

## Mutational Analysis of the MutH Protein from *Escherichia coli*\*

Received for publication, August 30, 2000, and in revised form, November 30, 2000  
Published, JBC Papers in Press, December 21, 2000, DOI 10.1074/jbc.M007935200

Tamalette Loh, Kenan C. Murphy, and Martin G. Marinus‡

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

Site-directed mutagenesis was performed on several areas of MutH based on the similarity of MutH and PvuII structural models. The aims were to identify DNA-binding residues; to determine whether MutH has the same mechanism for DNA binding and catalysis as PvuII; and to localize the residues responsible for MutH stimulation by MutL. No DNA-binding residues were identified in the two flexible loop regions of MutH, although similar loops in PvuII are involved in DNA binding. Two histidines in MutH are in a similar position as two histidines (His-84 and His-85) in PvuII that signal for DNA binding and catalysis. These MutH histidines (His-112 and His-115) were changed to alanines, but the mutant proteins had wild-type activity both *in vivo* and *in vitro*. The results indicate that the MutH signal for DNA binding and catalysis remains unknown. Instead, a lysine residue (Lys-48) was found in the first flexible loop that functions in catalysis together with the three presumed catalytic amino acids (Asp-70, Glu-77, and Lys-79). Two deletion mutations (MutH $\Delta$ 224 and MutH $\Delta$ 214) in the C-terminal end of the protein, localized the MutL stimulation region to five amino acids (Ala-220, Leu-221, Leu-222, Ala-223, and Arg-224).

MutH is the endonuclease in the methyl-directed mismatch repair system (1), which, along with MutS and MutL (and bound ATP), initiates the repair process (2). MutS binds specifically to mismatched bases or deletion/insertion loops of one to four nucleotides (3–5). MutL appears to be a helper protein; it interacts with other proteins to increase their DNA binding or converts them to the active form (6). MutL appears to activate MutH for DNA binding and cutting the unmethylated strand of hemimethylated DNA based on endonuclease assays (7, 8). In *Escherichia coli*, mismatch repair takes place shortly after replication before the DNA becomes fully methylated (the newly synthesized daughter strand is unmethylated). MutH in the ternary complex binds to a hemimethylated d(GATC) site, recognizes the asymmetry of the unmethylated N6 position of the d(A) (9, 10), and cleaves 5' to the d(G) in the unmethylated strand leaving ligatable ends. If both strands are unmethylated, MutH is able to cleave both sites independently (dissociating after the first d(GATC) cut, with a second binding and cleavage event) leaving a four-base overhang (2). Although MutS is thought to bind heteroduplex DNA first, followed by

MutL and MutH, it is unclear what the binding order is for these proteins. It is also unclear in the ternary complex whether MutH acts as a monomer or (like restriction endonucleases) as a dimer. The purified protein with a calculated molecular mass of 25.5 kDa is a monomer in solution.

Recently, the structure of the MutH protein was solved (11). It was described as resembling a clamp with a large cleft dividing the protein into two subdomains. From the three structures analyzed, it was shown that the two subdomains adopt multiple conformations that correlate with movement in the C-terminal end of the protein. This suggests that the C-terminal-exposed helix is a possible site through which MutL stimulation occurs. It was postulated that DNA lies in the bottom of the cleft in close proximity to residues Asp-70, Glu-77, and Lys-79, which would constitute the active site for catalysis. The catalytic sequence of D(X)<sub>6–30</sub>(E/D)XK was proposed for MutH based on the close resemblance to the structure of the restriction endonuclease PvuII (with a root mean square deviation of 2.3 over 83 pairs of C $\alpha$  atoms) (11). Both enzymes share a common core motif with a number of other restriction enzymes (*Bam*HI, *Eco*RV, and *Eco*RI) (12, 13). These restriction enzymes are grouped according to their similarity in structure as well as function (*Pvu*II and *Eco*RV are in one group and *Bam*HI and *Eco*RI are in another). It was surprising that MutH is closer in structure to *Pvu*II than *Bam*HI, because *Pvu*II approaches DNA from the minor groove and makes contacts in the major groove by reaching around the DNA with flexible loops to produce blunt ends. *Bam*HI, however, approaches DNA from the major groove where the base-specific contacts are made to produce cuts with a four-base overhang. In addition, *Bam*HI has a recognition site with d(GATC) as the central bases and cuts 5' to the d(G). There is no structure available of MutH complexed with its cognate DNA, so it is not known from which direction MutH approaches the DNA and whether it acts as a monomer or dimer in the repair complex. Given the close similarity in structure between MutH and *Pvu*II, we altered amino acids in MutH based on the corresponding functional regions of *Pvu*II.

### EXPERIMENTAL PROCEDURES

*Strains, Plasmids, and Media*—KM54 ( $\Delta$ *mutH461::Cam*) was constructed by electroporation of strain KM22 ( $\Delta$  (*recC ptr recB recD*))::P<sub>lac</sub> *red Kan* (20) with a recombinant PCR<sup>1</sup> product as described by Murphy *et al.* (14), using the following oligonucleotides: *mutH*U2, ATCATCGA-GCTCCACCAGCTGCAAGAGAAACCATTT; *mutH*N2, TGCCGATCAA-CGTCTCATGCGGCCGCTTGGGACATGTCATGATACCTTGA; *mutH*C2, AATGGCAGAAATTCGAAAGCGGCCGCT TTTCTGATCCAGTA-GCCATCGCTTT; *mutH*D2, TCATCAGCATGCCCGAAGGATGGCGTT-CGACAAAAT. GM7586 was generated by P1 transduction of GM4244 (CC106 (15)) with a lysate grown on KM54 and selecting for chloram

\* This work was supported by Grant RPG-97-127-01-GMC from the American Cancer Society. 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. North, Worcester, MA 01655. Tel.: 508-856-3330; Fax: 508-856-5080; E-mail: martin.marinus@umassmed.edu.

<sup>1</sup> The abbreviations used are: PCR, polymerase chain reaction; kb, kilobase(s); Amp, ampicillin; Cam, chloramphenicol; Rif, rifampicin; PAGE, polyacrylamide gel electrophoresis; bp, base pair(s); PNK, T4 polynucleotide kinase.

phenicol resistance. Isogenic strains GM4244 and GM7586 (GM4244  $\Delta$ mutH461::Cam) were used for the *in vivo* screen. Strain GM3856 (*hsdR17 endA1 thi-1 spoT1 rfbO1 supE44 mutH471::Tn5*) was also used for the overexpression of wild-type MutH and the mutant proteins. Strain GM5862 (16) was used for the overexpression of wild-type MutL.

Plasmid pMQ402 is a pBAD18 (17) derivative containing the *mutH* gene with an N-terminal polyhistidine sequence from the plasmid TX417 (18). The ~1-kb *mutH* fragment was obtained by digestion of pTX417 with the restriction enzymes *Xba*I and *Hind*III. The fragment also contains several hundred base pairs of chromosomal DNA distal to the *mutH* gene. pTX417 was kindly provided by M. Winkler (University of Texas Medical School, Houston TX). Plasmid pMQ393 contains the wild-type His-tagged *mutL* gene in a pACYC184 backbone (6).

The Luria-Bertani (LB) medium was made with 10 g of tryptone, 5 g of yeast extract, 5 g of NaCl, and 1 ml of 1 M sodium hydroxide per liter of water. The YT medium and TAE buffer were made according to Sambrook *et al.* (19). Ampicillin (Amp), chloramphenicol (Cam), and rifampicin (Rif) were added to media at 100, 30, and 100  $\mu$ g/ml, respectively.

**Site-directed Mutagenesis**—Two-stage PCR-mediated mutagenesis was carried out essentially as described by Murphy (14). After each PCR reaction, the products were electrophoresed on a 1% agarose gel and purified using the QIAquick gel extraction kit (Qiagen). The full-length product was cut with *Bgl*II and *Hind*III for 1 h at 37 °C and electrophoresed on a 1% agarose gel and purified before ligation to the similarly cut pMQ402 backbone. The ligation was carried out overnight at 14 °C, and the mixture was transformed into CC106 (14) for recovery. Transformant colonies were streaked out several times, and a single colony was inoculated in L ampicillin broth for overnight growth. Plasmid DNA was isolated using QIAprep Spin Miniprep (Qiagen) and sequenced to identify the mutations. When secondary mutations were present, the sequence containing the correct mutation was removed by restriction enzyme digestion and religated to a similarly cut pMQ402 backbone. The *mutH* gene of the resulting plasmid was then sequenced to confirm the mutational change. The two outside flanking primers for PCR and sequencing were MM247 (5'-CATCACAGCAGCGGCTGG-3') (Operon Technologies) for the forward reactions and MM248 (5'-CAGACCGCTTCTGCG-3') (Operon Technologies) for the reverse reactions.

Amber stop codons were placed into the *mutH* sequence at codons 215 and 225 using the Sculpture mutagenesis kit (Amersham Pharmacia Biotech). In addition, the plasmid containing the 215 mutation also incorporated an A to C change within codon 213, changing the amino acid at this position from leucine to valine. The MutH $\Delta$ 224 mutant has the last 5 amino acids of the C terminus removed (HFLIQ) and the MutH $\Delta$ 214 mutant has the last 15 amino acids of the C terminus removed (KNFTSALLARHFLIQ).

**DNA Sequencing**—All DNA sequencing was carried out by the DNA Sequencing Facility at Iowa State University. Both strands of wild-type MutH and the promoter region were sequenced. We used the *mutH* sequence in GenBank<sup>®</sup>, accession number U16361. The site-directed PCR mutants were identified by sequencing with the outside flanking primers used in the PCR reactions. Both strands and the promoter region of all mutant plasmids were sequenced.

**Preparation of Histidine-tagged MutH**—The wild-type and all mutant proteins were purified in the same manner. For each strain, a 20-ml LB-Amp culture was grown overnight from a single colony. One liter of LB-Amp medium was inoculated with the 20-ml overnight culture and allowed to grow with shaking at 37 °C to an  $A_{600}$  of 1.0. Arabinose (Difco) was added to a final concentration of 0.2%, and the culture was induced for 2 h at 37 °C. Cells were harvested by centrifugation for 20 min at 5000 rpm, washed with water, and stored frozen (-80 °C). The pellet was thawed and resuspended with 4 ml of reconstitution buffer (20 mM Hepes (pH 7.4), 300 mM NaCl, and 1 mM phenylmethylsulfonyl fluoride), and the cells were lysed using a French pressure cell. The extract was then sonicated (20 pulses from a Tekmar sonicator) and centrifuged for 30 min at 15,000 rpm to remove cell debris. The supernatant was applied to a 4-ml Fast Flow Chelating Sepharose (Amersham Pharmacia Biotech) column charged with 100 mM NiCl<sub>2</sub> and equilibrated with the reconstitution buffer. Step fractions of 100, 150, 250, and 400 mM imidazole (10 ml each) were used to elute the MutH protein. The 250 mM fraction contained the MutH protein and was dialyzed against two changes of 20 mM Hepes (pH 7.4), 300 mM NaCl, and 0.1 mM EDTA. The MutH concentration was determined by measuring the  $A_{280}$  (1 A = 0.67 mg/ml). MutH was at least 95% pure as determined by a SDS-PAGE Coomassie Brilliant Blue-stained gel.

**Preparation of Histidine-tagged MutL**—The His-tagged MutL pro-

tein was purified as described previously (6), except that the transformation mixture of GM5862 with pMQ393 was incubated for 90 min after heat shock and then added to 50 ml of YT-Cam medium for overnight incubation at 37 °C. The step fractions used for elution from the Sepharose column were one 100 mM, two 150 mM, and one 400 mM Imidazole (10 ml each). The 400 mM fraction contained the MutL protein, which was dialyzed against two changes of 20 mM Hepes (pH 7.4), 300 mM NaCl, and 0.1 mM EDTA. The concentration of MutL was determined by measuring the  $A_{280}$  (1 A = 1.24 mg/ml). The protein was at least 90% pure as determined by a Coomassie Brilliant Blue-stained SDS-PAGE gel. Native MutL was graciously provided by Dr. Paul Modrich and Dr. Claudia Spampinato (Duke University).

**In Vivo Screen**—The *in vivo* screen is based on the mutator phenotype, which occurs when the cells lack one or more of the Mut proteins (21). The bacteria accumulate mutations at a greater frequency than wild-type, and this can be monitored by selection on plates with rifampicin. The vector plasmid (pBAD18) and the wild-type MutH plasmid (pMQ402) were transformed into strains GM4244 and GM7586 and used as controls. The mutant plasmids were transformed into GM7586 for complementation testing and GM4244 for dominant-negative screening. A single colony was inoculated into 1 ml of L-Amp broth, grown overnight at 37 °C, diluted 10<sup>-6</sup>, and 50  $\mu$ l was spread on L plates with and without ampicillin. The rifampicin plates were spread with 200  $\mu$ l of undiluted wild-type culture (GM4244), 100  $\mu$ l of undiluted mutant culture (GM7586) and 50  $\mu$ l of undiluted *mutH* plasmid mutant cultures. Two cultures (from two separate colonies) were grown for each test and each culture was plated twice. The plates were incubated at 37 °C overnight, and colonies were counted the next day. The L plates were used to monitor loss of plasmids due to high cellular amounts of MutH. No arabinose was included in the plates, because this leads to loss of viability. The plasmid produces enough MutH protein to complement the *mutH* strain without adding arabinose to the plates.

**Endonuclease Assay**—Bacteriophage MR1 is a derivative of phage  $\phi$ IR229 that contains only one d(GATC) site (22). The covalently closed replicative form (RF) of this phage was purified as previously described from chloramphenicol-treated infected GM2807 (*dam*-16::Kan) cells (9). The MR1 phage was provided by Dr. R. Lahue (University of Nebraska). Enzyme activity (amounts are indicated in the figures) was tested at 37 °C in a 10- $\mu$ l reaction containing 20 mM Tris-HCl (pH 7.7), 5 mM MgCl<sub>2</sub>, and 25 fmol of MR1 for 1 h. MutL and ATP were added at 71 pmol and 1.25 mM, respectively, in the MutH endonuclease stimulation assays. The assays were carried out with native MutL or His-tagged MutL. The reaction was stopped with 5  $\mu$ l of a 50% glycerol solution containing 50 mM EDTA, 1% SDS, 0.05% bromophenol blue, and 0.05% Xylene Cyanol. The reactions were electrophoresed on a 1% agarose gel at 5 V/cm for 3 h. The gels were stained with *Vistra* Green (Amersham Pharmacia Biotech) following the manufacturer's instructions (1:10,000 dilution) for 1 h with shaking. The gels were then scanned on a PhosphorImager (Storm, Molecular Dynamics), and the amount of nicking was quantitated using the ImageQuant (Molecular Dynamics) software. Any nicked substrate detected in the negative control was subtracted from the amount seen with the proteins added.

**Preparation of Labeled Homoduplex**—A 36-base pair (bp) DNA oligonucleotide, MM40, 5'-GCATACGGAAGTAAAGTGC GGATCATCTCTAGCCA-3' (Operon Technologies) containing a single, centrally located d(GATC) site, was labeled using T4 polynucleotide kinase (PNK) (New England Biolabs). A concentration, 2  $\mu$ M, of the oligonucleotide was used in a 12- $\mu$ l reaction with 1.2  $\mu$ l of PNK buffer, 1  $\mu$ l of PNK enzyme, 1  $\mu$ l of [ $\gamma$ -<sup>32</sup>P]ATP (Amersham Pharmacia Biotech), and water to a final concentration of 0.5  $\mu$ M oligonucleotide. The reaction was incubated for 1 h at 37 °C and then 10 min at 75 °C. 2  $\mu$ M cold complementary oligonucleotide (containing the N6-methyl adenine in the d(GATC) site) was added with water to a final volume of 100  $\mu$ l. This was incubated for 5 min at 95 °C, 30 min at 37 °C, and 30 min at room temperature. The unincorporated and single-strand DNA were removed using a MicroSpin G-25 column (Amersham Pharmacia Biotech) and the labeled homoduplex DNA resuspended in a final volume of 100  $\mu$ l of water. A 5 $\times$  reaction mixture with 50  $\mu$ l of the labeled homoduplex and 50  $\mu$ l of 10 $\times$  AR buffer (200 mM Tris-HCl (pH 7.6), 50 mM MgCl<sub>2</sub>, 1 mM dithiothreitol, and 0.1 mM EDTA) was made for the band shift assay.

**Band Shift Assay**—MutH binding to DNA was assayed using the labeled homoduplex prepared as described previously (4). In a 5- $\mu$ l reaction volume, there was 22.5 fmol of the labeled homoduplex (1  $\mu$ l of the 5 $\times$  reaction mixture) and MutH, which was added in varying concentrations (water makes up the remainder of the reaction volume). MutL and ATP were added at 9 pmol and 1.25 mM, respectively, in the stimulation assays. The mixture was incubated for 30 min on ice, followed by the addition of 1.6  $\mu$ l of a 50% sucrose solution and loaded

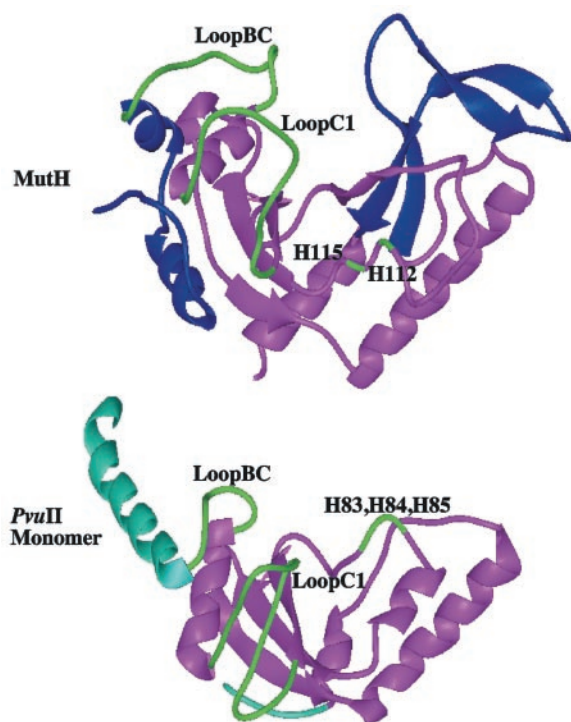


FIG. 1. Comparison of MutH and *PvuII* structures for the construction of MutH mutants. The ribbon diagram of MutH was made from the B monomer with the coordinates submitted by W. Yang (Protein Data Bank code 2AZO) using MIDAS. The ribbon diagram of *PvuII* was made from the B monomer with the coordinates submitted by X. Cheng (Protein Data Bank code 1PVI) using MIDAS. Six N termini residues from the *PvuII* monomer have been deleted for clarity. The magenta-colored areas of each protein represent structurally similar areas. The two loops and histidines highlighted in green for each model show similar areas that were targeted for mutagenesis in MutH (residues 41–49, residues 59–69, His-112, and His-115).

on a prerun (10 min) 6% native polyacrylamide gel in TAE buffer (pH 7.5) and electrophoresed at 50 mA for 2.5 h. The gel was then laid on a PhosphorImager screen overnight for scanning on the Storm Imager (Molecular Dynamics). The free DNA and shifted band intensities were quantitated to determine binding ability using ImageQuANT software. Apparent dissociation constants ( $K_d$ ) were determined by fitting the quantitated experimental data to the equation  $1/r = 1 + K_d/[C]_{total}$  ( $f = 1-r$ , where  $f$  is the degree of dissociation) with the program KaleidaGraph (Synergy Software). The degree of binding or association is  $r$  and  $[C]_{total}$  is the protein concentration in the reaction.

## RESULTS

**Construction of the Mutant Plasmids**—After a comparison of the MutH and *PvuII* structural models, areas in MutH were found that resembled functional areas of *PvuII* (Fig. 1). There were three areas targeted for mutagenesis in MutH. The first area was the flexible loop between the second and third  $\alpha$  helices (loop BC). The second was the C1 loop between the third  $\alpha$  helix and the first  $\beta$  sheet (where the catalytic residues start). These areas target flexible loops of *PvuII* that make contact with DNA backbone phosphates and recognition bases (23, 24). MutH could make the same contacts if it approaches DNA from the minor groove. Loop BC residues Pro-41 to Gly-49 and loop C1 residues Leu-59 to Gln-69 (except for Ala-61 and Ala-63) in MutH were changed to alanines by site-directed mutagenesis. The third area of *PvuII* contains conserved histidines (His-84 and His-85) that signal binding of the DNA to allow catalysis to take place and serve as protein-protein contacts (25). These histidines are also present in MutH (His-112 and His-115) and may serve as the same signal for MutH. Histidine112 and histidine 115 were also changed to alanines to test this possibility.

Based on the MutH crystal structure, it was suggested that the C-terminal tail region is contacted by MutL to stimulate the MutH endonuclease activity (7, 11). To test this possibility, two deletion mutants were made with the creation of amber stop codons. MutH  $\Delta 214$  and MutH  $\Delta 224$  (the number of the last residue present in the protein) were made to narrow down the area where MutL might stimulate MutH activity.

**In Vivo Screening of the Mutant Plasmids**—After checking that there were no secondary mutations in each of the mutant MutH plasmids, they were transformed into strains GM4244 (wild-type) and GM7586 ( $\Delta mutH::Cam$ ). The null mutation in *mutH* was constructed, because the commonly used *mutH471::Tn5* mutation is not completely defective in our assays.<sup>2</sup>

The mutant MutH plasmids in GM4244 were tested for a dominant-negative phenotype (mutator phenotype in a wild-type strain) by the same method as the *in vivo* screen described under “Experimental Procedures.” None of the mutant plasmids gave a dominant-negative phenotype (data not shown).

For the complementation testing in strain GM7586, the controls (GM7586 versus pMQ402 in GM7586) show that the wild-type MutH protein causes a 200-fold reduction in the frequency to rifampicin resistance, indicating efficient complementation of the chromosomal  $\Delta mutH::Cam$  allele (Table I). The vector plasmid (pBAD18) causes a small reduction in the mutation frequency and doesn’t interfere with the complementation testing. No alterations in mutation frequency were seen in GM4244 with the control plasmids (pBAD18 and pMQ402). Very few of the MutH site-directed mutants showed a mutator phenotype. Plasmid  $\Delta 214$ -containing cells show a strong mutator phenotype along with mutants K48A and G49A. These three were chosen for further biochemical studies. A few others were also chosen to validate the screen and for the following specific reasons. The D47A mutant was used because of the Asp-34 analogous position in the *PvuII* structure that recognizes the d(A) in the recognition sequence for *PvuII* (23). The H115A and H112A mutants were used because of the signaling properties of the *PvuII* histidines (His-84 and His-85) described earlier. The  $\Delta 224$  mutant was used to test the hypothesis that the C-terminal end is where MutL stimulation of MutH activity occurs.

The steady-state level of MutH produced by wild-type and mutant constructs was determined by Western blotting (New England BioLabs). All of the mutant plasmids produced comparable amounts of MutH to the wild-type plasmid in strain GM3856 (data not shown).

**Endonuclease Assay**—This assay tests for the endonuclease function of the MutH protein. The product of the reaction is a nick in the MR1 covalently closed circular molecule, which has only one d(GATC) site. In the absence of MutL, increasing amounts of the wild-type MutH protein were added to the duplex until complete nicking was achieved (Fig. 2A and Fig. 3A). The addition of 2.5–3.0 pmol of wild-type protein resulted in complete nicking of the substrate. The results of the endonuclease assay for the remaining three mutants (K48A, G49A, and  $\Delta 214$ ) are shown below the wild-type in Fig. 2A. K48A has no detectable endonuclease activity and G49A activity is decreased  $\sim 30$ -fold (Fig. 3A). D47A, H112A, H115A, and  $\Delta 224$  MutH proteins gave the same result as wild-type and were not tested further (data not shown). These results were in agreement with the *in vivo* screen, because the mutation frequencies in cells with these proteins (D47A, H112A, and H115A) were similar to the mutation frequency of the wild-type MutH protein in strain GM7586 ( $\Delta mutH::Cam$ ). A circular dichroism

<sup>2</sup> T. Loh, K. C. Murphy, and M. G. Marinus, unpublished data.

TABLE I  
In vivo complementation screen for mutant MutH proteins

| Strain        | Colonies on media containing |                         | Mutation frequency <sup>c</sup> |
|---------------|------------------------------|-------------------------|---------------------------------|
|               | Ampicillin <sup>a</sup>      | Rifampicin <sup>b</sup> |                                 |
| GM4244        | 280                          | 42                      | 4 (±1)                          |
| GM7586        | 235                          | 1559                    | 675 (±121)                      |
| pBAD18/GM4244 | 261                          | 16                      | 1 (±1)                          |
| pBAD18/GM7586 | 362                          | 1447                    | 546 (±52)                       |
| pMQ402/GM4244 | 108                          | 0                       | 0                               |
| pMQ402/GM7586 | 232                          | 41                      | 5 (±1)                          |
| P41A          | 215                          | 24                      | 8 (±0)                          |
| E42A          | 203                          | 10                      | 3 (±0)                          |
| N43A          | 217                          | 14                      | 4 (±1)                          |
| L44A          | 239                          | 32                      | 8 (±0)                          |
| K45A          | 213                          | 9                       | 3 (±1)                          |
| R46A          | 223                          | 6                       | 2 (±1)                          |
| D47A          | 244                          | 9                       | 2 (±1)                          |
| K48A          | 242                          | 2178                    | 560 (±47)                       |
| G49A          | 269                          | 1502                    | 380 (±95)                       |
| L59A          | 303                          | 20                      | 5 (±1)                          |
| G60A          | 246                          | 15                      | 4 (±1)                          |
| S62A          | 347                          | 5                       | 1 (±1)                          |
| G64A          | 266                          | 18                      | 5 (±1)                          |
| S65A          | 214                          | 14                      | 4 (±1)                          |
| K66A          | 214                          | 23                      | 6 (±1)                          |
| P67A          | 252                          | 27                      | 7 (±1)                          |
| E68A          | 280                          | 18                      | 4 (±1)                          |
| Q69A          | 239                          | 16                      | 3 (±0)                          |
| H112A         | 171                          | 11                      | 2 (±0)                          |
| H115A         | 207                          | 21                      | 10 (±1)                         |
| Δ214          | 199                          | 2234                    | 1123 (±38)                      |
| Δ224          | 175                          | 6                       | 3 (±1)                          |

<sup>a</sup> The number of colonies in this column are from a 10<sup>-6</sup> dilution of the 1-ml overnight culture. These numbers are used to calculate the total number of cells present in the culture: colony number × (volume of dilution/volume plated) × 10<sup>6</sup>.

<sup>b</sup> The number of colonies in this column are from a percentage plated from the undiluted culture as described under "Experimental Procedures." The total number of rifampicin-resistant cells in the culture are then calculated: colony number × (volume of culture/volume plated).

<sup>c</sup> The numbers in this column represent the total rifampicin-resistant cells calculated per total number of cells calculated, in the overnight culture (rifampicin total/10<sup>8</sup> cells). The numbers in the columns represent the average of three independent experiments.

experiment was performed on the K48A mutant protein because of the lack of activity without MutL present in the endonuclease assay. The results showed that the K48A mutant protein was folded similarly to the wild-type MutH protein (data not shown). A circular dichroism experiment was also performed on the Δ214 mutant protein, because the entire last helix had been removed and may have resulted in large structural changes for the protein. The results showed that the Δ214 mutant protein was folded similarly to the wild-type MutH protein with the curve exhibiting more of the expected β sheet-like properties due to the loss of the large α helix in the protein (data not shown).

In the presence of MutL and ATP, 1000-fold less MutH is needed to achieve the same results as without MutL (Figs. 2B and 3B). The cleavage of mutants K48A and G49A are shown below the wild-type in Fig. 2B. Although G49A is partially stimulated by MutL (it is increased 10-fold from the unstimulated activity), its activity is reduced about 200-fold compared with the stimulated wild-type protein (Fig. 3B). K48A shows the most dramatic increase in MutL-stimulated nicking and is reduced only 3-fold relative to the wild-type protein (Fig. 3C). MutL is unable to increase the activity of the G49A mutant as much as that for K48A. There was no MutL stimulation seen for Δ214 in this assay, and its activity is decreased about 15-fold from the unstimulated wild-type MutH protein (Fig. 3D). It would appear then to have a 15,000-fold decrease in activity compared with the stimulated MutH protein. Because

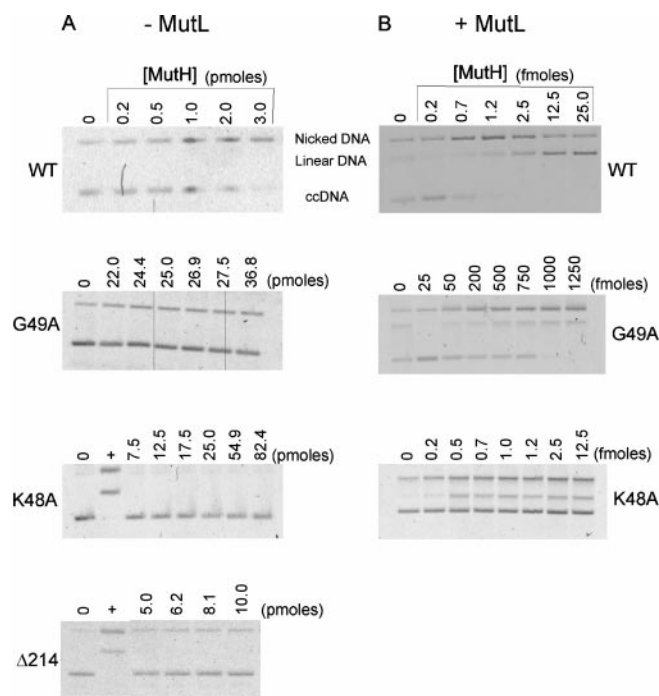


FIG. 2. Endonuclease activity of the wild-type and mutant MutH proteins. The endonuclease assays were performed as described under "Experimental Procedures." Increasing amounts of MutH (indicated at the top of each lane) were incubated with the MR1 DNA substrate. The DNA products were separated in a 1% agarose gel before staining and scanning for quantitation. A, all reactions are without MutL and contained 25 fmol of homoduplex DNA substrate. The 0 lane represents unreacted homoduplex DNA substrate. +, control for mutants K48A and Δ214 containing 2.5 fmol of wild-type MutH protein in the reaction. B, all reactions contain 71 pmol of MutL, 1.25 mM ATP, and 25 fmol of homoduplex DNA substrate. The 0 lane represents unreacted homoduplex DNA. ccDNA is the closed circular substrate used in the reaction.

the reduction in endonuclease activity of the mutant proteins could be due to decreased binding ability, binding assays were done as described below.

**Band Shift Assay**—In this assay the specific binding of the MutH protein to DNA is tested. The substrate is a linear 36-mer duplex of DNA with one d(GATC) site. The MutH binding is specific, because it is not seen with the same linear substrate that has a d(GGTC) site in place of the d(GATC) site (data not shown). Without MutL, the wild-type protein gives a complete band shift with the appearance of a transient band between the free DNA and bound DNA (Fig. 4A). Both mutant proteins, K48A and G49A, gave results similar to the wild-type (data not shown). The Δ214 mutant protein appeared to initiate a binding-like activity but failed to achieve a final shifted position (Fig. 4A).

The band shift assay for MutH in the presence of MutL was complicated, because MutL shows nonspecific DNA binding at concentrations above 9 pmol to approximately the same position in the gel as the MutH band. The amount of MutL, therefore, was lowered until it no longer shifted the DNA substrate. The assay with MutL and wild-type MutH caused the band shift to start a little sooner (compare at 38 pmol), but it finished at the same concentration as without MutL (Fig. 4B). The K48A mutant was like wild-type and the G49A mutant bound a little tighter than wild-type. The Δ214 mutant was not stimulated by MutL and gave the same result as without the added protein (not shown).

From these experiments, the apparent dissociation constants ( $K_d$ ) for the proteins and the substrate were calculated (Fig. 5). The binding affinity for the wild-type MutH protein increased

FIG. 3. Comparison of wild-type and mutant MutH endonuclease activity.

These graphs represent the quantitation of the data from the endonuclease assays. A, this graph is the activity of the wild-type and mutant MutH proteins without MutL added. The wild-type MutH protein (■), the G49A mutant (◆), and the K48A mutant (●) are shown together. B, the activity of the wild-type and G49A mutant MutH proteins with 9 pmol of MutL added in the reaction. The wild-type MutH protein (■) and G49A mutant (●) are shown. C, the activity of the wild-type and K48A mutant MutH proteins with 9 pmol of MutL added in the reaction. The wild-type MutH protein (■), and K48A mutant (●) are shown together. D,  $\Delta 214$  mutant. The  $\Delta 214$  mutant with 9 pmol of MutL added to the reaction (■) and the  $\Delta 214$  mutant without MutL in the reaction (●) are plotted. This shows the lack of stimulation by MutL for this mutant. Each point on all graphs represents the average of at least three independent measurements.

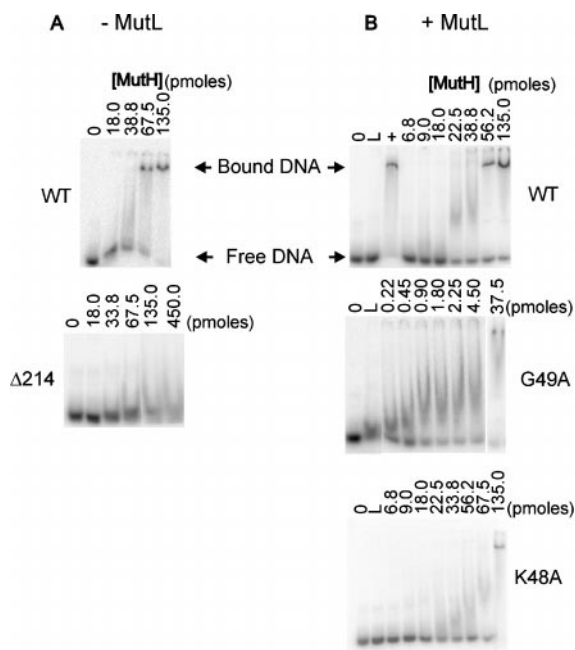
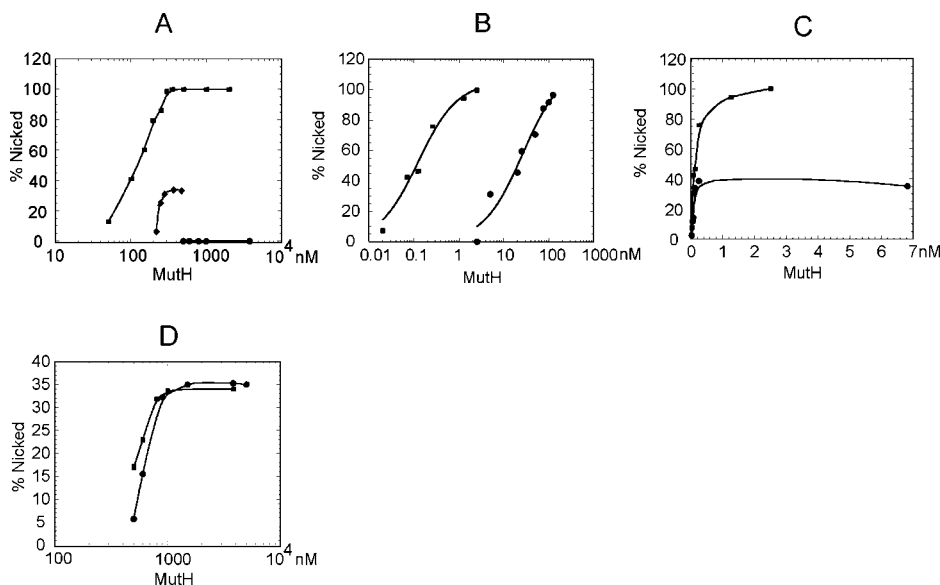


FIG. 4. Binding abilities of the wild-type and mutant MutH proteins. The binding ability of the wild-type and mutant MutH proteins were measured by a band shift assay using a  $^{32}\text{P}$ -labeled oligonucleotide as described under "Experimental Procedures." The DNA was incubated with increasing amounts of the MutH protein (indicated above each lane) in the reaction and separated by a 6% PAGE gel followed by scanning (Storm) for quantitation. A, all reactions contain 22 fmol of homoduplex DNA, 1.25 mM ATP, and do not contain MutL. The 0 lanes are without MutH and represent 22 fmol of unreacted homoduplex DNA. B, all reactions contain 22 fmol of homoduplex DNA, 1.25 mM ATP and 9 pmol MutL (except the + lane). The L lanes do not contain MutH. The + control lane contains 135.0 pmol of MutH, 1.25 mM ATP, and 22 fmol of homoduplex DNA.

at least 16-fold when the MutL protein was present (Fig. 5B). K48A was as proficient as the wild-type protein in binding ability, but the activity in the endonuclease assay was reduced, so it has a catalytic defect that MutL is able to overcome. G49A bound 2- to 4-fold tighter than wild-type, which may or may not be a significant contribution to the overall ability of the protein to function. Because  $\Delta 214$  did not achieve the full band shift position, the apparent dissociation constant was not calculated. Although MutL has an effect on MutH binding to the DNA

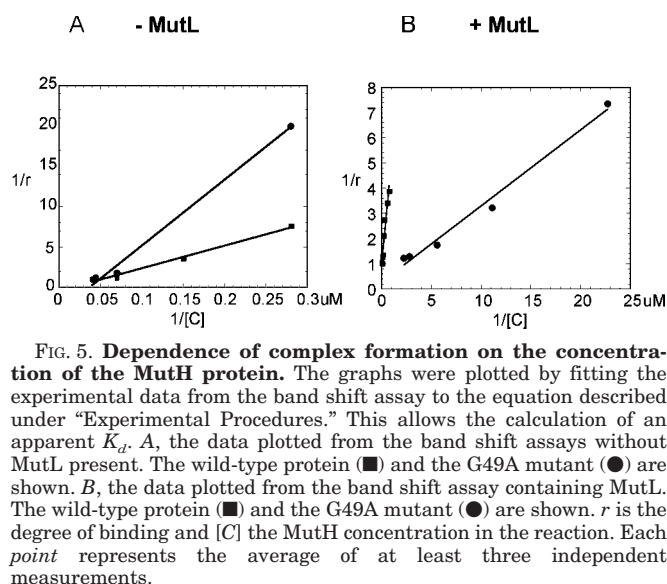


FIG. 5. Dependence of complex formation on the concentration of the MutH protein. The graphs were plotted by fitting the experimental data from the band shift assay to the equation described under "Experimental Procedures." This allows the calculation of an apparent  $K_d$ . A, the data plotted from the band shift assays without MutL present. The wild-type protein (■) and the G49A mutant (●) are shown. B, the data plotted from the band shift assay containing MutL. The wild-type protein (■) and the G49A mutant (●) are shown.  $r$  is the degree of binding and  $[C]$  the MutH concentration in the reaction. Each point represents the average of at least three independent measurements.

(about 16-fold), the greatest effect is on catalysis (1000-fold stimulation). The apparent  $K_d$  values of the wild-type and mutant proteins and their substrate are given in Table II.

## DISCUSSION

Methyl-directed mismatch repair is initiated by a mismatch in the DNA followed by formation of a ternary complex containing the MutS (with bound ATP), MutL, and MutH proteins. A defect in the ability to function correctly of any one of these three proteins causes a mutator phenotype. In this paper, that phenotype was used as a screen to identify defective MutH proteins after site-directed mutagenesis. The *in vivo* results indicated that only three of the site-directed mutations, K48A, G49A, and  $\Delta 214$ , exhibited a mutant phenotype in *mutH E. coli*. Subsequent biochemical assays confirm a defect in MutH activity in these mutant proteins. Four site-directed mutants having a wild-type mutation frequency to rifampicin resistance (D47A, H112A, H115A, and  $\Delta 224$ ) displayed normal MutH activity *in vitro*. These results indicate that the *in vivo* screen was sensitive enough to detect defective MutH proteins. It takes very few copies of the protein (estimated at 34 monomers per cell) to achieve full repair activity (26). The copies of protein from the plasmid are probably much higher than those normally in the cell. Under these conditions an

TABLE II  
Apparent equilibrium dissociation constants for wild-type  
and mutant MutH proteins

| MutL |      | $K_d$   |
|------|------|---------|
|      |      | $\mu M$ |
| +    | WT   | 4.2     |
| +    | K48A | 9.8     |
| +    | G49A | 1.0     |
| -    | WT   | 68.2    |
| -    | K48A | 68.7    |
| -    | G49A | 27.4    |

impaired protein might do much better in the *in vivo* screen than in the *in vitro* testing. Therefore an estimated 10-fold reduction in activity of MutH would probably not have been detected.

The qualitative results between the biochemical and *in vivo* assays agree fairly well. Mutant  $\Delta 214$  had the highest mutation frequency of the mutants tested in the *in vivo* screen and had the highest reduction in endonuclease activity in the presence of MutL. Mutants G49A and K48A somewhat follow this pattern with G49A having the lower mutation frequency and achieving complete nicking in the endonuclease assay (with increased protein amounts).

Although the stimulation of MutH activity by MutL has been shown at the endonuclease level (7), the ability of MutL to stimulate the DNA binding ability of MutH has not been demonstrated. The amount of MutL (9 pmol) in the band shift assay appears to be adequate for binding all of the MutH (0.1 pmol) monomers present. Yet the amounts of MutH needed to achieve a full band shift remain the same (Fig. 4). One conclusion from the data is that the MutL contribution to MutH DNA binding is marginal. The differences in the  $K_d$  values (Table II) were about 16-fold. This is significant, but it adds very little compared with the 1000-fold difference MutL makes in catalysis. MutL's presence makes it a little easier for MutH to start binding to the DNA, but overall binding isn't affected that much.

The mutational changes made in MutH were based upon a comparison with the *PvuII* enzyme where detailed structure-function data are available. Mutational analysis of the amino acids in the flexible arms would show whether or not MutH has any functional similarity to *PvuII* (Fig. 1). In *PvuII* there are a few residues present that are involved in DNA binding or the recognition of the binding site, such as an aspartate (Asp-34) that would be comparable to Asp-47 in MutH (23). No one residue in the flexible arms of MutH was found to make an important base recognition DNA contact so that, when it is disrupted, a mutator phenotype results. Minor contacts in the DNA such as phosphate binding residues may not result in a mutator phenotype and therefore may not be seen through this *in vivo* screen. The amino acids chosen for this study are shown in the structure of MutH (Fig. 6).

The conserved histidines found in both *PvuII* (His-84 and His-85) and MutH (His-112 and His-115) do not appear to have the same function. Changing the histidines (His-84 and His-85) in *PvuII* results in a loss of catalysis but not binding (25). Their movement upon DNA binding brings the catalytic residues into proper alignment for catalysis. When the DNA is methylated, the excess movement of these histidines ensures the catalytic residues cannot function. No loss of binding or catalysis resulted when the histidines were substituted with alanines in MutH. Therefore, from this work, the signal (for DNA binding and state of methylation) for MutH is unknown. The DNA binding domain for MutH is clearly not similar to the one in *PvuII*.

A lysine (Lys-48) was found in MutH that functions in catalysis. The lysine to alanine mutation (K48A) in MutH resulted

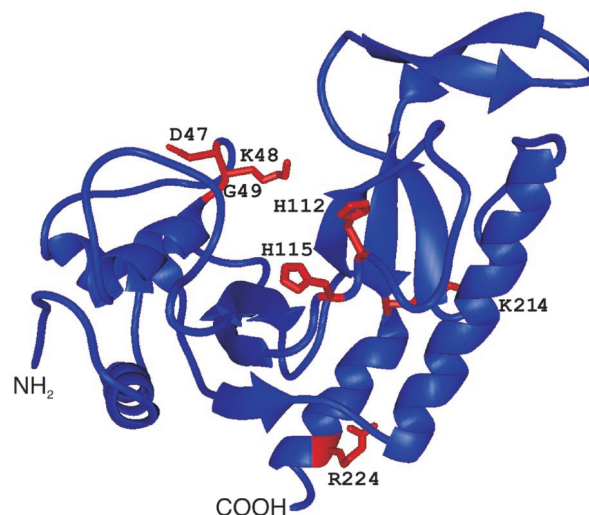


FIG. 6. Location of the mutated residues in the MutH structure. The ribbon diagram of MutH was made from the B monomer with the coordinates submitted by W. Yang (Protein Data Bank code 2AZO) using MIDAS. The N and C termini are labeled. The residues tested from the site-directed mutagenesis (Asp-47, Lys-48, Gly-49, His-112, His-115) are displayed in red. Also, the last residue remaining from each of the deletion mutants is shown in red (Lys-214 and Arg-224).

in a protein devoid of activity in the absence of MutL. MutL may have a lysine or other residue that is able to give the ternary complex partial function. The function of lysine 48 in MutH is unclear. It is positioned within the BC loop  $\sim 13.9$  Å from the lysine (Lys-79) presumed to be the catalytic amino acid. Upon the subsequent movement of MutH associated with DNA binding and activation of the protein, lysine 48 could be positioned within the catalytic core of the protein at the time of catalysis. This lysine may coordinate a critical water molecule for catalysis, be structurally important for the placement of the catalytic amino acids, be important for the placement of the metal binding residue of the protein, or may be the catalytic lysine in catalysis. The last possibility is very remote and a crystal structure with bound DNA and  $Mg^{2+}$  would be needed to confirm one of these hypotheses.

The G49A mutant may have defects in catalysis because of its proximity to lysine 48. The extra methyl group may cause slight structural changes in the lysine 48 side chain that slows catalysis. When MutL is present, the mutant is able to achieve full nicking of the substrate with larger amounts of protein than the wild-type MutH. The DNA binding of this glycine mutant is slightly tighter than the wild-type. The methyl group of G49A may be able to reach the phosphates of the DNA. If lysine 48 next to it is working in catalysis, then the DNA may be very close to these residues. This mutational change interferes with the proper functioning of the protein but is not informative about the mechanism of action.

Two C-terminal deletion mutations ( $\Delta 214$  and  $\Delta 224$ ) were made to help identify the residues that contact and cause the stimulation of MutH by MutL. This idea is based on predictions from the structural model of MutH (11). The  $\Delta 224$  functions similarly to the wild-type protein *in vivo* as well as *in vitro*, indicating that residues 225–229 at the C-terminal end of the protein are not required for MutL stimulation. The mutation frequency of the  $\Delta 214$  mutant is elevated *in vivo* and shows decreased cleavage activity with some initial DNA binding capabilities *in vitro*. However, it is not able to be stimulated by MutL. This places the residues between 215 and 224 as being involved in activation by MutL. It has been shown that a deletion of 10 amino acids in the C-terminal region of MutH ( $\Delta 219$ ) prevents MutL stimulation (7). That leaves the amino

acids 220 through 224 (sequence: ALLAR) as contacting MutL and stimulating MutH activity. Comparing the amino acid sequence of *E. coli* and *Hemophilus influenzae* MutH, there is one conserved amino acid: the leucine at position 222. Site-directed mutagenesis at each of the five positions (220–224) could reveal the most important residues for stimulation.

*Acknowledgment*—We thank Dr. Te Wu for helpful advice and discussions.

## REFERENCES

1. Modrich, P., and Lahue, R. S. (1996) *Annu. Rev. Biochem.* **65**, 101–133
2. Au, K. G., Welsh, K., and Modrich, P. (1992) *J. Biol. Chem.* **267**, 12142–12148
3. Su, S. S., and Modrich, P. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5057–5061
4. Parker, B. O., and Marinus, M. G. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 1730–1734
5. Su, S. S., Lahue, R. S., Au, K. G., and Modrich, P. (1988) *J. Biol. Chem.* **263**, 6829–6835
6. Drotschmann, K., Aronshtam, A., Fritz, H. J., and Marinus, M. G. (1998) *Nucleic Acids Res.* **26**, 948–953
7. Hall, M. C., and Matson, S. W. (1999) *J. Biol. Chem.* **274**, 1306–1312
8. Ban, C., and Yang, W. (1998) *Cell* **95**, 541–552
9. Lu, A. L., Clark, S., and Modrich, P. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 4639–4643
10. Pukkila, P. J., Peterson, J., Herman, G., Modrich, P., and Meselson, M. (1983) *Genetics* **104**, 571–582
11. Ban, C., and Yang, W. (1998) *EMBO J.* **17**, 1526–1534
12. Aggarwal, A. K. (1995) *Curr. Opin. Struct. Biol.* **5**, 11–19
13. Pingoud, A., and Jeltsch, A. (1997) *Eur. J. Biochem.* **246**, 1–22
14. Murphy, K. C., Campellone, K. G., and Poteete, A. R. (2000) *Gene* **246**, 321–330
15. Horst, J. P., Wu, T. H., and Marinus, M. G. (1999) *Trends Microbiol.* **7**, 29–36
16. Aronshtam, A., and Marinus, M. G. (1996) *Nucleic Acids Res.* **24**, 2498–2504
17. Guzman, L., Belin, D., Carson, M., and Beckwith, J. (1995) *J. Bacteriol.* **177**, 4121–4130
18. Feng, G., and Winkler, M. E. (1995) *BioTechniques* **19**, 956–965
19. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd. Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
20. Murphy, K. C. (1998) *J. Bacteriol.* **180**, 2063–2071
21. Wu, T.-H., and Marinus, M. G. (1994) *J. Bacteriol.* **176**, 5393–5400
22. Lahue, R. S., Su, S. S., and Modrich, P. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 1482–1486
23. Horton, N. C., Nastri, H. G., Riggs, P. D., and Cheng, X. (1998) *J. Mol. Biol.* **284**, 1491–1504
24. Nastri, H. G., Evans, P. D., Walker, I. H., and Riggs, P. D. (1997) *J. Biol. Chem.* **272**, 25761–25767
25. Horton, N. C., Bonventre, J., and Cheng, X. (1998) *Biol. Chem.* **379**, 451–458
26. Harris, R. S., Feng, G., Ross, K. J., Sidhu, R., Thulin, C., Longerich, S., Szigety, S. K., Winkler, M. E., and Rosenberg, S. M. (1997) *Genes Dev.* **11**, 2426–2437