THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 26, pp. 18413–18426, June 27, 2014

Specificity and Catalytic Mechanism in Family 5 Uracil DNA Glycosylase*

Received for publication, March 20, 2014, and in revised form, May 15, 2014 Published, JBC Papers in Press, May 16, 2014, DOI 10.1074/jbc.M114.567354

Bo Xia[‡], **Yinling Liu**[§], **Wei Li**[‡], **Allyn R. Brice**[§], **Brian N. Dominy**[§], **and Weiguo Cao**[‡1]

From the [‡]*Department of Genetics and Biochemistry, South Carolina Experiment Station and the* [§]*Department of Chemistry, Clemson University, Clemson, South Carolina 29634*

Background: Uracil DNA glycosylases are DNA repair enzymes involved in the removal of base damage.
Results: Family 5 UDGb is a uracil, hypoxanthine, and xanthine DNA glycosylase.
Conclusion: Family 5 UDGb adapts multiple catalytic amino acids for the excision of pyrimidine and purine deaminated DNA bases.
Significance: Family 5 UDGb exemplifies functional diversity in enzyme superfamilies.

UDGb belongs to family 5 of the uracil DNA glycosylase (UDG) superfamily. Here, we report that family 5 UDGb from *Thermus thermophilus* HB8 is not only a uracil DNA glycosyase acting on G/U, T/U, C/U, and A/U base pairs, but also a hypoxanthine DNA glycosylase acting on G/I, T/I, and A/I base pairs and a xanthine DNA glycosylase acting on all double-stranded and single-stranded xanthine-containing DNA. Analysis of potentials of mean force indicates that the tendency of hypoxanthine base flipping follows the order of G/I > T/I, A/I > C/I, matching the trend of hypoxanthine DNA glycosylase activity observed *in vitro*. Genetic analysis indicates that family 5 UDGb can also act as an enzyme to remove uracil incorporated into DNA through the existence of dUTP in the nucleotide pool. Mutational analysis coupled with molecular modeling and molecular dynamics analysis reveals that although hydrogen bonding to O_2 of uracil underlies the UDG activity in a dissociative fashion, Tth UDGb relies on multiple catalytic residues to facilitate its excision of hypoxanthine and xanthine. This study underscores the structural and functional diversity in the UDG superfamily.

DNA base deamination is a common mechanism of DNA damage caused by environmental and endogenous agents. Because of the reactivity of the exocyclic amino groups, DNA bases are subject to hydrolytic or oxidative deamination, in which the amino group in a DNA base is converted to keto group. Adenine (A), cytosine (C), and guanine (G) are deaminated to hypoxanthine (I), to uracil (U), and to xanthine (X) and oxanine (O), respectively (see Fig. 1*A*). Because of the altered base pair preferences, base deamination may result in mutations.

The base excision repair, initiated by DNA glycosylase, is a major pathway to repair damage caused by base deamination. The uracil DNA glycosylase (UDG)[2] superfamily, consisting of six families, is involved in the repair of deaminated base damage. Although family 1 UNGs show rather narrow specificity toward uracil and its derivatives (1, 2), family 2 MUG/TDG enzymes have much broader specificity with some members showing activity toward all deaminated bases (3–6). Most interestingly, recent studies have indicated that human TDG is a DNA glycosylase that is involved in the removal of formyl-C and carboxyl-C during enzymatic demethylation (7, 8), suggesting that some of the enzymes in the UDG superfamily have evolved functions beyond DNA repair. Family 3 SMUG1 enzymes, found in vertebrates and bacteria, act as both uracil DNA glycosylases and xanthine DNA glycosylases (XDGs) (9, 10). Family 4 UDGa was initially discovered in the hyperthermophilic bacterium *Thermotoga maritima* (11). Family 5 UDGb was first reported in the hyperthermophilic archaean *Pyrobaculum aerophilum* (12). Most recently, the family 6 enzymes were discovered as a class of enzymes with hypoxanthine DNA glycosylase (HDG) activity but not with uracil DNA glycosylase activity (13).

Family 5 UDGb exists in archaea and bacteria, many of which are hyperthermophiles or thermophiles. Biochemical characterization of *P. aerophilum* UDGb indicates that it can remove uracil, hydroxymethyluracil, or fluorouracil opposite from guanine base and hypoxanthine from T/I base pairs (12). The uracil DNA glycosylase and hypoxanthine DNA glycosylase activity prevents mutation resulting from cytosine and adenine deamination in cells (14). UDGb from *Mycobacterium tuberculosis* is also found to be active on ethenocytosine and the mutation rate in the absence of UDGb increases by 2-fold (15, 16). Similarly, biochemical and genetic analyses indicate that UDGb from thermophilic bacterium *Thermus thermophilus* (Tth) is a uracil DNA glycosylase that can reduce the mutation rate by 3-fold (17, 18). The crystal structure of UDGb from *T. thermophilus* confirms that UDGb adopts a structural fold similarly seen in other UDG enzymes (19).

* This work was supported, in whole or in part, by National Institutes of Health Grant GM090141. This work was also supported by South Carolina Experiment Station Grant SC-1700274 (technical contribution No. 6233), Department of Defense Grant W81XWH-10-1-0385, and National Science Foundation Career Award MCB-0953783 (to B. N. D.).

[1] To whom correspondence should be addressed: Dept. of Genetics and Biochemistry, Clemson University, Rm. 049 Life Sciences Facility, 190 Collings St., Clemson, SC 29634. Tel.: 864-656-4176; Fax: 864-656-0393; E-mail: wgc@clemson.edu.

[2] The abbreviations used are: UDG, uracil DNA glycosylase; XDG, xanthine DNA glycosylase; HDG, hypoxanthine DNA glycosylase; Tth, *T. thermophilus*; PMF, potential of mean force; MD, molecular dynamics.

Previous studies have provided valuable information on the structure-function relationship of family 5 UDGb enzymes and their physiological roles. However, some fundamental questions remain to be answered. How broad is the specificity of UDGb toward other deaminated bases? How does the active site in UDGb catalyze the cleavage of the glycosidic bond in deaminated DNA? To answer these questions, we conducted a comprehensive biochemical, genetic, and molecular dynamics analysis using UDGb from Tth as a model. The data presented here demonstrate that Tth UDGb can act as a uracil DNA glycosylase with enzymatic activity on double-stranded uracil-containing DNA, as a hypoxanthine DNA glycosylase with enzymatic activity on double-stranded hypoxanthine-containing DNA except for the C/I base pair, and as a xanthine DNA glycosylase with enzymatic activity on both double-stranded and single-stranded xanthine-containing DNA. This study also establishes the correlation between hypoxanthine DNA glycosylase activity and stability of hypoxanthine-containing base pairs; reveals the inverse correlation between the uracil DNA glycosylase activity against the A/U base pair and cell survival in an *Escherichia coli* cell deficient in *ung dut xth* genes; and identifies several catalytic residues that play an important role in the removal of deaminated bases in DNA. A model explaining the catalytic function of family 5 UDGb is proposed.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification of Tth UDGb—The UDGb gene from *T. thermophilus* HB8 (GenBank™ accession number YP_144415.1) was amplified by PCR using the forward primer Tth UDGbF (5′-TCAGGTGTGCATATGGACAGGG-AAGCCTTCGTCCAAACC-3′; the NdeI site is underlined) and the reverse primer Tth UDGbR (5′-TGAATCAAGCTTA-AGCCCGGCGAGGCGTTTAGCCTC-3′; the HindIII site is underlined). The PCR mixture (20 μl) consisted of 10 ng of *T. thermophilus* HB8 genomic DNA, 500 nM forward and reverse primers, 1× Phusion DNA polymerase buffer, 200 μM each dNTP, and 0.2 unit of Phusion DNA polymerase (New England Biolabs). The PCR procedure included a predenaturation step at 98 °C for 30 s; 30 cycles of three-step amplification with each cycle consisting of denaturation at 98 °C for 15 s, annealing at 60 °C for 15 s, and extension at 72 °C for 20 s; and a final extension step at 72 °C for 10 min. The PCR product was purified by gel DNA recovery kit (Zymo Research). The purified PCR product and plasmid pET21a were digested by NdeI and HindIII, purified by gel DNA recovery kit, and ligated according to the manufacturer's instructional manual. The ligation mixture was transformed into *E. coli* strain HB101 competent cells by electroporation. The sequence of the Tth UDGb gene in the resulting plasmid (pET21a-Tth-UDGb) was confirmed by DNA sequencing. Site-directed mutagenesis was performed similarly as previously described (13).

The pET21a-Tth-UDGb was transformed into *E. coli* strain BH214 (*ung*$^-$, *mug*$^-$) by the standard protocol to express the C-terminal His_6-tagged Tth UDGb protein. Induction, sonication, and purification were carried out as previously described with the following modifications (13). Prior to HiTrap chelating column chromatography, the sonicated solutions were incubated at 75 °C for 15 min. Denatured proteins were removed by centrifugation at 12,000 rpm for 20 min. Fractions (300–400 mM imidazole, 60–80% chelating buffer B) containing the Tth UDGb protein as seen on 12.5% SDS-PAGE were pooled and concentrated by Amicon YM-10 (Millipore).

Oligodeoxynucleotide Substrates—Oligodeoxynucleotides containing deoxyuridine (U), deoxyinosine (I), deoxyxanthosine (X), or deoxyoxanosine (O) were obtained or constructed as previously described (see Fig. 1*B*) (10).

DNA Glycosylase Activity Assay—DNA glycosylase cleavage assays for Tth UDGb were performed under optimized reaction conditions at 50 °C for 60 min in a 10-μl reaction mixture containing 10 nM oligonucleotide substrate, an indicated amount of glycosylase, 20 mM Tris-HCl (pH 7.6), 100 mM KCl, 1 mM DTT, and 1 mM EDTA. The resulting abasic sites were cleaved by incubation at 95 °C for 5 min after adding 1 μl of 1 N NaOH. The reaction mixtures (2 μl) were mixed with 7.8 μl of Hi-D formamide and 0.2 μl of GeneScan 500 LIZ Size Standard (Invitrogen) and analyzed by Applied Biosystems 3130xl sequencer with a fragment analysis module. The cleavage products and remaining substrates were quantified by GeneMapper software. For kinetics measurements, the reactions were carried out with 10 nM oligonucleotide substrate and 100 nM glycosylase. Increasing the enzyme concentration by 3-fold did not change the kinetic profile, suggesting that the reactions reached saturation condition. Samples were withdrawn at 0, 2.5, 5, 10, 20, 40, and 60 min. For some substrates, such as G/U and G/I with faster kinetics, samples were also withdrawn at the 1.0-min time point to capture the early reaction. The apparent rate constants were determined by curve fitting using the integrated first-order rate equation,

$$P = P_{max}(1 - e^{-kt}) \quad \text{(Eq. 1)}$$

where P is the product yield, P_{max} is the maximal yield, t is time, and k is the apparent rate constant.

Base Flipping Potential of Mean Force Calculations—Base flipping potentials of mean force (PMF) were constructed based on the dodecamer sequence d(GTCAGIGCATGG), where the hypoxanthine (I) was the base to be flipped out of the helix. Using the program 3DNA (20), the canonical B-form DNA structure of the sequence d(GTCAGCGCATGG) was constructed. The base complementary to I was systematically modeled as guanine, adenine, cytosine, and thymine. Starting from these four models of B-form DNA, umbrella sampling was performed to calculate the PMF associated with flipping deaminated DNA bases out of the double helix (21, 22). A detailed description of the computational methods can be found in our previous work (21).

Survival Analysis of E. coli BW276 Strain Complement with Tth udgb—The wild type and mutant Tth *udgb* genes were subcloned into plasmid pBluescript II SK (+). The WT Tth *udgb* gene is described as an example. The WT Tth *udgb* gene was amplified by PCR using the plasmid pET21a-Tth-UDGb as a template and with the forward primer TthUDGbF-pBS (5′-CCGGAATTCCCATATGGACAGGGAAGCCTTCG-3′; the EcoRI site is underlined) and the reverse primer Tth UDGbR-pBS (5′ ACGCGTCGACTCAGTGGTGGTGGTGGTGG-3′; the SalI site is underlined). The PCR conditions were the same

as the cloning of the Tth UDGb gene. The purified PCR products digested with a pair of EcoRI and SalI endonucleases were ligated to the cloning vector pBluescript II SK (+) treated with the same pair of restriction endonucleases. The recombinant plasmids were confirmed by DNA sequencing and transformed into *E. coli* BW276 (*dut*^{ts} *ung*$^{-}$ *xth*$^{-}$). *E. coli* BW276 (*thi-1*, *relA1*, *spoT1*, *dut-1*, *rfa-209*::Tn*10*, *ung-1* (*xth-pncA*)*90*) was a kind gift from Dr. Bernard Weiss (Emory University, Atlanta, GA) (23). Single colonies from each variant were inoculated into 1 ml of LB medium (with 100 μg/ml ampicillin and 125 μg/ml thymidine) and grown at 22 °C for 16 h. Overnight cultures (1 ml) were transferred to 4 ml of fresh LB medium (with 100 μg/ml ampicillin and 125 μg/ml thymidine) and grown at 22 °C until A_{600} reached 0.6. After adding isopropyl β-D-thiogalactopyranoside to a final concentration of 1 mM, the cultures were incubated at 22 °C until A_{600} reached 1.0. Before plating, 40 μl of 100 mM isopropyl β-D-thiogalactopyranoside was spread on the LB plates containing ampicillin and thymidine. Afterward, diluted cells (100 μl) were plated on LB plates. Cell numbers were scored after 24 h of incubation at 42 °C or after 72 h of incubation at 22 °C. The relative plating efficiencies were taken as the ratios of the cell numbers between 42 and 22 °C.

Molecular Modeling and Molecular Dynamics Simulations—The crystal structure of family 5 uracil DNA glycosylase (Tth UDGb) was acquired from the RCSB Protein Data Bank (accession code 2DEM), and used as a model for subsequent computational analysis (19). The initial structures of uridine, inosine, xanthosine, and oxanosine nucleotides bound to the family 5 UDG enzyme were obtained by manually modifying the structure of 2-deoxy-5-phosphono-ribose extracted from the crystal structure of Tth UDGb using the Swiss-Pdb Viewer program (24). Specifically, the 2-deoxy-5-phosphono-ribose was aligned to a cytidine, and the cytosine was changed to uracil, hypoxanthine, xanthine, and oxanine, respectively. After building the initial complex structures, an explicit solvent system using the TIP3P water model was constructed in the CHARMM c35b6 molecular mechanics package using a suitably sized box (25). The minimum distance between any of the atoms of the solvated UDG-DNA complex, and the box boundary was maintained to at least 9 Å. Sodium chloride ions were added to the system to achieve an electrically neutral system. The CHARMM 27 all hydrogen force field for proteins and nucleic acids were used (26, 27). Particle mesh Ewald summation was applied in the periodic boundaries condition for the efficient calculation of long range electrostatic interactions (28). Energy minimization was performed by using 4,000 steepest descent steps followed by the adopted basis Newton-Raphson (ABNR) method with the harmonic constraints from 10 to 1 kcal/(mol·Å^2) in decrements of 3 kcal/(mol·Å^2) every 1,000 steps to remove any unfavorable van der Waals clashes. Using a Langevin barostat (29), an isothermal-isobaric ensemble at 300 K was constructed using the NAMD program (30). An integration time step of 1 fs was used to avoid any significant structural deformation during equilibration and production run. Coordinates were saved every 2 ps. A total of 2 ns of equilibration and 3 ns of production simulation were performed for each structural analysis. VMD 1.9.1 was used for visualization (31).

RESULTS

Substrate Specificity of Tth UDGb—Previous studies have indicated that family 5 UDGb enzymes can act on some deaminated base damage. To gain a complete understanding of the specificity of the family 5 UDGb on deaminated bases, we measured the repair activity of Tth UDGb using all 20 possible uracil (U)-, hypoxanthine (I)-, oxanine (O)-, and xanthine (X)-containing double-stranded and single-stranded DNA substrates under the assay conditions in which the enzyme was in 10-fold excess over the substrate (Fig. 1*C*). Consistent with previous reports, Tth UDGb acted as a double-stranded uracil DNA glycosylase with a relatively low activity on the A/U base pair (Fig. 1*C*). No activity on single-stranded U-containing substrate was detected under the assay conditions. Tth UDGb could also act as a hypoxanthine DNA glycosylase with the strongest activity on the G/I base pair but no activity detected on the C/I base pair and the single-stranded I-containing substrate (Fig. 1*C*). The enzyme showed no activity on any of the oxanine-containing substrates, indicating that family 5 Tth UDGb is not an oxanine DNA glycosylase. On the other hand, Tth UDGb was able to incise xanthine in all xanthine-containing DNA, in which all five substrates were hydrolyzed to close to completion (Fig. 1*C*). Similar to some other enzymes in the UDG superfamily, no glycosylase activities were detected with the deaminated substrates under the assay conditions in which the enzyme:substrate ratio was 1:1 (data not shown). Consistent with previous observations, Tth UDGb was not inhibited by Ugi peptide and acted as a monofunctional glycosylase (Ref. 12 and data not shown).

To determine the catalytic efficiencies on the deaminated DNA, we measured the apparent rate constants. A representative time course analysis is shown in Fig. 1*D*, and complete data are summarized in Table 1. For the U-containing substrates, Tth UDGb was most active on three mismatched base pairs (T/U, C/U, and G/U) with apparent rate constants around 0.36 per min (Table 1). The apparent rate constant for the A/U pair was 5-fold slower than that for G/U pair. For the I-containing substrates, the apparent rate constant of 0.35 per min for the G/I base pair was comparable with the mismatched U-containing base pairs (Table 1). However, the apparent rate constants for the T/I and A/I base pairs were more than 10-fold lower. For the X-containing substrates, the apparent rates constants ranged from 0.057 per min for the C/X base pair to 0.075 per min for the A/X base pair (Table 1).

HDG Activity and Stability of I-containing Base Pairs—The base flipping tendencies for uracil- and xanthine-containing base pairs have been studied previously (21). The UDG and XDG activity of Tth UDGb followed the general trend of base pair stability. The kinetic analysis indicated that the hypoxanthine DNA glycosylase in Tth UDGb was strongest on the G/I base pair followed by the T/I and A/I base pairs. The enzyme was least active with the C/I base pair. To assess how the order of HDG activity is correlated with the stability of the I-containing base pairs, we analyzed PMF for the double-stranded I-containing base pairs. The PMF essentially represents the change in free energy associated with transitioning the DNA from the base paired state to base flipped-out state and is therefore capa-

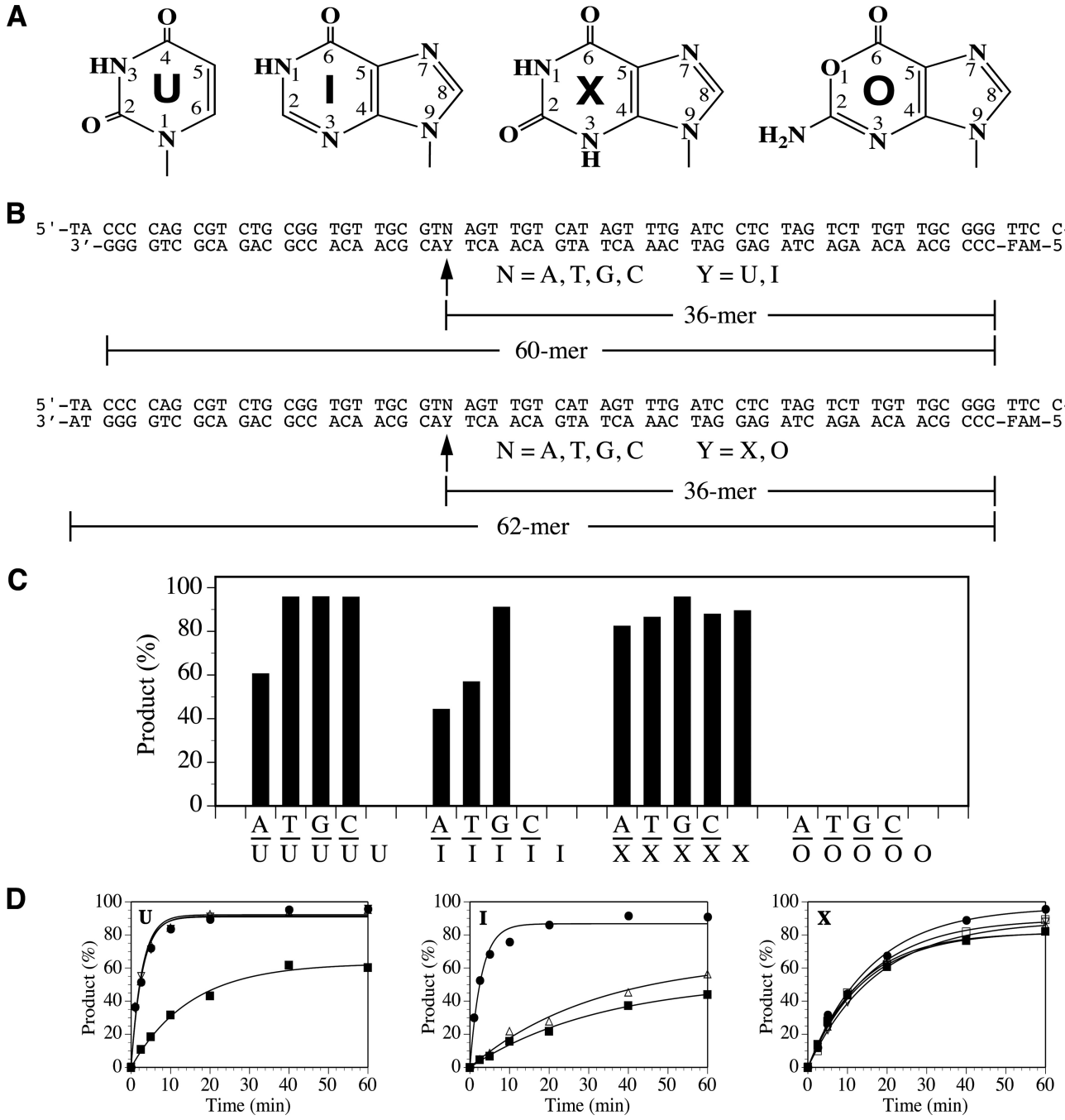


FIGURE 1. **Deaminated DNA repair activity in Tth UDGb.** *A*, chemical structures of deaminated DNA bases. *B*, sequences of xanthine (X)-, oxanine (O)-, hypoxanthine (I)-, and uracil (U)-containing oligodeoxyribonucleotide substrates. *FAM*, fluorophore. *C*, DNA glycosylase activity of Tth UDGb on U-, I-, O-, and X-containing substrates. Cleavage reactions were performed as described under "Experimental Procedures" with 100 nM WT Tth UDGb protein and 10 nM substrate. *D*, representative time course analysis of glycosylase activity of WT Tth UDGb on U-, I-, and X-containing DNA substrates. U-containing substrates: ■, A/U; △, T/U; ●, G/U; ▽, C/U. I-containing substrates: ■, A/I; △, T/I; ●, G/I. X-containing substrates: ■, A/X; △, T/X; ●, G/X; ▽, C/X; □, ss X.

ble of indicating the tendency of a deaminated base to flip out of an isolated B-form DNA double helix. Among the four base pairs, hypoxanthine demonstrates the greatest tendency (lowest PMF barrier) to flip when paired with guanine (Fig. 2). The thermodynamic stability of T/I and A/I base pairs were found to be similar to each other and more stable (higher PMF barrier) than the G/I pair. The most stable base pair is the C/I base pair, which can adopt a natural Watson-Crick base pair conformation. Interestingly, the PMF profile of the I-containing base pairs is quite consistent with activity pattern (Fig. 1 and Table 1), suggesting that the tendency of base flipping contributes to the recognition and consequently the catalysis.

Catalytic Residues in UDGb—To identify the amino acid residues in the family 5 Tth UDGb enzyme that may play an important role in its catalytic function, we took advantage of the Tth UDGb crystal structure complexed with an abasic site (19). The presence of the AP site in the structure helped us to locate potential catalytic residues that are in the vicinity of the scissile bond. Aspartate and asparagine are known catalytic residues in family 1 UNG and family 2 MUG, respectively. A close exami-

TABLE 1

Apparent rate constants of DNA glycosylase activity in wild type and mutant Tth UDGb enzymes

The reactions were performed as described under "Experimental Procedures" with 100 nM Tth UDGb protein and 10 nM substrate. The data are averages of three independent experiments. NA, no activity was detected under the assay conditions. ND, not determined because of a low level of activity.

		Rate constant				
	Bottom strand	Top strand A	Top strand T	Top strand G	Top strand C	
				min^{-1}		
WT	U	0.067 ± 0.0048	0.37 ± 0.031	0.35 ± 0.032	0.36 ± 0.032	NA
D75A		NA	0.028 ± 0.0040	0.031 ± 0.0063	ND	NA
N120A		NA	0.0065 ± 0.00053	0.095 ± 0.0052	0.029 ± 0.0037	NA
WT	I	0.030 ± 0.0042	0.031 ± 0.00059	0.35 ± 0.014	NA	NA
D75A		ND	ND	0.0061 ± 0.00043	NA	NA
N120A		NA	NA	0.0074 ± 0.00029	NA	NA
H190A		NA	NA	0.0022 ± 0.00063	NA	NA
WT	X	0.075 ± 0.0027	0.071 ± 0.0078	0.063 ± 0.0074	0.057 ± 0.0041	0.064 ± 0.0074

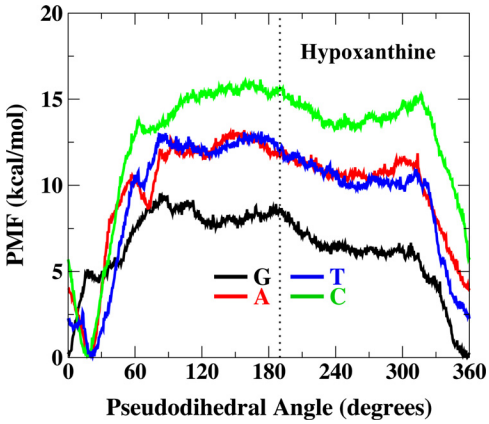


FIGURE 2. **PMFs of hypoxanthine-containing base pairs along the pseudodihedral angle coordinate.** Watson-Crick base pairing is ~10–30° pseudodihedral angle, and the flipped out state is ~190°. PMF profiles were generated in TIP3P explicit water solvent.

nation of the structure revealed several potential catalytic residues: Asp^{75}, Asn^{120}, His^{190}, and Asn^{195} (Fig. 3*A*). All these residues are highly conserved in family 5 UDGb homologs (Fig. 3*B*). To test the role of these residues in catalysis, Asp^{75} was substituted with Ala, Asn, Glu, and Gln; Asn^{120} and Asn^{195} were substituted with Ala, Asp, Glu, and Gln; and His^{190} was substituted with Ala, Asn, and Ser.

For the Asp^{75} position, substitution with an Ala residue exhibited the most profound effect on X-containing substrates because D75A completely lost all of its XDG activity (Fig. 4*A*). The UDG activity was also reduced, in particular for the A/U and C/U base pairs in which the activity was not detectable under the assay conditions or rather weak (Fig. 4*A*). For the T/U and G/U base pairs, the apparent rate constants were reduced by more than 10-fold (Table 1). For the I-containing substrates, the HDG activity on A/I and T/I was rather weak, and that on G/I was reduced by 57-fold (Fig. 4*A* and Table 1). The effect caused by substitution of Asp^{75} with Glu was in general much less severe than D75A because the D75E mutant still retained much of the activity (Fig. 4*B*). The substitutions of the carboxyl side chain with an amide side chain resulted in similar effects. Both D75N and D75Q mutants lost XDG activity while retaining much of their UDG and HDG activity (Fig. 4, *C* and *D*).

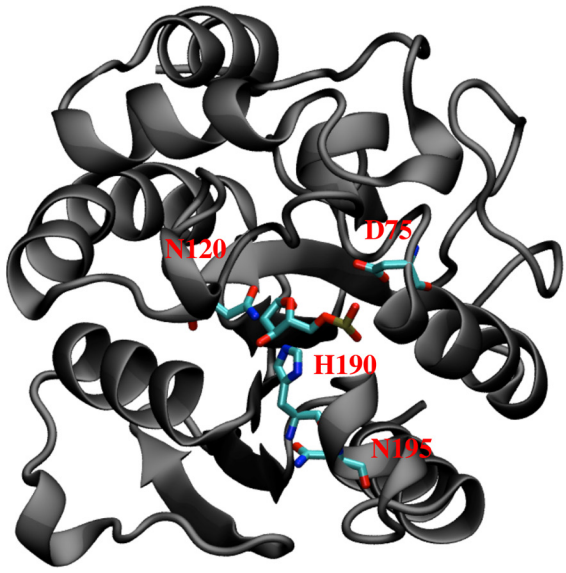
The effect caused by an Ala substitution at the N120 position was in general similar to D75A in which the XDG activity was lost, and the UDG and HDG activity was reduced (Fig. 5*A*). The main difference is that N120A retained a higher C/U and G/U activity but completely lost HDG activity on A/I and T/I (Fig. 5*A* and Table 1). The reduction of G/I activity, as judged by the apparent rate constants, was 47-fold (Table 1). Extension of the amide side chain by one methylene group, as shown in the N120Q mutant, exhibited the same effect as removal of the amide side chain as shown in N120A (Fig. 5*B*). Conversion of the amide group to a carboxyl group as shown in N120D mutant had a minimal effect on UDG and XDG activity except for a small reduction on the A/U base pair (Fig. 5*C*). The main effect of the Asp substitution was the reduction of HDG activity on A/I and T/I base pairs to below the level of detection under the assay conditions (Fig. 5*B*). N120E mutant amplified the effect of carboxyl substitution because UDG, HDG, and XDG activities were all reduced substantially (Fig. 5*D*). The effects observed in N120E and N120Q underscore the sensitivity of UDGb catalytic activity toward small length changes in the Asn^{120} side chain.

The mutational effect was most profound with the substitution of His^{190} with Ala. The H190A mutant lost all UDG, HDG, and XDG activity with the exception of retaining a low level of activity on G/I base pairs and a rather minor activity observed on G/X base pairs (Fig. 6*A*). The functional role of His^{190} could be accommodated by Asn or Ser to a very limited degree as indicated by the retention of weak UDG, HDG, and XDG activity (Fig. 6, *B* and *C*).

The effects of amino acid substitutions at the Asn^{195} position were less severe than those at the Asp^{75}, Asn^{120}, and His^{190} positions. In general, reduced UDG, HDG, and XDG activity was still observable in N195A, N195Q, N195D, and N195E mutants (Fig. 7). The most noticeable reduction of activity was observed using A/I and T/I substrates, in which the HDG activity was nearly or completely lost. Given that motif 2 serves as a wedge to insert into the space vacated by a flipped out base, mutations at the conserved Asn^{195} position may affect the wedging mediated by motif 2. The UDG activity on A/U base pairs was also reduced to varying degrees in N195A, N195Q, N195D, and N195E mutants. These mutants were used in the genetic analysis of A/U repair *in vivo* as described below.

Survival of UDGb in dut Temperature-sensitive E. coli Strain—Uracil may occur in DNA either through cytosine deamination or through incorporation of dUTP during DNA replication. In normal cells, incorporation of dUTP into DNA to form A/U base pairs is minimized because of the hydrolysis catalyzed by dUTPase (encoded by *dut*) (Fig. 8*A*). The dUMP thus produced is a precursor for the synthesis of dTMP by thymidylate syn-

A



B

			D75 N120 H190 N195
Family 5 (UDGb)	Tth	*N*--56-GLAPGAHGSNRTGRPFTGDASGAF-30-AAVRCAPPKN-69-HVSRQNT--23-*C*	
	Pae	*N*--65-GLAPAAHGGNRTGRMFTGDASAQN-31-SAVKCAPPKN-65-HPSPLNV--24-*C*	
	Sso	*N*--46-GLAPAGNGGNRTGRMFTGDESSNN-31-SAVKCAPPQN-76-HPSPRNM--25-*C*	
	Tvo	*N*--61-GLAPAATGGNRTGRVFTGDKSSDF-31-AAVKCVPPDN-71-HPSPRNV--23-*C*	
	Sco	*N*--60-GLAPAAHGGNRTGRMFTGDRSGDV-31-SPVHCAPPAN-79-HVSQRNT--26-*C*	
	Mtu	*N*-101-GLAPAAHGANRTGRMFTGDRSGDQ-31-APVRCAPPGN-71-HPSQQNM--24-*C*	
Family 1 (UNG)	Eco	*N*--61-GQDPYHGPGQAHGLAFSVRPGIAT-37-NTVLTVRAGQ-54-HPSPLSA--36-*C*	
Family 2 (MUG/TDG)	Eco	*N*--15-GINPGLSSAG-TGFPFAHPANRFW-29-KLVDRPTVQA-62-NPSGLSR--22-*C*	
Family 3 (SMUG1)	Gme	*N*--55-GMNPGPWGMAQTGVPFGEVAVVTE-56-NYCPLLFLTA-64-HPSPASP--21-*C*	
Family 4 (UDGa)	Tth	*N*--39-GEGPGEEEDK-TGRPFVGKAGQLL-17-NIVKCRPPQN-65-HPAYLLR--44-*C*	
Family 6 (HDG)	Mba	*N*--19-GSLPGDVSIR-KHQYYGHPGNDFW-31-DVFKAGKREG-52-SSSGANR--16-*C*	
		Motif 1	**Motif 2**

FIGURE 3. **Tth UDGb structure and sequence alignment.** *A*, Tth UDGb-AP site co-crystal structure (Protein Data Bank code 2DEM). The four residues (Asp[75], Asn[120], His[190], and Asn[195]) that were subjected to mutational analysis are shown in *red*. *B*, sequence alignment in family 5 UDGb and comparison with other UDG families. The alignment was based on BLAST and CLUSTALW analysis and constructed manually. Four conserved residues that are matched in *A* are shown in *red*. Family 5 (UDGb): *Tth*, *T. thermophilus* HB8, YP_144415.1; *Pae*, *P. aerophilum* str. IM2, NP_559226; *Sso*, *Sulfolobus solfataricus* P2, NP_344053.1; *Tvo*, *Thermoplasma volcanium* GSS1, NP_111346.1; *Sco*, *Streptomyces coelicolor* A3(2), NP_626251.1; *Mtu*, *M. tuberculosis* H37Rv, P64785 (Rv1259). Family 1 (UDG): *Eco*, *E. coli*, NP_289138. Family 2 (MUG/TDG): *Eco*, *E. coli*, P0A9H1. Family 3 (SMUG1): *Gme*, *G. metallireducens* GS-15,YP_383069. Family 3 (SMUG1): *Gme*, *G. metallireducens* GS-15,YP_383069; Family 4 (UDGa): *Tth*, *T. thermophilus* HB27, YP_004341.1.

thase (encoded by *tms*), which is then phosphorylated to form dTTP as one of the regular nucleoside triphosphates for DNA synthesis. The *E. coli* strain BW276 is deficient in *ung xth* and contains a temperature-sensitive *dut-1* gene (23). *E. coli* cells deficient in *xth* (encodes an AP endonuclease) is lethal, but the lethality is rescued by deletion of *ung* (23). At permissible temperature (22 °C), the active dUTPase prevents dUTP from incorporating into DNA during replication (Fig. 8*A*, *left panel*). At high temperature (42 °C), uracil is incorporated into DNA to form A/U base pairs because of the loss of much of the dUTPase activity. In the absence of the endogenous UNG, which removes uracil from A/U base pair, cells remain viable because a limited amount of uracil incorporated into the genome is tolerated (32). However, cells lose viability in the presence of active UNG, which constantly removes uracil from A/U base pairs and creates toxic AP sites (Fig. 8*A*, *right panel*).

We were interested in understanding how the ability of a UDG to remove uracil from an A/U base pair *in vitro* correlates to its ability to do so *in vivo*. In examining the UDG activity of Tth UDGb on the A/U base pair, we found a descending trend of the activity in the order of WT > D75Q > D75N > N120D > N195A > N195Q > N195D, N195E > D75E > N120Q, N120E, N120A, D75A in the Tth UDGb mutants (Figs. 4–7). We thought that this spectrum of UDG activity on the A/U base pair could provide an excellent opportunity to test how the *in vitro* activity correlates to *in vivo* uracil removal. As illustrated in Fig. 8*A*, the UDG activity on A/U base pairs in the wild type Tth UDGb was expected to render the *E. coli* BW 276 strain

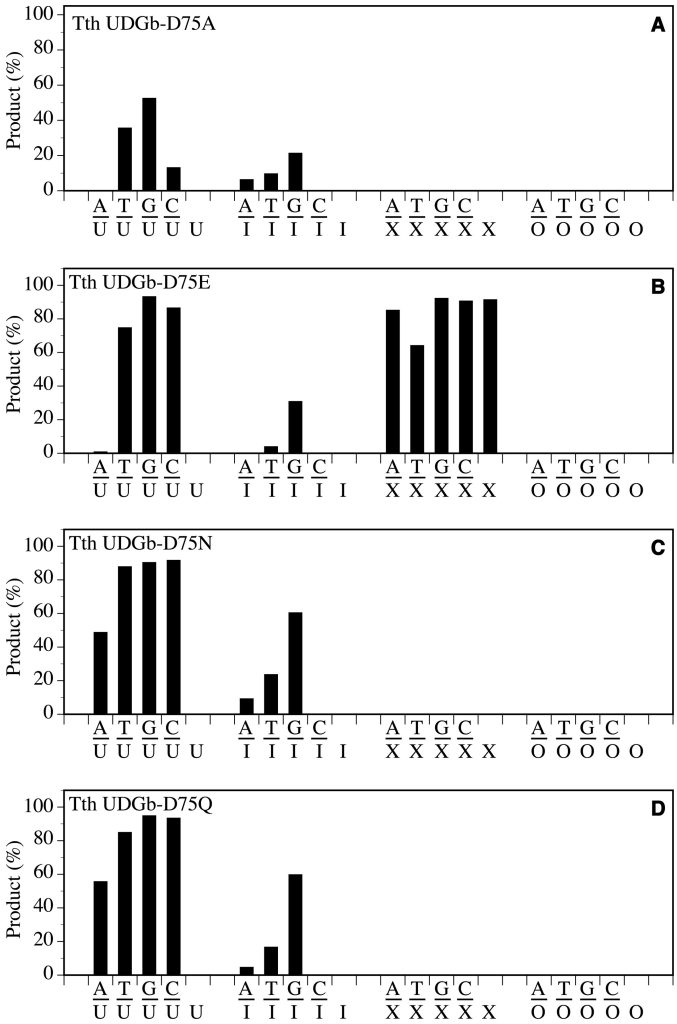


FIGURE 4. **Glycosylase activity of Tth UDGb D75 mutants on U-, I-, X-, and O-containing DNA substrates.** Cleavage reactions were performed as described under "Experimental Procedures" at 50 °C for 60 min with 100 nM protein and 10 nM substrate. *A*, Tth UDGb-D75A. *B*, Tth UDGb-D75E. *C*, Tth UDGb-D75N. *D*, Tth UDGb-D75Q.

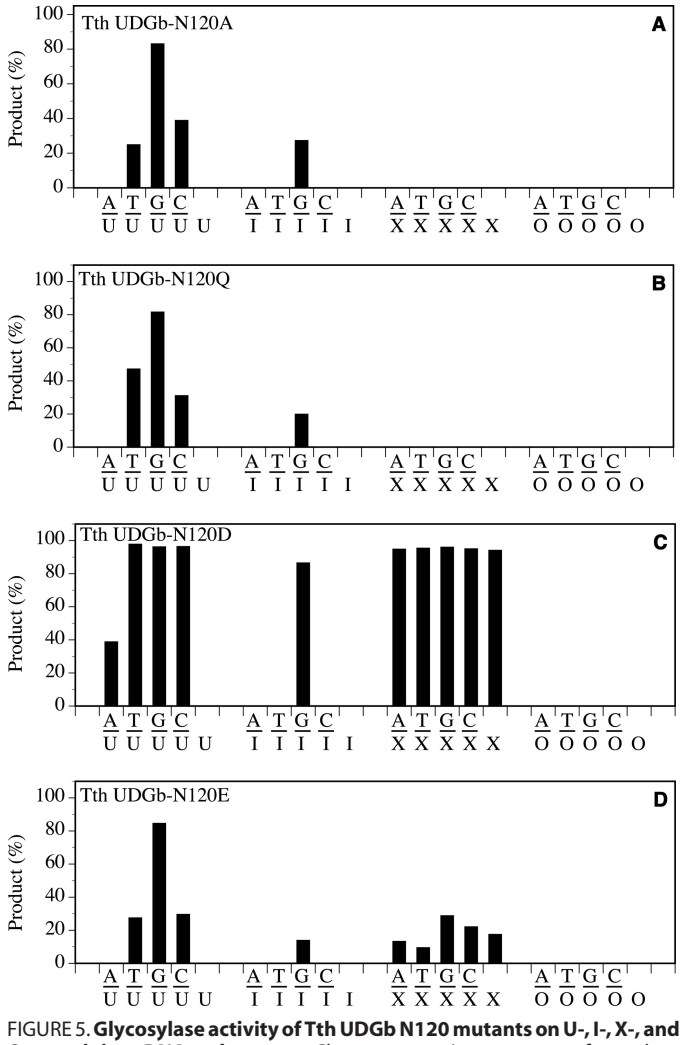


FIGURE 5. **Glycosylase activity of Tth UDGb N120 mutants on U-, I-, X-, and O-containing DNA substrates.** Cleavage reactions were performed as described under "Experimental Procedures" at 50 °C for 60 min with 100 nM protein and 10 nM substrate. *A*, Tth UDGb-N120A. *B*, Tth UDGb-N120Q. *C*, Tth UDGb-N120D. *D*, Tth UDGb-N120E.

much less viable. On the other hand, the lack of activity on A/U base pairs in the D75A mutant should allow the cells become much more viable, even at 42 °C. Mutants with intermediate levels of *in vitro* activity toward A/U base pairs should also exhibit intermediate cell viability. Indeed, we found an inverse correlation between the *in vitro* activity on A/U base pairs and the viability as measured by the relative plating efficiencies at 42 °C/22 °C (Fig. 8, *B* and *C*). These results indicate that a higher level of the activity on the A/U base pairs will lead to a better removal of uracil from A/U base pairs in the genome and thus limits the survival of the *dut*^{ts} cells.

DISCUSSION

Comparison of Substrate Specificity in UDG Superfamily—UDG superfamily encompasses six diverse families. This study reveals that family 5 UDGb is a deamination repair enzyme with rather broad specificity. Among the UDG enzymes we have investigated, only *Schizosaccharomyces pombe* TDG shows a broader specificity, because it acts on all deaminated bases (6). Family 5 UDGb can act as a uracil, hypoxanthine, and xanthine DNA glycosylase (Fig. 1). In comparison, family 1 UNG only works on uracil, whereas family 2 *E. coli* MUG and family 3 SMUG1 act on both uracil and xanthine. Therefore, family 5 UDGb is a versatile repair enzyme that can deal with multiple types of deaminated base lesions. Interestingly, family 5 UDGb is implicated in playing a role in the repair of hypoxanthine *in vivo* in addition to its role in the repair of uracil (14). Recently, we reported on a new family of enzymes in the UDG superfamily that act as hypoxanthine DNA glycosylases rather than uracil DNA glycosylases (13). Taken together, these data indicate that enzymes in the UDG superfamily can play a variety of roles in the repair of deaminated base lesions. Structurally, the active sites of the enzymes in the superfamily appear to be adaptable to accommodate both pyrimidine and purine deaminated products. Evolutionarily, different families have evolved a rather impressive spectrum of specificity, from family 1 UNG acting exclusively on one type of lesion (uracil) to family 6 HDG acting primarily on one type of lesion (hypoxanthine); from family 2 *E. coli* MUG and family 3 SMUG1 acting on two types of lesions (uracil and xanthine) to family 5 UDGb acting on three types of lesions (uracil, hypoxanthine, and xanthine) and

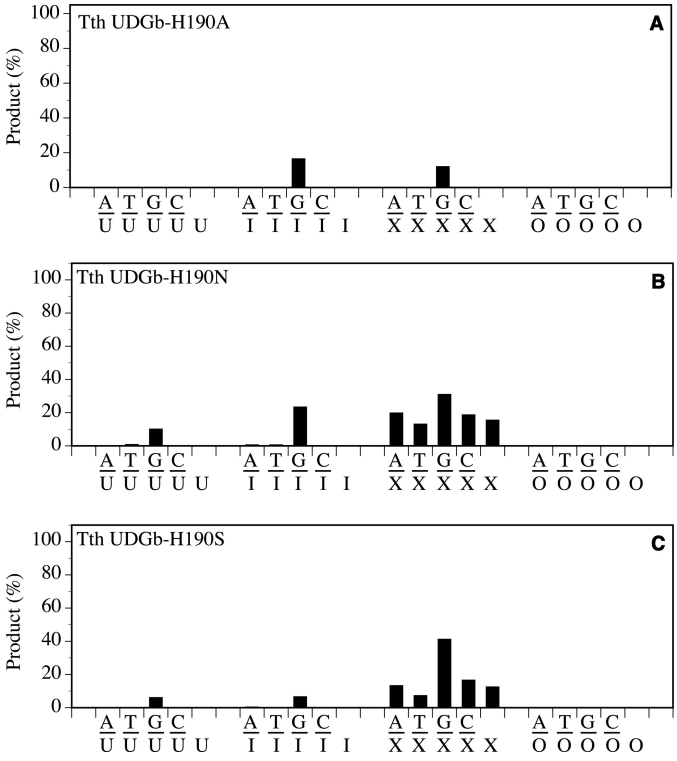


FIGURE 6. **Glycosylase activity of Tth UDGb H190 mutants on U-, I-, X-, and O-containing DNA substrates.** Cleavage reactions were performed as described under "Experimental Procedures" at 50 °C for 60 min with 100 nM protein and 10 nM substrate. *A*, Tth UDGb-H190A. *B*, Tth UDGb-H190N. *C*, Tth UDGb-H190S.

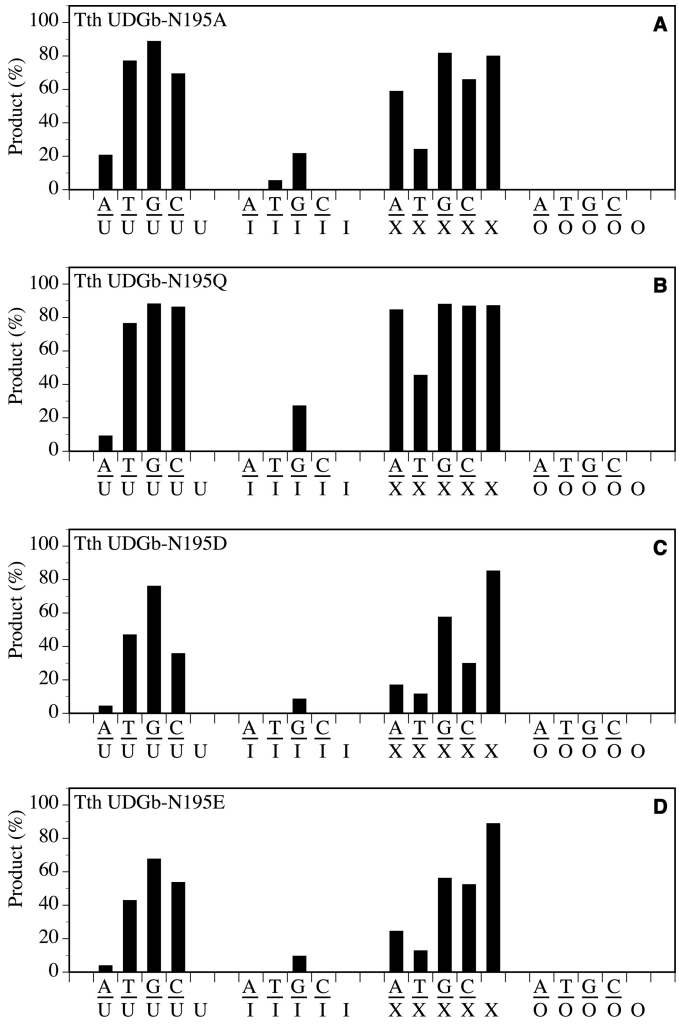


FIGURE 7. **Glycosylase activity of Tth UDGb N195 mutants on U-, I-, O-, and X-containing DNA substrates.** Cleavage reactions were performed as described under "Experimental Procedures" at 50 °C for 60 min with 100 nM protein and 10 nM substrate. *A*, N195A. *B*, N195Q. *C*, N195D. *D*, N195E.

family 2 *S. pombe* TDG acting on four types of lesions (uracil, hypoxanthine, xanthine, and oxanine). Additionally, the family 2 human TDG has evolved to act as a demethylase to remove formyl-C and carboxyl-C to initiate an enzymatic demethylation process in epigenetics.

HDG Activity and Base Flipping—The hypoxanthine DNA glycosylase activity from Tth UDGb showed a trend of G/I > A/I and T/I (Fig. 1 and Table 1). This trend, in general, is consistent with the observations we previously made in *Methanosarcina bakeri* HDG and human endonuclease V (13, 33). The PMF analysis indicates that the G/I base pair is the least stable among the four base pairs (Fig. 2). The T/I and A/I base pairs have a similar level of higher stability than the G/I base pair. The C/I base pair has the highest level of stability, in keeping with its structural similarity to a C/G base pair (34). The outcome of the computational analysis on the stability of the hypoxanthine-containing base pairs is consistent with experimental analyses previously reported (35–37). The stability of the base pairs as estimated by the PMF calculation suggests that the G/I base pair is more prone to base flipping. Indeed, the wild type Tth UDGb showed 10-fold higher activity with the G/I base pair than the A/I or the T/I base pair (Table 1). In addition, a large number of mutants still maintained some level of activity on the G/I base pair, whereas the activity on the A/I or T/I base pair was either negligible or not detected (Figs. 4–7). This is in accord with the notion that base flipping renders the glycosidic bond more reactive (38). The consistency between the hypoxanthine DNA glycosylase activity on different base pairs and the tendency of base flipping suggests that some of the repair enzymes rely on the instability of the base pairs to capture the deaminated base, hypoxanthine. Therefore, these data underscore the importance of base pair stability in promoting the catalysis of deaminated bases. This is an example in which the observed catalytic activity of an enzyme is perturbed not by chemically altering the substrate functional groups directly involved in enzyme recognition but by altering the conformational equilibrium of a substrate.

Inverse Correlation of Viability and the UDG Activity on A/U Base Pairs—The UDG enzymes may play two roles *in vivo*. First, when cytosine is deaminated, UDG can remove the deaminated cytosine, i.e., uracil, from the G/U base pair. Second, uracil may be incorporated into a DNA through dUTP to form A/U base pairs. The A/U base pair can be removed by a UDG with activity on A/U base pairs. The family 5 UDG enzymes possess repair activity on both G/U and A/U base pairs. The physiological role of the UDG in removing uracil from G/U base pairs has been studied (14, 18). However, whether the UDG activity of UDGb on A/U base pairs can remove uracil *in vivo* remained unclear. Based on the genetic properties of dut^{ts} *ung xth* genes in *E. coli*, we adopted strain

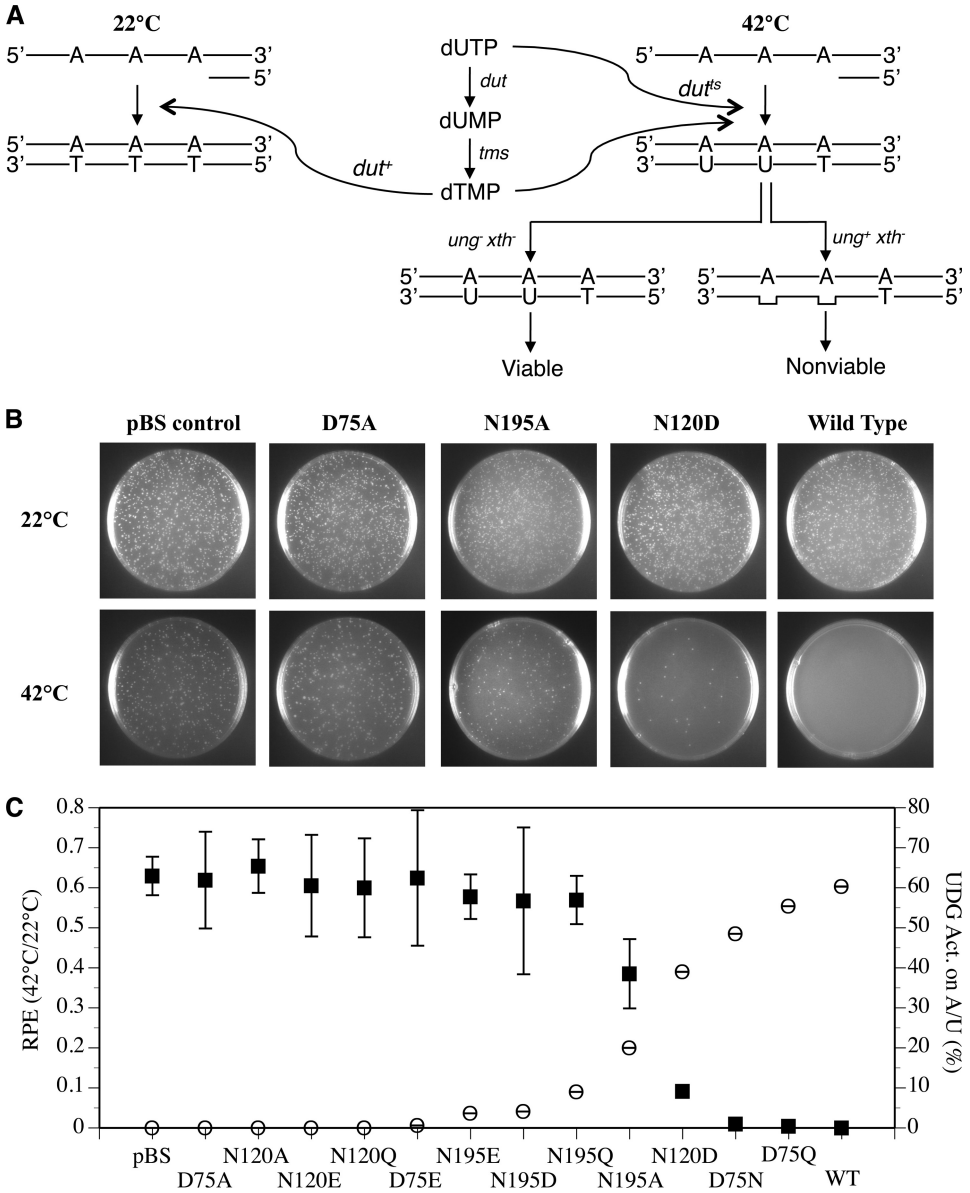


FIGURE 8. **The survival of *E. coli* BW276 containing Tth *udgb* genes at 42 and 22 °C and their correlation with enzymatic activity of Tth UDGb variants on the A/U base pair.** *A*, principle of the *in vivo* assay. *dut*, gene for dUTPase; *tms*, gene for thymidylate synthase; *xth*, AP endonuclease; *ts*, temperature-sensitive. *B*, the survival of *E. coli* BW276 containing pBS-Tth-UDGb variants vectors at 42 and 22 °C. Cells with A_{600} value of 1 were diluted 1×10^4 times, and 150 μl of diluted cells were spread on the LB plates containing ampicillin and thymidine and incubated for 24 h at 42 °C or 72 h at 22 °C. *pBS*, pBluescript. *C*, correlation between A/U enzymatic activities of Tth UDGb variants and the survival of *E. coli* BW276 containing pBS-Tth-UDGb variants vectors at 42 and 22 °C. The cleavage reactions of Tth UDGb variants were performed as described under "Experimental Procedures" at 50 °C for 60 min with 100 nM protein and 10 nM substrate. The data of cleavage activities were plotting according to the right *y* axis. For the convenience of counting colonies grown at 22 °C, *E. coli* cells were diluted 1×10^6 times prior to plating. For counting colonies grown at 42 °C, *E. coli* cells were diluted as follows prior to plating: pBluescript, D75A, N120A, N120E, N120Q, D75E, N195E, N195Q, N195D, and N195A, 1×10^6 times; N120D, 1×10^5 times; D75N, 2×10^3 times; D75Q, 1×10^3 times; wild type, 1×10^2 times. Cell numbers were counted on LB plates containing ampicillin and thymidine after incubation at 42 °C for 24 h or incubation at 22 °C for 72 h. The relative plating efficiencies were calculated by the ratios of the cell numbers between 42 and 22 °C. The data of relative plating efficiencies (*RPE*) are plotted as the left *y* axis. The data are the averages of at least three independent experiments. ■, relative plating efficiencies; ○, UDG activity on the A/U base pair.

BW 276 for our investigation for the purpose of understanding the physiological role of the UDG activity on A/U base pairs.

Since the transition from the RNA world to the DNA world, nature has chosen thymine over uracil as the fourth base in DNA. Because DNA polymerase can incorporate dUTP to DNA, the level of dUTP in a cell has to be kept minimal to prevent appearance of uracil in DNA genomes. The highly active dUTPase encoded by the *dut* gene efficiently hydrolyzes dUTP to dUMP to serve two important functions. First, it minimizes the dUTP in the nucleotide pool to prevent its incorporation into DNA. Second, it produces the dUMP, which is converted to dTMP by thymidylate synthase for DNA synthesis. Once a dUTP is accidentally incorporated into DNA to form an A/U base pair, the UDG with activity on A/U base pairs can remove it through the base excision repair pathway. Thus, dUTPase and the base excision repair pathway serve as preventative and corrective mechanisms to ensure a uracil-free DNA genome in the DNA world at both the nucleotide and DNA levels. The inverse correlation between the UDG activity on A/U base pairs and the viability of the BW 276 cells indicates

that Tth UDGb can remove uracil from A/U base pairs (Fig. 8). The increase of UDG activity on A/U base pairs accompanied by the corresponding decrease in viability suggests that the genetic assay correlates well with the *in vitro* biochemical assay. Therefore, the genetic system described here can be a useful tool to study other uracil DNA glycosylases with activity on A/U base pairs.

Unique Catalytic Mechanism—The catalytic mechanisms are most extensively studied in family 1 UNGs and to a lesser degree in family 2 enzymes. A common theme in the hydrolysis of *N*-glycosidic bonds is that an oxacarbenium ion intermediate is formed in the transition state (38, 39). The catalytic power to accelerate the hydrolysis reaction comes from a combination of activation of the leaving group, stabilization of the oxacarbenium ion, and activation of water as a nucleophile. For family 1 UNG enzymes, structural, biochemical, mutational, and kinetic investigations have identified a His residue (His^{187} in *E. coli* UNG) in motif 2 as a critical residue that forms a short hydrogen bond to O_2 of uracil to promote the departure of the uracil anion leaving group (40–44). In addition, a negatively charged Asp residue (Asp^{64} in *E. coli* UNG) is proposed to act as a general acid to activate a water molecule (40, 43, 45). For family 2 MUG/TDG enzymes, it is proposed that an Asn residue (Asn^{18} in *E. coli* MUG), which is located in an equivalent position as Asp^{64} in *E. coli* UNG, may activate a water molecule to initiate the nucleophilic attack at the glycosidic bond (46, 47). Family 5 UDGb enzymes lack a catalytic Asp/Asn residue in motif 1 as seen in family 1 and family 2 enzymes. However, a water molecule was observed near the conserved Asn^{120} in the Tth UDGb structures complexed with AP/G- or AP/A-containing DNA (19). The conserved His^{190} in motif 2 is also implicated in the removal of uracil by the Pae UDGb (12). In light of the broad deaminated base excision specificity demonstrated in this study (Fig. 1*B*), we set out to identify residues that are important for the excision of both pyrimidine base damage and purine base damage.

In the Tth UDGb crystal structure, Asp^{75} is involved in coordination of a water molecule through its side chain (19). The D75E mutant, which retained the negative charge on the side chain, still maintained UDG, HDG, and XDG activity (Fig. 4*B*). Other substitutions result in substantial reduction of XDG, HDG, and UDG activity (Fig. 4). To understand the role of Asp^{75} in the excision of both pyrimidine and purine base damage, we modeled U, I, and X into the crystal structure and carried out molecular dynamics simulation analysis. In the modeled Tth UDGb-U structure, the side chain of Asp^{75} is in close proximity to the O4 of uracil. A distance averaging 4.9 Å is consistent with a bridging water molecule mediating an interaction between Asp^{75} and the O4 of uracil (Fig. 9*A*). A closer look at the modeled Tth UDGb structures suggests that Asp^{75} is in close proximity to the 5′ phosphate of the uridine, which can raise the pK_a and facilitate protonation of the Asp^{75} side chain. This will allow the bridging water molecule to form a hydrogen bond with the O4, which facilitates the removal of uracil (Fig. 10, *A* and *D*). Mutations at the Asp^{75} position exhibited more profound effects on HDG and XDG activity (Fig. 4). In the modeled UDGb-I and UDGb-X structures, the protonated side chain carboxylate of Asp^{75} can interact with the N7 in the hypoxanthine and xanthine through a bridging water molecule in a similar fashion (Fig. 10, *B–E* and *C–F*). MD analysis indicates that the carboxyl side chain is within a distance capable of forming a water-mediated hydrogen bond with the N7 in the purine bases (Fig. 9, *B* and *C*). On the other hand, the lack of interaction to the N7 moiety in oxanine may be in part responsible for the lack of ODG activity, although the interaction with the N3 in oxanine appears feasible (Fig. 9, *D* and *E*). Activation of a purine base through hydrogen bonding or protonation has been proposed as a catalytic mechanism for the cleavage of glycosidic bonds in purine nucleotides. In acid-catalyzed hydrolysis of purine nucleosides, N7 and N3 protonations promote the departure of the purine base (48). In enzyme-catalyzed reactions, MutY, a DNA repair glycosylase involved in removal of adenine from G/A base pairs, catalyzes the excision of adenine by protonating the N7 position (49, 50). Biochemical and structural studies identify a Glu residue (Glu^{37} in *E. coli* MutY and Glu^{43} in *Bacillus stearothermophilus* MutY) in the active site that can serve as a general acid in promoting N7 protonation (51–55). The close proximity of a water molecule between the Glu^{43} in *B. stearothermophilus* MutY and N7 indicates that the proton for N7 protonation can come from water coordinated by Gu^{43} (56). Within the UDG superfamily, we identified an Ser^{23}-N7 interaction in the *E. coli* family 2 MUG enzyme and an Met^{64} main chain-N7 interaction in the *Geobacter metallireducens* family 3 SMUG1 enzyme that play important roles in the excision of xanthine bases (10, 21). Apparently, N7 interaction or protonation is a common catalytic mechanism for leaving group activation in the hydrolysis of purine deaminated bases.

According to the AP site cocrystal structure, Asn^{120} is involved in coordinating a water molecule that is located on the opposite side of the deoxyribose (19). Elimination of the amide group or lengthening of the amide group by one methylene carbon leads to the loss of XDG activity and a substantial reduction of HDG and UDG activity (Fig. 5). Based on the structural information and biochemical analysis, we speculate that Asn^{120} may perform a functional role similar to Asn^{18} in *E. coli* family 2 MUG, in which the Asn helps activate/position a water molecule for initiating a nucleophilic attack on the glycosidic bond. According to MD analysis, an average distance between OD1 of Asn^{120} and C1′ carbon of deoxyribose is ~5.0 Å (Fig. 9, *F–H*). A structural comparison between Asn^{18} in MUG and Asn^{120} in Tth UDGb is shown in Fig. 11. Although differences in the sequence lengths between the Tth UDGb and the *E. coli* MUG enzyme prevent a perfect alignment, strong structural similarity is noted in the core secondary structural elements including five β-sheets and four α-helices (Fig. 11*A*). The strong structural similarity is also noted in the active sites (Fig. 11*B*). The Asn^{18} in MUG and Asn^{120} in Tth UDGb hydrogen bond with water through their side chains, positioning the water molecule proximal to the anomeric carbon of the bound nucleotide. Thus, different families in the UDG superfamily have adopted the same amino acid residue in different structural locations to perform a similar function.

Among the six families within the UDG superfamily, families 1, 3, 4, and 5 contain a His residue at the beginning of motif 2 (Fig. 3*B*). Mutational studies in family 1 enzymes confirm that

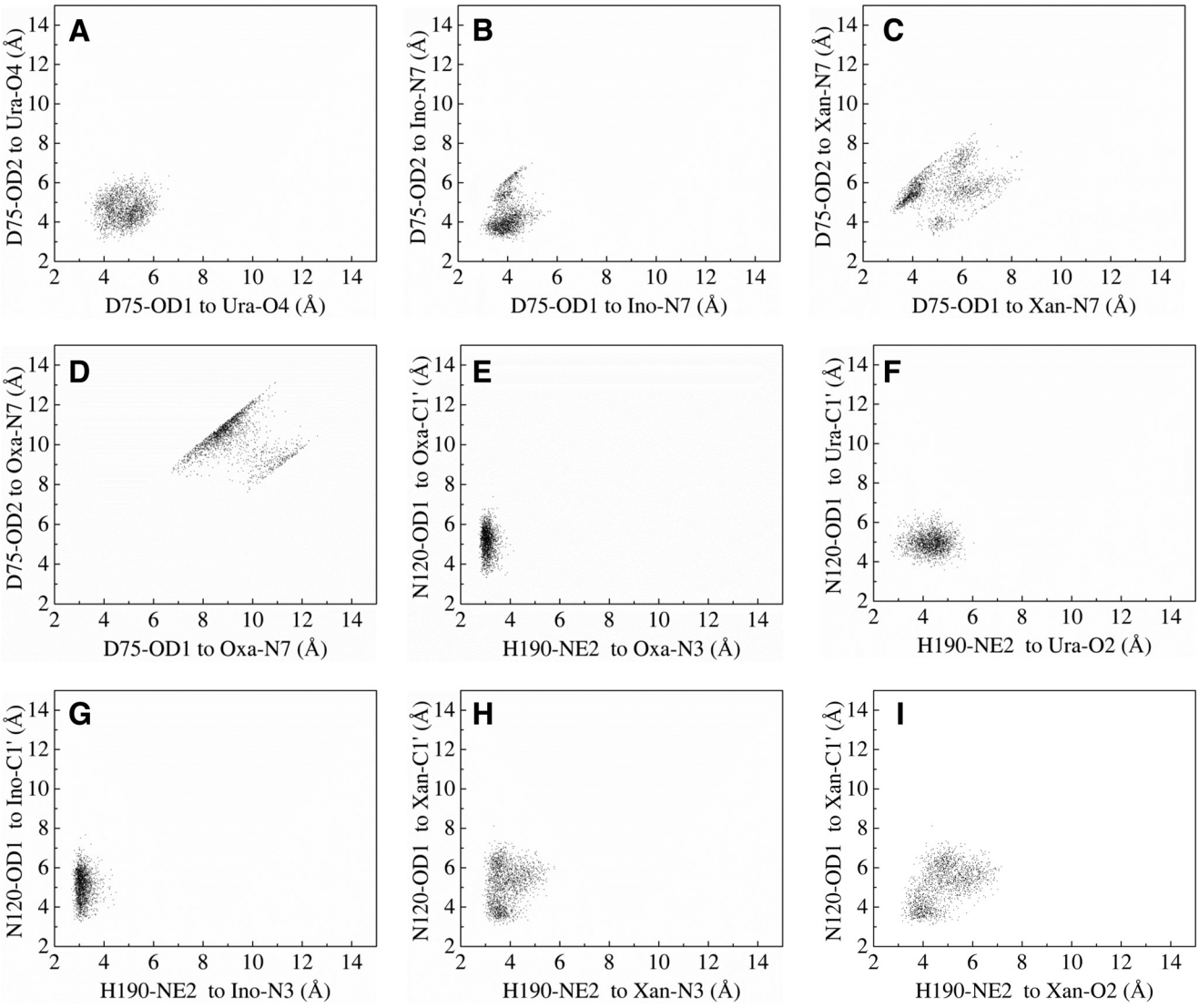


FIGURE 9. **Two-dimensional scatter plots of heavy atom distances in the active site of enzyme-DNA complexes obtained from MD trajectories.** *A*, distances for interactions between Asp^{75} and uridine. *B*, distances for interactions between Asp^{75} and inosine. *C*, distances for interactions between Asp^{75} and xanthosine. *D*, distances for interactions between Asp^{75} and oxanosine. *E*, distances for interactions between Asn^{120}, His^{190}, and oxanosine. *F*, distances for interactions between Asn^{120}, His^{190}, and uridine. *G*, distances for interactions between Asn^{120}, His^{190}, and inosine. *H* and *I*, distances for interactions between Asn^{120}, His^{190}, and xanthosine.

the His residue plays an important role in catalysis and suggest that His^{187} in *E. coli* UNG can form a hydrogen bond with O_2 of uracil (41–43). Spectroscopic analyses indicate that His^{187} in *E. coli* UNG is neutral and forms a short hydrogen bond with the O_2 group (44, 57). Mutational and biochemical investigations in family 3 SMUG1 enzymes also underscore the importance of the His residue in motif 2 in catalysis (10, 58). Data from this work indicate that His^{190} in family 5 Tth UDGb is critical for all deaminated base glycosylase activity (Fig. 6). The elimination of the imidazole side chain renders the enzyme essentially inactive except for a minor HDG activity on G/I base pairs and a barely detectable XDG activity on G/X base pairs (Fig. 6). H190N and H190S can rescue the activity to a very limited extent (Fig. 6). Evidently, UDG activity relies on His^{190} in motif 2 because H190A is the only mutant that results in a loss of UDG activity. D75A and N120A mutants reduce but do not completely eliminate the UDG activity. Similar to family 1 UNG enzymes, MD analysis suggests that His^{190} can form a hydrogen bond with O_2 of uracil (Fig. 9*F*). The predominant role His^{190} plays in UDG activity suggests that the catalysis of glycosidic bond hydrolysis during uracil excision is likely to follow a stepwise $D_N{*}A_N$ mechanism in which the uracil base departs as a uracil anion. Similar to the catalytic mechanism in family 1 UNG enzymes, the hydrogen bonding provided by His^{190} can promote the departure of the leaving group by stabilizing the uracil anion. This appears to be the main catalytic power endowed in family 5 Tth UDGb for uracil excision.

His^{190} is also important for hypoxanthine and xanthine excision (Fig. 6). Similar to nonenzymatic purine nucleoside hydrolysis (48), the hydrogen bonding provided by His^{190} to N3 of hypoxanthine could promote the departure of hypoxanthne. Biochemical studies using adenine analogs also indicate that *E. coli* MutY utilizes the N3 interaction to enhance adenine excision (59). In the modeled structure, the hydrogen in NE2 of His^{190} of Tth UDGb is within 2.1 Å of the N3 of hypoxanthine (Fig. 10, *B* and *E*), consistent with a moderately strong hydrogen

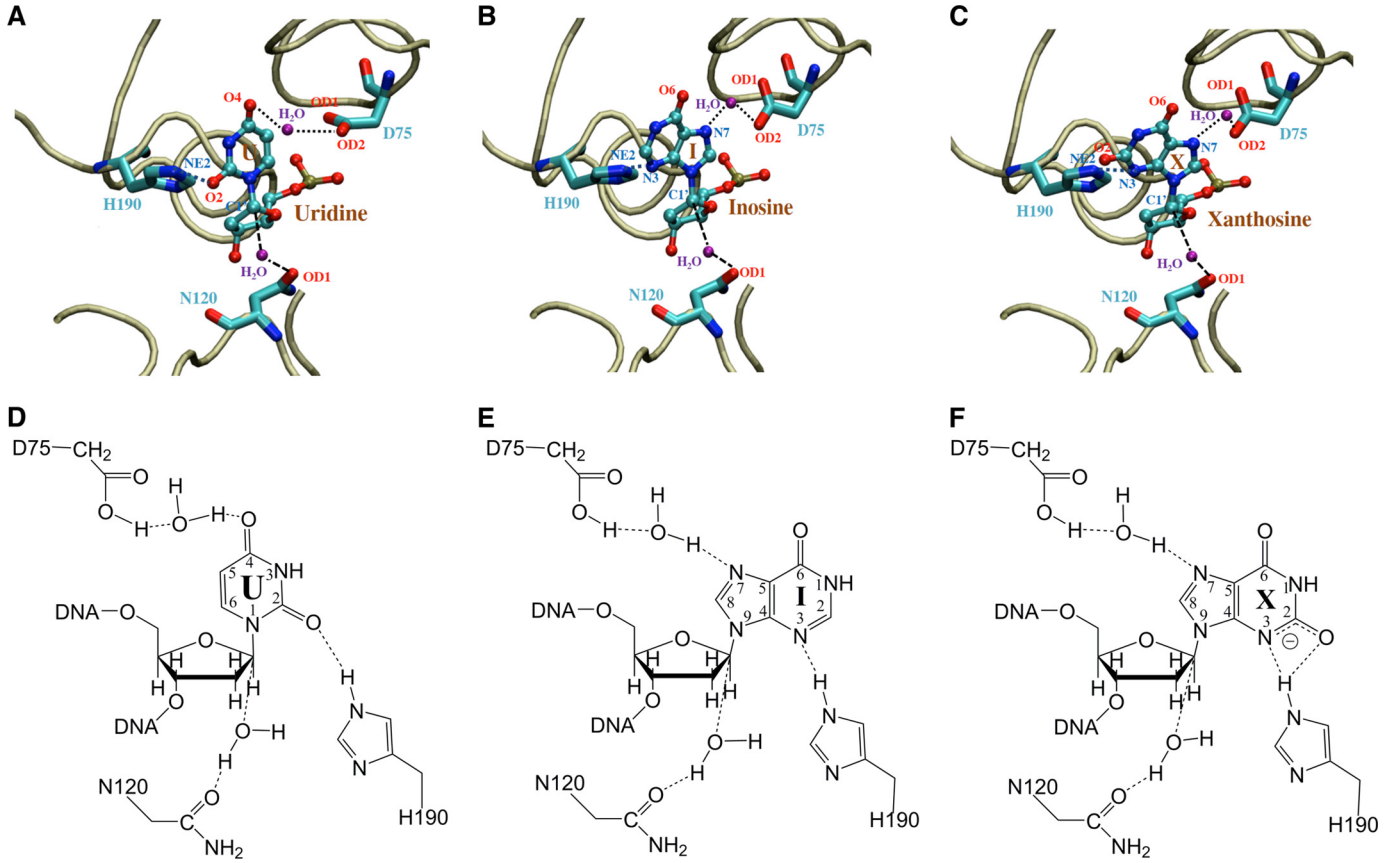


FIGURE 10. **Close-up views of the Tth UDGb-DNA active site interactions in the energy minimized structures.** Tube trace of the protein is colored in *tan*. Uridine (U), inosine (I), and xanthosine (X), as well as amino acids Asp^{75}, Asn^{120}, and His^{190}, are colored by atom type. Water molecules are labeled in *purple*. *Dashed lines* indicate inferred hydrogen bonds or water association. *A*, modeled Tth UDGb-U interactions. *B*, modeled Tth UDGb-I interactions. *C*, modeled Tth UDGb-X interactions. *D*, chemical illustration of Tth UDGb-U interactions. *E*, chemical illustration of Tth UDGb-I interactions. *F*, chemical illustration of Tth UDGb-X interactions.

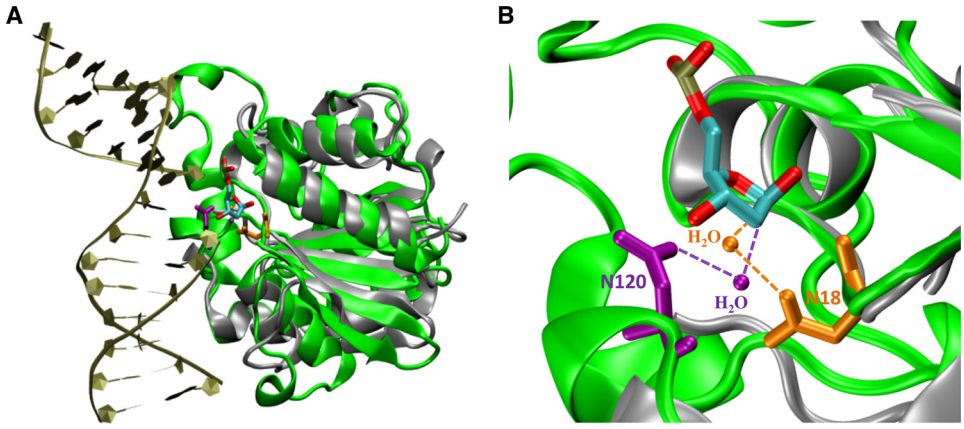


FIGURE 11. **Comparison of interactions between Tth UDGb-N120 with water and *E. coli* MUG-N18 with water.** *A*, superimposition of UDGb-AP structure (Protein Data Bank code 2DEM; *green*) with *E. coli* MUG structure (Protein Data Bank code 1MUG; *silver*). The AP site is colored by atom type. The two structures were superimposed using the program VMD. *B*, close-up view of Tth UDGb-N120-water and *E. coli* MUG-N18-water interactions using the same coloring scheme described in *A*. Asn^{120} and the interacting water in Tth UDGb structure are shown in *purple*. Asn^{18} and the interacting water in the MUG structure are shown in *orange*.

bond. MD analysis also supports the hydrogen bonding model between His^{190} and N3 of hypoxanthine (Fig. 9*G*). For the xanthine-containing base pairs, although normally the N3 position is shown as protonated, xanthosine exists as a monoanion under physiological conditions (60, 61). Thus, as supported by the modeled structure and MD analysis (Figs. 9*H* and 10, *C* and *F*), His^{190} can promote the departure of xanthine through hydrogen bonding with the N3 moiety of xanthosine. His^{190} is 3.5 and 4.2 Å from N3 and O_2 of xanthosine based on the energy minimized structure, raising the possibility of forming a bidentate hydrogen bond (Fig. 10, *C* and *F*). MD analysis also supports the possibility of a bidentate hydrogen bond (Fig. 9, *H* and *I*). The caveat here is that for the excision of deaminated purine bases, the potential water activation by Asn^{120} and the water-

mediated N7 contact by Asp^{75} are also important. Therefore, family 5UDGb enzymes may rely on the concerted action of multiple catalytic residues to excise hypoxanthine and xanthine and are likely to catalyze the *N*-glycosidic bond hydrolysis through a dissociative A_ND_N mechanism.

In summary, this study for the first time comprehensively investigated deaminated base repair by the family 5 Tth UDGb enzyme. The data presented here reveal that family UDGb enzymes are uracil, hypoxanthine, and xanthine DNA glycosylases. The inverse correlation between cellular viability and the UDG activity on A/U base pairs offers a tool to study the role of different UDG enzymes in keeping uracil out of DNA in the DNA world. In addition to taking advantage of greater tendency for base flipping that occurs in damaged DNA base pairs, family 5 UDGb enzymes have adapted multiple mechanisms to achieve deaminated base excision. Whereas the UDG activity heavily relies on the leaving group activation mediated by O_2 of uracil and His^{190} to excise a uracil base, the family UDGb enzymes combine an existing catalytic element (His^{190} in Tth UDGb) with additional nucleophile activation and purine-specific leaving group activation mechanisms to expand their DNA repair capacity. As such, family 5 UDGb offers an example of how an enzyme may strategically acquire catalytic elements to broaden its specificity.

Acknowledgments—We thank Dr. Bernard Weiss at Emory University for providing the BW 276 strain. We thank members of Cao laboratory for assistance and discussions.

REFERENCES

1. Parikh, S. S., Putnam, C. D., and Tainer, J. A. (2000) Lessons learned from structural results on uracil-DNA glycosylase. *Mutat. Res.* **460,** 183–199
2. Krokan, H. E., Drabløs, F., and Slupphaug, G. (2002) Uracil in DNA: occurrence, consequences and repair. *Oncogene* **21,** 8935–8948
3. Cortázar, D., Kunz, C., Saito, Y., Steinacher, R., and Schär, P. (2007) The enigmatic thymine DNA glycosylase. *DNA Repair* **6,** 489–504
4. Hardeland, U., Bentele, M., Lettieri, T., Steinacher, R., Jiricny, J., and Schär, P. (2001) Thymine DNA glycosylase. *Prog. Nucleic Acid Res. Mol. Biol.* **68,** 235–253
5. Pearl, L. H. (2000) Structure and function in the uracil-DNA glycosylase superfamily. *Mutat. Res.* **460,** 165–181
6. Dong, L., Mi, R., Glass, R. A., Barry, J. N., and Cao, W. (2008) Repair of deaminated base damage by *Schizosaccharomyces pombe* thymine DNA glycosylase. *DNA Repair* **7,** 1962–1972
7. He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C. X., Zhang, K., He, C., and Xu, G. L. (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333,** 1303–1307
8. Maiti, A., and Drohat, A. C. (2011) Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J. Biol. Chem.* **286,** 35334–35338
9. Haushalter, K. A., Todd Stukenberg, M. W., Kirschner, M. W., and Verdine, G. L. (1999) Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors. *Curr. Biol.* **9,** 174–185
10. Mi, R., Dong, L., Kaulgud, T., Hackett, K. W., Dominy, B. N., and Cao, W. (2009) Insights from xanthine and uracil DNA glycosylase activities of bacterial and human SMUG1: switching SMUG1 to UDG. *J. Mol. Biol.* **385,** 761–778
11. Sandigursky, M., and Franklin, W. A. (1999) Thermostable uracil-DNA glycosylase from *Thermotoga maritima* a member of a novel class of DNA repair enzymes. *Curr. Biol.* **9,** 531–534
12. Sartori, A. A., Fitz-Gibbon, S., Yang, H., Miller, J. H., and Jiricny, J. (2002) A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. *EMBO J.* **21,** 3182–3191
13. Lee, H. W., Dominy, B. N., and Cao, W. (2011) New family of deamination repair enzymes in uracil-DNA glycosylase superfamily. *J. Biol. Chem.* **286,** 31282–31287
14. Wanner, R. M., Castor, D., Güthlein, C., Böttger, E. C., Springer, B., and Jiricny, J. (2009) The uracil DNA glycosylase UdgB of *Mycobacterium smegmatis* protects the organism from the mutagenic effects of cytosine and adenine deamination. *J. Bacteriol.* **191,** 6312–6319
15. Srinath, T., Bharti, S. K., and Varshney, U. (2007) Substrate specificities and functional characterization of a thermo-tolerant uracil DNA glycosylase (UdgB) from *Mycobacterium tuberculosis*. *DNA Repair* **6,** 1517–1528
16. Malshetty, V. S., Jain, R., Srinath, T., Kurthkoti, K., and Varshney, U. (2010) Synergistic effects of UdgB and Ung in mutation prevention and protection against commonly encountered DNA damaging agents in *Mycobacterium smegmatis*. *Microbiology* **156,** 940–949
17. Starkuviene, V., and Fritz, H. J. (2002) A novel type of uracil-DNA glycosylase mediating repair of hydrolytic DNA damage in the extremely thermophilic eubacterium *Thermus thermophilus*. *Nucleic Acids Res.* **30,** 2097–2102
18. Sakai, T., Tokishita, S., Mochizuki, K., Motomiya, A., Yamagata, H., and Ohta, T. (2008) Mutagenesis of uracil-DNA glycosylase deficient mutants of the extremely thermophilic eubacterium *Thermus thermophilus*. *DNA Repair* **7,** 663–669
19. Kosaka, H., Hoseki, J., Nakagawa, N., Kuramitsu, S., and Masui, R. (2007) Crystal structure of family 5 uracil-DNA glycosylase bound to DNA. *J. Mol. Biol.* **373,** 839–850
20. Lu, X. J., and Olson, W. K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31,** 5108–5121
21. Lee, H. W., Brice, A. R., Wright, C. B., Dominy, B. N., and Cao, W. (2010) Identification of *Escherichia coli* mismatch-specific uracil DNA glycosylase as a robust xanthine DNA glycosylase. *J. Biol. Chem.* **285,** 41483–41490
22. Priyakumar, U. D., and MacKerell, A. D., Jr. (2006) Computational approaches for investigating base flipping in oligonucleotides. *Chem. Rev.* **106,** 489–505
23. Taylor, A. F., and Weiss, B. (1982) Role of exonuclease III in the base excision repair of uracil-containing DNA. *J. Bacteriol.* **151,** 351–357
24. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18,** 2714–2723
25. Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30,** 1545–1614
26. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–3616
27. MacKerell, A. D., and Banavali, N. K. (2000) All-atom empirical force field for nucleic acids: II. application to molecular dynamics simulations of DNA and RNA in solution. *J. Comput. Chem.* **21,** 105–120
28. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089–10092
29. Adelman, S. A., and Doll, J. D. (1976) Generalized Langevin equation approach for atom/solid-surface scattering: General formulation for classical scattering off harmonic solids. *J. Chem. Phys.* **64,** 2375–2388
30. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E.,

Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802
31. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38
32. Warner, H. R., Duncan, B. K., Garrett, C., and Neuhard, J. (1981) Synthesis and metabolism of uracil-containing deoxyribonucleic acid in *Escherichia coli*. *J. Bacteriol.* **145,** 687–695
33. Mi, R., Alford-Zappala, M., Kow, Y. W., Cunningham, R. P., and Cao, W. (2012) Human endonuclease V as a repair enzyme for DNA deamination. *Mutat. Res.* **735,** 12–18
34. Kumar, V. D., Harrison, R. W., Andrews, L. C., and Weber, I. T. (1992) Crystal structure at 1.5-A resolution of d(CGCICICG), an octanucleotide containing inosine, and its comparison with d(CGCG) and d(CGCGCG) structures. *Biochemistry* **31,** 1541–1550
35. Case-Green, S. C., and Southern, E. M. (1994) Studies on the base pairing properties of deoxyinosine by solid phase hybridisation to oligonucleotides. *Nucleic Acids Res.* **22,** 131–136
36. Martin, F. H., Castro, M. M., Aboul-ela, F., and Tinoco, I., Jr. (1985) Base pairing involving deoxyinosine: implications for probe design. *Nucleic Acids Res.* **13,** 8927–8938
37. Watkins, N. E., Jr., and SantaLucia, J., Jr. (2005) Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes. *Nucleic Acids Res.* **33,** 6258–6267
38. Berti, P. J., and McCann, J. A. (2006) Toward a detailed understanding of base excision repair enzymes: transition state and mechanistic analyses of *N*-glycoside hydrolysis and *N*-glycoside transfer. *Chem. Rev.* **106,** 506–555
39. Stivers, J. T., and Jiang, Y. L. (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem. Rev.* **103,** 2729–2759
40. Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L. (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature* **373,** 487–493
41. 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
42. Shroyer, M. J., Bennett, S. E., Putnam, C. D., Tainer, J. A., and Mosbaugh, D. W. (1999) Mutation of an active site residue in *Escherichia coli* uracil-DNA glycosylase: effect on DNA binding, uracil inhibition and catalysis. *Biochemistry* **38,** 4834–4845
43. Drohat, A. C., Jagadeesh, J., Ferguson, E., and Stivers, J. T. (1999) Role of electrophilic and general base catalysis in the mechanism of *Escherichia coli* uracil DNA glycosylase. *Biochemistry* **38,** 11866–11875
44. Drohat, A. C., and Stivers, J. T. (2000) *Escherichia coli* uracil DNA glycosylase: NMR characterization of the short hydrogen bond from His187 to uracil O2. *Biochemistry* **39,** 11865–11875
45. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature* **384,** 87–92
46. 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
47. Barrett, T. E., Schärer, O. D., Savva, R., Brown, T., Jiricny, J., Verdine, G. L., and Pearl, L. H. (1999) Crystal structure of a thwarted mismatch glycosylase DNA repair complex. *EMBO J.* **18,** 6599–6609
48. Zoltewicz, J. A., Clark, D. F., Sharpless, T. W., and Grahe, G. (1970) Kinetics and mechanism of the acid-catalyzed hydrolysis of some purine nucleosides. *J. Am. Chem. Soc.* **92,** 1741–1749
49. McCann, J. A., and Berti, P. J. (2008) Transition-state analysis of the DNA repair enzyme MutY. *J. Am. Chem. Soc.* **130,** 5789–5797
50. Michelson, A. Z., Rozenberg, A., Tian, Y., Sun, X., Davis, J., Francis, A. W., O'Shea, V. L., Halasyam, M., Manlove, A. H., David, S. S., and Lee, J. K. (2012) Gas-phase studies of substrates for the DNA mismatch repair enzyme MutY. *J. Am. Chem. Soc.* **134,** 19839–19850
51. Wright, P. M., Yu, J., Cillo, J., and Lu, A. L. (1999) The active site of the *Escherichia coli* MutY DNA adenine glycosylase. *J. Biol. Chem.* **274,** 29011–29018
52. Brinkmeyer, M. K., Pope, M. A., and David, S. S. (2012) Catalytic contributions of key residues in the adenine glycosylase MutY revealed by pH-dependent kinetics and cellular repair assays. *Chem. Biol.* **19,** 276–286
53. Fromme, J. C., Banerjee, A., Huang, S. J., and Verdine, G. L. (2004) Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase. *Nature* **427,** 652–656
54. Lee, S., and Verdine, G. L. (2009) Atomic substitution reveals the structural basis for substrate adenine recognition and removal by adenine DNA glycosylase. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 18497–18502
55. Guan, Y., Manuel, R. C., Arvai, A. S., Parikh, S. S., Mol, C. D., Miller, J. H., Lloyd, 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
56. Brunk, E., Arey, J. S., and Rothlisberger, U. (2012) Role of environment for catalysis of the DNA repair enzyme MutY. *J. Am. Chem. Soc.* **134,** 8608–8616
57. Drohat, A. C., Xiao, G., Tordova, M., Jagadeesh, J., Pankiewicz, K. W., Watanabe, K. A., Gilliland, G. L., and Stivers, J. T. (1999) Heteronuclear NMR and crystallographic studies of wild-type and H187Q *Escherichia coli* uracil DNA glycosylase: electrophilic catalysis of uracil expulsion by a neutral histidine 187. *Biochemistry* **38,** 11876–11886
58. Matsubara, M., Tanaka, T., Terato, H., Ohmae, E., Izumi, S., Katayanagi, K., and Ide, H. (2004) Mutational analysis of the damage-recognition and catalytic mechanism of human SMUG1 DNA glycosylase. *Nucleic Acids Res.* **32,** 5291–5302
59. Francis, A. W., Helquist, S. A., Kool, E. T., and David, S. S. (2003) Probing the requirements for recognition and catalysis in Fpg and MutY with nonpolar adenine isosteres. *J. Am. Chem. Soc.* **125,** 16235–16242
60. Kulikowska, E., Kierdaszuk, B., and Shugar, D. (2004) Xanthine, xanthosine and its nucleotides: solution structures of neutral and ionic forms, and relevance to substrate properties in various enzyme systems and metabolic pathways. *Acta Biochim. Pol.* **51,** 493–531
61. Poznanski, J., Kierdaszuk, B., and Shugar, D. (2003) Structural properties of the neutral and monoanionic forms of xanthosine, highly relevant to their substrate properties with various enzyme systems. *Nucleosides Nucleotides Nucleic Acids* **22,** 249–263