The three-dimensional structure of an oxidative-damaged DNA repair enzyme, MutM, from Extremely Thermophile

Doctoral Thesis

January 2000

Mitsuaki Sugahara

Department of Biology, Graduate School of Science

Osaka University

CONTENTS

ABBREVIATIONS

Fpg	Formamidopyrimidine DNA glycosylase
AP	apurinic/apyrimidinic
GO	8-oxoguanine
AO	8-oxoadenine
Fapy	Formamidopyrimidine
5OHC	5-hydroxycytosine
5OHU	5-hydroxyuracil
EndoIII	endonuclease III
T4 EndoV	bacteriophage T4 endonuclease V
AtMMH	*Arabidopsis thaliana* MutM homologues
HhH	helix-hairpin-helix
H2TH	helix-two turn-helix
EndoVIII	endonuclease VIII
PEG	polyethyleneglycol
CHES	2-(cyclohexylamino)ethanesulfonicacid
XAFS	X-ray absorption fine structure
MAD	multiwavelength anomalous diffraction
NCS	non crystallographic symmetry
MD	molecular dynamics
dsDNA	double-stranded DNA

ABSTRUCT

MutM (Fpg) protein is a trifunctional DNA base-excision-repair enzyme, which removes a wide range of oxidatively damaged bases (N-glycosylase activity) and cleaves both the 3'- and 5'-phosphodiester bonds of the resulting apurinic/apyrimidinic site (AP-lyase activity). The crystal structure of MutM from an extreme thermophile, *Thermus thermophilus* HB8, was determined at 1.9 Å resolution. MutM was composed of two distinct and novel domains separated by a large, positively charged and sequence-conserved cleft. The locations of the proposed active-site residues and the two DNA-binding motifs suggest that the oxidized base is flipped-out from double-stranded DNA and excised by a mechanism similar to that of bifunctional base-excision-repair enzymes. On the basis of the three-dimensional structure and previous biochemical results, a reaction mechanism for MutM is proposed.

INTRODUCTION

In aerobic organisms, cellular DNA is frequently damaged by reactive oxygen species from aerobic energy metabolism or oxidative stress. Reactive oxygen accelerates the rate of spontaneous mutation and has therefore been implicated in aging and in the pathogenesis of disease, including cancer (Breimer, 1990; Ames et al., 1995). In particular, the 8-oxoguanine (GO) lesion is one of the most stable products of oxidative DNA damage (Dizdaroglu, 1985). As GO is able to pair with adenine as well as cytosine, it causes a G:C to T:A transversion (Wood *et al.*, 1990; Shibutani et al., 1991). In *Escherichia coli*, the GO repair system, consisting of the MutM protein [EC 3.2.2.23], also called Fpg protein, accompanied by MutY and MutT, removes the GO lesion pair and prevents the mutation (Michaels et al., 1992)(Figure 1).

MutM was originally isolated and characterized from *E. coli* as a DNA glycosylase that removes 2,6-diamino-4-hydroxy-5-N-methylformamido-pyrimidine (Fapy), the imidazole ring-open form of N^7-methylguanine (Boiteux et al., 1987). More recently, MutM has been reported to bind to a wide range of oxidatively damaged bases (see Figure 19), such as GO paired to cytosine (Tchou et al., 1991), formamidopyrimidine (FapyG or FapyA) (Boiteux et al., 1992) and 5-hydroxycytosine (5OHC) (Hatahet et al., 1994), and to apurinic/apyrimidinic (AP) sites (Castaing et al., 1992). MutM shows at least

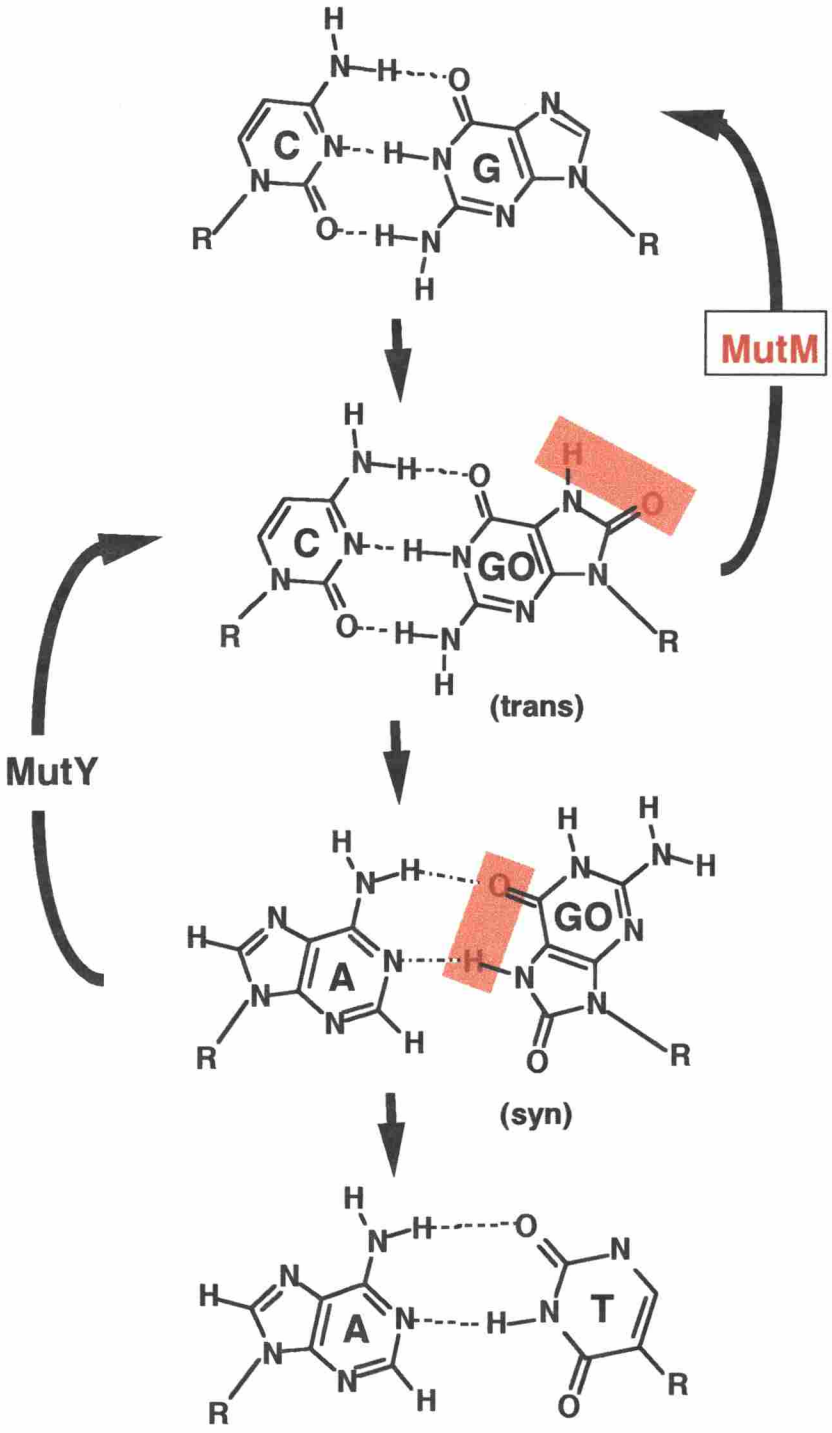


Figure 1. C:G → A:T transversion caused by GO base, and the role of MutM and MutY. Oxidative processes can lead to GO lesions in DNA. The MutM protein removes GO lesions and subsequent completion of base-excision repair can restores the original G:C base pair. If GO lesion is not removed by MutM, MutY removes the misincorporated adenine from the A:GO mispairs that result from error-prone replication past the GO lesion. Repair polymerases are less error-prone during translesion synthesis and can lead to a C:GO pair – a substrate for MutM (Michaels et al. , 1992)

three enzymatic activities: (1) a DNA glycosylase activity that excises various damaged bases from DNA to produce an aldehydic abasic site, (2) an AP lyase activity that cleaves the 3'-phosphodiester bond at AP sites through β-elimination, and (3) a similar AP lyase activity that cleaves the 5'-phosphodiester bond through δ-elimination (O'Connor et al., 1989; Boiteux et al., 1990; Bhagwat and Gerlt, 1996)(see Figure 2). In terms of the mechanism of MutM glycosylase/AP lyase activity, the initial step has been proposed to be formation of a covalent Schiff base intermediate by nucleophilic attack of the imino group of Pro1 on deoxyribose C1' of the damaged site (see Figure 22) (Tchou and Grollman, 1995; Bhagwat and Gerlt, 1996; Castaing et al., 1999). This mechanism is common with that of other N-glycosylase/AP-lyase base-excision-repair enzymes such as *E. coli* endonuclease III (EndoIII) (Kow and Wallace, 1987; Purmal et al., 1996) and bacteriophage T4 endonuclease V (T4 EndoV) (Schrock and Lloyd, 1991; Purmal et al., 1996). However, the activity of MutM N-glycosylase/AP-lyase is quite different from that of EndoIII and T4 EndoV, which only cleave the 3'-phosphodiester bond through β-elimination. For EndoIII and T4 EndoV (see Figure 21), the nucleophiles are the N-terminal α-amino groups of Lys120 and Thr2, respectively, and the residues that activate the nucleophile and participate in the β-elimination reaction are Asp138 and Glu23, respectively. In the case of MutM, the Schiff base intermediate can be trapped by reduction with sodium borohydride, allowing identification of the α-

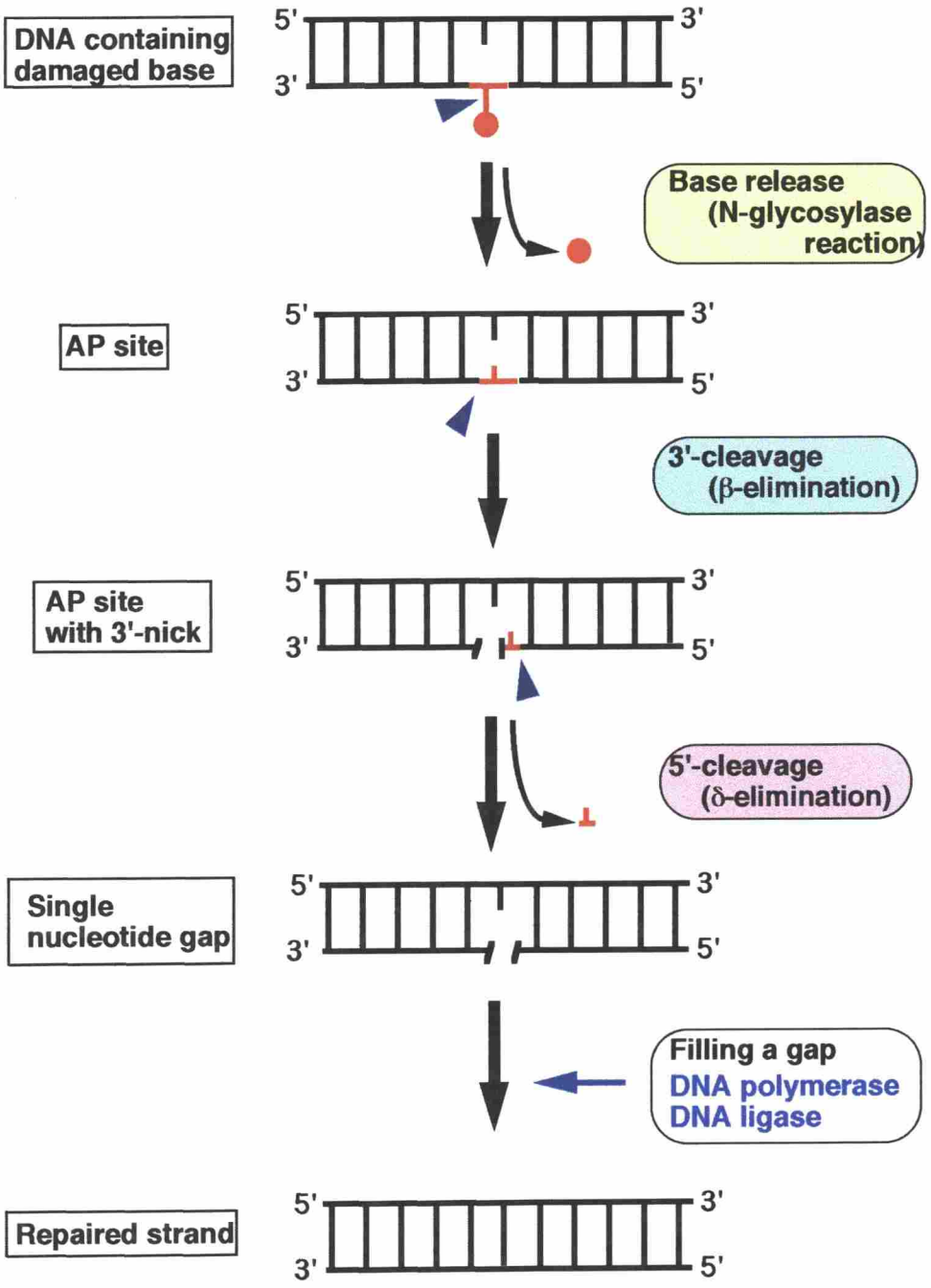


Figure 2. The Function of DNA Glycosylases in Base-Excision Repair.

imino group of Pro1 as the nucleophile (Zharkov et al., 1997). Site-directed mutagenesis of the conserved Lys52 and Lys147 residues suggested their interaction with GO (Rabow and Kow, 1997; Sidorkina and Laval, 1998). However, the residues that activate the nucleophile and participate in the β- and δ-elimination reactions remain to be elucidated.

The *mutM* (*fpg*) gene encoding MutM protein is highly conserved across a wide range of aerobic bacteria. These enzymes, of molecular mass ~30 kDa, possess a conserved N-terminal sequence Pro-Glu-Leu-Pro-Glu-Val-, two lysine residues (Lys52 and Lys147), and a zinc finger motif (-Cys-X_2-Cys-X_{16}-Cys-X_2-Cys-) near the C-terminus (Figure 11). Identification of the trapped reactive intermediate by reduction with sodium borohydride and mutational analyses have shown that the N-terminal proline (Pro1) acts as a nucleophile (Zharkov et al., 1997) and that the two conserved lysine residues (Rabow and Kow, 1997; Sidorkina et al., 1998) and the zinc finger motif (Tchou et al., 1993; O'Connor et al., 1993) participate in the catalytic reaction. The primary structures of bacterial MutM proteins are similar to those of MutM homologues in *Arabidopsis thaliana* (*At*MMH), a eukaryotic plant (Ohtsubo et al., 1998, Murphy et al., 1998), and *E. coli* endonuclease VIII (EndoVIII) (Jiang et al., 1997) (Figure. 11). These enzymes cleave both 5'- and 3'-phosphodiester bonds of the DNA strand by β, δ-elimination, as do MutM proteins (Melamede et al., 1994; Ohtsubo et al., 1998). However, *At*MMH has no zinc-finger motif

and EndoVIII shows a different substrate specificity from that of MutM proteins. Recently, OGG1 proteins from several eukaryotes, including humans, and *Drosophila* ribosomal protein S3 (Krokan et al., 1997) have been characterized as GO:C-specific N-glycosylase/AP-lyases with activities identical to those of MutM.

In recent years, the crystal structures of several DNA glycosylases and their complexes with dsDNA have been solved, and a few common structural features such as a helix-hairpin-helix (HhH) motif and nucleotide flipping have been reported (see Figure 3). The HhH motif was first reported in *E. coli* endonuclease III (EndoIII) (Thayer et al., 1995). This motif was also found in *E. coli* AlkA (Yamagata et al., 1996; Labahn et al., 1996), *E. coli* MutY (Guan et al., 1998) and in OGG1 proteins (Krokan et al., 1997). Comparisons of amino acid sequences and three-dimensional structures among these DNA glycosylases showed that these proteins share common motifs and active-site residues. In spite of the similarity in catalytic properties, it has been difficult to identify the HhH motif, which interacts with the DNA phosphodiester backbone, in MutM. The putative active-site residues of MutM also seem to be different from those of the other glycosylases.

In order to understand the structure-function relationships of MutM at the atomic level, I have determined the X-ray crystallographic structure of the enzyme from an extremely thermophilic bacterium, *Thermus thermophilus* HB8, at 1.9Å resolution, allowing me to propose a reaction mechanism for this enzyme.

Base release (N-glycosylase reaction)

3-methyl aenine DNA glycosylase			Auther	Year	DNA	HhH
AlkA	*E.coli*	2.3Å	Yamagata *et al.*	1996	×	○
		1.8Å	Labahn *et al.*	1996	×	○
AAG	*human*	1.01Å	Lau *et al.*	1998	○	×
Uracil DNA glycosylase						
UDG	*HSV-1*	1.75Å	Sarva *et al.*	1995	×	×
	human	2.0Å	Mol *et al.*	1995	×	×
	human	2.9Å	Slupphaug *et al.*	1996	○	×
	human	1.9Å	Parikh *et al.*	1998	○	×
G:T(U) mismach specific DNA glycosylase						
MUG	*E.coli*	1.8Å	Barnet *et al.*	1998	○	×
		2.85Å	Barnet *et al.*	1999	○	×
GO(G):A mismach adenine DNA glycosylase						
MutY	*E.coli*	1.4Å	Guan *et al.*	1998	×	○

Base release + 3'-cleavage (β-elimination)

Thymine glycol／5-hydroxycytosine DNA glycosylase/AP-lyase						
EndoIII	*E.coli*	1.84Å	Kuo *et al.*	1992	×	○
		1.85Å	Thayer *et al.*	1995	×	○
Pylimidine-dimer specific DNA glycosylase/AP-lyase						
T4 EndoV	*bacteriophage T4*					
		1.6Å	Morikawa *et al.*	1992	×	×
		2.75Å	Vassylyev *et al.*	1995	○	×
		1.45Å	Morikawa *et al.*	1995	×	×

Base release + 3'-cleavage + 5'-cleavage (δ-elimination)

Formamidopylimidine/8-oxoguane DNA glycosylase/AP-lyase						
MutM	*T.thermophilus HB8*					
		1.9Å	Sugahara *et al.*	2000	×	○

Figure 3. The Classification of DNA Glycosylases

EXPERIMENTAL PROCEDURES

Crystallization

Thermus thermophilus HB8 MutM protein was purified as described previously (Mikawa et al., 1998). The purified enzyme in 50 mM Tris-HCl buffer (pH 7.6), containing 5 mM 2-mercaptoethanol, was filtered and concentrated by ultrafiltration up to about 10 mg ml^{-1}. Crystallization of the MutM protein was performed by the hanging-drop vapor-diffusion method at 20°C. The screening for initial crystallization conditions was performed with 50 stock solutions (pH 4.6, 5.5, 6.5, 7.5, 8.5 and 9.5) including precipitants of purified polyethyleneglycols (PEG-400, 2000, 4000 and 8000), salts (ammonium sulfate, sodium chloride, sodium formate and lithium sulfate) and organic solvents (2-methyl-2,4-pentanediol and 2-propanol). MutM protein formed crystals in a few days in solutions containing PEG and sodium formate at pH 9.5. The crystals, however, consisted of many thin plates (Figure 4). By optimizing the crystallization conditions, I was able to grow single crystal regions large enough for an X-ray diffraction experiment (Figure 5). The reservoir solution that gave the best crystals contained 5% (w/v) polyethyleneglycol monomethylether 2000, 100 mM 2-(cyclohexylamino) ethanesulfonic acid (CHES) buffer (pH 9.8), and 5 mM 2-mercaptoethanol.

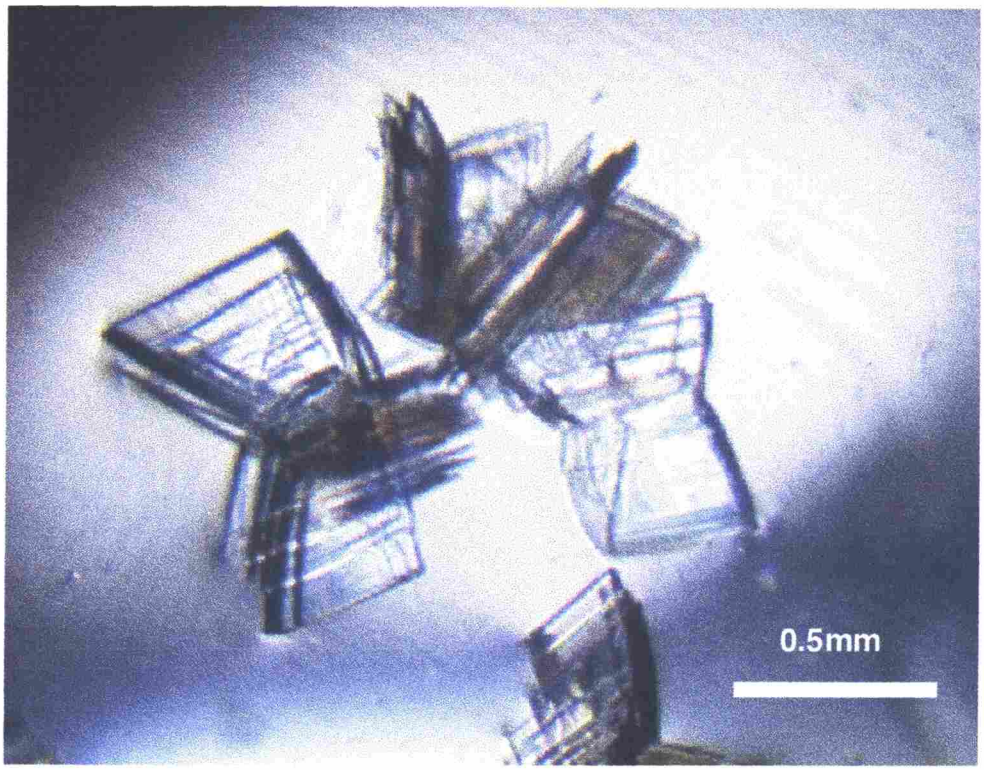


Figure 4. Typical Crystal of the MutM from *T. thermophilus* HB8. The crystal consists of many thin plates. A single plate crystal extruding from these blocks was used for the X-ray diffraction experiment. The reservoir solution that gave the best crystal contained 5% (w/v) polyethyleneglycol monomethylether 2000, 100 mM 2-(cyclohexylamino) ethanesulfonic acid (CHES) buffer (pH 9.8) and 5 mM 2-mercaptoethanol.

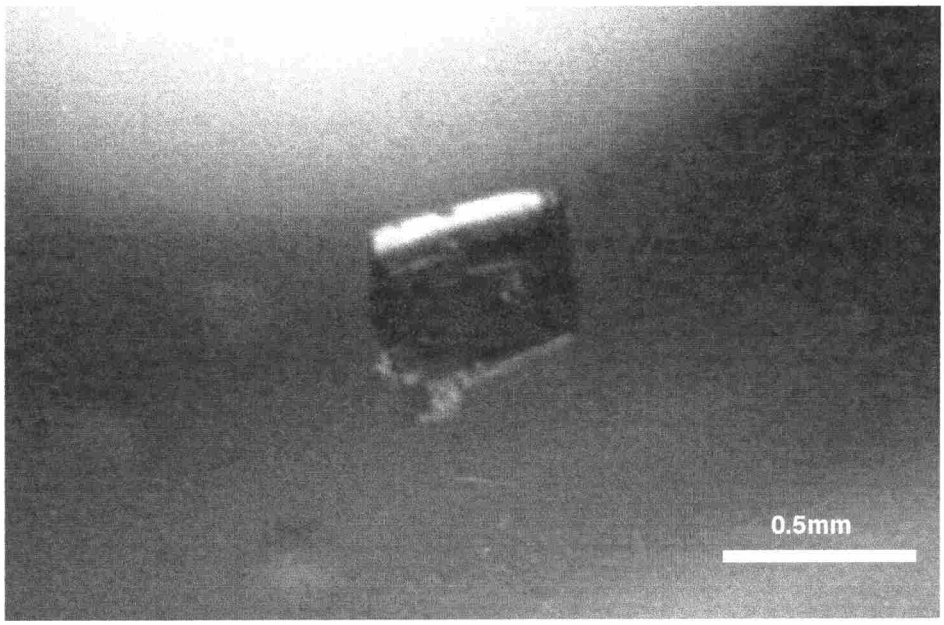


Figure 5. A Single Plate Crystal Used for the X-ray Diffraction Experiment. The approximate size of the plate crystal was 0.5 × 0.3 × 0.02 mm3.

Preliminary X-ray Experiment

Initial intensity data were collected with the laboratory X-ray source at room temperature. A crystal was sealed in a glass capillary tube with a small amount of the mother liquid. X-ray diffraction studies were carried out with a Rigaku R-AXIS IV imaging-plate system using Ni-filtered Cu K_ radiation. The X-rays, generated with a Rigaku rotating anode at 40 kV and 100 mA, were focused with Ni-coated double-bent mirrors. Diffraction data recorded on the imaging plates were processed and scaled with DENZO and SCALEPACK (Otwinowski and Minor, 1997). The crystals belonged to a monoclinic system, with unit-cell dimensions a = 45.4 Å, b = 62.0 Å, c = 99.7 Å and β = 90.8°. Assuming that two MutM protein molecules are present in an asymmetric unit, the V_m value and solvent content were calculated to be 2.35 $Å^3$ Da^{-1} and 48.8%, respectively (Matthews et al., 1968). The systematic absence of diffraction data is compatible with space group $P2_1$. I collected the data set up to 3.2 Å resolution (crystal 1 in Table I). In order to stabilize the crystal against X-ray irradiation, it was mounted in the loop and flash-cooled to 100 K using a Cryostream Cooler (Oxford Cryosystems). Before flash-cooling, the crystal was soaked for several seconds in a cryoprotectant solution prepared by the addition of glycerol to the mother liquid to a final concentration of 20% (v/v). Upon freezing of the crystal, the cell dimensions changed slightly to a = 44.4 Å, b = 61.3 Å, c = 98.3 Å, and β = 91.7°. The diffraction data set of the frozen

crystals were collected up to 3.0 Å resolution (crystal 2 in Table I).

I achieved higher-resolution data collection for frozen crystals using synchrotron radiation at the BL45XU of the SPring-8 (Yamamoto et al, 1998). The diffraction intensities recorded on IP were collected with a Rigaku R-AXIS IV imaging-plate system, and then processed and scaled with DENZO and SCALEPACK (Otwinowski and Minor, 1997). The results of preliminary intensity measurement are presented in Table I (see crystal 3). Table I shows that the crystals yield diffraction data over 2.1 Å resolution. and are suitable for X-ray analysis at atomic resolution.

X-ray Diffraction Data Collection

Multiwavelength anomalous diffraction (MAD) data were collected at two wavelengths near the zinc absorption edge and at 1.04 Å wavelength as remote data (Figure 6) using a single crystal at beamline BL45XU at SPring-8 (Yamamoto et al., 1998) (Figure 7). Prior to data collection, the crystal was soaked for several seconds in a cryoprotectant solution prepared by adding glycerol to the mother liquid to a final concentration of 20% (v/v) and flash-cooled in an N_2 stream at 100 K. All the data were processed and scaled using the programs DENZO and SCALEPACK (Otwinowski and Minor, 1997).

TABLE I. Summary of preliminary intensity measurements of MutM protein crystals.

	Crystal 1	Crystal 2	Crystal 3
Temperature (K)	296	100	100
Wavelength (Å)	1.5418	1.5418	1.04
Oscillation angle (°)	1.5	1.5	2.0
Resolution limit (Å)	3.2	3.0	2.1
Measured reflections	33991	47474	301427
Independent reflections	9319	10715	31112
Completeness (%) †	87.0 (75.3)	72.6 (53.8)	91.5 (85.8)
Mean I/σ(I) †	7.9 (2.5)	6.9 (2.3)	32.3 (8.3)
R_{merge} (%) †‡	9.5 (24.5)	8.7 (25.8)	5.8 (23.6)

†Reflections with F < σ(F) were rejected. Numerals in parentheses indicate the data in the highest-resolution shell.

‡ $R_{merge} = \Sigma_h \Sigma_i |I_{hi} - \langle I_h \rangle| / \Sigma_h \Sigma_i I_{hi}$

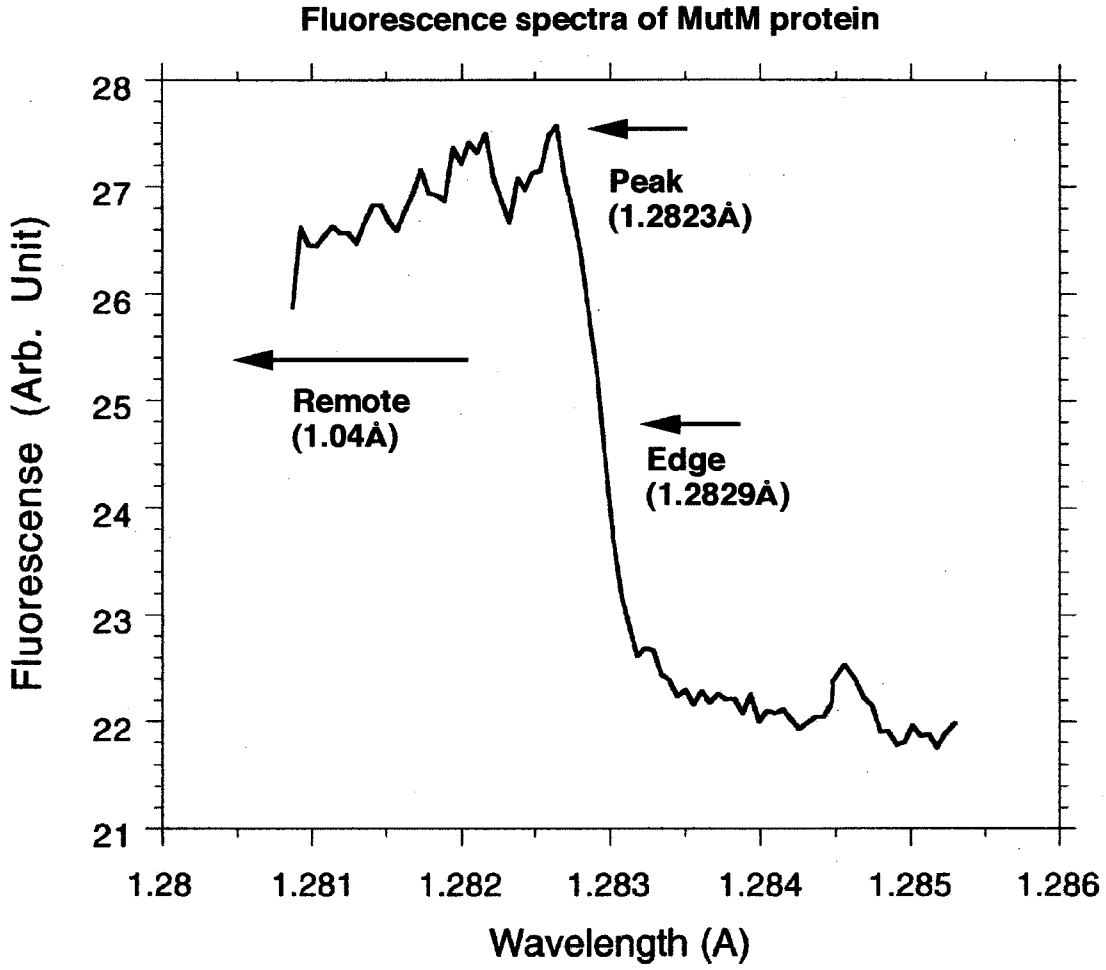


Figure 6. X-ray Absorption Fine Structure (XAFS) Spectra Measurement for MAD Data Collection. The XAFS technique provides information about the atomic structure in the immediate vicinity of an X-ray absorbing atom. In order to select two wavelengths near the zinc absorption edge (Peak and Edge) for MAD data collection, fluorescence spectra of the MutM protein was measured at beamline BL45XU at SPring-8.

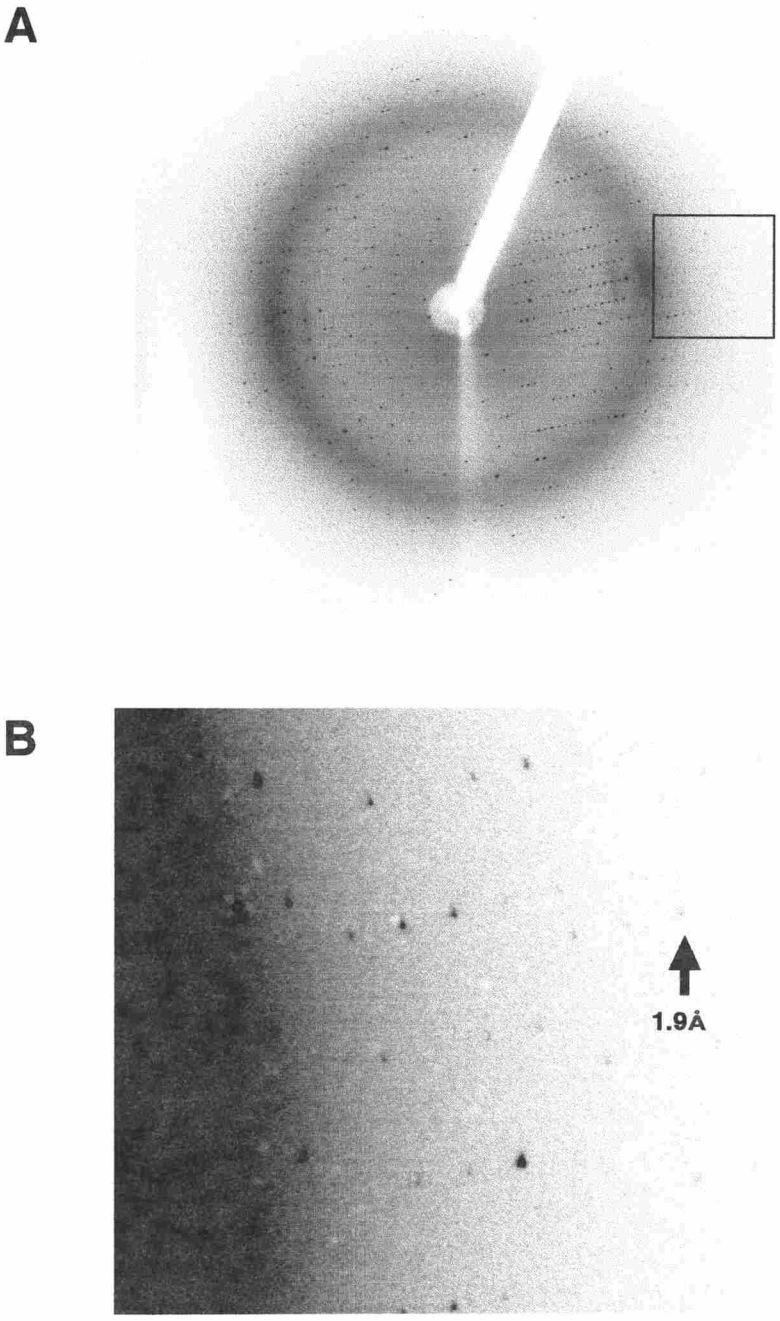


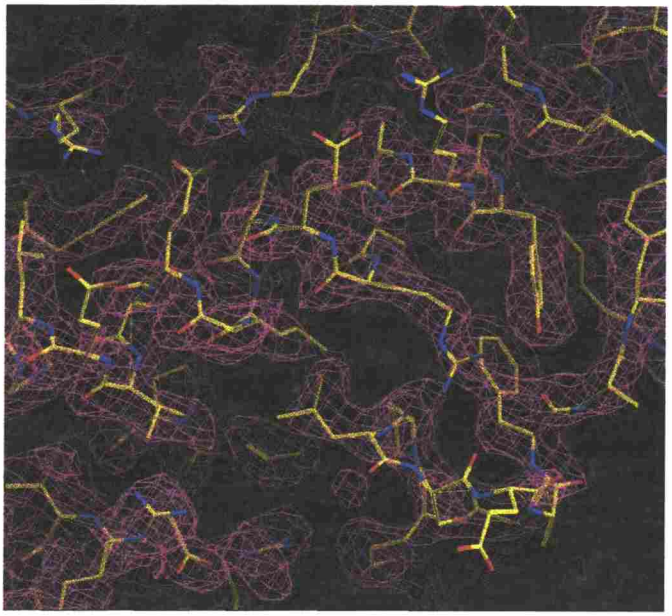
Figure 7. X-ray Diffraction Image of MutM Crystal. I achieved higher-resolution data collection for frozen crystals using synchrotron radiation at the BL45XU of the SPring-8 (Yamamoto et al., 1998). The diffraction image was taken from a RIGAKU R-AXIS IV imaging-plate system, with a crystal-to-film distance of 240mm. The oscillation range is 2 degree/frame. The detector edge corresponds to 1.87Å. (A) One diffraction image; (B) enlargement of the indicated portion of the images in (A)

MAD Phasing, Model Building, and Refinement

The structure of MutM was solved with MAD techniques. The phasing was performed with SOLVE (Terwilliger and Berendzen, 1999) to give a 2.2 Å resolution phase set with an overall figure of merit of 0.40. The MAD phases were further improved by solvent flattening (47% solvent content) by program DM in PERT mode (CCP4, 1994). At this stage, the map showed a solvent boundary and was partially interpretable. The partial poly-alanine model encompassing some of the secondary structural elements was built using the program "O" (Jones et al., 1991). This alanine model indicated that the two MutM molecules in an asymmetric unit (Sugahara et al., 2000) had different conformations. Therefore, the phases were improved by combination of density modification and non-crystallographic symmetry (NCS) multi-domain averaging (program "DM" (CCP4, 1994)) using two separate masks and matrices (CCP4, 1994). The NCS matrix was determined with program LSQKAB (CCP4, 1994), and the molecular mask for NCS averaging was generated using programs O (Jones et al., 1991) and RAVE (Kleywegt and Jones, 1994). The excellent quality of the 2.2 Å MAD map (Figure 8A) enabled us to assign more than 90% of the amino acid residues. The domain orientation of the two molecules in an asymmetric unit was slightly different from each other. This initial model was first refined from 50 to 1.9 Å resolution by Powell conjugate gradient minimization and torsion angle-

restrained molecular dynamics without NCS restraint using the Crystallography and NMR System package (CNS version 0.9; Brünger et al., 1998). The remaining residues were fitted to the electron density maps calculated from experimental and model phases. The success of model refinement was evaluated at each stage by the change in the free R factor (Brünger, 1992) and by inspection of stereochemical parameters with the program PROCHECK (Laskowski et al., 1993). The Ramachandran plot indicated that 86.6% of the residues lay in the favorable regions, 12.7% in the allowed regions, and 0.7% in the generously allowed regions. The current model consists of the entire 266 amino acid residues coded by the *T. thermophilus* HB8 *mutM* gene, a zinc atom, and 292 water molecules. All of the main chain atoms and almost all of the side chain atoms are clearly visible in the 2Fo-Fc electron density map (Figure 8B) contoured at the 1σ level. The refined R-values and root-mean-square deviations from ideality are described in Table II.

A. Fo map after NCS averaging



B. 2Fo-Fc map of final model

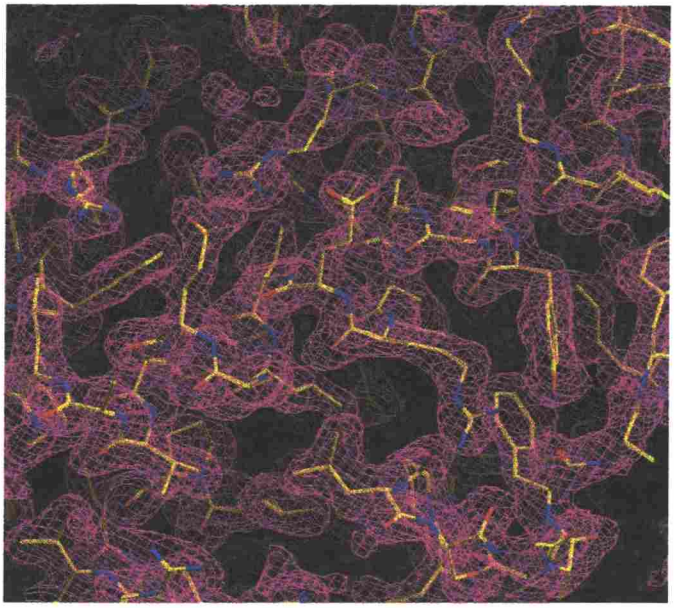


Figure 8. Electron-Density Maps After Multi-Domain NCS Averaging and final model. (A) The excellent quality of the 2.2 Å MAD map after multi-domain NCS averaging (CCP4, 1994) enabled us to assign more than 90% of the amino acid residues. (B) All of the main chain atoms and almost all of the side chain atoms are clearly visible in the 2Fo-Fc electron density map contoured at the 1σ level.

Table II. MAD Data Collection and Refinement Statistics

Parameters	λ1(Native)	λ2(Peak)	λ3(Edge)
Data collection			
Resolution (Å)	1.9	2.2	2.2
Wavelength (Å)	1.04	1.2823	1.2829
Observations	402,397	260,102	260,684
Independents	37,892	25,080	25,137
Completeness (%)	90.0	92.7	92.7
R_{merge} (%)[a]	6.3	6.2	6.5
I/σI	29.1	32.1	31.6
Refinement			
Resolution (Å)	50-1.9		
Reflections	36,536		
R_{cryst}[b]	0.214		
R_{free}[c]	0.258		
Number of non-hydrogen atoms	4335		
R.m.s. deviations from ideality			
Bond length (Å)	0.006		
Angles (°)	1.35		
Dihedral angles (°)	23.37		
Improper angles (°)	0.84		

[a] $R_{merge} = \Sigma_h \Sigma_i | I_{hi} - \langle I_h \rangle | / \Sigma \langle I_h \rangle$.

[b] $R_{cryst} = \Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$.

[c] R_{free} is monitored with 10% of the reflection data excluded from refinement.

Molecular Dynamic Simulation

The model of MutM complexed with GO-flipped DNA was obtained from molecular dynamic (MD) simulation using the CNS version 0.9 program (Brünger et al., 1998). The dielectric constant was fixed at 1.0, and a 13 Å cutoff was used for the nonbonded interactions.

All molecular images were produced using Raster3D (Merrit and Murphy, 1994) and Molscript (Kraulis, 1991), except for Figure 13, which was produced with GRASP (Nicholls et al., 1993).

RESULT and DISCUSSION

Overall Structure

MutM was approximately 50 × 30 × 30 $Å^3$ and consisted of two domains (Figure 9). The N-terminal domain (residues 1-109, blue color) was a two-layered β-sandwich structure composed of nine antiparallel β-strands (β1-β9), except for strand β2, with an α-helix (αA and αB) on each side (Figures 9 and 10). The C-terminal domain (residues 129-266, red, orange and green colors) had four helices (αC to αF) and two β-strands (β10 and β11) (Figures 9 and 10). A search of the protein structural database for the whole protein and for each domain did not reveal any protein with similar folding (the DALI server; Holm and Sander, 1993). Thus, MutM forms a distinct structural group from other base-excision-repair enzymes.

From sequence comparison among the sixteen MutM proteins, the conserved amino acid residues were located in (1) the N-terminal six residues and the loop between β3 and β4 (colored magenta in Figure 9 and the colored boxes in Figure 11), (2) the two loops between β5 and β6 and between β8 and β9 (colored magenta), (3) the helix-two turns-helix (H2TH) motif (αD to αE, colored orange), which resembles the HhH motif and interacts with the DNA phosphodiester backbone (Thayer et al., 1995; Mullen et al., 1997), and (4) the zinc finger motif (β10 and β11, colored green) forming a β-hairpin loop

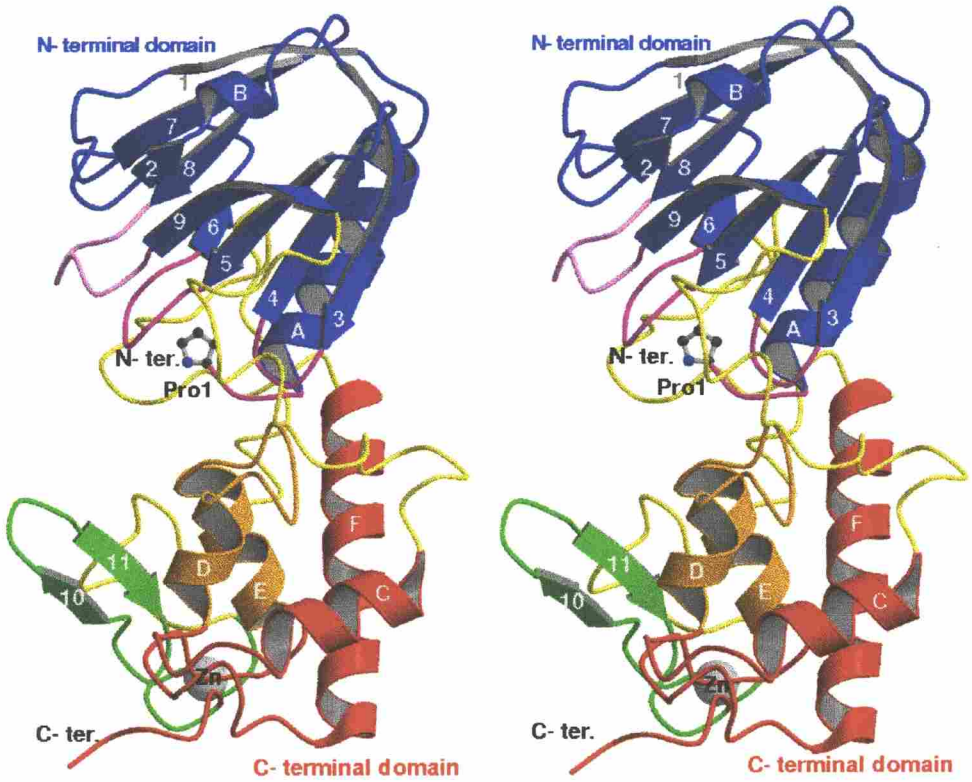


Figure 9. MutM Overall Structure. A stereo ribbon diagram of MutM. The MutM molecule consists of an N-terminal domain (residues 1-109, blue), a C-terminal domain (residues 131-203 and 248-266, red) and two long loops (11 0-130 and 204-237, yellow). The N-terminal domain (shown in blue) consists of a two-layered beta-sandwich with two alpha helices. The C-terminal domain consists of four alpha helix bundles (red and orange) and a beta hairpin loop of the zinc finger motif (green). There were two conformers in an asymmetric unit (see "Experimental Procedures"). The conformations of the two long loops (in yellow) in the interdomain cleft were different between the two conformers. These two long loops will work as a hinge in domain movement. Pro1 (shown by a ball-and-stick model), which works as a nucleophile in the glycosylase/AP lyase reaction, was situated at the bottom of the cleft. The highly conserved regions (purple) in the N-terminal domain, the H2TH motif (orange), and the zinc finger motif (green) were located around the cleft (see Figure 11).

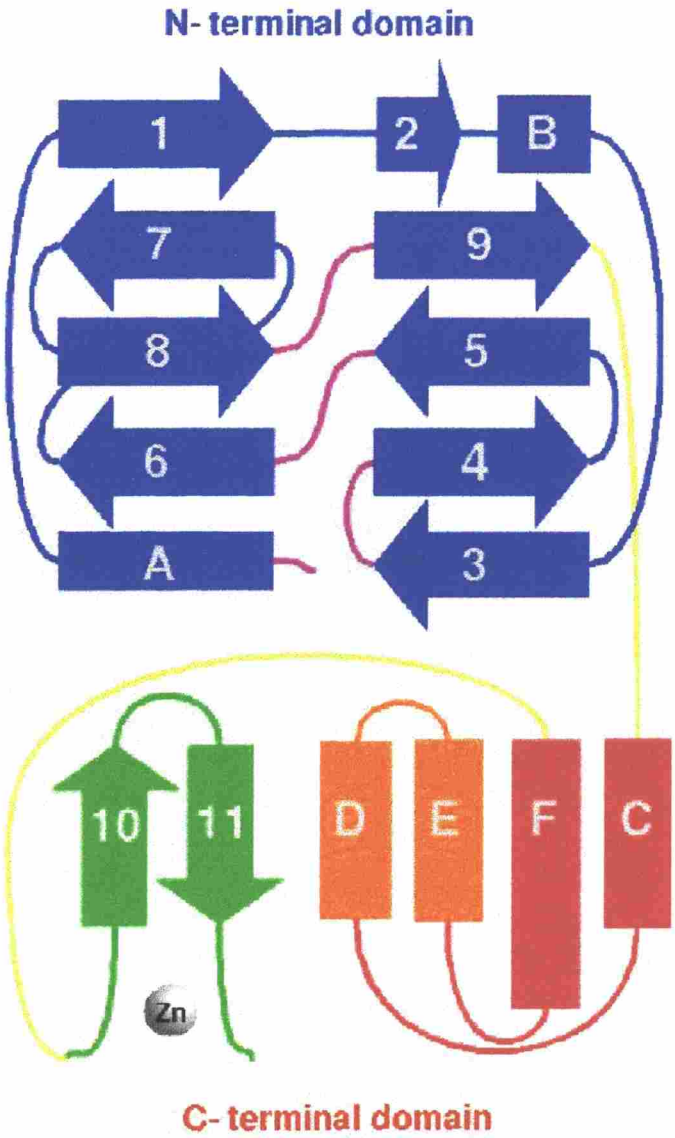


Figure 10. Topology of the MutM Structure. α helices are shown as bars and β strands as arrows as coloured in Figure 9.

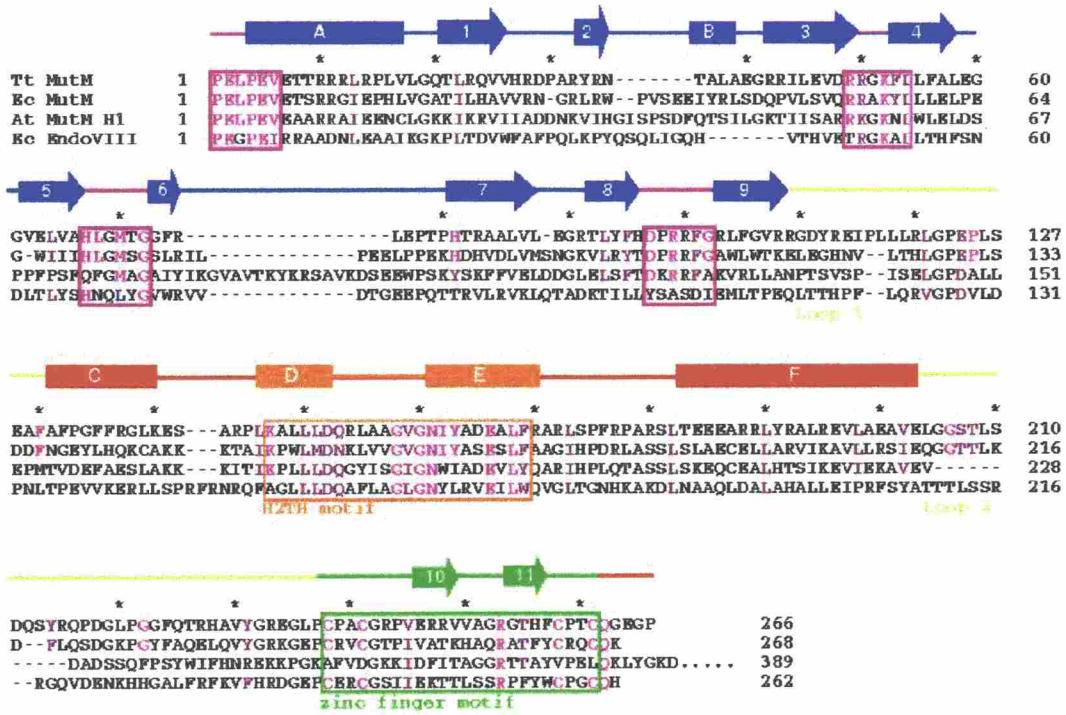


Figure 11. Amino Acid sequences of MutM homologues. Aligned sequences of bacterial *T. thermophilus* MutM (Tt MutM) (Mikawa et al., 1998) and *E. coli* MutM (Ec MutM) (Boiteux et al., 1987), eukaryotic *A. thaliana* MutM homolog 1 (At MutM H1) (Ohtsubo et al., 1998; Murphy et al., 1998), and *E. coli* endonuclease VIII (Ec EndoVIII) (Jiang et al., 1997). The α-helices (bars) and β-strands (arrows) are shown above the aligned sequence. Colored letters represent the residues conserved among 16 MutM proteins (see also Figure 12). The asterisks represent every tenth amino acid residue of the *T. thermophilus* sequence. The highly conserved regions in the N-terminal domain (magenta), the H2TH motif (orange) and the zinc finger motif (green) are boxed.

extending into the cleft. Almost all of these conserved amino acid residues cluster in the large cleft between the two domains (Figures 9 and 10). The catalytically essential residues and the two DNA-binding motifs were included in these regions, as described below (Figures 9 to 22). In particular, the N-terminal proline identified as a nucleophile in the DNA glycosylase/AP lyase reaction (Zharkov et al., 1997) is in the bottom of the cleft (Figure 12). In addition, the molecular surface of the cleft is positively charged (Figure 13). The conserved residues in the cleft are essential for DNA binding and catalysis.

MutM-DNA Interactions

The DNA footprinting experiment showed that MutM interacts with six nucleotides of the damaged strand, two and three nucleotides on the 3' and 5' side of the lesion, respectively (Castaing et al., 1999). In the enzymatic reaction process, the α-imino group of Pro1 forms a Schiff base with C1' of ribose (Zharkov et al., 1997). Lys52 has been suggested to interact with the C8 oxygen of the flipped GO base (Sidorkina et al., 1998). As mentioned above, the MutM molecule has a positively charged region, the H2TH motif, and the zinc finger motif in the large cleft. These regions are expected to interact with DNA.

On the basis of these suggestions, I proposed an initial conformation for the MutM-DNA complex and started the molecular dynamic calculations. For the initial structure, GO-flipped dsDNA bent by 22° (Lau et al., 1998) was used.

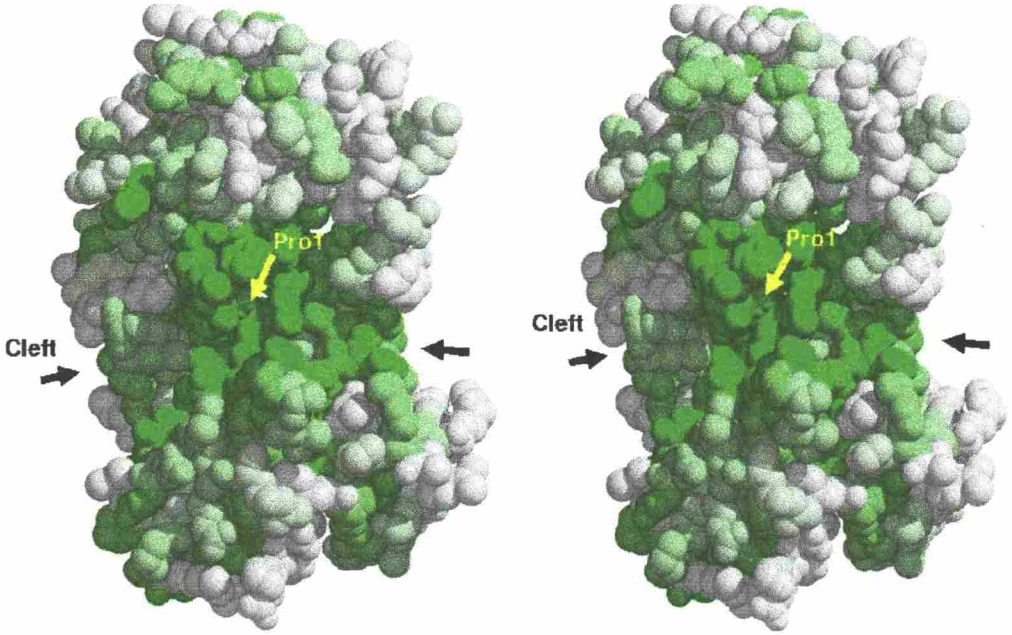


Figure 12. Conserved Residues of MutM

Stereo CPK model of MutM looking towards the interdomain large cleft. The surface is shown by color gradient according to the residue conservation, with absolutely conserved residues in dark green. Homology was based on the multiple alignment of the MutM from *T. thermophilus* HB8 (Mikawa et al., 1998), *E. coli* (Boiteux et al., 1987), *Neisseria meningitidis* (Swartley and Stephens, 1995), *Bacillus firmus* (Ivey, 1990), *Synechococcus elongatus naegeli* (Floss et al., 1997), *Bacillus subtilis* (Lapidus et al., 1997), *Salmonella typhimurium* (Suzuki et al., 1997), *Lactococcus lactis subsp. cremoris* (Duwat et al., 1995), *Haemophilus influenzae* (sp|P44948), *Synechocystis PCC6803* (sp|P74290), *Mycobacterium leprae* (emb|CAA19197.1), *Mycobacterium tuberculosis* (sp|Q10959), *Streptococcus mutans* (sp|P55045), *Streptomyces coelicolor* A3(2) (emb|CAB63194.1), *Mycoplasma pneumoniae* (sp|P75402), and *Mycoplasma genitalium* (sp|P55825). Homologous sequences were identified with PSI-BLAST (Altschul et al., 1997), and the sequences were aligned with CLUSTAL W (Thompson et al., 1994).

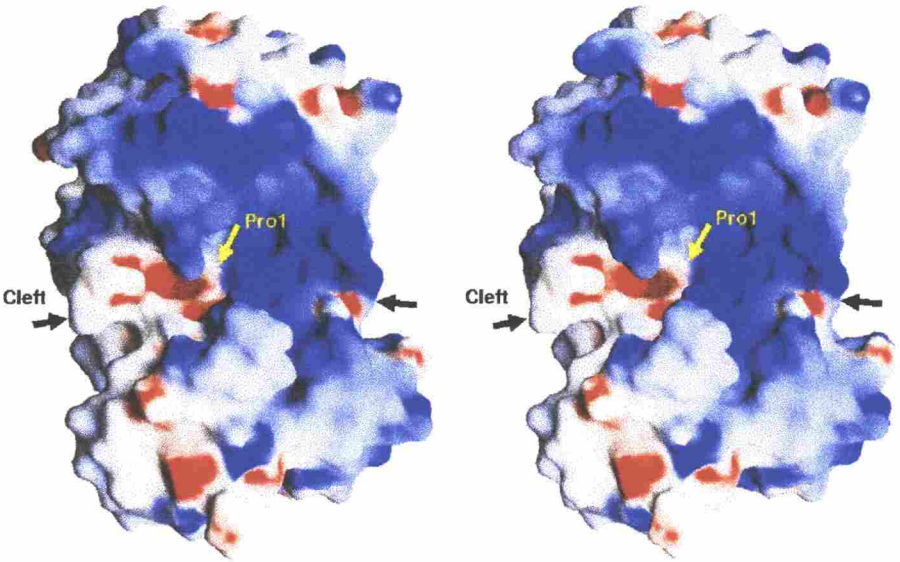


Figure 13. Electrostatic Potential of MutM. The solvent-accessible surface is colored according to the electrostatic potential, with positive colored blue and negative red. The deep cleft between the two domains was positively charged. Damaged DNA will bind to this large cleft.

The calculated model showed that the GO-flipped DNA fitted well (Figures 14 and 15) to MutM. The DNA interacted with all the following four conserved regions (Figures 9 - 11): (1) the H2TH motif, (2) the zinc finger motif, (3) the turns of β5-β6 and β8-β9, and (4) the catalytic groups. I will discuss the role of these regions.

H2TH Motif

The αD-αE region (residues 146-169) in the C-terminal domain is highly conserved among MutM homologues and EndoVIII (Figure 11). Similar consensus sequences are also found in *E. coli* EndoIII (Thayer et al. 1995) and in *Pyrococcus furiosus* flap endonuclease 1 (FEN-1) (Hosfield et al., 1998) (Figure 16).

The helix-hairpin-helix (HhH) motif is found in *E. coli* EndoIII (Figure 16, green) (Thayer et al., 1995) as a common non-specific DNA-binding motif and is implicated in binding to the DNA backbone phosphate. The HhH motif is also observed in the X-ray crystal structure of *E. coli* AlkA (Yamagata et al., 1996; Labahn et al., 1996) and that of *E. coli* MutY (Guan et al., 1998), and in the primary structure of OGG1 proteins (Krokan et al., 1997). These common HhH motifs in monofunctional glycosylases and bifunctional glycosylase/AP-lyases suggest similar roles in DNA recognition, although the catalytic mechanisms of these enzymes are somewhat different (Labahn et al., 1996).

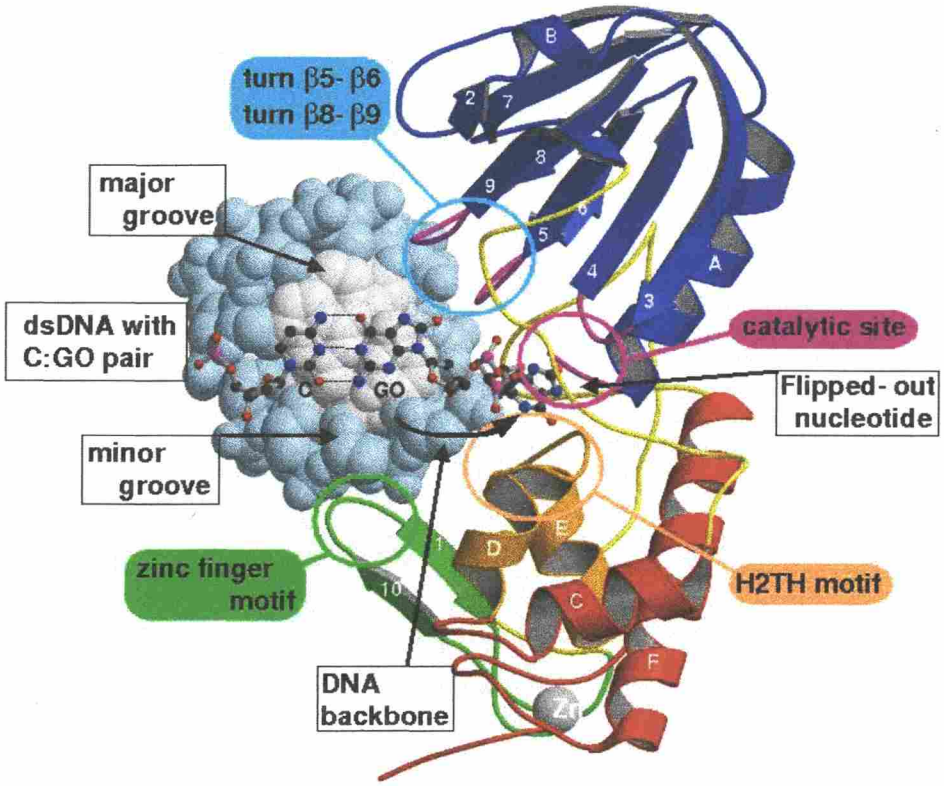


Figure 14. The MutM - DNA Complex Model. The complex model between the flipped-out DNA and MutM was obtained by molecular dynamic calculation. The viewing direction is similar to that in Figure 9. The kinked DNA is drawn with the CPK model with backbones colored in steel-blue and with bases in steel-gray. C and GO bases and their sugar residues before and after flipping out are shown by ball-and-stick models. All the four conserved regions ((1) turns β5-β6 and β8-β9, (2) the catalytic site, (3) the H2TH motif, and (4) the zinc finger motif) are in the large cleft of the MutM molecule. The N-terminal domain can access the major groove of DNA and the zinc finger motif of the C-terminal domain can access the minor groove. The H2TH motif of the C-terminal domain is situated near the active site, and will interact with the DNA backbone with the damaged base

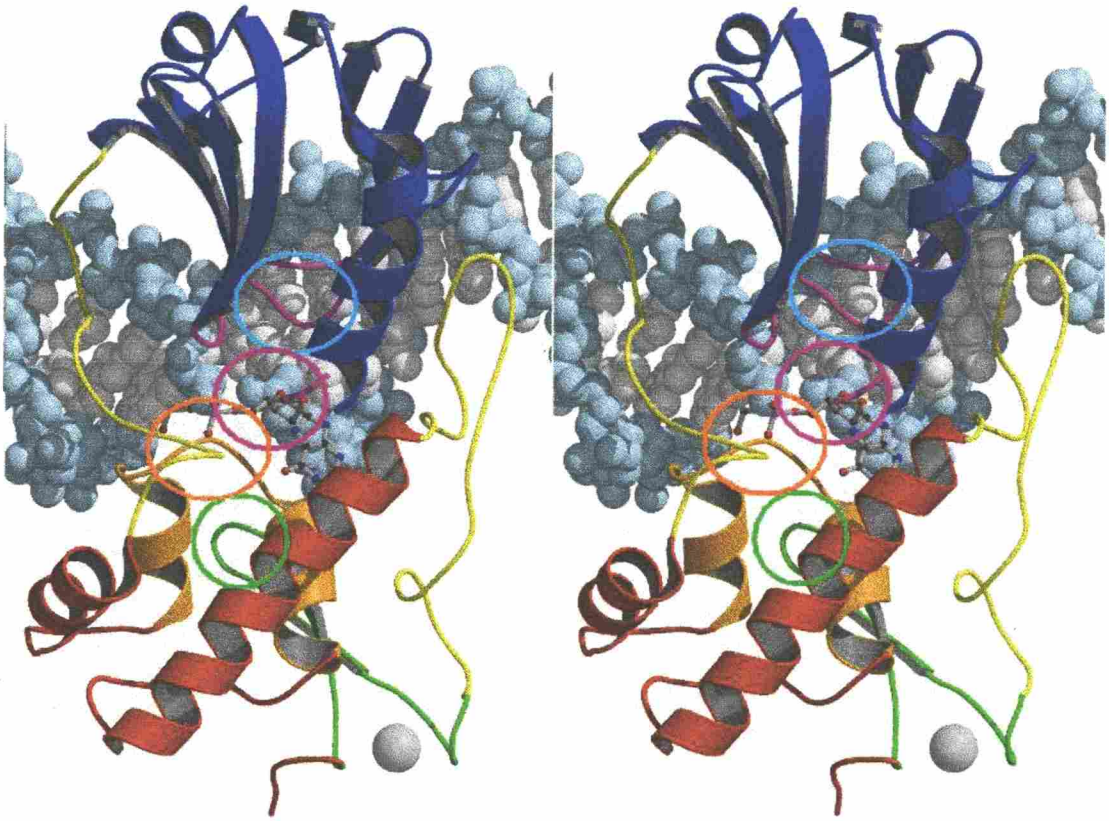


Figure 15. The Stereo View from the Left-Hand Side of Figure 14.

The helix-three turns-helix (H3TH) motif has three turns between the two helices, and is found in the *P. furiosus* FEN-1 crystal structure (Figure 16, blue) (Hosfield et al., 1998). Both the HhH and H3TH motifs are found in a wide range of non-sequence-specific DNA binding proteins, such as nucleases, N-glycosylases, ligases, helicases, topoisomerases, and polymerases, that are essential for protein-mediated DNA synthesis and DNA repair. By analogy with the DNA polymerase β structure, it is suggested that each of these motifs binds to DNA in a non-sequence-specific manner via the formation of hydrogen bonds between protein backbone nitrogens and DNA phosphate groups (Doherty et al., 1996; Artymiuk et al., 1997).

The MutM-dsDNA complex model (Figures 14 and 15) shows that the H2TH motif, which has two turns between the two helices, can access the DNA backbone at the bottom of the cleft. Therefore, it will play a role similar to the HhH and H3TH motifs, which participate in non-sequence-specific recognition.

Zinc Finger Motif

MutM has another DNA-binding motif, the zinc finger motif near the C-terminus (Figures 11, 14, and 17). The zinc finger motif, a well known DNA-binding motif, has been found in various sequence-specific DNA-binding proteins (O'Connor et al., 1993). The zinc finger motif of MutM (-Cys-X_2-Cys-X_{16}-Cys-X_2-Cys-) was a unique and rather simple β-hairpin loop that

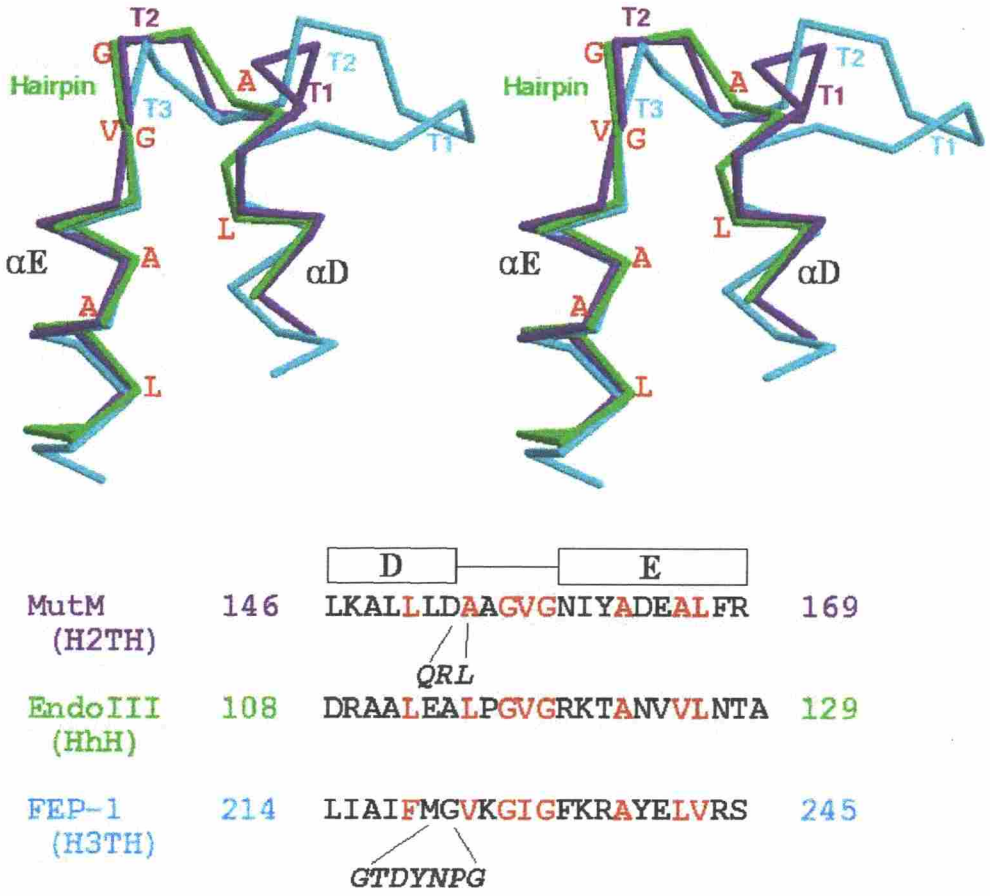


Figure 16. The H2TH Motif Found in the MutM Structure. Superposition of the MutM helix-two turns-helix (H2TH) motif (in purple) upon the helix-hairpin-helix (HhH) motif of *E. coli* endonuclease III (in green; PDB code 2ABK) (Thayer et al., 1995) and the helix-three turns-helix (H3TH) motif of *M. jannaschii* Flap endonuclease-1 (in cyan; PDB code 1A76) (Hwang et al., 1998). In the H2TH motif of MutM, two turns exist between α-helix D (αD) and α-helix E (αE). The amino acid sequences were aligned based on structural superposition. The amino acid residues in α-helix D (αD) and α-helix E (αE) were conserved among MutM, EndoIII, and FEP-1 (letters in red). The HhH motif in EndoIII and the H3TH motif in FEN-1 have been suggested to access the dsDNA backbone. Therefore, the H2TH motif of MutM is expected to play a role similar to the HhH and H3TH motifs, and will contribute to non-sequence-specific recognition and/or nucleotide-flipping.

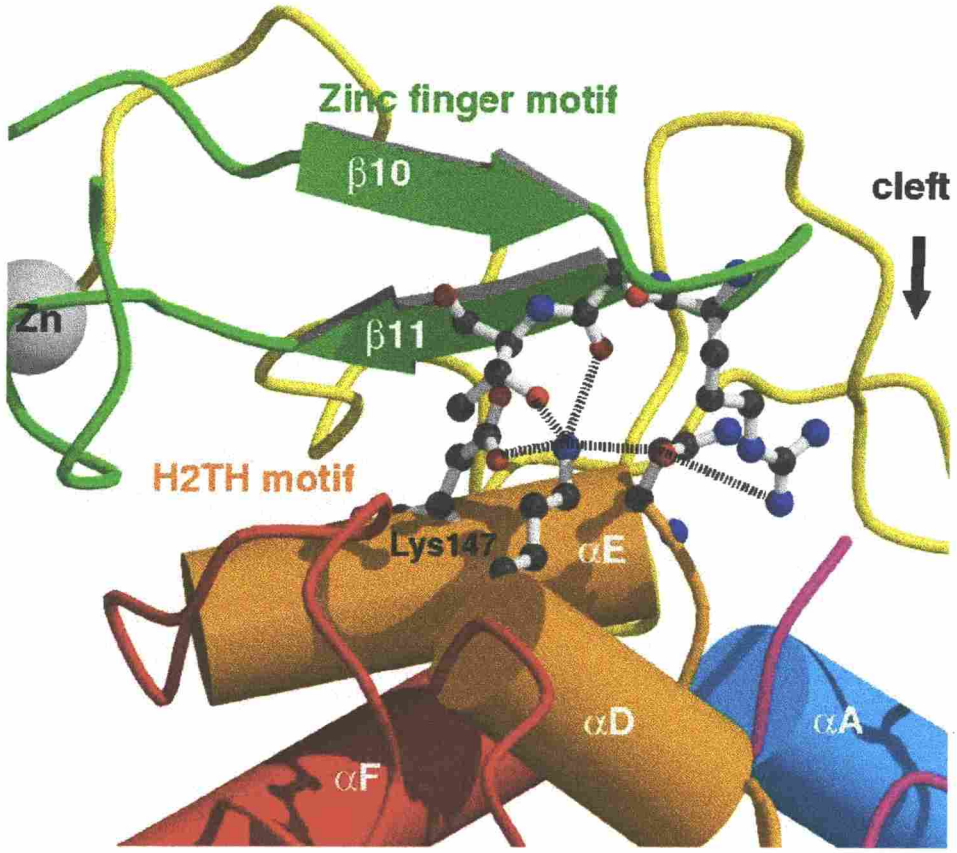


Figure 17. Conserved Lys147 residue between the H2TH motif and the zinc finger motif. The ε-amino group of Lys147 forms hydrogen bonds (dashed lines) to the conserved side chain and main chain (ball-and-stick) in the H2TH motif (orange) and in the zinc finger motif (green). The decrease in the catalytic efficiency caused by replacement of Lys147 with alanine (Rabow and Kow, 1997) will be due to the disruption of these hydrogen bonds, which in turn weaken the interaction between MutM and DNA. These results support the hypothesis that the H2TH and zinc finger motifs of MutM will interact with DNA.

extended into the interdomain cleft of MutM. Such a β-hairpin loop has also been reported in the crystal structure of human 3-methyladenine DNA glycosylase (AAG) complexed with DNA (Lau et al., 1998). In that enzyme, the tyrosine residue at the top of the hairpin loop interacts directly with the DNA lesion in the minor groove and is essential for base-flipping. The [4Fe-4S] cluster loop motif of EndoIII also extends into the interdomain cleft and is suggested to intercalate into the minor groove of DNA (Thayer et al., 1995). This [4Fe-4S] cluster loop motif has a lysine residue and an arginine residue near the edge of the loop.

The β-hairpin loop of MutM also has the conserved Arg253 on its edge (Figures 18). The interaction of the positively charged residue with the minor groove of DNA may be similar in MutM a

nd EndoIII.

Lys147 plays a central role in forming hydrogen bond networks between the H2TH motif and the zinc finger motif (Figure 18). The replacement of Lys147 in MutM by alanine decreases the activity against GO:C substrates by ~50-fold, but does not affect AP site substrates (Rabow and Kow, 1997). The replacement of Lys147 by alanine will cause disruption of these hydrogen bond networks. This may disrupt the conformation of the zinc finger motif, which interacts with the minor-groove side of the GO:C pair, and will decrease activity against GO:C substrates. In contrast, the conformation of the H2TH motif, which interacts with the phosphate backbone of DNA, will not be changed by the mutation, and the excision of the sugar moiety from the flipped-out AP site will not be affected.

Implications of Recognition of a Wide Range of Oxidatively Damaged Bases

MutM has a broad specificity for oxidatively damaged lesions, such as GO paired to cytosine (Tchou et al., 1991), formamidopyrimidine (Fapy) (Boiteux et al., 1992), and 5-hydroxycytosine (5OHC) paired to guanine (Hatahet et al., 1994), and apurinic/apyrimidinic (AP) sites (Castaing et al., 1992) (Figure 19).

The crystal structure of MutM (Figures 14, 18A, and 18B) explains the specificity for these substrates. The C8 oxygen of GO is located in the major

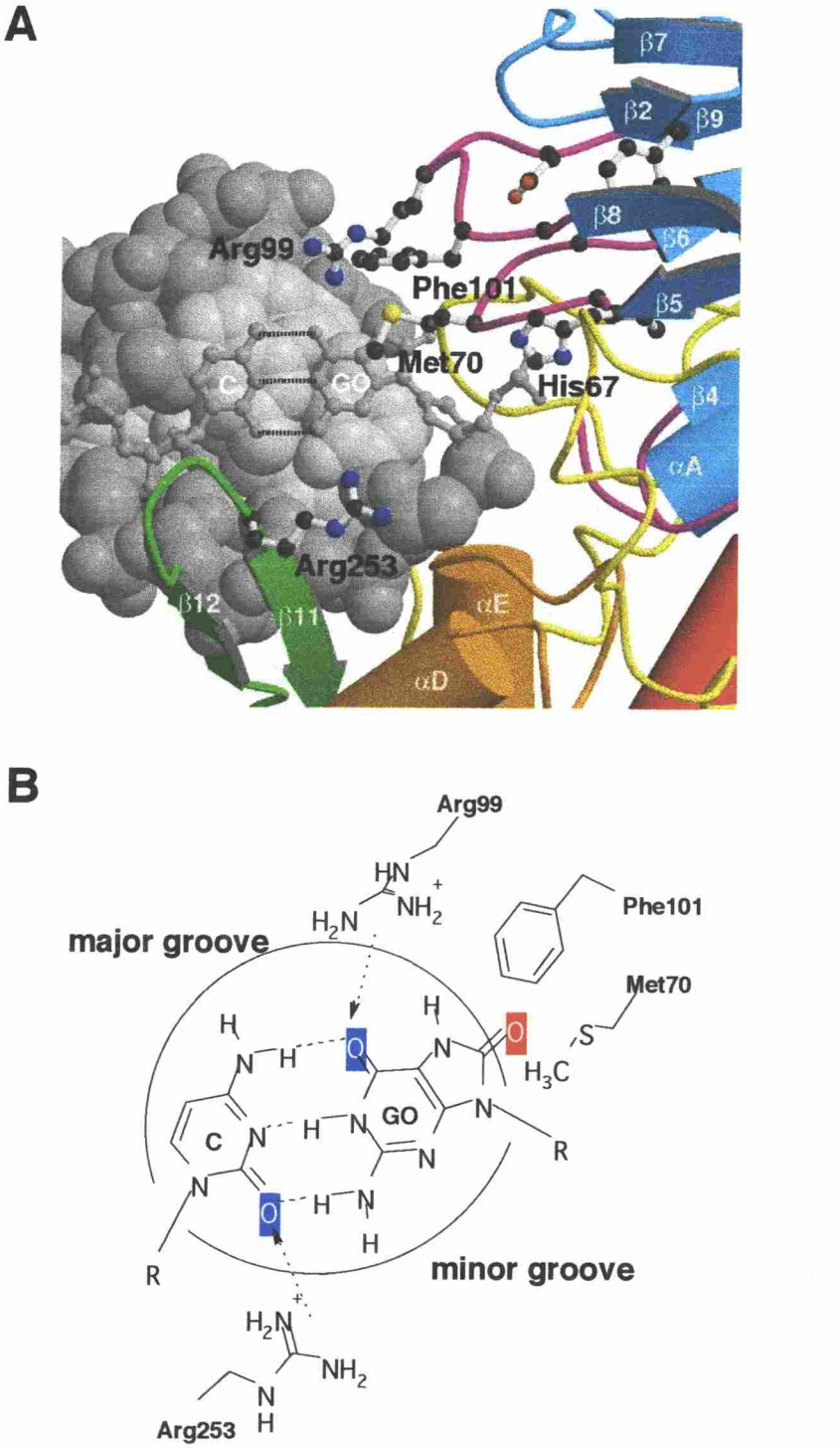


Figure 18. Implications for Lesion Recognition.
(A) Model of recognition of GO:C-paired DNA (gray) by conserved amino acid residues (ball-and-stick) of MutM. **(B)** Schematic representation of the proposed lesion-recognition mechanism for the GO:C pair in double-strand DNA. Two arginine residues, Arg99 and Arg253, interact with the keto group (white letters on blue) in both grooves of DNA to disrupt the three hydrogen bonds of the GO:C pair. The C8 oxo group (white letters on red) in the DNA major groove will be recognized by either Met70 or Phe101 in the N-terminal domain. Phe101 will be inserted into the abasic site after nucleotide-flipping to stabilize the base stacking.

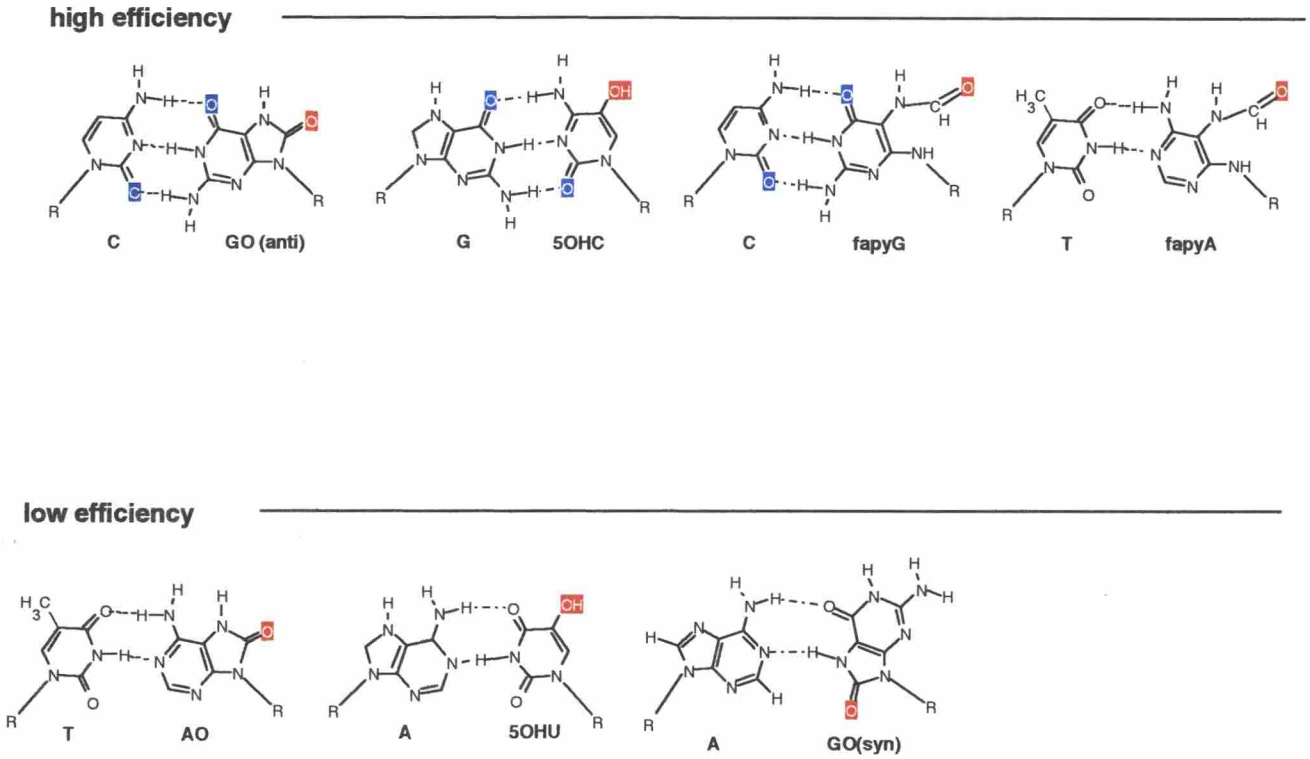


Figure 19. Pairing of the Oxidized Base in High- and Low-Efficiency MutM Substrates

groove of the model B-form dsDNA (Lipscomb et al., 1995) (Figures 18A and 18B). The damaged base interacts with turns β5-β6 and β8-β9 in the N-terminal domain, which contain the conserved residues His67, Met70, Arg99, and Phe101 and participate in lesion recognition. In particular, Arg99 could form a hydrogen bond with either the C6 or C8 oxo group.

The minor groove of DNA interacts with the β-hairpin loop of the zinc finger motif. As Arg253 is the only conserved residue in the β-hairpin loop, it may interact with the C2 oxo group of cytosine in the GO:C pair (Figure 18B). The two arginine residues, Arg99 and Arg253, will interact with both sides of the DNA to cleave the three hydrogen bonds of the GO:C pair. The C8 oxo group in the major groove may be repelled by a large hydrophobic residue, either Met70 or Phe101. Domain movement may also contribute to the cleavage of GO:C hydrogen bonds. After nucleotide-flipping of the GO residue, Phe101 may be inserted into the abasic site to stabilize the base stacking.

This model is applicable to another GO:C base-excision repair enzyme with different substrate specificity, EndoVIII (Figure 11). The residues Met70, Arg99, and Phe101, which contribute to substrate specificity, are not conserved in EndoVIII, although the catalytic groups Pro1, Glu2, and Glu5 are conserved between these enzymes.

Postulated Catalytic Groups in the Active Site

The crystal structure of MutM reveals that the nucleophile Pro1 is situated at the center of the cleft and surrounded by the conserved charged residues Glu2, Glu5, and Lys52 (Figure 20).

Monofunctional DNA glycosylases (AlkA, AAG, UDG, MUG, and MutY in Figure 21) hydrolyze the glycosidic bond between the modified base and deoxyribose, and bifunctional DNA glycosylases/β-lyases (EndoIII and T4 EndoV in Figure 21) hydrolyze the glycosidic bond and perform β-elimination. MutM, a trifunctional DNA glycosylase/β-lyase/δ-lyase, hydrolyzes the glycosidic bond and performs β,δ-elimination.

When the first catalytic group in T4 EndoV (the α-amino group of Thr2) was superimposed on the imino group of Pro1 in MutM (these residues are known to form a Schiff base with the reducing C1' atom of the deoxyribonucleotidyl residue) and the bound DNA was also superimposed, the second catalytic groups (Glu5 in MutM and Glu23 in T4 EndoV) were situated at the same position (Figure 20); we assign the role of the three catalytic groups as shown in Figure 21, left. According to this model, the first catalytic group (the imino of Pro1) cleaves the glycosidic bond between the damaged base and C1', the second catalytic group (Glu5) abstracts the H(C2') followed by β-elimination, and the third catalytic group (Glu2) abstracts the H(C4') followed by δ-elimination.

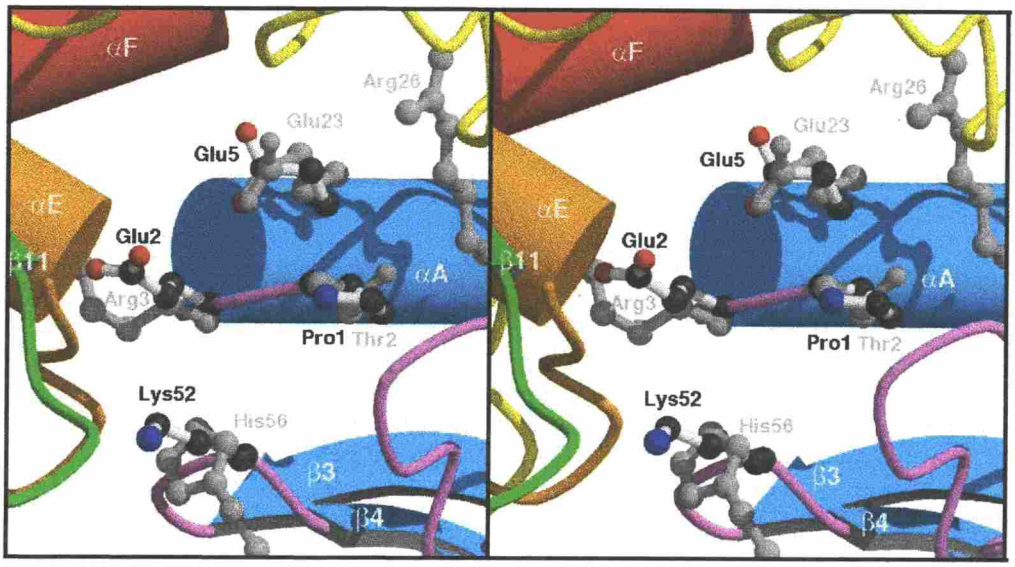
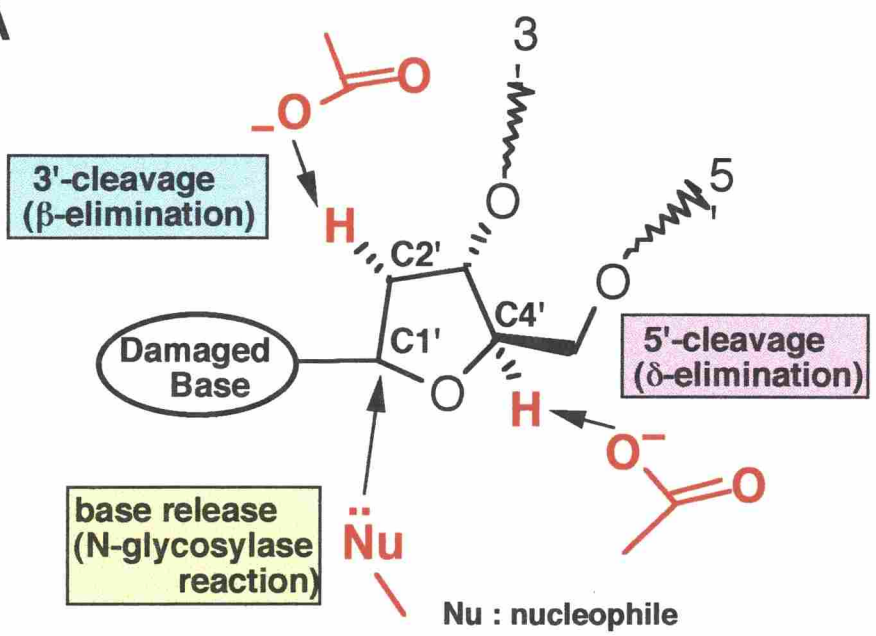


Figure 20. The Catalytic Site of MutM.
The postulated active-site residues of MutM (ball-and-stick model) were superimposed upon the residues of T4 EndoV (ball-and-stick in gray). The nucleophile Pro1 of MutM is surrounded by conserved charged residues, Glu2, Glu5 and Lys52.

A



B

	nucleophile	3'-cleavage	5'-cleavage
AlkA	H_2O	-	-
AAG	H_2O	-	-
UDG	H_2O	-	-
MUG	H_2O	-	-
MutY	H_2O	-	-
EndoIII	Lys120	Asp138	-
EndoV	Thr2	Glu23	-
MutM	Pro1	Glu5	Glu2

Figure 21. Comparison of MutM and Other DNA glycosylases.
(A) A schematic representation of the relative positions of postulated catalytic groups and substrate. (B) A list of putative catalytic residues in various mono-, bi-, and tri-functional DNA glycosylases; AlkA (Yamagata et al., 1996, Labahn et al., 1996), AAG (human 3-methyladenine DNA glycosylase) (Lau et al., 199 8), UDG (uracil DNA glycosylase) (Mol et al., 1995; Savva et al., 1995), MUG (*E. coli* G:T/U mismatch-specific DNA glycosylase) (Barrett et al., 1998), MutY (Guan et al., 1998), EndoIII (Thayer et al., 1995), and T4 EndoV (Vassylyev et al., 1995).

The assignment of the third catalytic group was as follows. The conserved Glu2 in MutM was near Arg3 in T4 EndoV, a residue which cannot be a catalytic group because its pK_a value is too high. Lys52 of MutM was near His56 of T4 EndoV. In general, the pK_a value of the catalytic group of the enzyme is near 7. Since the pK_a value of Arg is too high for a catalytic group, the candidates for the catalytic group are Asp, Glu, His, Cys, Lys, Tyr, the α-amino group, and the α-carboxyl group. Lys52 of MutM cannot be a catalytic group with pK_a around 7, because (1) the intrinsic pK_a value of Lys52 is near 10, (2) there is no positive charge near Lys52 that might lower the pK_a value, and (3) a hydrophobic environment in the active site cannot change its pK_a value by more than ±1 pH unit (Edsall and Wyman, 1958). The conserved Lys52 will help the reaction through hydrogen bonding with the substrate, as shown in Figure 7.

Figures 20 and 21 (left) explain the catalytic specificities of the DNA glycosylases listed in this Figure.

Catalytic Mechanism of MutM

On the basis of the biochemical studies on the reaction mechanism (Bhagwat and Gerlt, 1996; Castaing et al., 1999) and on our three-dimensional structure, I propose the following catalytic mechanism for the MutM N-glycosylase/AP-lyase reaction (Figure 22): (I) binding of MutM to GO-flipped DNA, (II) nucleophilic attack of the imino group of Pro1 on C1' to form a

Schiff base, with the help of Lys52 and Glu2 by electron withdrawal, (III) release of the GO base, C2' deprotonation by Glu5 with the help of Lys52, (IV) β-elimination, (V) C4' deprotonation by Glu2, (VI) δ-elimination, and (VII) release of the product.

The proposed mechanism explains the activity of T4 EndoV in which Glu23 participates in the β-elimination reaction (Vassylyev et al., 1995). According to the mechanism proposed by Castaing et al. (1999), a residue forms a hydrogen bond to the C4' hydroxyl group. A candidate for this residue would be Lys52. In order to confirm the detailed mechanism of the MutM N-glycosylase/AP-lyase reactions, mutational analyses of the active-site residues are underway.

Conclusions

The crystal structure and sequence alignment of MutM revealed some important conserved residues and motifs for MutM catalysis. On the basis of these structural studies, as well as the biochemical results, I have proposed mechanisms for N-glycosylase/AP-lyase activity and recognition of oxidative damage.

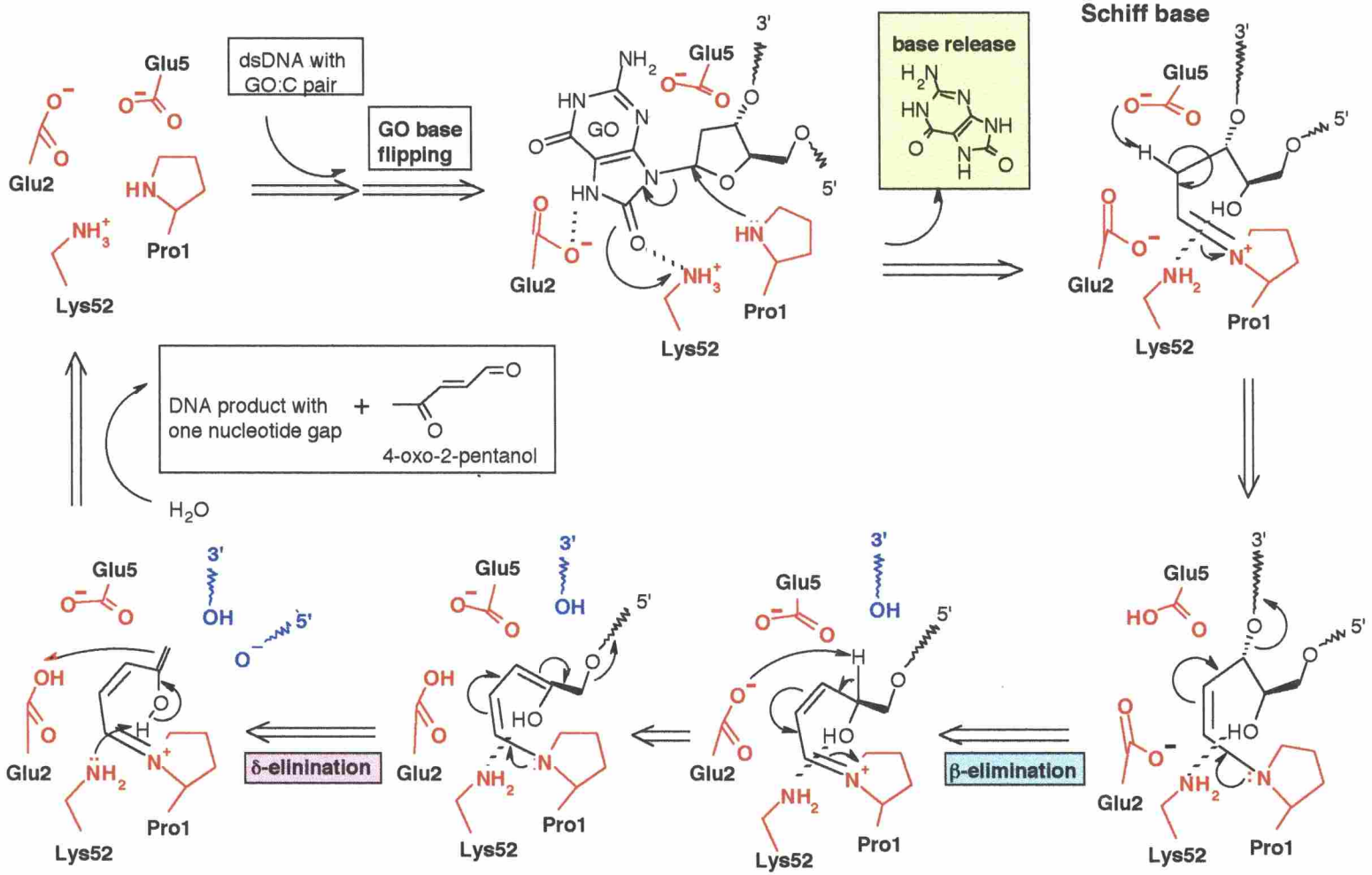


Figure 22. Schematic Representation of the Proposed Reaction Mechanism of MutM N-Glycosylase/ AP-Lyase. This reaction scheme was based on the mechanism of Castaing et al. (1999). The residues that will contribute to each reaction step were deduced from the crystal structure and are shown in red. The substrate, oxidized DNA, and the products are shown in black.

REFERENCES

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389-3402.

Ames B.N., Gold L.S., and Willett W.C. (1995). The causes and prevention of cancer. Proc. Natl. Acad. Sci. USA. *92*, 5258-5265.

Artymiuk, P.J., Ceska, T.A., Suck, D., and Sayers, J.R. (1997). Prokaryotic 5'-3' exonucleases share a common core structure with gamma-delta resolvase. Nucleic Acids Res. *25*, 4224-4229.

Barrett, T.E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J., and Pearl, L.H. (1998). Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. Cell *92*, 117-129.

Bhagwat, M., and Gerlt, J.A. (1996). 3'-and 5'-strand cleavage reactions catalyzed by the Fpg protein from *Escherichia coli* occur via successive beta- and delta-elimination mechanisms, respectively. Biochemistry *35*, 659-665.

Boiteux, S., Gajewski, E., Laval, J., Dizdaroglu, M. (1992). Substrate specificity of the *Escherichia coli* Fpg protein (formamidopyrimidine-DNA glycosylase): excision of purine lesions in DNA produced by ionizing radiation or photosensitization. Biochemistry *31*, 106-110.

Boiteux, S., O'Connor, T.R., and Laval, J. (1987). Formamidopyrimidine-DNA glycosylase of *Escherichia coli*: cloning and sequencing of the *fpg* structural gene and overproduction of the protein. EMBO J. *6*, 3177-3183.

Boiteux, S., O'Connor, T.R., Lederer, F., Gouyette, A., and Laval, J. (1990). Homogeneous *Escherichia coli* FPG Protein. A DNA Glycosylase which excises imidazole ring-opened purines and nicks DNA at apurinic/apyrimidinic sites. J. Biol. Chem. *265*, 3916-3922.

Breimer, L.H. (1990). Molecular mechanisms of oxygen radical carcinogenesis and mutagenesis: the role of DNA base damage. Mol. Carcinogenesis *3*, 188-197.

Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472-475.

Brünger, A.T., Adams, P.D., Clore, G.M., Delano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. *D54*, 905-921.

Castaing, B., Boiteux, S., and Zelwer, C. (1992). DNA containing a chemically reduced apurinic site is a high affinity ligand for the *E. coli* formamidopyrimidine-DNA glycosylase. Nucleic Acids Res. *20*, 389-394

Castaing, B., Fourrey, J.L., Hervouet, N., Thomas, M., Boiteux, S., and Zelwer, C. (1999). AP site structural determinants for Fpg specific recognition. Nucleic Acids Res. *27*, 608-615.

Collaborative Computational Project No.4 (CCP4). (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. *D 50*, 760-763.

Dizdaroglu, M. (1985). Formation of an 8-hydroxyguanine moiety in deoxyribonucleic acid on gamma-irradiation in aqueous solution. Biochemistry *24*, 4476-4481.

Doherty, A.J., Serpell, L.C., and Ponting, C.P. (1996). The helix-hairpin-helix DNA-

binding motif: a structural basis for non-sequence-specific recognition of DNA. Nucleic Acids Res. *24*, 2488-2497.

Duwat, P., de Oliveira, R., Ehrlich, S.D., and Boiteux, S. (1995). Repair of oxidative DNA damage in gram-positive bacteria: the *Lactococcus lactis* Fpg protein. Microbiology *141*, 411-417.

Edsall, J.T., and Wyman, J. (1958). Biophysical Chemistry, Volume I, (New York: Academic Press), pp. 406-476.

Floss, B., Igloi, G., Cassier-Chauva, T,C., and Muehlenhoff, U. (1997). Molecular cloning and overexpresion of the *petF* gene from the thermophilic cyanobacterium *Synechococcus elongatus*. Photosyn. Res. *54*, 63-71.

Guan, Y., Manuel, R.C., Arvai, A.S., Parikh, S.S., Mol, C.D., Miller, J.H., Lloyd, R.S., and Tainer, J.A. (1998). MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. Nature Struct. Biol. *5*, 1058-1064.

Hatahet, Z., Kow, Y.W., Purmal, A.A., Cunningham, R.P., Wallace, S.S. (1994). New substrates for old enzymes. 5-Hydroxy-2'-deoxycytidine and 5-hydroxy-2'-deoxyuridine are substrates for *Escherichia coli* endonuclease III and formamidopyrimidine DNA N-glycosyalse, while 5-hydroxy-2'-deoxyuridine is a substrate for uracil DNA N-glycosylase. J. Biol. Chem. *269*, 18814-18820.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123-138.

Hoseki, J., Yano, T., Koyama, Y., Kuramitsu, S., and Kagamiyama, H. (1999). Directed evolution of thermostable kanamycin-resistant gene products in an extremely thermophilic bacterium, *Thermus thermophilus*. J. Biochem. (Tokyo) *126*, 951-956.

Hosfield, D.J., Mol, C.D., Shen, B., and Tainer, J.A. (1998). Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. Cell *95*, 135-146.

Hwang, K.Y., Baek, K., Kim, H.Y., and Cho, Y. (1998). The crystal structure of flap endonuclease-1 from *Methanococcus jannaschii*. Nature Struct. Biol. *5*, 707-713.

Ivey, D.M. (1990). Nucleotide sequence of a gene from alkaliphilic *Bacillus firmus* RAB that is homologous to the fpg gene of *Escherichia coli*. Nucleic Acids Res. *18*, 5882.

Jiang, D., Hatahet, Z., Melamede, R.J., Kow, Y.W., and Wallace, S.S. (1997). Characterization of *Escherichia coli* Endonuclease VIII. J. Biol. Chem. *272*, 32230-32239.

Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A47, 110-119.

Kleywegt, G.J., and Jones, T.A. (1994). Halloween ... masks and bones. In "From First Map to Final Model", edited by S. Bailey, R. Hubbard and D. Waller. SERC Daresbury Laboratory, Warrington, UK, pp. 59-66.

Kow, Y.W., and Wallace, S.S. (1987). Mechanism of action of *Escherichia coli*

endonuclease III. Biochemistry *26*, 8200-8206.

Kraulis, P.J. (1991). MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures. J. Appl. Crystallogr. *24*, 946-950.

Krokan, H.E., Standal, R., and Slupphaug, G. (1997). DNA glycosylases in the base excision repair of DNA. Biochem. J. *325*, 1-16.

Labahn, J., Schärer, O.D., Long, A., Ezaz-Nikpay, K., Verdine, G.L., and Ellenberger, T.E. (1996). Structural basis for the excision repair of alkylation-damaged DNA. Cell *86*, 321-329.

Lapidus, A., Galleron, N., Sorokin, A., and Ehrlich, S.D. (1997). Sequencing and functional annotation of the *Bacillus subtilis* genes in the 200 kb *rrnB-dnaB* region. Microbiology *143*, 3431-3441.

Laskowski, R.A., McArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHEK-A program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283-291.

Lau, A.Y., Schrer, O.D., Samson, L., Verdine, G.L., and Ellenberger, T. (1998). Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. Cell *95*, 249-258.

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). X-ray structure of a DNA decamer containing

7,8-dihydro-8-oxoguanine. Proc. Natl. Acad. Sci. USA *92*, 719-723.

Matthews, B.W. (1968). Solvent content of protein crystals. J. Mol. Biol. *33*, 491-497.

Melamede, R.J., Hatahet, Z., Kow, Y.W., Ide, H., and Wallace, S.S. (1994). Isolation and characterization of endonuclease VIII from *Escherichia coli*. Biochemistry *33*, 1255-1264.

Merritt, E.A., and Murphy, M.E.P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. Acta Crystallogr. *D50*, 869-873.

Michaels, M.L., Tchou, J, Grollman, A.P., and Miller, J.H. (1992). A repair system for 8-oxo-7,8-dihydrodeoxyguanine. Biochemistry *31*, 10964-10968.

Mikawa, T., Kato, R., Sugahara, M., and Kuramitsu, S. (1998). Thermostable repair enzyme for oxidative DNA damage from extremely thermophilic bacterium, *Thermus thermophilus* HB8. Nucleic Acids Res. *26*, 903-910.

Mol, C.D., Arvai, A.S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H.E., and Tainer, J.A. (1995). Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. Cell *80*, 869-878.

Mullen, G.P., and Wilson, S.H. (1997). DNA polymerase beta in abasic site repair: a structurally conserved helix-hairpin-helix motif in lesion detection by base excision repair enzymes. Biochemistry *36*, 4713-4717.

Murphy, T.M., and Gao, M.J. (1998). Two cDNAs (Accession Nos. AF099970 and AF099971) encoding *Arabidopsis* homologs of bacterial formamidopyrimidine – DNA glycosylase genes are produced by alternative processing (PGR98-204). Plant Physiol. *118*, 1535.

Nakagawa, N., Sugahara, M., Masui, R., Kato, R., Fukuyama, K., and Kuramitsu, S. (1999). Crystal structure of *Thermus thermophilus* HB8 UvrB protein, a key enzyme of nucleotide excision repair. J. Biochem. (Tokyo) *126*, 986-990.

Nicholls, A., Bharadwaj, R., and Honig, B. (1993). GRASP-graphical representation and analysis of surface properties. Biophys. J. *64*, A116.

O'Connor, T.R., Graves, R.J., de Murcia, G., Castaing, B., and Laval, J. (1993). Fpg protein of *Escherichia coli* is a zinc finger protein whose cysteine residues have a structural and/or functional role. J. Biol. Chem. *268*, 9063-9070.

O'Connor, T.R., and Laval, J. (1989). Physical association of the 2,6-diamino-4-hydroxy-5*N*-formamidopyrimidine-DNA glycosylase of *Escherichia coli* and an activity nicking DNA at apurinic/apyrimidinic sites. Proc. Natl. Acad. Sci. USA *86*, 5222-5226.

Ohtsubo, T., Matsuda, O., Iba, K., Terashita, I., Sekiguchi, M, and Nakabeppu, Y. (1998). Molecular cloning of AtMMH, an *Arabidopsis thaliana* ortholog of the *Escherichia coli mutM* gene, and analysis of functional domains of its product. Mol. Gen. Genet. *259*, 577-590.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307-326.

Purmal, A.A., Rabow, L.E., Lampman, G.W., Cunningham, R.P., and Kow, Y.W. (1996). A common mechanism of action for the N-glycosylase activity of DNA N-glycosylase/AP lyases from *E. coli* and T4. Mutat. Res. *364*, 193-207.

Rabow, L.E., and Kow, Y.W. (1997). Mechanism of action of base release by *Escherichia coli* Fpg protein: role of Lysine 155 in catalysis. Biochemistry *36*, 5084-5096.

Savva, R., McAuley-Hechet, K., Brown, T., and Pearl, L. (1995). The structural basis of specific base-excision repair by uracil-DNA glycosylase. Nature *373*, 487-493.

Schrock, R.D., III, and Lloyd, R.S. (1991). Reductive methylation of the amino terminus of endonuclease V eradicates catalytic activities. Evidence for an essential role of the amino terminus in the chemical mechanisms of catalysis. J. Biol. Chem. *266*, 17631-17639.

Shibutani, S., Takeshita, M., and Grollman, A.P. (1991). Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. Nature *349*, 431-434.

Sidorkina, O.M., and Laval, J. (1998). Role of lysine-57 in the catalytic activities of *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg protein). Nucleic Acids Res. *26*, 5351-5357.

Sugahara, M., Mikawa, T., Kato, R., Kumasaka, T., Yamamoto, M., Fukuyama, K., Inoue, Y., and Kuramitsu, S. (2000). Crystallization and preliminary X-ray Crystallographic Studies of *Thermus thermophilus* HB8 MutM protein involved in repairs oxidative DNA damage. J. Biochem. (Tokyo) 127, 9-11.

Suzuki, M., Matsui, K., Yamada, M., Kasai, H., Sofuni, T., and Nohmi, T. (1997). Construction of mutants of *Salmonella typhimurium* deficient in 8-hydroxyguanine DNA glycosylase and their sensitivities to oxidative mutagens and nitro compounds. Mutat. Res. *393*, 233-246.

Swartley, J.S., and Stephens, D.S. (1995). Co-transcription of a homologue of the formamidopyrimidine-DNA glycosylase (fpg) and lysophosphatidic acid acyltransferase (nlaA) in *Neisseria meningitidis*. FEMS Microbiol. Lett. *134*, 171-176.

Thayer, M.M., Ahern, H., Xing, D., Cunningham, R.P., and Tainer, J.A. (1995). Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. EMBO J. *14*, 4108-4120.

Tchou, J., and Grollman, A.P. (1995). The catalytic mechanism of Fpg protein. Evidence for a Schiff base intermediate and amino terminus localization of the catalytic site. J. Biol. Chem. *270*, 11671-11677.

Tchou, J., Kasai, H., Shibutani, S., Chung, M.H., Laval, J., Grollman, A.P., and Nishimura, S. (1991). 8-Oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity.

Proc. Natl. Acad. Sci. USA *88*, 4690-4694.

Tchou, J., Michaels, M.L., Miller, J.H., and Grollman, A.P. (1993). Function of zinc finger in *Escherichia coli* Fpg protein. J. Biol. Chem. *268*, 26738-26744.

Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr. *D55*, 849-861.

Thompson J.D., Higgins D.G., and Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673-4680.

Vassylyev D.G., Kashiwagi, T., Mikami, Y., Ariyoshi, M., Iwai, S., Ohtsuka, E., and Morikawa, K. (1995). Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: structural basis for damaged DNA recognition. Cell *83*, 773-782.

Wood M.L., Dizdaroglu M., Gajewski E. and Essigmann J.M. (1990). Mechanistic studies of ionizing radiation and oxidative mutagenesis: genetic effects of a single 8-hydroxyguanine (7-hydro-8-oxoguanine) residue inserted at a unique site in a viral genome. Biochemistry *29*, 7024-7032.

Yamagata, Y, Kato, M, Odawara, K., Tokuno, Y., Nakashima, Y., Matsushima, N., Yasumura, K., Tomita, K., Ihara, K., Fujii, Y., Nakabeppu, Y., Sekiguchi, M., and Fujii, S.

(1996). Three-dimensional structure of a DNA repair enzyme, 3-methyladenine DNA glycosylase II, from *Escherichia coli*. Cell *86*, 311-319.

Yamamoto, M., Kumasaka, T., Fujisawa, T., and Ueki, T. (1998). Trichromatic Concept at SPring-8 RIKEN Beamline I. J. Synchrotron Rad. *5*, 222-225.

Zharkov, D.O., Rieger, R.A., Iden, C.R., and Grollman, A.P. (1997). NH_2-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein. J. Biol. Chem. *272*, 5335-5341.

ACKNOWLEDGEMENTS

I would like to express great appreciation to Professors Y. Inoue, K. Fukuyama and S. Kuramitsu, and Drs. T. Kumasaka, M. Yamamoto, R. Masui and R. Kato for their continuing guidance and many valuable discussions. I am very grateful to my colleagues Dr. T. Mikawa, N. Nakagawa and J. Ishijima for providing protein samples and helping experiments. I am grateful to Dr. Tetsuya Ishikawa, RIKEN Harima Institute, and Dr. Tatsuo Ueki, Japan Synchrotron Radiation Research Institute (JASRI), for provision of X-ray facilities. Finally, I thank my colleagues in Professor Kuramitsu's laboratory for their kind help in this work.

LIST OF PUBLICATIONS

Sugahara, M., Suzuki, N., Iwashita-Ohno, T., and Miki, K. (1996). Crystallization and preliminary X-ray diffraction Studies of θ-toxin (Perfringolysin O), a pore-forming cytolysin of *Clostridium perfringens*. Journal of Crystal Growth *168*, 288-291.

Mikawa, T., Kato, R., Sugahara, M., and Kuramitsu, S. (1998). Thermostable repair enzyme for oxidative DNA damage from extremely thermophilic bacterium, *Thermus thermophilus* HB8. Nucleic Asids Res. *26*, 903-910.

Shibata, A., Nakagawa, N., Sugahara, M., Masui, R., Kato, R., Kuramitsu, S., and Fukuyama, K. (1999). Crystallization and preliminary X-ray diffraction studies of a DNA excision repair enzyme, UvrB, from *Thermus thermophilus* HB8. Acta Crystallogr. *D55*, 704-705.

Nakagawa, N., Sugahara, M., Masui, R., Kato, R., Fukuyama, K., and Kuramitsu, S. (1999). Crystal structure of *Thermus thermophilus* HB8 UvrB protein, a key enzyme of nucleotide excision repair. J. Biochem. (Tokyo) *126*, 986-990.

Sugahara, M., Mikawa, T., Kato, R., Kumasaka, T., Yamamoto, M., Fukuyama, K., Inoue, Y., and Kuramitsu, S. (2000). Crystallization and preliminary X-ray Crystallographic Studies of *Thermus thermophilus* HB8 MutM protein involved in repairs oxidative DNA damage. J. Biochem. (Tokyo) *127*, 9-11.