A New Family of Deamination Repair Enzymes in the Uracil DNA Glycosylase Superfamily

Hyun-Wook Lee1, Brian N. Dominy2 and Weiguo Cao1,*

From 1Department of Genetics and Biochemistry, South Carolina Experiment Station, Clemson University, Room 219 Biosystems Research Complex, 51 New Cherry Street, Clemson, SC 29634, USA
2Department of Chemistry, Clemson University, 100 Hunter Hall, Clemson, SC 29634, USA
* Corresponding Author: email: wgc@clemson.edu; Tel.: (864) 656-4176; Fax: (864) 656-0393.
Running Title: A new family of hypoxanthine DNA glycosylases

DNA glycosylases play a major role in the repair of deaminated DNA damage. Previous investigations identified five families within the uracil DNA glycosylase (UDG) superfamily. All enzymes within the superfamily studied thus far exhibit uracil DNA glycosylase activity. Here we identify a new class of DNA glycosylases in the UDG superfamily that lacks UDG activity. Instead, these enzymes act as hypoxanthine DNA glycosylases in vitro and in vivo. Molecular modeling and structure-guided mutational analysis allowed us to identify a unique catalytic center in this class of DNA glycosylases. Based on unprecedented biochemical properties and phylogenetic analysis, we propose this new class of DNA repair glycosylases that exists in bacteria, archaea, and eukaryotes as family 6 and designate it as the HDG family. This study demonstrates the structural evolvability that underlies substrate specificity and catalytic flexibility in the evolution of enzymatic function.

DNA is subject to chemical damage under normal and stressed biological conditions (1). A common type of DNA base damage involves deamination of cytosine to uracil, adenine to hypoxanthine, and guanine to xanthine and oxanine. DNA glycosylases play a major role in the repair of deaminated DNA damage (2). The uracil DNA glycosylase (UDG) superfamily consists of five families based on conserved motifs and structural similarity (3). Structurally, they are organized by a four-stranded β-sheet surrounded by α-helices, while functionally, all of the DNA glycosylases within the superfamily studied thus far are proven biochemically to be uracil DNA glycosylases. Due to the need for removing deaminated cytosine from the genome, almost all organisms contain at least one uracil DNA glycosylase. Uracil-N-glycosylases (UNG) are family 1 UDGs first discovered in Escherichia coli (4). All characterized UNGs show exquisite specificity towards uracil in both double-stranded (ds) and single-stranded (ss) DNA (5). Family 2 enzymes are represented by human thymine DNA glycosylase (TDG), E. coli mismatch-specific uracil-DNA glycosylase (MUG), and a broad substrate specificity TDG from the fission yeast Schizosaccharomyces pombe. This family exhibits more sequence and functional diversity than family 1. The family 3 SMUG1 enzymes (single-strand-selective monofunctional uracil-
DNA glycosylase) are considered hybrids with a MUG/TDG-like motif 1 and an UNG-like motif 2, which are active as uracil DNA glycosylases and xanthine DNA glycosylases (6,7). While the prokaryotic family 4 UDG enzymes act on both single-stranded and double-stranded uracil-containing DNA, family 5 UDGs are found to be active towards G/U and, to a lesser degree, towards T/I substrates. Regardless of which family an enzyme belongs to, a common biochemical feature is that they all have uracil DNA glycosylase activity.

In the course of studying deaminated DNA repair, we identified a new group of DNA glycosylases in the UDG superfamily. Unlike all previously known families, enzymes from this new family possess no uracil DNA glycosylase activity, and instead, exhibit repair activity towards hypoxanthine, a deamination product of adenine. The catalytic center is located on a completely conserved asparagine situated in a distinct orientation relative to the glycosidic bond. Based on the distinct repair capacity and active site architecture, we propose this class of repair enzymes as family 6 in the UDG superfamily. The discovery of this new family enzyme, which exists in all three domains of life, underlies the functional and catalytic diversity in the UDG superfamily and has implications for the evolution of repair enzymes.

**Experimental Procedures**

**Strains and Genomic DNA**

*E. coli* host strain BH214 (thr-1, ara-14, leuB6, tonA31, lacY1, tsx-78, galK2, galE2, dcm-6, hisG4, rpsL, xyl-5, mtl-1, thi-1, ung-1, tyrA::Tn10, mug::Tn10, supE44, (DE3)) is a kind gift from Dr. Ashok Bhagwat (Wayne state university, Detroit, MI) and JM109 is from our laboratory collection. *E. coli* BW1506 (CC106 nfi-1::cat) are a kind gift of Dr. Bernard Weiss (Emory University, Atlanta, GA) (8,9). Genomic DNA from *Methanosarcina barkeri* strain Fusaro and *Methanosarcina acetivorans* C2A is a kind gift from Dr. William Metcalf (University of Illinois, Urbana, IL).

**Cloning, Expression and Purification of Mba DNA Glycosylase**

The *Methanosarcina barkeri* DNA glycosylase gene (GenBank accession number: YP304295.1) was amplified by PCR using the forward primer Mba.F (5'-GGG AAT TCC ATA TGA AAA AAC AAG GTT TCC CAC CAG TCA TT -3'; the NdeI site is underlined) and the reverse primer Mba.R (5'-CCG CTC GAG ACG CTT GAA AGC TTC TGC CCA TCC CGA TTT -3'; the XhoI site is underlined). The PCR reaction mixture (50 µl) consisted of 8 ng of *Mba* genomic DNA, 200 nM forward primer and reverse primer, 1 x Taq PCR buffer (New England Biolabs), 200 µM each dNTP, and 5 units of Taq DNA polymerase (New England Biolabs). The PCR procedure included a predenaturation step at 94°C for 3 min, 30 cycles of three-step amplification with each cycle consisting of denaturation at 94°C for 40 sec, annealing at 60°C for 40 sec and extension at 72°C for 1 min, and a final extension step at 72°C for 10 min. The PCR product was purified with Gene Clean 2 Kit (Qbiogene). Purified PCR product and plasmid pET21a were digested with NdeI and XhoI, and followed by purifying gene fragments with Gene Clean 2 Kit and ligated according to the manufacturer’s instruction manual. The ligation mixture
was transformed into *E. coli* strain JM109 competent cells prepared by electroporation (10). The sequence of the Mba DNA glycosylase gene in the resulting plasmid (pET21a-Mba) was confirmed by DNA sequencing.

To express the C-terminal His-6-tagged Mba glycosylase, the plasmid pET21a-Mba was transformed into *E. coli* strain BH214 (ung- mug-) by electroporation (10). An overnight *E. coli* culture was diluted 100-fold into LB medium supplemented with 100 µg/mL ampicillin. The *E. coli* cells were grown at 37°C while being shaken at 250 rpm until the optical density at 600 nm reached approximately 0.6. After adding IPTG to a final concentration of 1 mM, the culture was grown at room temperature for an additional 16 h. The cells were collected by centrifugation at 4,000 rpm at 4°C and washed once with pre-cooled sonication buffer.

To purify the Mba DNA glycosylase protein, bacterial cells from a 500-ml culture grown to late exponential phase were harvested by centrifugation at 4,000 rpm for 10 min. The cell pellet was suspended in 7 ml of lysis buffer (20 mM Tris-HCl (pH 7.5), 1 mM EDTA (8.0), 0.1 mM DTT, 0.15 mM PMSF and 50 mM NaCl) and followed by sonication at output 5 for 3 x 1 min with 5 min rest on ice between intervals using a Sonifier Cell Disruptor 350 (Branson). The lysate was clarified by centrifugation at 12,000 rpm for 20 min and filtered through a 25 mm GD/X syringe filter (Whatman). The supernatant was transferred into a fresh tube and loaded onto a 1-ml HiTrap chelating column (GE Healthcare). The column was washed with 5 ml of chelating buffer A (20 mM sodium phosphate (pH 7.4), 500 mM NaCl and 2 mM imidazole). The bound protein in the column was eluted with a linear gradient of 0-100 % chelating buffer B (chelating buffer A and 500 mM imidazole).

Fractions of the eluate were analyzed by 12% SDS-PAGE and those fractions containing Mba glycosylase protein (40% chelating buffer B) were pooled. The partially purified protein was then loaded onto a 1-ml HiTrap SP column, washed with 5 ml of HiTrap SP buffer A (20 mM HEPES (pH 8.0), 1 mM EDTA and 0.1 mM DTT) and eluted with a linear gradient of 0-100 % HiTrap SP buffer B (HiTrap SP buffer A and 1 M NaCl). Fractions containing Mba glycosylase (30-50% HiTrap SP buffer B) were pooled and concentrated through Microcon YM 10 (Millipore). The protein concentration was quantified by SDS-PAGE analysis using bovine serum albumin as a standard. The Mba glycosylase protein was stored in aliquots at -80°C. Prior to use, the protein was diluted in equal volume of 2 x storage buffer (20 mM Tris-HCl (pH 8.0), 2 mM DTT, 2 mM EDTA, 400 µg/ml BSA, 100% Glycerol).

**Cloning, Expression and Purification of Mac DNA Glycosylase**

The *Methanosarcina acetivorans* DNA glycosylase gene (GenBank accession number: NP615428.1) was amplified by PCR using the forward primer Mac.F (5’-GGGAATTCCATATGATAAAGCGAGTTTTCCTGCAGTCCTT-3’; the NdeI site is underlined) and the reverse primer Mac.R (5’-CCGCTCGAGATGCCTGAAAACAGCTCCACTCGGATCC-3’; the XhoI site is underlined). The cloning, expression and purification steps were the same as described above.
Site-Directed Mutagenesis

Site-directed mutagenesis of Mba glycosylase was performed using an overlapping extension PCR procedure (11). The first round of PCR was carried out using pET21a-Mba as template DNA with two pairs of primers, Mba.F and N39A-R (5’-CCTCCAGAAATCAGCGCCTGGGATG CCCATAGTATTG-3’) pair and and N39A-F (5’-TATGGGCATCCAGGCCTGATTTCC TGGAGGCTGCT-3’) and Mba.R pair; Mba.F and N39D-R (5’-CCTCCAGAAATCATC GCCTGGATG CCCATA GTATTG-3’) pair and N39D-F (5’-TATGGGCATCCAGGCCTGGCTTTC TGGAGGC TGCTT-3’) and Mba.R pair; Mba.F and N39Q-R (5’-CCTCCAGAAATCTTG GCCTGGATG CCCATA GTATTG-3’) pair, and N39Q-F (5’-TATGGGCATCCAGGGCAAGA TTTCTGGAGGCTGCTT-3’) and Mba.R pair, respectively. PCR and subsequent molecular cloning procedures were performed similarly as previously described (12).

DNA Glycosylase Activity Assay

The preparation of fluorescently labeled deaminated base-containing oligodeoxynucleotide substrates were described previously (12). DNA glycosylase cleavage assays were performed at 37°C for 60 min in a 10-µl reaction mixture containing 10 nM oligonucleotide substrate, 100 nM glycosylase protein, 20 mM Tris-HCl (pH 7.5), 5 mM EDTA and 2 mM 2-mercaptoethanol. The resulting abasic sites were cleaved by incubation at 95°C for 5 min after adding 1-µl of 1 N NaOH. Reactions were quenched by addition of an equal volume of GeneScan stop buffer. Samples (4 µl) were loaded onto a 7 M urea-10% denaturing polyacrylamide GeneScan gel (acrylamide : bisacrylamide = 19:1, 1 x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA). Electrophoresis was conducted at 1500 V for 1.5 h using an ABI 377 sequencer (Applied Biosystems). Cleavage products and remaining substrates were quantified using GeneScan analysis software.

Reversion of lacZ Gene in E. coli CC106

The plasmids pET21a-Mba and pET21a-Mba-N39A were digested with EcoRI and XhoI and the fragment containing the Mba DNA glycosylase gene was cloned to pBluescript SK(+) to generate pBS-Mba and pBS-Mba-N39A. The resulting plasmids were transformed to E. coli BW1506 (CC106 nfi-1::cat). Tester cultures, inoculated as a single colony, were grown in Luria-Bertani (LB) medium at 37°C for 16 hr. Overnight cultures (1 ml) were transferred to 4 ml fresh LB medium and incubated at 37°C for 4-5 hr until OD<sub>600</sub> reached 0.6. After adding IPTG to a final concentration of 1 mM, the cultures were incubated at 37°C for an additional 5 hr. E. coli cells (2 x 10<sup>8</sup>) were incubated with 30 µl of 1 M NaNO<sub>2</sub> (30 mM final concentration) and 970 µl of 100 mM sodium acetate buffer (pH 4.6) at 37°C for 30 min. Top agar consisting of 0.5 % NaCl, 0.6% agar (Difco) and 0.2 mg/ml nutrient broth (Difco) was autoclaved and then maintained at 45°C in a water bath. After centrifugation, the treated cells were suspended with 1 ml of 10 mM MgSO<sub>4</sub>. Molten top agar prepared above (8 ml) was added to the suspended cells and the mixtures were immediately overlaid onto a minimum lactose (ML) plate prepared according to
the recipe as previously described (13). Cells were incubated at 37°C for four days to allow a few cycles of cell division to fix the mutations in the presence of nutrient broth (14).

Molecular Modeling

Sequence alignment was performed using CLUSTAL W. pairwise alignment of the amino acid sequence from the Mba glycosylase (YP_304295.1) and chain A of 2rba.pdb (human TDG). This alignment resulted in 19% identity between the two sequences (Supplemental Fig. S1). Based on the sequence alignment and the 2rba pdb structure, a homology model was constructed for the Mba enzyme using the NEST program (15). The NEST program utilizes an artificial evolution based algorithm to generate a homology model of a protein based on an input amino acid sequence alignment and corresponding template structure. Default settings were used to develop the model structure of the Mba glycosylase including: no tuning or adjustment of the sequence alignment, a short minimization to remove clashes between atoms, and a single round of structural refinement.

Phylogenetic Analysis

The phylogenetic tree was generated with a neighbor-joining algorithm within the MEGA 5 software package (http://www.megasoftware.net/) applied to a multiple sequence alignment produced with the CLUSTAL W program. The parameters for pairwise alignment are: gap opening penalty, 10; gap extension penalty 0.1. The parameters for multiple alignment are: gap opening penalty, 10; gap extension penalty 0.2. Other parameters are: protein weight matrix, Blosum; residue-specific penalties: on; hydrophilic penalties: on; gap separation distance: 4; end gap separation: off.

Results

Deaminated Base Repair Activities

Family 2 enzymes (MUG/TDG) in the UDG superfamily were previously found only in bacteria and eukaryotic systems. In a search (Psi BLAST) for uracil DNA repair enzymes in archaea, we found a hypothetical protein (YP_502486.1) in archaea Methanospirillum hungatei that showed an insignificant E-value (E = 0.41) with E. coli MUG. However, querying GenBank with this protein led to the identification of a large number of genes in archaea, eubacteria, and eukaryotes with significant homology to UDG superfamily enzymes. Alignment of proteins from this family to the UDG family indicated significant sequence divergence in active site motifs I and II (Fig. 1A). Most notably, the position equivalent to the catalytic residue in E. coli UNG (family 1, D64), human TDG (Family 2, N140) and E. coli MUG (family 2, N18) is occupied by a leucine residue at position 22. To investigate the repair activities of this class of enzymes, we expressed the homologous gene from Methanosarcina barkeri (Mba) in an E. coli strain lacking both ung and mug and purified the protein to homogeneity. Interestingly, the recombinant protein completely lacked any detectable uracil DNA glycosylase activity on all four uracil-containing base pairs and single-stranded uracil-containing DNA (Fig. 1B). Alternatively, hypoxanthine DNA glycosylase activity was found in all four double-stranded substrates (Fig. 1B). Some minor xanthine DNA glycosylase activity was also detected in this Mba
enzyme but no oxanine DNA glycosylase activity was detected (Fig. 1B). To verify that the lack of UDG activity and existence of HDG activity is not unique to the Mba enzyme, we also investigated the deaminated DNA glycosylase activity in the homologous gene from *Methanosarcina acetivorans*. A similar glycosylase pattern was observed (supplemental Fig. S2), indicating that this is a common repair function in this class of enzymes.

**Kinetic Analysis and Catalytic Center**

Kinetic analysis showed that this glycosylase was most active on the G/I base pair with an apparent rate constant of 0.085 per min, followed by T/I, A/I and C/I base pairs (Table 1). No activity was detected on single-stranded inosine-containing DNA (Fig. 1B). The catalytic mechanisms of enzymes from families 1 and 2 have been extensively studied. D64 in *E. coli* UNG has been identified as the key catalytic residue that activates a water molecule for subsequent nucleophilic attack on the anomeric carbon (16). In *E. coli* MUG, N18 utilizes the mainchain and sidechain oxygen to activate the water and the mainchain amino group to interact with uracil (17). However, the residue (Asp or Asn) that can perform the catalytic function is notably missing in the Mba enzyme. Instead, the equivalent position is occupied by a hydrophobic leucine residue at position 22 (Fig. 1A and supplemental Fig. S3). We constructed a model of the Mba enzyme based on the crystal structure of human TDG because they share moderate sequence homology (Fig. 2A) (18). The availability of the AP site in the human TDG structure allowed us to approximate the distances of amino acid residues to the anomeric carbon. The modeled structure is further confirmed by modeling of the Mba DNA glycosylase sequence based on a recently solved NMR structure of the homologous *M. acetivorans* DNA glycosylase (supplemental Fig. S4).

Within a 10 Å radius of the AP site, the following potential catalytic residues were identified: N39 (7.08 Å), N84 (9.89 Å), N113 (8.62 Å), D74 (5.55 Å), D25 (9.46 Å), D86 (8.60 Å), E82 (3.05 Å), E85 (8.28 Å). Among the eight residues identified, only N39 is completely conserved in the Mba enzyme and its homologs (supplemental Fig. S3). D74, D86 and N113 are highly conserved. To assess the role of these residues in catalysis, D74, D86 and N113 were substituted by either Ala or Asn. A significant portion of HDG activity on the G/I substrate was retained, suggesting that D74, D86 and N113 did not play a key role in catalysis (Table 1). Given the completely conserved nature, N39 was substituted with Ala, Asp and Gln. The Ala substitution, which removed the functional group of the N39 sidechain, completely abolished the glycosylase activity (Fig. 2B and Table 1). Conversion of the sidechain from an amide group to a carboxyl group (N39D) rendered the enzyme active only on the G/I substrate (Fig. 2B and Table 1). Addition of a methylene group into the sidechain (N39Q) allowed the enzyme to retain significant activity on the G/I substrate but negligible activity on the T/I substrate (Fig. 2B and Table 1). These results indicate that N39 serves as a catalytic center in the active site environment.

**In vivo Repair**

To assess the role of the new group of DNA glycosylase may play in vivo, we adopted the lacZ-based genetic assay (8). The *E. coli* CC106 strain
contains a G to A transition mutation at codon 461 that allows for detection of A to G revertants, which has been used to provide evidence to implicate bacterial endonuclease V (encoded by \textit{nfi} gene) as an enzyme for the repair of adenine deamination (9). It is known that DNA polymerases predominantly incorporate dCMP to pair with inosine in DNA templates (19). To assess the role that the hypoxanthine DNA glycosylase activity may play \textit{in vivo}, both a plasmid containing the wt Mba glycosylase gene and a plasmid containing the N39A mutant gene were transformed into an \textit{nfi} deficient CC106 strain. While the N39A mutant failed to suppress A to G mutations in the CC106 \textit{nfi} cells under nitrosative stress, the wt Mba DNA glycosylase reduced the number of revertants by more than five-fold (Fig. 2C). These results suggest that the hypoxanthine DNA glycosylase activity in the Mba repair enzyme is responsible for the removal of hypoxanthine \textit{in vivo}.

\section*{Discussion}

\subsection*{UDG-less Enzymes in UDG Superfamily}

Since the discovery of the first uracil DNA glycosylase in \textit{E. coli} in 1974, all the uracil DNA glycosylases studied possess UDG activity. While repair activities on other deaminated bases have been found in different families in the UDG superfamily, UDG activity is always present. This work for the first time reports a class of UDG enzymes without UDG activity. Instead, they exhibit significant repair activity on hypoxanthine, the deamination product of adenine. Data from the \textit{in vivo} reversion assay are consistent with the notion that this class of enzymes acts as a hypoxanthine DNA glycosylase (Fig. 2C). These observations suggest that different UDG families have evolved to possess distinct repair specificities during evolution. Similar to the family 1 UNG enzymes, the repair activity of this class of enzymes appears to be limited to hypoxanthine, indicating that it may have a narrow specificity. The HDG activity follows the order of G/I > T/I > A/I > C/I, which is consistent with the stability of inosine-containing base pairs (20-22). Therefore, the tendency of spontaneous base flipping appears to play an important role in determining the repair activity of this class of enzymes. Hypoxanthine differs from uracil in that it lacks a C\textsubscript{2}-keto and has an imidazole ring. It remains to be answered how this class of enzyme specifically recognizes hypoxanthine but not uracil. More structural studies are needed to address this question.

\subsection*{Unique Catalytic Center}

The catalytic mechanisms in families 1 and 2 have been extensively studied. In family 1 UNG enzymes, a completely conserved Asp residue (D145 in human UNG) rotates 120° once bound to a uracil-containing DNA and acts as a general base to activate a bound water molecule (23-25). In family 2 enzymes, the corresponding position is occupied by an Asn residue, which is proposed to activate a bound water molecule through sidechain and mainchain interactions (17,26). In the Mba enzyme, the equivalent position is occupied by a Leu residue (Fig. 1A). In the modeled structure, the L22 is not positioned toward the scissile bond (Fig. 2A). Among residues in the vicinity of the scissile bond, N39 emerged as the key catalytic residue that may perform a catalytic function similar to Asp in family 1 and Asn in family 2. The
evidence supporting this notion includes, 1) N39 is completely conserved among the new class of enzymes; 2) N39 is better positioned in the modeled structure to activate a water molecule to potentiate an in-line attack on the glycosidic bond; 3) the N39A mutant completely loses its catalytic activity.

**A New Family in UDG Superfamily**

Based on the unprecedented lack of uracil DNA glycosylase activity, the presence of HDG activity, and the unique location of the catalytic center, we performed a phylogenetic analysis within the UDG superfamily. As shown in Fig. 3, this new class of enzymes emerged as a distinct group within the superfamily. Therefore, we propose this class of enzymes as family 6 in the UDG superfamily and designate it as the HDG family. The information provided for the family 6 enzymes could help us understand the multiplicity of UDG superfamily enzymes in a variety of organisms. For example, while a family 4 enzyme in *M. barkeri* may carry out the repair function for cytosine deamination damage by acting as a UDG, the family 6 enzyme described here could perform the necessary repair for adenine deamination by acting as an HDG. In addition to the surprising lack of uracil DNA repair function, the use of different catalytic center is particularly interesting, as it exemplifies the ability of enzymes to adapt to different catalytic environment to perform a catalytic function. Therefore, this study demonstrates the structural and catalytic flexibility that underlies the functional diversity in enzyme evolution.

**References**


**Footnotes**

This project was supported in part by the Cooperative State Research, Education, and Extension Service/United States Department of Agriculture (SC-1700274, technical contribution No. 5905), the Department of Defense (W81XWH-10-1-0385 to W. C.), the National Institutes of Health (GM090141 to W. C.), and the National Science Foundation (MCB-0953783 to B. N. D.). We thank Drs. Ashok Bhagwat, Bernard Weiss, and William Metcalf for providing strains or genomic DNA.

**The abbreviations used are:** HDG, hypoxanthine DNA glycosylase; UNG, uracil N-glycosylase; MUG, mismatch-specific uracil DNA glycosylase; TDG, thymine DNA glycosylase; SMUG1, single-strand-selective monofunctional uracil DNA glycosylase; UDG, uracil DNA glycosylase; XDG, xanthine DNA glycosylase; U, uridine; X, xanthosine; I, inosine; O, oxanosine.
Figure Legends

Figure 1. Sequence motifs and deaminated base repair activity from Mba DNA glycosylase. 


B. Deaminated base repair activity from Mba DNA glycosylase. DNA glycosylase activity assays were performed as described in Experimental Procedures with 100 nM wt Mba glycosylase protein and 10 nM oligonucleotide substrate. A control reaction was performed with *E. coli* UNG (lane 21). X, U, I, O: xanthosine-, uridine-, inosine-, oxanosine-containing DNA. Lane 1, C/X; 2, G/X; 3, A/X; 4, T/X; 5, single-stranded X; 6, C/U; 7, G/U; 8, A/U; 9, T/U; 10, single-stranded U; 11, C/I; 12, G/I; 13, A/I; 14, T/I; 15, single-stranded I; 16, C/O; 17, G/O; 18, A/O; 19, T/O; 20, single-stranded O; 21, G/U.

Figure 2. Catalytic center and role of Mba hypoxanthine DNA glycosylase activity in vivo. 

A. Modeled structure of Mba DNA glycosylase. The AP site, L22, D25, N39, D74, D86 and N113 are shown in color. The conserved N95 is also shown in color, where the distance between the sidechain oxygen (O) and the anomeric carbon is 17.70 Å. 

B. Hypoxanthine DNA glycosylase activity of N39A, N39D and N39Q mutants. DNA glycosylase activity assays were performed as described in Experimental Procedures with 100 nM wt Mba glycosylase protein and 10 nM oligonucleotide substrate. Lane 1, C/I; 2, G/I; 3, A/I; 4, T/I; 5, single-stranded I. 

C. Role of Mba DNA glycosylase in vivo. Numbers of revertants were scored on minimal lactose plates after incubation at 37°C for four days. The data are the means ± SD of three independent experiments. Open bar, N39A mutant of Mba DNA glycosylase; hatched bar: wt Mba DNA glycosylase.

Figure 3. Phylogenetic analysis of UDG superfamily. 

### Table 1. Apparent rate constants of hypoxanthine DNA glycosylase activity in Mba DNA glycosylase (min$^{-1}$)$^a$

<table>
<thead>
<tr>
<th></th>
<th>Bottom Strand</th>
<th>Top Strand</th>
<th>C</th>
<th>G</th>
<th>A</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td></td>
<td></td>
<td>0.0092</td>
<td>0.085</td>
<td>0.014</td>
<td>0.026</td>
</tr>
<tr>
<td>D25I</td>
<td></td>
<td></td>
<td>0.0052</td>
<td>0.034</td>
<td>0.0099</td>
<td>0.024</td>
</tr>
<tr>
<td>N39A</td>
<td></td>
<td></td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>N39D</td>
<td></td>
<td></td>
<td>n.a.</td>
<td>0.0011</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>N39Q</td>
<td></td>
<td>I</td>
<td>n.a.</td>
<td>0.012</td>
<td>n.a.</td>
<td>n.d.$^c$</td>
</tr>
<tr>
<td>D74A</td>
<td></td>
<td></td>
<td>n.a.</td>
<td>0.0029</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>D74N</td>
<td></td>
<td></td>
<td>n.a.</td>
<td>0.0018</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>D86A</td>
<td></td>
<td></td>
<td>n.a.</td>
<td>0.011</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>N113A</td>
<td></td>
<td></td>
<td>n.a.</td>
<td>0.061</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

$^a$: The reactions were performed as described in Experimental Procedures with 100 nM Mba DNA glycosylase and 10 nM substrate except that samples were withdrawn at 10-min intervals over 60 min. Data are an average of at least two independent experiments.

$^b$: n.a.: No activity was detected under assay conditions.

$^c$: n.d.: Not determined due to a very low level of activity.
Family 6 (HDG) Mba N--19-GSLPGDVSRKHY-YGHPGNDFW-34-D-62-SSSGANR--16-C
Family 1 (UNG) Eco N--61-GQDPYHGPGQAHGLAFSRPIAT-37-N-63-HPSPLSA-36-C
Family 2 (MUG/TDG) Hsa N-137-GINPGMAAYKGGH-YPGPQNHFW-30-N-77-MPSSSAR-135-C
Family 3 (SMUG1) Gme N--55-GMNPGPWGMAQTGVPGAVVTE-56-N-73-HPSPLSA-23-C
Family 4 (UDGa) Pae N--39-GEAPGASEDEAGPR-FVGAAGQLL-17-N-81-HPAAVLRT-28-C
Family 5 (UDGb) Pae N--65-GLAPAHQQRTGRMTGDASAQN-31-S-74-HPSPLNV-24-C

**A**

**Motif 1**

**Motif 2**

**B**

|   | C | X | X | A | T | X | X | C | U | G | U | A | U | T | U | C | G | I | A | I | T | I | C | O | G | O | A | O | T | O | U |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

Substrate

Product

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Product

Lee et al. Fig. 2
A new family of deamination repair enzymes in the uracil DNA glycosylase superfamily

Hyun-Wook Lee, Brian N. Dominy and Weiguo Cao

J. Biol. Chem. published online June 3, 2011

Access the most updated version of this article at doi: 10.1074/jbc.M111.249524

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2011/06/15/M111.249524.DC2

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2011/06/03/jbc.M111.249524.full.html#ref-list-1