MutS-MutL interaction, suggesting that amino acids 556–630, which include the P-loop motif, are required.

The methods we have used for protein interaction do not allow a quantitative measurement of the interaction affinities. We have, therefore, used a qualitative scale to express the results (Fig. 1).

The MutS family of proteins from prokaryotes and eu-

### Acknowledgments
We thank Dr. Kendall Knight for synthesis of oligonucleotides and for advice on the ATPase assays and Dr. R. Lahue for advice and suggestions in the early part of this project. We also thank Patricia Hess, who constructed some of the plasmids used here and helped devise protein expression conditions during a graduate student research rotation.

### REFERENCES
1. Modrich, P., and Lahue, R. (1996) *Annu. Rev. Biochem.* 65, 101–133
2. Horst, J. P., Wu, T. H., and Marinus, M. G. (1999) *Trends Microbiol.* 7, 29–36
3. Modrich, P. (1991) *Annu. Rev. Genet.* 25, 229–253
4. Kolodner, R. (1996) *Genes Dev.* 10, 1435–1442
5. Umar, A., and Kunkel, T. A. (1996) *Eur. J. Biochem.* 238, 297–307
6. Karran, P., and Bignami, M. (1994) *Bioessays* 16, 833–839
7. Su, S. S., and Modrich, P. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 83, 5057–5061
8. Parker, B. O., and Marinus, M. G. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 1730–1734
9. Allen, D. J., Makhov, A., Grilley, M., Taylor, J., Thresher, R., Modrich, P., and Griffith, J. B. (1997) *EMBO J.* 16, 4467–4476
10. Gradia, S., Acharya, S., and Fishel, R. (1997) *Cell* 91, 995–1005
11. Yamaguchi, M., Dao, V., and Modrich, P. (1998) *J. Biol. Chem.* 273, 9197–9201
12. Wu, T. H., and Marinus, M. G. (1994) *J. Bacteriol.* 176, 5393–5400
13. Weinstock, G. M., McEntee, K., and Lehman, I. R. (1981) *J. Biol. Chem.* 256, 8829–8834
14. Wu, T. H., Clarke, C. H., and Marinus, M. G. (1990) *Gene* (Amst.) 87, 1–5
15. Alani, E., Sokolosky, T., Studamire, B., Miret, J. J., and Lahue, R. S. (1997) *Mol. Cell. Biol.* 17, 2436–2447
16. Grilley, M., Welsh, K. M., Su, S. S., and Modrich, P. (1989) *J. Biol. Chem.* 264, 1000–1004
17. Malkov, V. A., Biwas, I., Camerini-Otero, R. D., and Hsieh, P. (1997) *J. Biol. Chem.* 272, 23811–23817
18. Tachiki, H., Kato, R., Masui, R., Hasegawa, K., Itakura, H., Fukuyama, K., and Kuramitsu, S. (1998) *Nucleic Acids Res.* 26, 4153–4159
19. Alani, E. (1996) *J. Mol. Biol.* 26, 5604–5615
20. Drotschmann, K., Arnothman, A., Fritz, H. J., and Marinus, M. G. (1998) *Nucleic Acids Res.* 26, 948–953
21. Eisen, J. A. (1998) *Nucleic Acids Res.* 26, 4291–4300