Identification of the Structural and Functional Domains of MutY, an Escherichia coli DNA Mismatch Repair Enzyme*

(Received for publication, December 11, 1995, and in revised form, March 16, 1996)

Raymond C. Manuel†, Edmund W. Czerwinski§, and R. Stephen Lloyd¶¶

From the †Sealy Center for Molecular Science, §Sealy Center for Structural Biology, and ¶¶Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch, Galveston, Texas 77555

The linear amino acid sequences of the Escherichia coli DNA repair proteins, MutY and endonuclease III, show significant homology, even though these enzymes recognize entirely different substrates. In this study, proteolysis and molecular modeling of MutY were used to elucidate its domain organization. Proteolysis by trypsin cleaved the enzyme into 26- and 13-kDa fragments. NH₂-terminal sequencing showed that the p13 domain begins at Gin226, indicating that the COOH-terminal portion of MutY, absent in endonuclease III, is organized as a separate domain. The large p26 domain is almost equivalent to the size of endonuclease III. Binding activity of the p26 domain to a DNA substrate containing an A-G mismatch was comparable with that of the intact enzyme. In vitro studies show that the p26 domain retains adenine glycosylase and AP lyase activity on DNA containing undamaged adenine opposite guanine or 8-oxo-7,8-dihydro-2'-deoxyguanine (8-oxo-dG). Although the activity was somewhat reduced, the above results show that the critical amino acid residues involved in substrate binding and catalysis are present in this domain. The structure predicted by molecular modeling indicates that the region of MutY (Met1-Trp316), which is homologous to endonuclease III exhibits a two domain structure, even though this portion is resistant to proteolysis by trypsin.

Proteins are constructed on a modular basis, and frequently these modules or domains are known to have unique functions. Isolation and characterization of these domains provide significant insight into the relationship between particular structural elements of the enzyme and its various activities. The mutY gene of Escherichia coli encodes a 39.1-kDa DNA mismatch repair protein. A significant portion of this protein is homologous to the 26.3-kDa E. coli endonuclease III. These two proteins are 66.3% similar and 23.8% identical over a 181-amino acid region (1). Another enzyme with sequence similarity to MutY is the product of the pdg gene in Micrococcus luteus, which recognizes and incises DNA containing cyclobutane pyrimidine dimers (2). The three enzymes mentioned above contain a [4Fe-4S] cluster, coordinated by four cysteine residues which are perfectly conserved in all three proteins.

In contrast to their structural similarities, the functional properties of the above three DNA repair proteins are very different. MutY recognizes and removes the undamaged adenine mispaired with guanine, cytosine, 8-oxo-7,8-dihydro-2'-deoxyguanine (8-oxo-dG) and 8-oxo-7,8-dihydro-2'-deoxyadenine (3) and is also reported to have AP endonuclease activity (4, 5). In vitro experiments show that MutY also recognizes and removes adenine analogs when they are paired with guanine (5). Endonuclease III primarily removes oxidized pyrimidines from DNA (6), and the pdg gene product in M. luteus repairs UV-induced pyrimidine dimer lesions in DNA (2). Thus, despite their structural similarities, these DNA repair enzymes have different substrate specificity. The catalytic mechanism of endonuclease III involves the glycosylase activity, followed by the β-elimination reaction which cleaves the phosphodiester bond 3' to an AP site (7). Results indicate that the M. luteus pyrimidine dimer glycosylase also follows a similar reaction mechanism (2).

The mutagenic nucleotide 8-oxo-dG primarily is incorporated opposite template adenine by several DNA polymerases (8). E. coli has multiple enzymes that act in different pathways to remove this lesion and maintain the integrity of DNA. In E. coli, the first line of defense is with MutT and MutM (Fapy DNA glycosylase), while MutY provides an alternative strategy (reviewed in Ref. 9). Another major substrate for MutY is an A-G mismatch containing DNA, where it removes the mismatched adenine. Failure to remove this mismatched adenine has been shown to increase C-G to A-T transversion in vivo (10).

Previous studies have shown that an A-G mismatch is stable within the DNA helix, although the flanking sequences can play a major role in their stability (11). The structural features of a DNA duplex containing an A-G mismatch have been studied by NMR and x-ray crystallography, and it has been shown that an A-G mismatch can exist in different structural conformations within the DNA helix (12–15). DNA duplexes containing 8-oxo-dG have also been studied through x-ray crystallography (16, 17). Since mismatch recognition and repair in DNA have not been characterized at the molecular level, it is vital to dissect the structural properties of the enzymes catalyzing these reactions. In this study, domain mapping studies were performed to define the structural and functional domains of MutY protein.

MATERIALS AND METHODS

Purification of MutY Protein—MutY was overexpressed from the plasmid pKYE with the mutY structural gene downstream of the tac promoter (gift from Drs. M. Michaels and J. H. Miller) in E. coli R. C. Manuel and R. S. Lloyd, unpublished results.

* This work was supported by United States Public Health Service Grants ES50491 and ES06676 and American Cancer Society Grant AFA381. 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.

† Recipient of the American Cancer Society Research Award FRA381. To whom correspondence should be addressed. Tel.: 409-772-2179; Fax: 409-772-1790; E-mail: rslloyd@scms.utmb.edu.

‡ The abbreviations used are: 8-oxo-dG, 8-oxo-7,8-dihydro-2'-deoxyguanine; AP, apurinic/apyrimidinic; dhU, dihydrouracil; Fapy, 2,6-dihydroxy-5-N-formamidopyrimidine; FPLC, fast protein liquid chromatography; PAGE, polyacrylamide gel electrophoresis; ssDNA, single-stranded DNA; DTT, dithiothreitol.

§ Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch, Galveston, Texas 77555.
GX1200. A five liter culture was grown in LB medium containing 100 μg/ml ampicillin in a New Brunswick Scientific BioFlo III fermenter at 37°C and induced with isopropyl-β-D-thiogalactopyranoside at a concentration of 1 mM when A600 reached 0.8. Growth of the culture was terminated after 4 h and the cells harvested by centrifugation at 5000 × g for 4°C for 20 min. The cells were resuspended in buffer A (50 mM Tris-HCl (pH 7.5), 80 mM NaCl, 1 mM EDTA, 0.6 M sucrose, 0.1% Triton X-100, 5% glycerol) and lysed in a French press. The lysate was clarified by centrifugation at 10,000 × g for 30 min at 4°C. The supernatant was diluted with buffer B (50 mM sodium phosphate (pH 7.5), 1 mM EDTA, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride) to a final NaCl concentration of 125 mM and loaded onto Q-Sepharose and SP-Sepharose columns (Pharmacia Biotech Inc.) in tandem, which were previously equilibrated with buffer B containing 125 mM NaCl. MutY protein was eluted from the SP-Sepharose column with a linear salt gradient in buffer B (0.1–1.0 mM NaCl). Fractions containing the protein were pooled and diluted with buffer C (10 mM sodium phosphate (pH 7.5), 0.25 mM DTT, and 0.1 mM EDTA) to a final NaCl concentration of 100 mM NaCl. This was loaded onto a single-stranded DNA (ssDNA) cellulose column (U. S. Biochemical Corp.) previously equilibrated with buffer C containing 100 mM NaCl, and MutY was eluted with a linear salt gradient of buffer C (0.1–1.0 mM NaCl). Fractions containing the enzyme were combined and diluted to a final salt concentration of 100 mM NaCl and further purified on a phosphocellulose (P11) column (Whatman) previously equilibrated with buffer C containing 100 mM NaCl. The protein was eluted as with the ssDNA cellulose column. Each chromatographic step, fractions containing MutY were identified by Coomassie Brilliant Blue and silver staining. The purity of the p26 domain was established by SDS-PAGE followed by silver staining.

The purity of the p26 domain was isolated after digesting 100 μg of the full-length MutY protein with 10 μg of trypsin for 60 min at 25°C. The digest was loaded on a ssDNA cellulose column (1 ml) previously equilibrated with buffer C containing 100 mM NaCl. The fractions generated by proteolysis were eluted with a step gradient of buffer C with increasing concentration of NaCl. The volume of each fraction was 500 μl, and they were analyzed on a 12.5% SDS-polyacrylamide gel followed by silver staining. Since the intact MutY and p26 domain coeluted, fractions were combined and diluted with buffer C to a final NaCl concentration of 100 mM before loading onto a Mono S HR 5/5 column (Pharmacia). A linear salt gradient with buffer C (0.1–1.0 mM NaCl) was used to separate the full-length MutY and the p26 domain. The purity of the p26 domain was established by SDS-PAGE followed by silver staining.

Gel Mobility Shift Assay to Determine DNA Binding Activity—A DNA substrate (30 base pairs) with a site-specific A-G mismatch was used in this assay. The sequence of the DNA substrate was as follows. The oligonucleotide containing the mismatched adenine was 5’-end-labeled with [γ-32P]ATP before annealing to its complementary strand.

The oligonucleotide containing the mismatched adenine was 5’-end-labeled with [γ-32P]ATP before annealing to its complementary strand. The binding reaction mixture contained 20 mM Tris-HCl (pH 7.5), 80 mM NaCl, 1 mM EDTA, and 2.5% glycerol and was incubated at two different temperatures (37 or 25°C) for 30 min. The protein-DNA complexes were separated by electrophoresis in a non-denaturing polyacrylamide gel (8%) and subjected to autoradiography and PhosphorImager analysis (Molecular Dynamics).

To identify the proteolytic fragments, the peptide was gel-purified, and the strand containing the mismatched adenine was 5’-end labeled with [γ-32P]ATP before annealing to its complementary strand. Different amounts of MutY or the p26 domain were added to the DNA substrate in a reaction buffer containing 25 mM sodium phosphate (pH 6.8), 1 mM EDTA, 50 mM NaCl, and 100 μM bovine serum albumin. The 20-μl reaction mixture was incubated at 37°C for 30 min before terminating the reaction in a dry ice ethanol bath. To monitor the inherent glycosylase activity, another set of reactions was terminated after reacting with piperidine for 15 min at 85°C. Treatment of the reaction products with piperidine results in the cleavage of phosphodiester bonds by 3'-endoelimination at residual AP sites. The frozen samples were dried in vacuum and resuspended in loading buffer (95% formamide, 20 mM EDTA, 0.05% xylene cyanol, and 0.05% bromphenol blue), and the products were separated through a 15% SDS-PAGE gel and stained with Coomassie Brilliant Blue.

The cleaved products and the remaining substrate were quantified on a PhosphorImager.

The above oligonucleotides were gel-purified, and the strand containing the mismatched adenine was 5’-end labeled with [γ-32P]ATP before annealing to its complementary strand. The binding reaction mixture contained 20 mM Tris-HCl (pH 7.5), 80 mM NaCl, 1 mM EDTA, and 2.5% glycerol and was incubated at two different temperatures (37 or 25°C) for 30 min. The protein-DNA complexes were separated by electrophoresis in a non-denaturing polyacrylamide gel (8%) and subjected to autoradiography and PhosphorImager analysis (Molecular Dynamics).

The above oligonucleotides were gel-purified, and the strand containing the mismatched adenine was 5’-end labeled with [γ-32P]ATP before annealing to its complementary strand. The reaction conditions and analysis of the products were similar to the assay for A-G mismatch incision activity.

Molecular Modeling—Homology modeling of a portion of MutY (Met1 to Trp232) was performed on an Evans and Sutherland (PS390) graphics workstation using CHAIN (20). The structural coordinates of the x-ray crystal structure of endonuclease III (21) served as the basis for the modeling. These amino acid residues of endonuclease III, which were neither identical nor similar to MutY, were replaced. The side chain atoms on the substituted amino acid residues were moved by manually rotating about the bonds to obtain the best visual conformation. The resultant "mutant" structure was then refined by energy minimization using X-PLOR (22). The structure presented in this report converged after 120 cycles. The superimposed images were produced using Midos-Plus in a Silicon Graphics workstation.

RESULTS

Purification of MutY—To perform limited proteolysis, MutY was purified to electrophoretic homogeneity using four different chromatographic steps (Fig. 1). The purity of the enzyme
was determined by SDS-PAGE gels stained with either Coomassie Brilliant Blue or silver. The combined chromatographic steps using Q-Sepharose and SP-Sepharose (lane 2) removed a majority of the soluble \textit{E. coli} proteins from the crude lysate (lane 1). Following ssDNA cellulose chromatography (lane 3), pure MutY was obtained from a phosphocellulose (P11) column (lane 4). At each chromatographic step, fractions were assayed for MutY, based on adenine glycosylase activity using a heteroduplex DNA containing a single A-G mismatch (data not shown). Although there are reports that MutY has only DNA glycosylase activity, complete phosphodiester backbone scission was observed with every fraction at each chromatographic step that contained MutY activity. MutY can be purified up to a concentration of 7 mg/ml. However, the enzyme does not remain in solution at such high concentrations. The conditions have been optimized to allow storage of the enzyme at a concentration of up to 3.5 mg/ml without any precipitation. Precipitation was also observed when the salt concentration of the storage buffer was below 400 mM or when the enzyme was stored in the absence of glycerol. Therefore, fractions combined after each chromatographic step were diluted rather than being dialyzed with low salt buffer before loading on to the next column for further purification. The fractions containing MutY were yellow in color due to the presence of the iron-sulfur cluster, and this facilitated a rapid way to monitor the fractions from various column chromatographic steps for the presence of the enzyme. The intensity of the color was directly proportional to both enzyme concentration and catalytic activity.

Domain Structure of MutY—Conditions were surveyed to generate optimal yields of stable domain fragments of the MutY. Although several proteases yielded polypeptide fragments, digestion with trypsin yielded two stable domains: the 26- and 13-kDa fragments, which are designated as p26 and p13, respectively (Fig. 2). These fragments were resistant to further proteolysis for at least 2 h, even though the trypsin to MutY ratio was 1:10, a condition which is much harsher than standard proteolytic digests. NH$_2$-terminal sequence analysis revealed the location of the fragments in the primary structure. The p26 domain starts at Met1 and the p13 domain at residue Gln226 (Table I). The NH$_2$-terminal methionine remained intact in MutY and the p26 domain, indicating that there is no post-translational modification of the NH$_2$ terminus in vivo. The COOH-terminal ends of the two proteolytic domains are estimated based on the electrophoretic mobility, and it is likely that there are no truncations at the COOH-terminal end, although this has not been specifically tested. The single site that is sensitive to trypsin indicates that the p26 and p13 domains are stable and tightly folded. These studies suggest that MutY has a two domain structure: a large NH$_2$-terminal domain (p26) comparable in size to \textit{E. coli} endonuclease III, which is a 211-amino acid DNA repair enzyme and a relatively small COOH-terminal domain (p13).

Nonspecific DNA Binding Activity of p26 Domain—Removal of the p13 domain did not compromise the general DNA binding activity of MutY. This was determined based on identical binding affinity of intact MutY and the truncated p26 domain to the ssDNA cellulose column (Fig. 3). This binding was tight and thus both intact protein and p26 domain eluted at NaCl concentrations of 450 and 500 mM (lanes 4 and 5, respectively). The p13 domain by itself did not have any DNA binding activity, as this fragment was eluted in the low salt (100 mM NaCl) wash.
of undigested intact MutY (Fig. 4, graphic steps). Although the first step using ssDNA cellulose through proteolysis with trypsin was purified into two chromatographic steps (Fig. 4, fractions collected with the elution buffer containing 450 and 500 mM NaCl, respectively, showing the simultaneous elution of intact MutY and the p26 domain.

Isolation of p26 Domain—The stable p26 domain, obtained through proteolysis with trypsin, was purified in two chromatographic steps. Although the first step using ssDNA cellulose column did not separate the p26 domain from the small amount of undigested intact MutY (Fig. 4, lane 2), the p13 domain was separated due to its reduced binding affinity to this column (Fig. 3, lane 3). The trypsin present in the proteolytic digest was also eluted, as it does not bind to the ssDNA cellulose column. Isolation of the homogeneous 26-kDa fragment was achieved using a Mono S FPLC column. The full-length MutY eluted at a lower salt concentration (not shown) relative to the p26 domain; lanes 4 and 5, fractions collected with the elution buffer containing 450 and 500 mM NaCl, respectively, showing the simultaneous elution of intact MutY and the p26 domain.

Binding Affinity of MutY and p26 Domain on DNA Containing an A-G Mismatch—Using synthetic heteroduplex DNA (50-mer) containing a single A-G mismatch, we compared the efficiency of adenine glycosylase and AP lyase activities between native MutY and the p26 domain (Fig. 6A). PhosphorImager analysis demonstrated that 12 and 74% of the substrate can be cleaved by 46 and 230 fmol of intact MutY, respectively. This can be compared with approximately 11% cleavage effected by the p26 domain at 230 fmol (Fig. 6B). With either the intact MutY or the p26 domain, the amounts of cleaved products generated were almost equal, whether or not the reaction products were treated with piperidine (Fig. 6B). This demonstrates that the glycosylase activity by MutY and p26 domain is accompanied by an efficient AP lyase activity. It is also evident from our experiments that another product with faster electrophoretic mobility is seen due to cleavage of the C5-N-O-P bond 5'-elimination), releasing the ring-opened deoxyribose sugar moiety and leaving a 3'-phosphate end (Fig. 6A). Since the 5'-elimination product accumulates to a greater degree at higher enzyme concentrations, it was concluded that the reaction is not a concerted β-δ reaction. Incision activity was not observed in DNA that contained a thymine opposite adenine (data not shown).

Removal of Adenine Opposite a 8-Oxo-dG Lesion by MutY and p26 Domain—The adenine glycosylase activity followed by AP lyase activity was clearly observed by treating the substrate with intact MutY (Fig. 7A) or the p26 domain (Fig. 7B). The autoradiograph obtained to produce Fig. 7B was exposed for a longer period of time compared with Fig. 7A to enable the visualization of the cleaved products. The percentage of products formed by the p26 domain was significantly less than that observed for the intact protein. The maximum conversion to product was 4%, which represents about 5% of the activity seen
with the wild type enzyme (Fig. 7C). However, when the substrate and enzyme were in equimolar amounts the percentage of activity of the p26 domain was 28 and 13%, respectively (Fig. 7C), when compared with that of the intact MutY with and without piperidine. The \( \delta \)-elimination reaction product was present in reactions with intact MutY and the p26 domain (Fig. 7, A and B).

Absence of Activity for MutY and the p26 Domain on DNA Substrates for Fapy-DNA Glycosylase and Endonuclease III—MutY and the p26 domain did not have any catalytic activity on a DNA substrate containing an 8-oxo-dG lesion opposite cytosine and a DNA substrate containing 5,6-dihydrouracil paired with guanine (data not shown). The above substrates are recognized and catalyzed by Fapy-DNA glycosylase and endonuclease III of E. coli.

Molecular Modeling—Alignment of the primary amino acid sequence of the molecular modeled portion of MutY, with the corresponding region of the native MutY, is shown in Table II. The energy minimized structure of the portion of MutY (Met1 to Trp216), when superimposed on the x-ray crystal structure of endonuclease III, indicated a high structural homology at the tertiary level (Fig. 8). The molecular modeled portion of MutY shows an elongated two domain structure with a deep cleft separating them. However, this structure is resistant to proteolysis by trypsin. The secondary structure primarily is \( \alpha \)-helical, similar to endonuclease III. The spatial geometry of the 4 cysteine residues (residues 192, 199, 202, and 208 in native MutY) implicated in coordinating the Fe-S cluster is well conserved in the modeled structure.

DISCUSSION

Domain Organization—Controlled proteolysis of native proteins can generate large polypeptide fragments that reflect independently folding domains. This can provide information on the global structure of the protein and also permit the isolation of discrete domains for further structural and biochemical characterization. In this study, proteolysis of MutY with trypsin yielded two distinct domains (Fig. 2): a large NH\(_2\)-terminal domain (p26) and a COOH-terminal domain (p13). Amino-terminal sequence analyses of the proteolytic fragments have led to the identification of only one cleavage site for trypsin. The NH\(_2\)-terminal fragment beginning at Met\(^1\) comprises the p26 domain, and the remaining portion of MutY beginning at Gln226 constitutes the COOH-terminal domain. The absence of multiple protease-sensitive sites also indicates that the enzyme is globular and compact. These two domains are extremely stable and were separated by chromatographic methods (Figs. 3 and 4).

Fig. 5. Proteolysis of MutY with trypsin and binding of intact MutY and the p26 domain generated by trypsin to A-G mismatch containing DNA. A, SDS-PAGE analysis of intact MutY and domain fragments generated by trypsin. MutY (1 \( \mu \)g) was proteolysed with 20 ng trypsin (lane 2) and 100 ng trypsin (lane 3) as described under "Materials and Methods." The amount of MutY cleaved is proportional to the amount of trypsin used. Intact MutY used for the trypsinic digest is shown in lane 1. Prestained molecular weight markers are shown in the margin. B, autoradiograph showing the binding of MutY and p26 domain to A-G mismatch containing DNA. The oligonucleotide (30-mer) containing the mismatched adenine was 5'-end labeled with \( \gamma \)-\( ^{32} \)P]ATP and annealed to its complementary strand. The heteroduplex DNA (15 fmol) was allowed to bind with the proteolyzed mixture of MutY at 37°C (lanes 3 and 4) and 25°C (lanes 5 and 6). Lanes 3 and 5 show the protein-DNA complexes formed from the proteolyzed mixture (A, lane 3) in which 100 ng of trypsin was used. Lanes 4 and 6 show the protein-DNA complexes formed from the proteolyzed mixture (A, lane 2) in which 20 ng of trypsin was used. In each binding reaction 20 ng of poly(dI-dC) only, as there was no protein added in the binding reaction, and lane 2 shows the MutY-DNA complex formed from undigested MutY. In all the binding reactions p13-DNA complex was not formed. C, graphic representation of MutY and the p26 domain present in the proteolytic digest with 100 and 20 ng of trypsin and compared with the MutY-DNA and p26-DNA complexes formed at 37°C in the reaction conditions described in the methods section. The protein-DNA complexes were quantified by PhosphorImager analysis and the amounts of intact MutY and p26 domain in the SDS-PAGE gel were quantified using the Imager (Appligene) and VISAGE electrophoresis gel analysis system (Millipore).
Intact MutY and the p26 domain bind to the ssDNA cellulose column with equal affinity, suggesting that the critical amino acid residues or localized region(s) involved in the electrostatic interaction with DNA are present in this domain. This view is further strengthened by the absence of DNA binding activity in the p13 domain (Fig. 3). In addition intact MutY and the p26 domain show almost equal binding affinity to duplex DNA containing a single A-G mismatch. This has been demonstrated by varying the amount of intact MutY and the p26 domain in the same binding reaction mixture (Figs. 5, B and C).

In MutY, a potential role for the Fe-S cluster has been suggested in previous studies which show that renaturation of denatured enzyme requires iron and sulfur to recover the catalytic activity (4). The retention of specific DNA binding activity by the p26 domain indicates that the structural integrity of the [4Fe-4S]$_2^+$ cluster coordinated by four cysteine residues is conserved, thereby providing a surface to interact with DNA, and this region may even serve as a portion of the binding pocket. The above hypothesis is proposed because in endonuclease III, the region containing the [4Fe-4S]$_2^+$ cluster has the highest positive electrostatic potential, and this region is implicated in positioning the basic residues for interaction with the phosphate backbone of the DNA substrate (21, 23). In addition, based on the location of the iron-sulfur cluster in endonuclease III, models have been built to illustrate its role in DNA binding (24). Another DNA repair enzyme, Fapy DNA glycosylase, which removes 8-oxo-dG when paired with cytosine, shows reduced DNA binding and enzymatic activities when the four cysteine residues coordinating the zinc metal ion are disturbed (25, 26). Thus, it is very likely that these metal coordinated complexes in the above mentioned DNA repair enzymes play a role in maintaining the structural integrity and provide an appropriate surface to interact with the substrate, rather than being involved in catalysis.

Catalytic Activities—The size similarity and sequence homology between endonuclease III and the p26 domain, along with the observation that the p26 domain had not lost its substrate-specific DNA binding activity, led us to investigate the catalytic activity of the p26 domain on substrates which are recognized and cleaved by MutY and other closely related DNA repair glycosylase/AP lyases. Comparison of the DNA mismatch repair activity of full-length MutY and the p26 domain on a substrate with an A-G mismatch shows that the p26 domain retains approximately 15% of the adenine glycosylase and AP lyase activities (Fig. 6). When MutY and the p26 domain were allowed to react with DNA containing an adenine opposite 8-oxo-dG, a smaller amount of activity was recovered from the p26 domain (Fig. 7). Although the amount of activity recovered was less than that of the native MutY, it is clear that the p26 domain is able to recognize and catalyze the removal of mismatched bases in DNA, as seen with intact MutY. The truncated p26 domain also retains a rigorous substrate specificity equivalent to that of the intact MutY. The p26 domain does not have catalytic function on DNA containing 8-oxo-dG-C or dhU-G base pairs, which are substrates for two other closely related E. coli proteins, namely Fapy DNA glycosylase and endonuclease III, respectively (data not shown).

Since the p26 domain retains at least reduced levels of catalytic function, the key residues involved in catalysis must be present in this domain, in addition to the amino acids involved in substrate recognition and DNA binding. Thus, we can conclude that the p26 domain is catalytically competent, even though the efficiency is compromised. We can only speculate about the reasons for the reduced levels of catalytic function by p26 domain at this time. Although the COOH-terminal domain may not contain the critical residues for catalysis, it could play

**Fig. 6.** A-G mismatch repair activity of intact MutY and the p26 domain. A, autoradiograph showing cleaved products after incubating MutY and the p26 domain at different concentrations with an oligonucleotide duplex (50 base pairs) containing a single A-G mismatch. The strand containing the mismatched adenine was 5'-end-labeled with [γ-32P]ATP. The reaction products were separated by electrophoresis in a 15% denaturing polyacrylamide gel. Oligonucleotide markers are indicated in the margin. In all the reactions 46 fmol of substrate was used. Lane 1, substrate only; lanes 2-4, MutY at 1×, 5×, 10× molar substrate concentration; lanes 5-7, MutY at 1×, 5×, 10× molar substrate concentration, followed by piperidine treatment; lanes 8-10, p26 domain at 1×, 5×, 10× molar substrate concentration; lanes 11-13, p26 domain at 1×, 5×, 10× molar substrate concentration, followed by piperidine treatment. B, quantitation of A-G mismatch repair activity by MutY and the p26 domain. The percentage of products formed at different substrate to enzyme ratio. 1 and 3 show percent products cleaved by MutY and p26 domain, respectively; 2 and 4 show percent products cleaved by MutY and p26 domain, respectively, followed by piperidine treatment.
a role in domain movements during catalysis.

Mechanistic Considerations—Earlier reports suggest that MutY is an adenine glycosylase and lacks detectable AP lyase activity (3, 27). However, there have also been several studies that show MutY to have adenine glycosylase and 3’ AP endonuclease activities (4, 5, 28). In a recent report, a putative mammalian homolog of E. coli MutY mismatch repair protein has been shown to recognize similar substrates with the nicking activity observed at the phosphodiester bond 3’ to the AP site following adenine glycosylase activity (29). Our results are consistent with the glycosylase activity accompanying AP lyase activity. This has been exhibited by MutY and the p26 domain on at least two primary substrates in which the undamaged adenine is paired with guanine or 8-oxo-dG. The absence of any catalytic activity by MutY and the p26 domain on 8-oxo-dG-C and dU-G lesions, which are substrates for Fapy DNA glycosylase and endonuclease III, respectively, confirms that the AP lyase reaction observed in our reaction conditions is not due to any contamination by the above E. coli DNA glycosylase/AP lyases. The AP lyase activity is also apparent in other reaction conditions with different buffers (Hepes and Tris), different pH and varying salt concentrations (data not shown). Presently, we do not have a clear explanation for this discrepancy at the mechanistic level. Further studies are under way to resolve this question.

A covalent imino intermediate formation between the enzyme and the substrate is characteristic of all enzymes that show glycosylase and AP lyase activity at the same rate (reviewed in Ref. 30). It has been demonstrated that this intermediate formation is due to a primary amine involved as a nucleophile (31–34). Previous studies have demonstrated that the above intermediate is not formed between MutY and a DNA substrate containing an A-G mismatch (35). The reaction mechanism of MutY could be similar to other DNA glycosylases/AP lyases with a nucleophile other than a primary amine.

The initial action of all DNA glycosylases involves a nucleophilic attack at the deoxyribose sugar C-1' of the mismatched/damaged nucleoside, resulting in the hydrolysis of the N-glycosyl bond. When these glycosylases also catalyze the breakage of the sugar phosphate backbone by β-elimination, they leave a 3'-α,β-unsaturated aldehyde, and a 5'-phosphate (36, 37). In some DNA glycosylases with associated AP lyase activity, there is also a δ-elimination that is observed at the AP site leaving a 3'-phosphate end (38, 39). The δ-elimination reaction product results from the cleavage of the C5-O-P bond...

---

3 R. C. Manuel and R. S. Lloyd, personal observation.
TABLE II
Amino acid sequence homology between a region of native MutY (Met1 to Trp216) and the molecular modeled portion of MutY. Amino acid residues that are identical between MutY and the molecular modeled portion of MutY are shown in bold type, and the remaining residues are similar. Homology modeling was based on the x-ray crystal structure of E. coli endonuclease III.

| MutY   | MQASQFSAQVLDWYDKYRTLPQIDETPKYKVLSEVLQQTQVAATIVYPFERFMRFPFTVDDNAPLDLDEVLYH | 75 |
|--------|---------------------------------------------------------------------------------|----|
| Model  | -QASQFSAQVLDWYDKYRTLPQIDETPKYKVLSEVLQQTQVAATIVYPFERFMRFPFTVDDNAPLDLDEVLYH     | 74 |
| MutY   | LWTGLGY-ARARNLHKAQQVAIHLHGKFPEFEEVAALPGVGRSTAGAILLLSLGKIP1LDGNVKLVARCYA       | 150|
| Model  | LWTGLGYNSNARLHKAQQVAIHLHGKFPEFEEVAALPGVGRSTAGAILLLSLGKIP1LDGNVKLVARCYA       | 150|
| MutY   | VSGRFLKEVESKILWLSSEQTSAVGVSRFAQMDAOGICGTRKPSLCLKQNGC1AATNNW                  | 216|
| Model  | APEGNPKQVESKILV-YVAEAVGRFQDFAAILHGAUGMC1RPFRPGKCPFEDGC1AAANH                  | 211|

5' to the AP site. From this study, it is clear that intact MutY and the p26 domain exhibit a δ-elimination reaction following β-elimination. This result can be verified by comparing the δ-elimination product with the product generated by piperidine treatment. The δ-elimination product is shown to increase as the enzyme concentration increases (Figs. 6A and 7, A and B) as was demonstrated with another DNA repair enzyme, T4 endonuclease V (39). Creating a gap by δ-elimination leaves 3'-phosphate and 5'-phosphate ends, which are repaired by sequential reactions catalyzed by an AP endonuclease, DNA polymerase and ligase.

Homology Modeling with Endonuclease III—Homologous proteins with a strong sequence similarity are known to have a similar secondary and tertiary structure. In the absence of atomic coordinates, molecular modeling of a polypeptide, based on its sequence homology to another protein for which the structure is known by x-ray crystallography or NMR, is a useful approach to predict its structure. Comparative modeling is known to yield successful results when the known structure has at least 40% sequence homology with the unknown structure (40). Native MutY and endonuclease III have 66% similarity and 23% identity over a 181 amino acid region. Although there is no direct proof that the modeled portion of MutY and endonuclease III share structural homology at the tertiary level, molecular modeling studies performed in this study indicate that this is likely to be the case. Based on the model developed in this study, we predict that a portion of MutY (Met1 to Trp216), which shares sequence homology with endonuclease III, could have the same overall conformation as endonuclease III (Fig. 8).

The above model neglects the structural conformation of MutY in the presence of the COOH-terminal region (Ala217 to Val359). However, the proteolytic map of MutY indicates that the p26 and p13 domains are extremely stable by themselves and can exist as separate domains with independent structures. The modeled portion of MutY also suggests that deletion of amino acids at the NH2 terminus is likely to change the coordination geometry of the iron-sulfur cluster due to the
proximity of this region to the NH₂ terminus of the protein. The model predicted in this report provides an opportunity to resolve key questions about the structure and function of MutY. Future uses of the model will include guidance to mutagenesis experiments to identify the amino acid residues involved in substrate binding and catalysis.

Acknowledgments—We are indebted to Dr. M. L. Michaels (Amgen Corp.) and J. H. Miller (UCLA) for their generous gift of the overexpressing plasmid pKYEC containing the mutY gene. We thank Dr. R. Cunningham for providing the coordinates of the crystal structure of endonuclease III. We also thank Dr. M. L. Dodson for valuable discussions. The oligonucleotide containing 8-oxo-dG was a generous gift from Prof. E. Ohtsuka to Dr. S. Mitra. Synthesis of other oligonucleotides and amino acid sequence analysis was performed by the Recombinant DNA Laboratory and the Protein Chemistry Core Facility, respectively, at the Sealy Center for Molecular Science, University of Texas Medical Branch, Galveston, TX. Molecular graphics images were produced using MidasPlus obtained from the Computer Graphics Laboratory, University of California, San Francisco.

REFERENCES
1. Michaels, M. L., Pham, L., Nghiem, Y., Cruz, C., and Miller, J. H. (1990) Nucleic Acids Res. 18, 3841–3845
2. Piersen, C. E., Prince, M. A., Augustine, M. L., Dodson, M. L., and Lloyd, R. S. (1995) J. Biol. Chem. 270, 23475–23484
3. Michaels, M. L., Tchou, J., Grolimund, A. P., and Miller, J. H. (1992) Biochemistry 31, 10964–10968
4. Tsai-Wu, J.-J., Liu, H.-F., and Lu, A.-L. (1992) Biochemistry 31, 89, 8779–8781
5. Lu, A.-L., Tsai-Wu, J.-J., and Cillo, J. (1995) J. Biol. Chem. 270, 23582–23588
6. Kow, Y. W., and Wallace, S. (1987) Biochemistry 26, 8200–8206
7. Bailly, V., and Verly, W. G. (1987) Biochem. J. 242, 565–572
8. Pavlo, Y. I., Minnick, D. T., Izuta, S., and Kunkel, T. A. (1994) Biochemistry 33, 4695–4703
9. Tchou, J., and Grollimund, A. P. (1993) Mutat. Res. 299, 277–287
10. Nghiem, Y., Cabrera, M., Cupples, C. G., and Miller, J. H. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 2709–2713
11. Li, Y., and Agrawal, S. (1995) Biochemistry 34, 10056–10062
12. Patel, D. J., Kodowski, S. A., Ikuta, S., and Itakura, K. (1984) Biochemistry 23, 3207–3227
13. Gao, X., and Patel, D. J. (1988) J. Am. Chem. Soc. 110, 5178–5182
14. Nikonenowicz, E. P., Meadows, R. P., Fagan, P., and Gorenstein, D. G. (1991) Biochemistry 30, 1323–1334
15. Nikonenowicz, E. P., and Gorenstein, D. G. (1992) J. Am. Chem. Soc. 114, 7494–7503
16. McAuley-Hecht, K. E., Leonard, G. A., Gibson, N. J., Thomson, J. B., Watson, W. P., Hunter, W. N., and Brown, T. (1994) Biochemistry 33, 10266–10270
17. Lipscomb, L. A., Peek, M. E., Morningstar, M. L., Verghis, S. M., Miller, E. M., Rich, A., Essigmann, J. M., and Williams, L. D. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 719–723
18. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
19. Manuel, R. C., Latham, K. A., Dodson, M. L., and Lloyd, R. S. (1995) J. Biol. Chem. 270, 2652–2661
20. Sadik, S. (1988) J. Mol. Graphics 6, 224–225
21. Kuo, C.-F., McRee, D. E., Fisher, C. L., O’Handley, S. F., Cunningham, R. P., and Tainer, J. A. (1992) Science 258, 434–440
22. Brumberg, J. T. (1992) X-PLOR Manual, Version 3.0, Yale University, New Haven, CT
23. Xing, D., Dorr, R., Cunningham, R. P., and Schles, C. P. (1995) Biochemistry 34, 2537–2544
24. Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P., and Tainer, J. A. (1995) EMBO J. 14, 4108–4120
25. O’Connor, T. R., Graves, R. J., de Murcia, G., Castaing, B., and Laval, J. (1993) J. Biol. Chem. 268, 9063–9070
26. Tchou, J., Michaels, M. L., Miller, J. H., and Grollimund, A. P. (1993) J. Biol. Chem. 268, 26738–26744
27. Au, K. G., Clark, S., Miller, J. H., and Modrich, P. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 8877–8881
28. Lu, A.-L., and Chang, D.-Y. (1988) Cell 54, 805–812
29. McGoldrick, J. P., Yeh, Y.-C., Solomon, M., Essigmann, J. M., and Lu, A.-L. (1995) Mol. Cell. Biol. 15, 989–996
30. Dodson, M. L., Michaels, M. L., and Lloyd, R. S. (1994) J. Biol. Chem. 269, 32709–32712
31. Schrock, R. D., III, and Lloyd, R. S. (1991) J. Biol. Chem. 266, 17631–17639
32. Schrock, R. D., III, and Lloyd, R. S. (1993) J. Biol. Chem. 268, 880–886
33. Dodson, M. L., Schrock, R. D., III, and Lloyd, R. S. (1993) Biochemistry 32, 8284–8290
34. Tchou, J., and Grollimund, A. P. (1995) J. Biol. Chem. 270, 11671–11677
35. Sun, B., Latham, K. A., Dodson, M. L., and Lloyd, R. S. (1995) J. Biol. Chem. 270, 19501–19508
36. Mazumder, A., and Gerlt, J. A. (1989) J. Am. Chem. Soc. 111, 8029–8030
37. Gerlt, J. A. (1993) in Nucleases (Linn, S. M., Lloyd, R. S., and Roberts, R. J., eds) Vol. II, pp. 263–316, Cold Spring Harbor Laboratory Press, Plainview, NY
38. Bailly, V., Verly, W. G., O’Connor, T., and Laval, J. (1989) Biochem. J. 262, 581–589
39. Latham, K. A., and Lloyd, R. S. (1995) Biochemistry 34, 8796–8803
40. Sali, A., and Blundell, T. L. (1993) J. Mol. Biol. 234, 779–815