Specific Recognition of A/G and A/7,8-Dihydro-8-oxoguanine (8-oxoG) Mismatches by *Escherichia coli* MutY: Removal of the C-Terminal Domain Preferentially Affects A/8-oxoG Recognition†

Arhonda Gogos,‡ Jason Cillo,§ Neil D. Clarke,‡ and A-Lien Lu*,§

Department of Biophysics and Biophysical Chemistry, Johns Hopkins School of Medicine, Baltimore, Maryland 21205, and Department of Biochemistry and Molecular Biology, School of Medicine, University of Maryland at Baltimore, Baltimore, Maryland 21201

Received April 8, 1996; Revised Manuscript Received August 20, 1996⊗

ABSTRACT: *Escherichia coli* MutY is a 39 kDa adenine DNA glycosylase and 3′ apurinic/apyrimidinic (AP) lyase that is active on DNA substrates containing A/G, A/C, or A/8-oxoG mismatches. 8-oxoG (7,8-dihydro-8-oxoguanine or GO) is a major stable product of oxidative damage, and A/GO mismatches may be particularly important biological substrates for MutY. Proteolytic digestion of MutY using thermolysin was found to produce two relatively stable fragments of 25 and 12 kDa. The 25 kDa fragment begins at the N terminus of MutY and spans the region homologous with *E. coli* endonuclease III, a DNA glycosylase/AP lyase that repairs oxidatively damaged pyrimidines. The 12 kDa fragment, which consists of much of the rest of MutY, had no detectable activity. The purified 25 kDa fragment (M25) had nearly wild-type binding and cleavage activities with A/G-mismatched substrates. Binding to A/GO-mismatched DNA, however, was dramatically reduced in M25 compared to that in intact protein. Borohydride-dependent enzyme−DNA cross-linking, which is a hallmark of the reaction of several DNA glycosylases that possess concomitant AP lyase activity, was also substantially reduced when M25 was allowed to react with A/GO-mismatched DNA. The significant differences in M25 recognition and reactivity with A/G and A/GO mismatches suggest that the C-terminal region of MutY, a region with no homologous counterpart in *E. coli* endonuclease III, plays an important role in the repair of mismatched DNA arising from oxidation damage.

MutY is a 39 kDa DNA-repair protein from *Escherichia coli* that has both adenine glycosylase and 3′ apurinic/apyrimidinic (AP[1]) lyase/endonuclease activities (Michaels et al., 1990; Tsai-Wu et al., 1991, 1992). It removes adenine residues that have been misincorporated opposite guanine (A/G), cytosine (A/C), or 7,8-dihydro-8-oxoguanine (A/GO) (Au et al., 1988, 1989; Lu & Chang, 1988b; Michaels & Miller, 1992; Michaels et al., 1992; Radicella et al., 1988; Su et al., 1988; Tsai-Wu et al., 1992) and then cleaves the phosphodiester bond that is immediately 3′ to the AP site (Lu & Chang, 1988a; Tsai-Wu et al., 1992). The combined glycosylase and AP lyase activities result in the cleavage of the A strand in A/G- or A/GO-mismatched substrates. The AP site created by MutY opposite GO is repaired by DNA polymerase I (Radicella et al., 1993; Tsai-Wu & Lu, 1994), which preferentially incorporates dCMP at such sites (Shibutani et al., 1991). The resulting C/GO base pair is then a substrate for a second repair enzyme, 8-oxoguanine DNA glycosylase (Fpg protein or MutM). Removal of GO by MutM, followed by gap repair of the abasic site, results finally in the restoration of the original C/G base pair (Chetsanga & Lindahl, 1979; Grollman, 1993; Tchou et al., 1991).

GO is a significant oxidation product in terms of mutagenicity (Cheng et al., 1991; Moriya, 1993; Moriya et al., 1991; Wood et al., 1990), and mutant strains that lack MutY activity have elevated rates of G/C to T/A transversions (Nghiem et al., 1988; Radicella et al., 1988). Although the endonuclease activity of MutY is slightly greater with A/G substrates than with A/GO, the protein binds A/GO-mismatched DNA about 80-fold more tightly than it does A/G (Lu et al., 1995, 1996). Binding of MutY to either A/GO or its own AP/GO product prevents MutM from reacting with the same substrates. The competitive inhibition of MutM activity by MutY binding to AP/GO may be important since GO glycosylase activity on an AP/GO substrate results in a total loss of the original coding information and probably a double strand break (Lu et al., 1995; Michaels et al., 1992). Although most biochemical and genetic studies of A/GO glycosylase activity have been performed on *E. coli* MutY, a similar activity has been identified in nuclear extracts of human HeLa cells and in calf thymus (McGoldrick et al., 1995; Yeh et al., 1991). Recently, a human gene homologous to *mutY* has been identified (Slupska et al., 1996).

MutY is homologous to *E. coli* endonuclease III (endo III), a DNA glycosylase/AP lyase with broad specificity

† This work was supported at Johns Hopkins by the Markey Center for Macromolecular Structure and Function and by an American Cancer Society Junior Faculty Award (N.D.C.) and by Public Health Service Grant GM 35132 from the National Institute of General Medical Science (A.L.).

* Corresponding author: Department of Biochemistry and Molecular Biology, School of Medicine, University of Maryland, 108 N. Greene St., Baltimore, MD 21201. Phone: (410) 706-4356. Fax: (410) 706-1787. E-mail: aluchang@umabnet.ab.umd.edu.

‡ Johns Hopkins School of Medicine.

§ University of Maryland at Baltimore.

⊗ Abstract published in *Advance ACS Abstracts,* December 1, 1996.

[1] Abbreviations: AP, apurinic/apyrimidinic; GO, 8-oxoG or 7,8-dihydro-8-oxoguanine; endo III, *E. coli* endonuclease III; M25, the N-terminal 25 kDa proteolyzed MutY fragment; M12, the C-terminal 12 kDa proteolyzed MutY fragment.

S0006-2960(96)00843-4 CCC: $12.00

toward oxidatively damaged pyrimidines. The structure of endo III has been determined by crystallographic methods, and a possible substrate binding site was identified by soaking the crystals with free thymine glycol (Kuo et al., 1992). A model for how endo III might bind to DNA substrates has been developed on the basis of these and other data (Kuo et al., 1992; Thayer et al., 1995). Among the residues that may be involved in DNA binding are residues 113−119 that form a β-hairpin and some that are close to an iron−sulfur cluster ([4Fe-4S]$^{2+}$). The function of the iron−sulfur cluster may be purely structural as there is no evidence that it has any redox activity relevant to the enzymatic activity of the protein (Cunningham et al., 1989; Fu et al., 1992; Thayer et al., 1995). MutY also binds iron (approximately four per polypeptide chain) (Lu et al., 1995), and the enzymatic activity of denatured MutY can be recovered only if the protein is renatured in the presence of ferrous iron and sulfide (Tsai-Wu et al., 1992). The sequence pattern of the cysteines that ligate the iron−sulfur cluster is conserved between MutY and endo III (Michaels et al., 1990; Tsai-Wu et al., 1991). The sequence similarity between the two proteins (23% identity) extends over the length of endo III (Michaels et al., 1990); however, MutY has an additional 138 amino acids at its C terminus that have no counterpart in endo III.

In this paper, we report the limited proteolytic digestion of MutY by thermolysin to create two stable fragments. The larger fragment, which we denote as M25, is 25 kDa, corresponds closely to the endo III homology region, and retains A/G binding and cleavage activities. The combined glycosylase and AP lyase activity of this fragment, as assayed with an A/G-mismatched substrate, is nearly as great as that of the full-length protein. However, the binding and biochemical activities of M25 with A/GO-mismatched substrate are reduced substantially as measured by gel retardation and sodium borohydride-mediated trapping assays. This difference between full-length protein and M25 suggests that the C-terminal domain plays a role in the recognition of DNA containing A/GO, but not A/G, mismatches.

MATERIALS AND METHODS

MutY Protein. MutY was purified to near homogeneity from an overproducing strain of *E. coli* as described previously (Tsai-Wu et al., 1992). The specific activity of the enzyme was approximately 13×10^6 units/mg, where 1 unit of activity is equivalent to the cleavage of 0.018 fmol of labeled A/G−DNA in 30 min at 37 °C. The protein concentration was measured using the Bradford assay (Bradford, 1976).

Proteolytic Digestion and Identification of Fragments. Proteolytic digestion experiments were performed using several ratios (w:w) of thermolysin (Boehringer Mannheim) to MutY. Digestion was carried out in 20 mM KH_2PO_4 (pH 7.4), 0.5 mM dithiothreitol, 0.3 M KCl, 0.1 mM EDTA, 10% glycerol, and 2 mM $CaCl_2$. Following incubation for 30 min at room temperature, EDTA was added to the reaction mixture to a final concentration of 200 μg/mL. The mixture was denatured in SDS for analysis on 17% SDS/polyacrylamide gels (Laemmli, 1970). Gels were stained with Coomassie blue or were used to transfer protein to PVDF paper (Millipore) for N-terminal sequence analysis.

Purification of the 25 kDa Proteolytic Fragment. The 25 kDa proteolytic product was purified on a heparin column. Following sample application, the gel was washed with 10 column volumes of binding buffer [20 mM KH_2PO_4 (pH 7.4), 0.5 mM dithiothreitol, 0.1 mM EDTA, and 0.1 M KCl] to remove unbound or nonspecifically bound material. Protein was eluted using a series of buffers of increasing KCl concentration. A 0.5 mL column was sufficient to purify M25 from mixtures containing 600 μg of protein.

N-Terminal Sequencing. The following procedure is based on a previously published protocol (Speicher, 1989). SDS/polyacrylamide gels were cast 1−3 days before use, and 0.1 M thioglycolate was present in the running buffer to scavenge reactive compounds left in the gel. Samples (10 μg) were heated at 37 °C prior to loading, with 5X solubilizing buffer [0.5 M sucrose, 15% SDS, 312.5 mM Tris-HCl (pH 6.9), 10 mM EDTA, 5% β-mercaptoethanol, and 0.0025% bromophenol blue]. Electrophoretic transfer to PVDF membrane (Millipore) was performed using a Mini Trans-Blot cell running at 30 V overnight and at 4 °C. The membrane was prewet in 100% methanol before adding to the transfer buffer. The transfer buffer was 0.5X Towbin buffer [12.5 mM Tris-HCl (pH 8.3) and 96 mM glycine] in 10% methanol. PVDF membranes were washed with water and stained in 0.025% Coomassie blue in 40% methanol and destained in 50% methanol. Standard Edman degradation was performed using an Applied Biosystems 492 Precise Sequencing System.

Mass Spectrometry. The protein digest was desalted by reverse phase chromatography using Waters Sep-Pak cartridges, and concentrated to 7 pmol/mL. The sample was analyzed by matrix-assisted laser desorption/ionization (MALDI), using a COMPACT MALDI III (Kratos, Manchester, England). Digest (0.3 μL) was deposited on a sample site, and 0.3 μL of matrix was added. The mixture was dried at room temperature. The matrix was a saturated solution of α-cyano-4-hydroxycinnamic acid (Aldrich) in 50% ethanol.

Oligonucleotide Substrates. The sequences of 19-mer oligonucleotide duplexes used were

```
5' CCGAGGAATTAGCCTTCTG  3'
3'  GCTCCTTAAXCGGAAGACG 5'
```

where X represents guanine (G) or 7,8-dihydro-8-oxoguanine (GO). The synthesis and annealing of the oligonucleotides were as previously described (Lu et al., 1995; McGoldrick et al., 1995). The annealed duplexes were radiolabeled at the 3′ end of the upper strand with Klenow fragment of DNA polymerase I for 30 min at 25 °C in the presence of [α-^{32}P]-dCTP (50 μCi at 3000 Ci/mmol), 20 μM dTTP, and 20 μM dGTP (Maniatis et al., 1982). The resulting blunt-ended duplex DNAs were 20 base pairs in length. The reaction mixture was then passed through a Sephadex G-25 QuickSpin column (Boehringer Mannheim).

MutY Binding Assay. The binding of MutY to A/G or A/GO-containing 20-mer oligonucleotides was assayed by gel retardation. The standard reaction mixture contained 20 mM Tris-HCl (pH 7.6), 80 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 2.9% glycerol, 20 ng of poly(dI-dC), and 1.8 fmol of labeled 20 bp oligonucleotide in a total volume of 20 μL. MutY protein, diluted in a buffer containing 20 mM potassium phosphate (pH 7.4), 1.5 mM dithiothreitol, 0.1 mM EDTA, 50 mM KCl, 200 μg/mL bovine serum albumin,

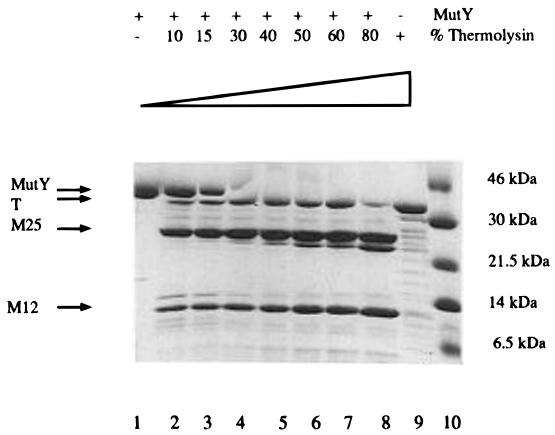

FIGURE 1: Limited proteolytic digestion of MutY. MutY was incubated with increasing amounts of thermolysin for 30 min at room temperature, and the products were analyzed by 17% SDS/PAGE (Laemmli, 1970). The relative amounts of thermolysin (w:w) were as follows: 10% (lane 2), 15% (lane 3), 30% (lane 4), 40% (lane 5), 50% (lane 6), 60% (lane 7), and 80% (lane 8). Intact MutY and thermolysin used were in lanes 1 and 9, respectively. Molecular mass standards (Rainbow, Amersham) were run in lane 10. The positions of MutY, thermolysin (T), and the main proteolytic fragments, 25 kDa (M25) and 12 kDa (M12), are indicated by arrows.

and 50% glycerol, was added to the reaction mixture and incubated at 37 °C for 30 min. Protein−DNA complexes were analyzed on 8% polyacrylamide gels in 50 mM Tris-borate (pH 8.3) and 1 mM EDTA.

MutY Cleavage Assay. The cleavage activity of MutY, which is the combined action of the glycosylase and AP lyase activities, was assayed as described previously (Lu & Chang, 1988a). 3′-End-labeled 20 bp oligonucleotides (1.8 fmol) were incubated with various concentrations of MutY protein as described in the binding assay, except that poly(dI-dC) was omitted from the reaction. The reaction products were fractionated on 14% polyacrylamide DNA-sequencing gels.

Formation of the MutY−DNA Covalent Complex (Borohydride-Mediated Trapping Assay). Reactions were carried out as described in the MutY cleavage assay except that the reactions were performed in the presence of 100 mM NaCl or 100 mM $NaBH_4$. A $NaBH_4$ stock solution (1 M) was freshly prepared immediately prior to use. After incubation at 37 °C for 30 min, the products were separated on a 12 or 15% polyacrylamide gel in the presence of SDS (SDS/PAGE) according to Laemmli (1970), and the gel was dried and autoradiographed.

RESULTS

Proteolytic Digestion. MutY was treated, in preliminary experiments, with trypsin, subtilisin, papain, and thermolysin to identify conditions under which stable fragments of the protein could be produced. We found thermolysin to be especially effective in producing a limited number of fragments, and these fragments were substantially resistant to further digestion. All subsequent experiments were therefore done with thermolysin.

The thermolysin concentration was varied from 10 to 80% (w:w) of MutY, and digestion was allowed to proceed for 30 min. At higher thermolysin concentrations, the 39 kDa MutY band disappeared and two major bands appeared, one around 25 kDa and one around 12 kDa (Figure 1). At a thermolysin:MutY ratio of 50% (w:w), these two fragments were stable for at least 4 h (data not shown). At concentrations greater than 50%, additional bands appeared that were only slightly smaller than the 25 kDa product (Figure 1, lanes 6−8).

The 25 and 12 kDa fragments were identified by N-terminal sequence analysis and mass spectrometry. The 12 kDa fragment (M12) represents most of the C-terminal third of MutY, starting at residue 227 and ending approximately 20 amino acids before the C terminus of the protein (Figure 2). The 25 kDa fragment (M25) has the same N terminus as the intact MutY protein. The size of M25, as determined by mass spectrometry, is consistent with the C terminus being at the cleavage site between Q226 and T227 (Figure 2). We cannot rule out a second cleavage site that might produce a fragment one or two residues shorter, however (Figure 2).

Amino acid sequencing of MutY had previously shown that the protein starts with the initiator methionine that is expected from the gene sequence (Michaels et al., 1990; Tsai-Wu et al., 1991, 1992). These earlier sequencing experiments were performed on protein purified from cells that contained only the chromosomal allele of *mutY*. However, amino acid sequencing of the MutY protein used in the experiments discussed here, which was produced by expression of the cloned gene, showed that about 70% of the protein had an additional threonine residue preceding the initiator methionine. There is a threonine codon just before the usual initiator methionine AUG codon, and inspection of the gene sequence suggests that translation in this case is probably initiated from a GUG codon immediately preceding the threonine codon. The methionine residue that initiates translation at this GUG is evidently removed post-translationally. Ample precedents exist both for translation initiation from GUG and for post-translational removal of initiator methionine residues (Gren, 1984; Hirel et al., 1989; Ringquist et al., 1992). The preference for GUG initiation versus AUG initiation just two codons away may be a consequence of the distance from the Shine−Dalgarno sequence in the cloned gene as compared to that in the chromosomal allele.

Purification of M25. M25 was readily purified from other proteolytic products using a heparin column. The column was washed with 0.1 M KCl protein buffer followed by elution with buffer containing 0.3−0.6 M KCl. Thermolysin and M12 were not bound to the heparin column at 0.1 M KCl (Figure 3, lane 4). M25 eluted at 0.4 M KCl (Figure 3, lane 5). Proteolytic fragments slightly smaller than M25 became apparent at higher thermolysin concentrations (Figure 1, lanes 5−8), but these had different properties on the heparin column and could be readily purified away from M25 (Figure 3, lane 4).

A/G Mismatch Binding Activity of M25. Heparin fractions were assayed for binding to mismatch-containing DNA to determine which MutY proteolytic fragments, if any, retained activity. The fraction that was not bound to the heparin column at 0.1 M KCl contained thermolysin, M12, and fragments slightly smaller than M25 (Figure 3, lane 4) and had no binding activity to mismatch-containing DNA (data not shown). M25 eluted at 0.4 M KCl from the heparin column (Figure 3, lane 5) and retained specific binding to A/G-containing DNA (Figure 4, lanes 6−10). Gel retardation binding experiments demonstrated that the amount of complex formed between M25 and A/G-mismatched DNA was about 4-fold lower than that for intact MutY at the same

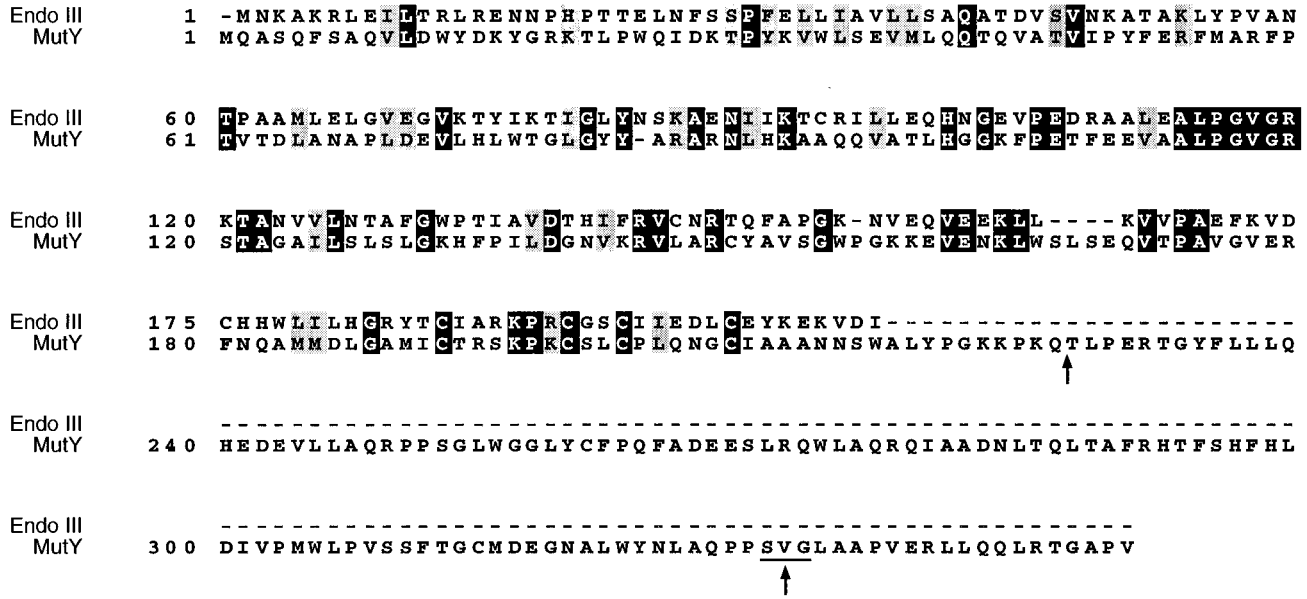

FIGURE 2: Alignment of E. coli MutY and endo III sequences and identified thermolysin cleavage sites. Sequences are *E. coli* MutY (accession no. P17802; Michaels et al., 1990; Tsai-Wu et al., 1991) and endo III (accession no. J02857; Asahara et al., 1989). Black boxes indicate identical amino acids, and gray boxes indicate conserved amino acids. The comparison was done by using Clustal W (Thompson et al., 1994). Arrows indicate the potential cleavage sites of thermolysin. Thermolysin cleaved MutY between Q226 and T227. Mass spectrometry was consistent with this cleavage site being the C-terminal end of the 25 kDa fragment, although the accuracy of this mass determination was such that a second cleavage site at residue 224 or residue 225 could not be ruled out. The N-terminal end of the 12 kDa proteolytic fragment was T227; the C-terminal end was, by mass spectrometry, equally likely to be S330 or V331. The underscore centered at that position indicates a 1% error in mass determination.

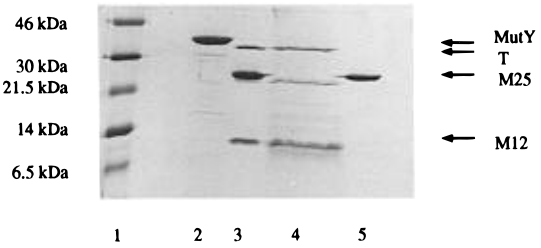

FIGURE 3: Purification of the 25 kDa proteolytic fragment. M25 was purified on a heparin column, after proteolytic digestion by 50% (w:w) thermolysin for 4 h at room temperature. The products were analyzed on a 17% SDS/PAGE (Laemmli, 1970): lane 3, proteolytic digestion products; lane 4; 0.1 M KCl wash of heparin column; and lane 5, 0.4 M KCl elution. MutY was digested to more than 99% completion by thermolysin as judged by scanning the Coomassie blue-stained gel by a densitometer. Molecular mass standards (Rainbow, Amersham) were run in lane 1, and intact MutY was run in lane 2. The positions of MutY, thermolysin (T), and the main proteolytic fragments, 25 kDa (M25) and 12 kDa (M12), are indicated by arrows.

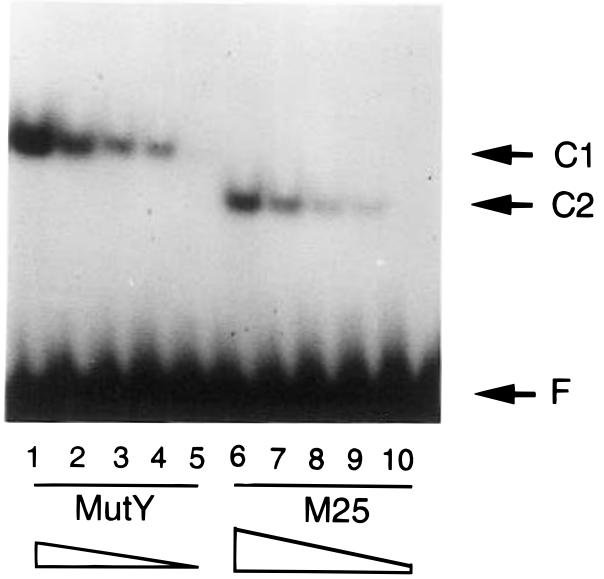

FIGURE 4: Binding of A/G-containing oligonucleotides by the 0.4 M KCl elution fraction from the heparin column. The 0.4 M KCl elution fraction from the heparin column, which was 99% M25, was assayed for binding activity with A/G-containing 20-mer oligonucleotides (lanes 6−10) and compared to intact MutY (lanes 1−5) for different enzyme concentrations at 37 °C for 30 min. The enzyme amounts used were as follows: 288 fmol (lane 6), 144 fmol (lanes 1 and 7), 72 fmol (lanes 2 and 8), 36 fmol (lanes 3 and 9), 18 fmol (lanes 4 and 10), and 9 fmol (lane 5). The products were analyzed on an 8% native gel. The positions of the MutY−DNA complex (C1), the M25−DNA complex (C2), and free DNA (F) are indicated by arrows.

protein concentration (see, for example, lanes 1 and 7 in Figure 4).

A/GO Mismatch Binding Activity of M25. Although the binding activity of M25 with A/G-containing DNA was only slightly lower than that of intact MutY, its binding affinity for A/GO-containing DNA was substantially reduced. In gel shift experiments, the binding affinities of the two proteins could be directly compared in the same binding reaction since the DNA−M25 complex had a faster mobility than the DNA−MutY complex. As shown in Figure 5 (lanes 4−9), when A/GO-mismatched DNA was incubated with an M25 preparation, almost all the A/GO-containing DNA that was shifted appeared to be due to intact MutY binding. As judged by scanning a Coomassie blue-stained gel with a densitometer, this M25 preparation contained a small amount (about 1%) of contaminating full-length protein (Figure 3, lane 5). This small contamination in our preparation of M25 fortuitously provided a useful internal control. Normalization of the bound DNA to the approximate amount of each protein present in the M25 preparation suggested that full-length MutY bound A/GO-mismatched DNA about 2 orders of magnitude more tightly than does the 25 kDa fragment

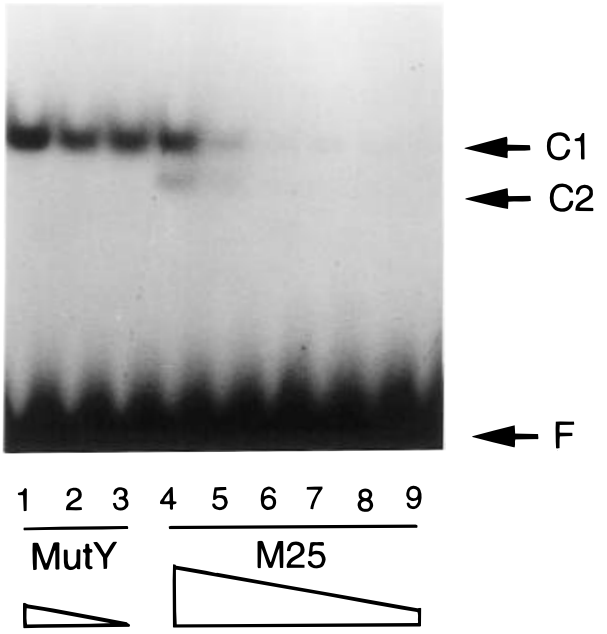

FIGURE 5: Binding of A/GO-containing oligonucleotides by the 0.4 M KCl elution fraction from the heparin column. Intact MutY (lanes 1−3) or the 0.4 M KCl elution fraction from the heparin column (lanes 4−9) was assayed for binding activity with A/GO-containing 20-mer oligonucleotides at 37 °C for 30 min. The enzyme amounts used were as follows: 1.8 fmol (lanes 1 and 9), 0.9 fmol (lane 2), 0.45 fmol (lane 3), 72 fmol (lane 4), 36 fmol (lane 5), 18 fmol (lane 6), 9 fmol (lane 7), and 4.5 fmol (lane 8). The products were analyzed on an 8% native gel. The positions of the MutY−DNA complex (C1), the M25−DNA complex (C2), and free DNA (F) are indicated by arrows.

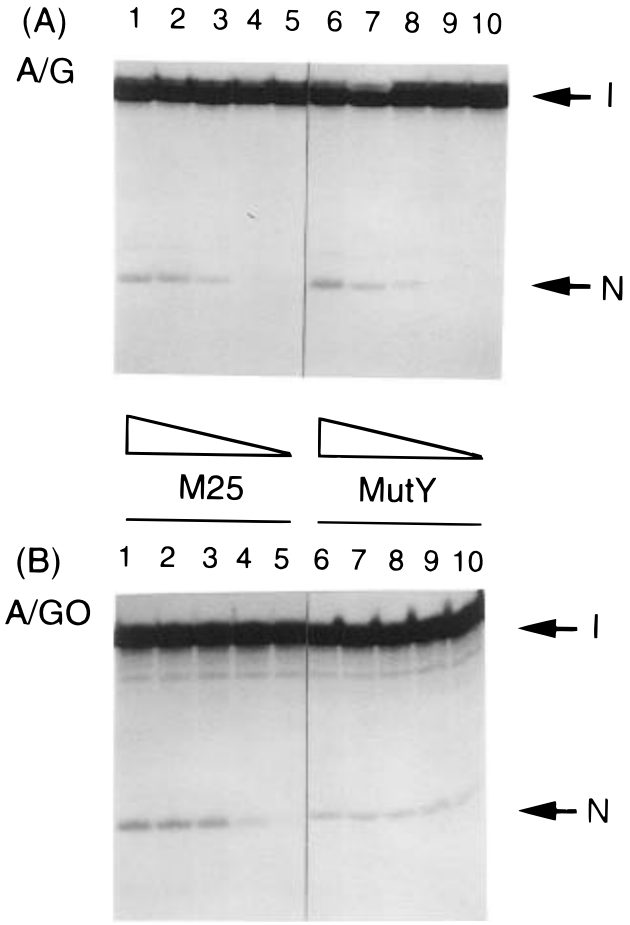

FIGURE 6: A/G and A/GO mismatch cleavage activity of M25. 3′-End-labelled A/G- or A/GO-containing (panels A and B, respectively) 20-mer oligonucleotides (1.8 fmol) were incubated for 30 min at 37 °C in the presence of MutY (lanes 6−10) or the 0.4 M KCl elution fraction from the heparin column (lanes 1−5), to assay for glycosylase/AP lyase activity. The products were analyzed on a 14% DNA sequencing gel. The enzyme amounts used were as follows: 72 fmol (lanes 1 and 6), 36 fmol (lanes 2 and 7), 18 fmol (lanes 3 and 8), 9 fmol (lanes 4 and 9), and 4.5 fmol (lanes 5 and 10). The positions of the intact oligonucleotide (I) and the nicked product (N) are indicated by arrows.

(Figure 8). In contrast, the tight binding of M25 to A/G-mismatched DNA meant that the contaminating MutY−A/G−DNA complex could barely be detected (Figures 4 and 8).

Biochemical Activities of M25 with A/G and A/GO Mismatches. The combined effects of the DNA glycosylase and AP lyase activities were assayed by monitoring the nicking of a 20-mer DNA substrate containing an A/G or A/GO mismatch. Under standard assay conditions using A/G-mismatched 20-mer substrate DNA, the amount of nicked DNA produced by M25 was similar to that produced by intact MutY at the same protein concentration (Figure 6A). This level of cleavage activity was consistent with our observation that the binding affinity of the truncated protein and A/G mismatches was only a few-fold lower than that of the wild type.

The apparent cleavage activity of M25 with A/GO mismatches was also quite similar to that of the full-length protein (Figure 6B), which was surprising given the decreased binding affinity for A/G mismatches. Because the 1% MutY contamination of the M25 preparation constituted the majority of the observed binding complexes, however, it was unclear to what extent the cleavage products were produced by intact MutY or M25. Unlike the protein−DNA complexes that were observed in the gel shift experiments, the DNA cleavage products formed by reaction with M25 and by MutY are indistinguishable. Although, in principle, it is possible from kinetic analyses to establish an upper bound for the contribution of the contaminating full-length protein, in practice, the enzyme has low turnover under standard assay conditions and cannot be used for this purpose. As an alternative means of analyzing the reactivity of M25 and full-length protein with A/G- and A/GO-mismatched DNA, we used $NaBH_4$ to trap a presumptive catalytic intermediate. A number of DNA glycosylases that have AP lyase activity are believed to act through a Schiff base covalent intermediate between the sugar and an amine nucleophile, and reduction with $NaBH_4$ stably cross-links the protein to DNA (Sun et al., 1995). We recently showed that MutY can also be cross-linked with DNA by $NaBH_4$ (Lu et al., 1996). Thus, we used this ability to be cross-linked as a measure of the reactivity of M25 and full-length protein with A/G- and A/GO-mismatched DNA (Figure 7).

M25 formed covalent intermediate complexes with A/G- and A/GO-mismatched DNA (Figure 7, C2) that were smaller in size than complexes with intact MutY (Figure 7, C1); the contributions of M25 and intact MutY were therefore resolved electrophoretically. The ratio of the amount of covalent complexes of C2 to C1 with A/G-mismatched DNA (Figure 7, lane 3) was nearly proportional to the ratio of the concentrations of M25 and intact MutY in the heparin fraction (Figure 3, lane 3). To the extent that $NaBH_4$ cross-linking represents the trapping of a catalytic intermediate, this assay showed that the truncated protein

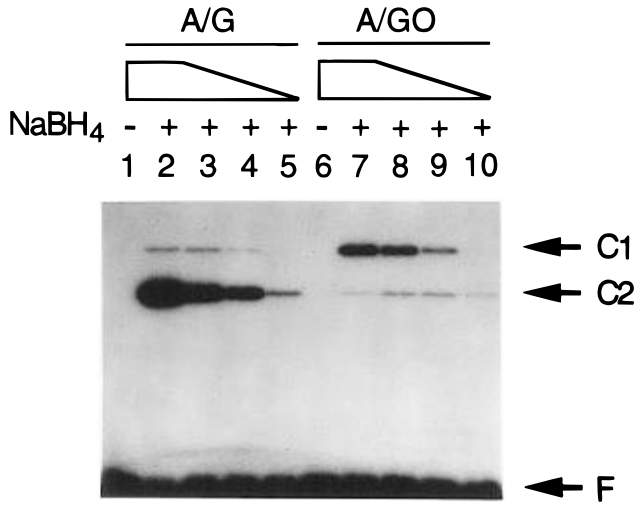

FIGURE 7: Formation of covalent complexes of M25 and A/G- or A/GO-containing DNA in the presence of $NaBH_4$. Oligonucleotide substrates (3′-end-labeled 20-mer, 1.8 fmol) containing A/G (lanes 1−5) or A/GO (lanes 6−10) mismatches were incubated with the 0.4 M KCl elution fraction from the heparin column in the presence of 0.1 M NaCl or $NaBH_4$. Lanes 1 and 6 show the results of reaction mixtures containing 288 fmol of M25 and A/G- or A/GO-mismatched DNA, respectively, in the presence of 0.1 M NaCl. Lanes 2−5 and 7−10 show the results of reaction mixtures containing a decreasing amount of M25 in the presence of 0.1 M $NaBH_4$. The amounts of M25 used were 288 fmol (lanes 2 and 7), 72 fmol (lanes 3 and 8), 18 fmol (lanes 4 and 9), and 4.5 fmol (lanes 5 and 10). The products, after heating at 90 °C for 2 min, were electrophoresed in a 12% SDS−polyacrylamide gel followed by autoradiography. The positions of the free oligonucleotide (F), the covalent MutY−DNA complex (C1), and the covalent M25−DNA complex (C2) are indicated by arrows. Unincorporated [α-^{32}P]-dCTP migrated very close to the free oligonucleotide.

possessed considerable A/G glycosylase activity, consistent with the cleavage assay (Figure 6A). In contrast, the ratio of the amount of covalent complexes of C2 to C1 with A/GO-mismatched DNA (Figure 7, lanes 7−10) was out of proportion to the ratio of the concentrations of M25 and intact MutY in the heparin fraction (Figure 3, lane 5). The ratio of MutY-cross-linked DNA (band C1) to M25-cross-linked DNA (band C2) was concentration-dependent, with the amount of full-length cross-linked product decreasing more rapidly with decreasing protein concentration than was the case for the truncated protein (Figure 7, lanes 7−9). This may be related to the fact that the contaminating full-length protein became substoichiometric to the DNA over this concentration range, but the exact reason for these differences is not understood.

The results of the binding and $NaBH_4$-induced cross-linking assays (at an enzyme to DNA molar ratio of 40) are summarized in Figure 8. Using A/G-mismatched DNA and a 0.4 M KCl eluent from the heparin column, the covalent M25−DNA complex represented 93% of the total covalent protein−DNA complex, the remainder being MutY−DNA. Using A/GO-mismatched DNA, only 15% of the total covalent complex formed was M25−DNA. These results were in agreement with the observed reduced binding affinity of M25 and A/GO-containing DNA.

DISCUSSION

Two relatively stable domains are produced by digesting MutY with thermolysin. The larger domain (M25) has the same N termini as our unproteolyzed MutY preparations and

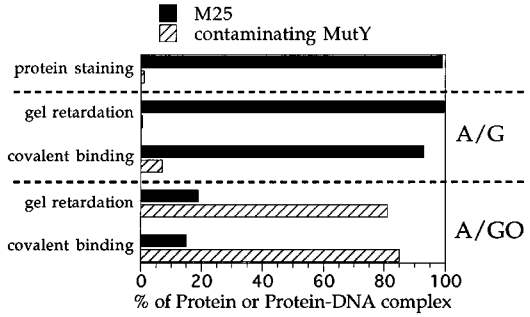

FIGURE 8: Summary of the percentages of intact MutY and M25 and the percentages of A/G- or A/GO-mismatched DNA that were shifted by these two proteins in the heparin 0.4 M elution fraction. As judged from scanning of a Coomassie blue-stained SDS/PAGE with a densitometer, the 0.4 M KCl elution fraction from the heparin column contained about 1% full-length MutY and 99% M25 (data derived from lane 5 of Figure 3). From gel retardation analysis of this fraction, the M25−A/G complex (data derived from lane 8 of Figure 4) and the M25−A/GO complex (data derived from lane 4 of Figure 5) comprised nearly 100 and 19%, respectively, of the total protein−DNA complexes. Covalent M25−A/G complex (data derived from lane 3 of Figure 7) and M25−A/GO complex (data derived from lane 8 of Figure 7) comprised 93 and 15%, respectively, of the total covalent protein−DNA complexes in the heparin 0.4 M elution fraction.

probably has its C-terminal end at Q226 (Figure 2). N-Terminal sequencing of the 12 kDa proteolytic fragment suggests that there is a cleavage site between Q226 and T227, but from mass spectrometry of M25, we cannot rule out a C-terminal end one or two amino acids shorter. About 70% of the MutY preparation used in these experiments contains an extra threonine residue preceding the normal initiator methionine; however, since the M25 proteolytic fragment also contains this same mixture of N termini, our conclusions about the relative activities of MutY and M25 are not affected by the existence of this residue.

The binding affinity and cleavage activity of M25 are comparable to those of intact MutY when A/G-mismatched substrates are used. It appears that the N-terminal 25 kDa domain of MutY has all of the structural and functional features required to specifically recognize and cleave mispaired adenine from A/G-containing DNA. Our results are consistent with those of Manuel et al. (1996), who found that a tryptic fragment of MutY consisting of residues 1−225 retained A/G substrate binding. We suspect that M25, which is nearly the same length as endo III, is the smallest thermolysin digestion fragment capable of glycosylase/AP lyase activity, since the products slightly smaller than M25 that appear after more extensive thermolysin digestion do not bind to a heparin column (Figure 3) and do not cause gel retardation of mismatched DNA (data not shown).

M25 is 23% identical to endo III (Figure 2) (Michaels et al., 1990). A model for DNA binding by endo III has been proposed (Kuo et al., 1992; Thayer et al., 1995), and many of the sequence features that are presumed to be important in that model are conserved in MutY. In particular, residues 113-119 in the endo III crystal, which adopt a β-hairpin and bind thymine glycol (Kuo et al., 1992), are highly conserved between endo III and MutY. Since endo III and MutY have very different substrate specificities, binding to this highly conserved region presumably reflects interactions that are common to the substrates of the two proteins. Residues in the vicinity of the iron−sulfur cluster have also been proposed to be involved in DNA binding, and the conserva-

tion of residue spacing for the cysteines that ligate the cluster in endo III and MutY suggests that this structure is very similar in the two proteins (Thayer et al., 1995).

Full-length MutY binds A/GO-containing DNA about 80-fold better than it does A/G-containing DNA (Lu et al., 1995); however, MutY cleaves A/GO-containing DNA about 3-fold less efficiently than it does A/G-containing DNA (Lu et al., 1996). The surprising result of this study is that the strong binding preference for A/GO mismatches over A/G mismatches is lost when the C-terminal 124 amino acids of MutY are removed. M25 also was much more poorly cross-linked to A/GO-mismatched substrates than was intact MutY in the presence of borohydride (Figure 7), consistent with reduced glycosylase activity toward this substrate. Further experiments will be required to reconcile this result with similar reactivity toward A/G- and A/GO-containing substrates as measured by the cleavage assay (Figure 6).

Our discovery that the C-terminal domain of MutY confers A/GO binding specificity is especially interesting in light of the observation that MutY efficiently competes with MutM for binding to A/GO and AP/GO mismatches. This binding may serve to prevent removal of 8-oxoguanine by MutM before the MutY-initiated repair of the mismatched adenine is complete. We have shown that the specificity for A/GO binding by MutY can be mainly ascribed to the C-terminal domain. In the absence of the C-terminal domain, the binding preference for A/GO over A/G is lost; however, the 12 kDa fragment does not detectably bind mismatched DNA by itself (data not shown). Since M25 binds A/G with an affinity only a few-fold lower than full-length MutY, most of the substrate binding energy likely comes from interactions between DNA and the N-terminal domain. The C-terminal domain could affect A/GO recognition by altering the conformation of the endo III-like portion of MutY, or perhaps it contains additional functional groups that directly interact with A/GO mismatches. Structural studies will likely be required to resolve this issue.

ACKNOWLEDGMENT

Amino acid sequencing was performed at the Protein Facility of the Johns Hopkins School of Medicine. Mass spectral measurements were carried out at the Middle Atlantic Mass Spectrometry Facility at Johns Hopkins University, an NSF Shared Instrumentation Facility.

REFERENCES

Asahara, H., Wistort, P. M., Bank, J. F., Bakerian, R. H., & Cunningham, R. P. (1989) *Biochemistry 28*, 4444−4449.
Au, K. G., Cabrera, M., Miller, J. H., & Modrich, P. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9163−9166.
Au, K. G., Clark, S., Miller, J. H., & Modrich, P. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8877−8881.
Bradford, M. (1976) *Anal. Biochem. 72*, 248−254.
Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., & Loeb, L. A. (1991) *J. Biol. Chem. 267*, 166−172.
Chetsanga, C. J., & Lindahl, T. (1979) *Nucleic Acids Res. 6*, 3673−3683.
Cunningham, R. P., Asahara, H., Bank, J. F., Scholes, C. P., Salerno, H. C., Surerus, K., Munck, E., McCraken, J., Peisach, J., & Emptage, M. H. (1989) *Biochemistry 28*, 4450−4455.
Fu, W., O'Handley, S. F., Cunningham, R. P., & Johnson, M. K. (1992) *J. Biol. Chem. 267*, 16135−16137.
Gren, E. G. (1984) *Biochimie 66*, 1−29.
Grollman, A. P., & Moriya, M. (1993) *Trends Genet. 9*, 246−249.
Hirel, P.-H., Schmitter, J., Dessen, P., Fayat, G., & Blanquet, S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8247−8251.
Kuo, C. F., McRee, D. E., Fisher, C. L., O'Handley, S. F., Cunningham, R. P., & Tainer, J. A. (1992) *Science 258*, 434−440.
Laemmli, U. K. (1970) *Nature 227*, 680−685.
Lu, A., & Chang, D. Y. (1988a) *Cell 54*, 805−812.
Lu, A., & Chang, D. Y. (1988b) *Genetics 118*, 593−600.
Lu, A., Tsai-Wu, J. J., & Cillo, J. (1995) *J. Biol. Chem. 270*, 23582−23588.
Lu, A., Yuen, D. S., & Cillo, J. (1996) *J. Biol. Chem. 271*, 24138−24143.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Plainview, NY.
Manuel, R. C., Czerwinski, E. W., & Lloyd, R. S. (1996) *J. Biol. Chem. 271*, 16218−16226.
McGoldrick, J. P., Yeh, Y. C., Solomon, M., Essigmann, J. M., & Lu, A. (1995) *Mol. Cell Biol. 15*, 989−996.
Michaels, M. L., & Miller, J. H. (1992) *J. Bacteriol. 174*, 6321−6325.
Michaels, M. L., Pham, L., Nghiem, Y., Cruz, C., & Miller, J. H. (1990) *Nucleic Acids Res. 18*, 3841−3845.
Michaels, M. L., Tchou, J., Grollman, A. P., & Miller, J. H. (1992) *Biochemistry 31*, 10964−10968.
Moriya, M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1122−1126.
Moriya, M., Ou, C., Bodepudi, V., Johnson, F., Takeshita, M., & Grollman, A. P. (1991) *Mutat. Res. 254*, 281−288.
Nghiem, Y., Cabrera, M., Cupples, C. G., & Miller, J. H. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2709−2713.
Radicella, J. P., Clark, E. A., & Fox, M. S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9674−9678.
Radicella, J. P., Clark, E. A., Chen, S., & Fox, M. S. (1993) *J. Bacteriol. 175*, 7732−7736.
Ringquist, S., Shinedling, S., Barrick, D., Green, L., Binkley, J., Stormo, G. D., & Gold, L. (1992) *Mol. Microbiol. 6*, 1219−1229.
Shibutani, S., Takeshita, M., & Grollman, A. P. (1991) *Nature 349*, 431−434.
Slupska, M. M., Baikalov, C., Luther, W. M., Chiang, J.-H., Wei, Y.-F., & Miller, J. H. (1996) *J. Bacteriol. 178*, 3885−3892.
Speicher, D. W. (1989) in *Techniques in Protein Chemistry* (Hugli, T., Ed.) pp 24−35, Academic Press, San Diego.
Su, S. S., Lahue, R. S., Au, K. G., & Modrich, P. (1988) *J. Biol. Chem. 263*, 6829−6835.
Sun, B., Latham, K. A., Dodson, M. L., & Lloyd, R. S. (1995) *J. Biol. Chem. 270*, 19501−19508.
Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Grollman, A. P., & Nishimura, S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4690−4694.
Thayer, M. M., Ahern, H., Cunningham, R. P., & Tainer, J. A. (1995) *EMBO J. 14*, 4108−4120.
Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
Tsai-Wu, J. J., & Lu, A. (1994) *Mol. Gen. Genet. 244*, 444−450.
Tsai-Wu, J. J., Radicella, J. P., & Lu, A. (1991) *J. Bacteriol. 173*, 1902−1910.
Tsai-Wu, J. J., Liu, H. F., & Lu, A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8779−8783.
Wood, M. L., Dizdaroglu, M., Gajewski, E., & Essigmann, J. M. (1990) *Biochemistry 29*, 7024−7032.

BI960843W