J|A|C|S
ARTICLES
Published on Web 08/31/2006

Specificity of Human Thymine DNA Glycosylase Depends on *N*-Glycosidic Bond Stability

Matthew T. Bennett,[‡] M. T. Rodgers,[†] Alexander S. Hebert,[§] Lindsay E. Ruslander,[§] Leslie Eisele,[§] and Alexander C. Drohat*[,‡]

Contribution from the Department of Biochemistry and Molecular Biology and Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore Maryland 21201, Department of Chemistry, Wayne State University, Detroit, Michigan 48202, and the Wadsworth Center, New York State Department of Health, Albany, New York 12201

Received May 18, 2006; E-mail: adroh001@umaryland.edu

Abstract: Initiating the DNA base excision repair pathway, DNA glycosylases find and hydrolytically excise damaged bases from DNA. While some DNA glycosylases exhibit narrow specificity, others remove multiple forms of damage. Human thymine DNA glycosylase (hTDG) cleaves thymine from mutagenic G·T mispairs, recognizes many additional lesions, and has a strong preference for nucleobases paired with guanine rather than adenine. Yet, hTDG avoids cytosine, despite the million-fold excess of normal G·C pairs over G·T mispairs. The mechanism of this remarkable and essential specificity has remained obscure. Here, we examine the possibility that hTDG specificity depends on the stability of the scissile base−sugar bond by determining the maximal activity (k_{max}) against a series of nucleobases with varying leaving-group ability. We find that hTDG removes 5-fluorouracil 78-fold faster than uracil, and 5-chlorouracil, 572-fold faster than thymine, differences that can be attributed predominantly to leaving-group ability. Moreover, hTDG readily excises cytosine analogues with improved leaving ability, including 5-fluorocytosine, 5-bromocytosine, and 5-hydroxycytosine, indicating that cytosine has access to the active site. A plot of log(k_{max}) versus leaving-group pK_a reveals a Brønsted-type linear free energy relationship with a large negative slope of $\beta_{lg} = -1.6 \pm 0.2$, consistent with a highly dissociative reaction mechanism. Further, we find that the hydrophobic active site of hTDG contributes to its specificity by enhancing the inherent differences in substrate reactivity. Thus, hTDG specificity depends on *N*-glycosidic bond stability, and the discrimination against cytosine is due largely to its very poor leaving ability rather than its exclusion from the active site.

Introduction

The chemically reactive bases in DNA are continuously modified by agents of cellular metabolism and from exogenous sources, producing lesions that threaten genetic integrity and play a role in aging and diseases including cancer.[1] Counteracting this inevitable damage is the base excision repair (BER) pathway, initiated by a damage-specific DNA glycosylase. Using a base-flipping mechanism, these enzymes recognize damaged bases and remove them by catalyzing the hydrolysis of the *N*-glycosidic bond connecting the nucleobase to the sugar, producing an abasic (AP) site in the DNA. While some DNA glycosylases possess significant catalytic power, they are perhaps more impressive in their specificity *for* certain damaged bases and *against* normal bases, in keeping with the threat to genomic integrity posed by the aberrant removal of normal bases from DNA. However, the challenge of distinguishing damaged from normal is daunting, given that many damaged bases differ modestly from the normal counterpart and that the lesions are hidden within the vast excess of normal DNA.

To obtain specificity, some DNA glycosylases employ a highly discriminating active site that accommodates only certain lesions. For example, uracil DNA glycosylase (UDG) uses steric exclusion and specific electrostatic interactions to recognize uracil and prevent other bases from docking in its active site.[2] Other enzymes with a highly selective active site include human 8-oxoguanine (OG) DNA glycosylase (hOGG1), which removes OG from OG·C pairs, but not G from G·C pairs,[3] and *Escherichia coli* 3-methyladenine DNA glycosylase I (eTAG), which binds 3-methyladenine but not adenine.[4,5] In contrast, many DNA glycosylases allow a broad range of damaged and even some normal bases access to the active site, as exemplified by 3-methyladenine DNA glycosylase II from *E. coli* (AlkA) and mammalian alkyladenine DNA glycosylase (AAG).[6−8] For these more promiscuous enzymes, certain properties of the

[‡] University of Maryland School of Medicine.
[†] Wayne State University.
[§] New York State Department of Health.

(1) Lindahl, T. *Nature* **1993**, *362*, 709−715.
(2) Parikh, S. S.; Walcher, G.; Jones, G. D.; Slupphaug, G.; Krokan, H. E.; Blackburn, G. M.; Tainer, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5083−5088.
(3) Banerjee, A.; Yang, W.; Karplus, M.; Verdine, G. L. *Nature* **2005**, *434*, 612−618.
(4) Drohat, A. C.; Kwon, K.; Krosky, D. J.; Stivers, J. T. *Nat. Struct. Biol.* **2002**, *9*, 659−664.
(5) Cao, C.; Kwon, K.; Jiang, Y. L.; Drohat, A. C.; Stivers, J. T. *J. Biol. Chem.* **2003**, *278*, 48012−48020.
(6) Berdal, K. G.; Johansen, R. F.; Seeberg, E. *Embo. J.* **1998**, *17*, 363−367.
(7) O'Brien, P. J.; Ellenberger, T. *J. Biol. Chem.* **2004**, *279*, 26876−26884.
(8) O'Brien, P. J.; Ellenberger, T. *J. Biol. Chem.* **2004**, 9750-9757.

10.1021/ja0634829 CCC: $33.50

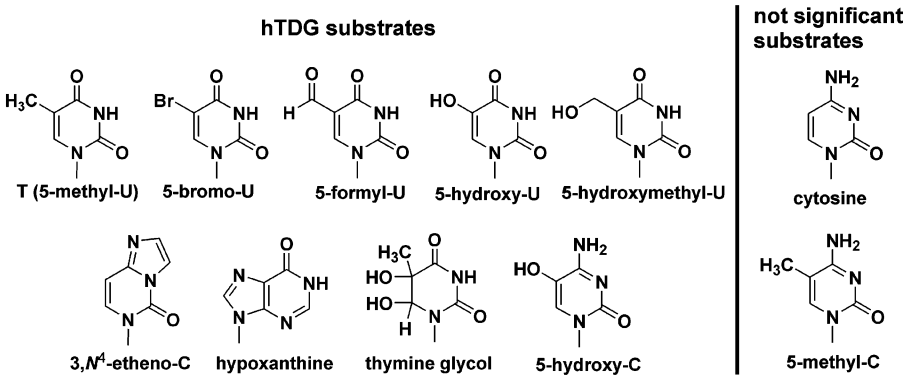

Figure 1. Known hTDG substrates are shown.[13−17] Cytosine and 5-methyl-C are not significant substrates.[13,18] We report several new hTDG substrates here, including 5-chloro-U, 5-iodo-U, 5-fluoro-C, and 5-bromo-C.

damaged base contribute to specificity. For example, alkylated bases form a more labile *N*-glycosidic bond and are more easily removed compared to the normal counterpart.[6,7] Additionally, some damaged bases have a diminished capacity for H-bonding and/or stacking interactions. As a result, such bases are more likely to flip out of the DNA duplex, which can lead to tighter binding to a DNA glycosylase and faster base excision.[7,8]

Another enzyme with broad specificity is human thymine DNA glycosylase (hTDG), which excises thymine from G·T mispairs, and removes many additional lesions, with a strong preference for bases that are paired with guanine and located at CpG sites (i.e., 5′-CpG/5′-XpG, where X is the target base).[9] In vertebrates, cytosine methylation (m^5C) at CpG serves as a mark for transcriptional silencing and is central to many important cellular processes. CpG sites are also associated with disease; improper CpG methylation plays a role in carcinogenesis,[10] and the deamination of m^5C to T produces G·T mispairs, contributing to the high mutational frequencies observed at CpG sites in humans.[11,12] Although hTDG is highly specific for removing bases that are paired with guanine as opposed to adenine, it avoids the removal of cytosine despite the million-fold excess of G·C pairs over G·T mispairs. The challenge of attaining such remarkable specificity would appear to be compounded by the broad range of bases that are excised by hTDG (Figure 1).

One proposal holds that specificity occurs at the base-flipping step, whereby the higher stability of G·C pairs relative to that of G·T mispairs would preclude the enzymatic flipping of cytosine into the active site.[19] Another possibility is that hTDG uses selective steric and electrostatic interactions to recognize multiple substrates while excluding cytosine. However, examination of the various bases removed by hTDG (Figure 1) indicates no obvious handle that would allow for substrate recognition *and* cytosine rejection. Moreover, the related mismatch-specific uracil DNA glycosylase from *E. coli* (eMUG) has a permissive and nonspecific active site which provides no hydrogen bonds that account for its specificity against cytosine,[20] and a recent structure of the hTDG catalytic core, conjugated to the ubiquitin-like modifier SUMO−1, reveals strong similarity to eMUG.[21] Another possibility is that specificity is obtained at the chemical step of the enzymatic reaction, cleavage of the *N*-glycosidic bond connecting the nucleobase to the sugar. Thus, hTDG could potentially flip a broad range of bases, including cytosine, into its active site, with specificity determined predominantly by the heterolytic stability of the scissile C−N bond.

Here, we examine the dependence of hTDG specificity on *N*-glycosidic bond stability. Our strategy stems from previous studies which found that nonenzymatic pyrimidine hydrolyses proceed through a highly dissociative mechanism, where the reaction rate (hence, glycosidic bond stability) depends on the leaving ability of the nucleobase.[22,23] Importantly, these studies found that 2′-deoxycytidine (dC) is significantly more stable than 2′-deoxyuridine (dU) and 2′-deoxythymidine (dT) at pH 7.4,[24] raising the possibility that hTDG specificity depends on substrate reactivity. To test this idea directly, we determined the maximal activity (k_{max}) of hTDG against a series of 5-substituted uracil and cytosine bases which varied in leaving-group ability. We show quantitatively that hTDG specificity is determined largely by *N*-glycosidic bond stability, and we find that a hydrophobic active site enhances the inherent differences in substrate reactivity. Thus, specificity against cytosine can largely be explained by its poor leaving ability rather than its exclusion from the hTDG active site.

Materials and Methods

Duplex DNA Substrates. The 19 bp duplex DNA substrates were made by mixing the target strand (5′-CACTGCTCAxGTACAGAGC, x = substrate base) and 5% excess of the complementary strand (5′-GCTCTGTACGTGAGCAGTG) in 0.01 M Tris, pH 8.0; 0.1 M NaCl; and 0.1 mM EDTA, heating to 80 °C, and slowly (>3 h) cooling to 22

(9) Waters, T. R.; Swann, P. F. *J. Biol. Chem.* **1998**, *273*, 20007−20014.
(10) Jones, P. A.; Baylin, S. B. *Nat. Rev. Genet.* **2002**, *3*, 415−428.
(11) Coulondre, C.; Miller, J. H.; Farabaugh, P. J.; Gilbert, W. *Nature* **1978**, *274*, 775−780.
(12) Rideout, W. M., III; Coetzee, G. A.; Olumi, A. F.; Jones, P. A. *Science* **1990**, *249*, 1288−1290.
(13) Hardeland, U.; Bentele, M.; Jiricny, J.; Schar, P. *Nucleic Acids Res.* **2003**, *31*, 2261−2271.
(14) Neddermann, P.; Jiricny, J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1642−1646.
(15) Liu, P.; Burdzy, A.; Sowers, L. C. *DNA Repair (Amst)*. **2003**, *2*, 199−210.
(16) Saparbaev, M.; Laval, J. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8508−8513.
(17) Yoon, J. H.; Iwai, S.; O'Connor, T. R.; Pfeifer, G. P. *Nucleic Acids Res.* **2003**, *31*, 5399−5404.
(18) Sibghat, U.; Gallinari, P.; Xu, Y. Z.; Goodman, M. F.; Bloom, L. B.; Jiricny, J.; Day, R. S., III. *Biochemistry* **1996**, *35*, 12926−12932.
(19) Barrett, T. E.; Savva, R.; Panayotou, G.; Barlow, T.; Brown, T.; Jiricny, J.; Pearl, L. H. *Cell* **1998**, *92*, 117−129.
(20) Barrett, T. E.; Scharer, O. D.; Savva, R.; Brown, T.; Jiricny, J.; Verdine, G. L.; Pearl, L. H. *Embo. J.* **1999**, *18*, 6599−6609.
(21) Baba, D.; Maita, N.; Jee, J.-G.; Uchimura, Y.; Saitoh, H.; Sugasawa, K.; Hanaoka, F.; Tochio, H.; Hiroaki, H.; Shirakawa, M. *Nature* **2005**, *435*, 979−982.
(22) Shapiro, R.; Kang, S. *Biochemistry* **1969**, *8*, 1806−1810.
(23) Shapiro, R.; Danzig, M. *Biochemistry* **1972**, *11*, 23−29.
(24) Berti, P. J.; McCann, J. A. *Chem. Rev.* **2006**, *106*, 506−555.

°C. The oligonucleotides were synthesized at the W.M. Keck Biotechnology Resource Laboratory of Yale University and at the Biopolymer Genomics Core Facility, University of Maryland Baltimore. Nucleoside phosphoramidites with modified bases were purchased from Glen Research (Sterling, VA), and were incorporated using standard phosphoramidite chemistry and the deprotection method recommended by the manufacturer. The phosphoramidite for 5-chloro-dU was generously provided by Dr. Yu Lin Jiang (East Tennessee State University). The oligonucleotides were purified by anion-exchange HPLC using a Zorbax Oligo column (Agilent Technologies), desalted by gel filtration, and stored at −20 °C. The purity was verified by analytical anion-exchange HPLC under denaturing conditions (pH 12) using a DNAPac PA200 column (Dionex Corp.). The molecular weight was verified by ESI mass spectrometry (Integrated DNA Technologies); the observed and calculated masses differed by <0.03%. The possibility that the 5-halogenated cytosine bases had spontaneously deaminated to the 5-halouracil base during purification or storage was discounted using anion-exchange HPLC at pH 12, which allowed the oligonucleotides (i.e., FC versus FU) to be fully resolved (not shown). The oligonucleotides were quantified by absorbance at 260 nm using pairwise extinction coefficients.[25]

Expression and Purification of hTDG. A pET-28-based expression plasmid for hTDG[26] was transformed into BL21(DE3) "Rosetta" cells (Novagen). The cells were grown in Luria Broth (LB) at 37 °C to an $OD_{600} = 0.8$, the temperature was reduced to 15 °C, and expression of hTDG was induced with 0.25 mM IPTG (isopropyl β-D-thiogalactoside) and continued for about 15 h. The cells were harvested and stored at −80 °C. The pellet was thawed and suspended in lysis buffer (0.05 M sodium phosphate, pH 8.0; 0.3 M NaCl; 0.02 M imidazole; 0.01 M β-mercaptoethanol), incubated with 1 mg/mL lysozyme and DNase (Novagen) for 30 min on ice, and then sonicated. The lysate was cleared by centrifugation, incubated with 4 mL of Ni-NTA resin (Qiagen) for 1 h at 4 °C, and the mixture was added to an empty column with the flow-through collected. The bound hTDG was washed using 30 mL of lysis buffer with 20 mM imidazole and 1 M NaCl, washed again using 30 mL of lysis buffer with 20 mM imidazole and 0.3 M NaCl, and eluted using 3 × 5 mL volumes of lysis buffer with 0.15 M imidazole. Purification was continued using a Q sepharose HP column (Amersham) with buffers IE-A (25 mM Tris pH 7.5, 75 mM NaCl, 1 mM DTT, 0.2 mM EDTA, 1% glycerol) and IE-B (IE-A with 1 M NaCl) and a gradient of 0−50% IE-B over 60 min at 2.5 mL/min. hTDG was purified further using a SP sepharose HP column (Amersham) with the same buffers and a gradient of 0−100% IE-B over 60 min at 2.5 mL/min. The purity of hTDG was >99% as judged by a Coomassie stained gel, and its molecular weight was confirmed by mass spectrometry. The yield of hTDG was typically >5 mg per L of culture. Purified hTDG was dialyzed overnight in 1 L storage buffer (20 mM HEPES pH 7.5, 0.1 M NaCl, 1 mM DTT, 0.5 mM EDTA, and 1% glycerol), concentrated to about 0.1 mM, flash frozen, and stored at −80 °C. The concentration of hTDG was determined by absorbance using $\epsilon^{280} = 31.5\ \mathrm{mM^{-1}\ cm^{-1}}$.[27]

Single Turnover Kinetics. Because hTDG is strongly inhibited by its abasic DNA product,[9] we used single turnover kinetics and saturating enzyme conditions. To ensure that maximal rate constants were obtained (i.e., $k_{obs} = k_{max}$), the experiments were collected with a large excess of enzyme, at least 100-fold greater than the reported $K_D = 41$ nM for hTDG binding to DNA containing a G·T mispair.[28] Saturating enzyme conditions were verified by collecting the experiments with at least two hTDG concentrations, typically 5 and 10 μM, which yielded rate constants that were equivalent within experimental uncertainty (±10%). The DNA substrate concentrations were typically 500 nM, although equivalent k_{max} values were obtained for 250 and 1000 nM substrate concentrations (not shown). The reactions were performed either manually or using a three-syringe rapid chemical quenched-flow instrument (RQF-3, Kintek Corp.). For some substrates, both methods were used, yielding data that were in very good agreement. Samples taken at specific time points were quenched with 50% (v:v) quench solution (0.3 M NaOH, 0.03 M EDTA), incubated at 85 °C for 15 min to induce alkaline cleavage of the DNA backbone at abasic sites, and then analyzed by HPLC to determine the reaction progress (see below). For the DNA substrates containing 5-hydroxyuracil (hoU), 5-hydroxycytosine (hoC), and 5-hydroxymethyluracil (hmU), the reactions were quenched using 0.1 M piperidine with 0.03 M EDTA (final concentration) and heated at 85 °C for 15 min, because a small amount of DNA backbone cleavage was observed (in the absence of hTDG) at the site of the modified base when the NaOH quench was used. The single turnover reactions proceeded to full completion for all substrates (except G·C19). Rate constants were determined by fitting the data to a single-exponential equation using nonlinear regression with Grafit 5.[29] The reactions were conducted at 22 °C in HEMN.1 buffer (20 mM HEPES, pH 7.50; 0.2 mM EDTA; 2.5 mM $MgCl_2$; 0.1 M NaCl) with 0.1 mg/mL bovine serum albumin.

HPLC Assay for Monitoring the hTDG Reaction. We have developed a HPLC assay for monitoring the kinetics of the hTDG-catalyzed reaction. Samples taken from the kinetic reactions at various time points contained a mixture of four oligonucleotides: the full-length (19mer) target strand, its complement, and the two smaller product strands resulting from alkaline cleavage of the abasic strand (produced by hTDG activity). As shown in Figure 2, these four strands are fully resolved by anion-exchange HPLC using denaturing (pH 12.0) conditions with a DNAPac PA200 column (Dionex Corp.). The alkaline conditions serve to suppress hybridization and to increase the resolution, because thymine and guanine are negatively charged at pH 12. The elution buffer was 0.02 M sodium phosphate, pH 12.0, containing either (A) 0.03 M $NaClO_4$ or (B) 0.50 M $NaClO_4$. The protocol employed: 8% B for 1 min, linear gradient of 8−25% B over 14 min, 100% B for 2 min, 8% B for 6 min, and a flow rate of 1.2 mL/min throughout. The oligonucleotides were detected by absorbance (260 nm). Assignment of the elution peaks was confirmed by running individually the oligonucleotides corresponding to the target and product strands. The fraction product was determined from the integrated peak areas for the target strand (A^S) and the two product strands (A^{P1} and A^{P2}) using eq 1:

$$F = \text{fraction product} = \frac{(A^{P1} + A^{P2})}{(A^S + A^{P1} + A^{P2})} \qquad (1)$$

Using multiple injections of an identical sample, we found that the fraction product was reproducible to within 1% (not shown). This HPLC assay is amenable to automation, as we routinely analyze dozens of samples overnight using an auto-sampling device. We anticipate that this approach will be applicable to other DNA glycosylases and other enzymes that act upon DNA. Indeed, we have used it to monitor the kinetics of human AP endonuclease (not shown).

Theoretical Calculations. To obtain model structures, vibrational frequencies, and energetics for the neutral and N1 deprotonated nucleobases, theoretical calculations were performed using Gaussian 03.[30] For the systems excluding 5-iodouracil, geometry optimizations and vibrational analyses were performed at the MP2(full)/6-31G* level. When used to calculate thermal energy corrections, the MP2(full)/6-31G* vibrational frequencies are scaled by a factor of 0.9646. Single-point energy calculations and electrostatic potential surfaces were performed at the MP2(full)/6-311+G(2d,2p) level using the MP2(full)/6-31G* geometries. Parameters for iodine are not available in the above

(25) Fasman, G. *CRC Handbook of Biochemistry and Molecular Biology*, 3rd ed.; CRC Press: Boca Ratton, FL, 1975.
(26) Hardeland, U.; Steinacher, R.; Jiricny, J.; Schar, P. *Embo. J.* **2002**, *21*, 1456−1464.
(27) Gill, S. C.; von Hippel, P. H. *Anal. Biochem.* **1989**, *182*, 319−326.
(28) Abu, M.; Waters, T. R. *J. Biol. Chem.* **2003**, *278*, 8739−8744.
(29) Leatherbarrow, R. J.; Erithacus Software Ltd.: Staines, U.K., 1998.
(30) Frisch, M. J. et al.; Gaussian, Inc.: Pittsburgh, PA, 2003.

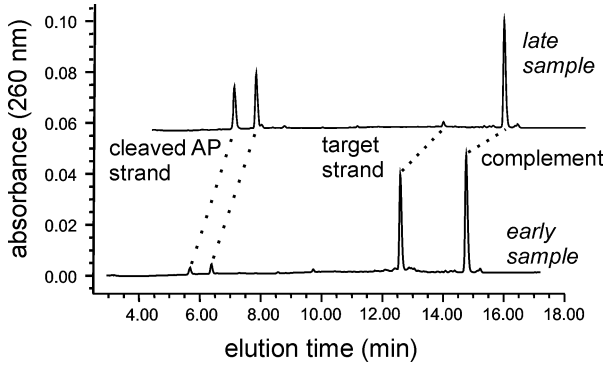

Figure 2. HPLC assay for monitoring DNA glycosylase reactions. Shown are chromatograms for samples taken from a single turnover reaction at an "early" time point where there was little product formation, and a "late" time point where the reaction was nearly complete. The ion-exchange HPLC under denaturing (pH 12) conditions gives excellent resolution of the target strand, its complement, and the two shorter product strands resulting from alkaline-induced cleavage of the nascent abasic strand produced by hTDG activity.

basis sets; therefore, calculations for the 5-iodouracil systems were performed using the modified LANL2DZ[32,33] effective core potentials (ECPs) and valence basis sets for I, while the standard basis sets described above were used for all other atoms. To obtain accurate deprotonation enthalpies and free energies, zero-point energy (ZPE) and basis set superposition error (BSSE) corrections[34,35] were included in the determination of these values.

To determine the relative size of uracil and cytosine and the influence of the 5-substituents upon the size of these nucleobases, we calculated molar volumes with Gaussian 03 at the MP2(full)/6-311+G(2d,2p) level using the MP2(full)/6-31G* geometries. To determine accurate volumes, the MP2 density was used, and numerical integration was employed as the default Monte Carlo algorithm produced spurious and irreproducible results.

To more clearly visualize the influence of the 5-substituents upon the electronic properties of the nucleobases, we calculated electrostatic potential maps for the neutral species. The process involves calculation of the interaction of a +1 probe charge and every part of the electron density cloud of these species calculated at the MP2(full)/6-311+G(2d,2p) level. The electrostatic potential was then mapped onto an isosurface of 0.002 electrons/Å^3; of the total SCF electron density for the species of interest. The electrostatic potential maps were then color-coded according to their potential with the regions of negative electrostatic potential shown in red, neutral regions shown in green, and positive regions shown in blue. The electrostatic potential range used for the maps generated in this work varied from −31 to +31 kcal/mol.

Results

Single Turnover Kinetics. Like many DNA glycosylases, hTDG binds with high affinity to its reaction product, abasic DNA.[36−38] For hTDG, the resulting product inhibition is so potent that it exhibits almost no turnover in vitro. We therefore used single turnover kinetics and saturating enzyme conditions to obtain rate constants (k_{max}) that were not impacted by product release or the bimolecular association of enzyme and substrate, and thus correspond to the maximal catalytic activity of hTDG (Figure 3).[39]

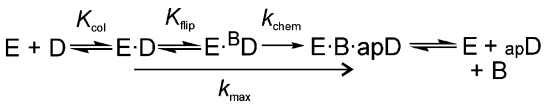

Figure 3. Minimal kinetic mechanism for the hTDG-catalyzed reaction. The initial association of hTDG (E) and substrate DNA (D) produces the collision complex (E·D) followed by a base-flipping step to form the reactive complex (E·^{B}D, where ^{B}D is DNA with an extrahelical base). Cleavage of the base−sugar bond (k_{chem}) yields the ternary product complex (E·B·apD). Release of the base (B) likely precedes very slow release of AP DNA (apD).[36,38] The k_{max} values report on the reaction steps from the initial hTDG·DNA collision complex to the ternary product complex.

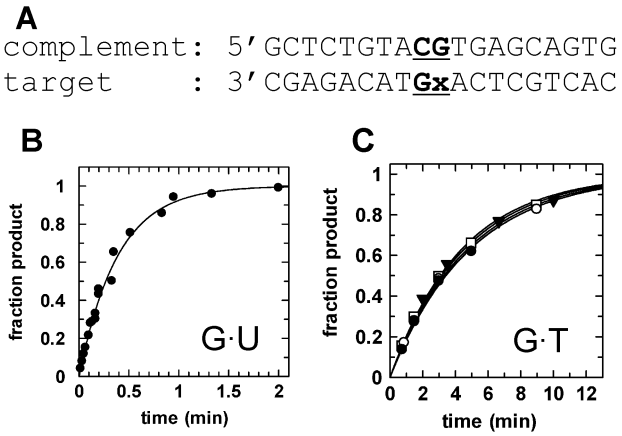

Figure 4. Single turnover kinetics for hTDG. (A) Structure of the 19 bp DNA substrate used for the kinetics experiments, where the target base (x) is located at a CpG site (bold and underlined). A substrate in which x = uracil is referred to as G·U19 (B) Representative data for the maximal activity of hTDG (5 μM) against G·U19 (500 nM) gives a rate constant of $k_{max} = 2.6$ min^{-1}. Data for G·U19 were collected using manual sampling and a rapid chemical quench-flow instrument (for $t < 8$ s). (C) Confirming the saturating enzyme conditions, k_{max} values for G·T19 are independent of enzyme for hTDG concentrations ≥1.0 μM: $k_{obs} = 0.20$ min^{-1} for 1.0 μM hTDG (●); $k_{obs} = 0.21$ min^{-1} for 2.5 μM hTDG (○); $k_{obs} = 0.23$ min^{-1} for 5 μM hTDG (▼); $k_{obs} = 0.22$ min^{-1} for 10 μM hTDG (□).

The kinetics experiments were conducted with 19 bp duplex DNA substrates that varied only in the identity of the target base (Figure 4A). Because hTDG is specific for bases that are paired with guanine and located in a CpG site,[18,41] the target base was placed in this context for all substrates. Representative data for the single turnover activity of hTDG against the G·U19 substrate is shown in Figure 4B. To confirm saturating enzyme conditions, the single turnover experiments were collected with multiple hTDG concentrations. The k_{max} values

(31) Foresman, J.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods: A Guide to Using Gaussian*, 2nd ed.; Gaussian: Pittsburgh, 1996.
(32) Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 10/21/03, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, U.S.A., and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multiprogram laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830.
(33) Check, C. E.; Faust, T. O.; Bailey, J. M.; Wright, B. J.; Gilbert, T. M.; Sunderlin, L. S. *J. Phys. Chem. A* **2001**, *105*, 8111−8116.
(34) Boys, S. F.; Bernardi, R. *Mol. Phys.* **1979**, *19*, 553.
(35) Van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873−1885.
(36) Waters, T. R.; Gallinari, P.; Jiricny, J.; Swann, P. F. *J. Biol. Chem.* **1999**, *274*, 67−74.
(37) Noll, D. M.; Gogos, A.; Granek, J. A.; Clarke, N. D. *Biochemistry* **1999**, *38*, 6374−6379.
(38) McCann, J. A.; Berti, P. J. *J Biol. Chem.* **2003**, *278*, 29587−29592.
(39) For the single turnover experiments under saturating enzyme conditions, $k_{max} = k_{chem}(K_{flip}/1 + K_{flip})$,[8] where K_{flip} is the equilibrium constant for the base-flipping step. Our findings indicate that k_{max} depends on k_{chem} for most of the substrates examined, except for those with bulky 5-substituents such as BrU, IU, and BrC (see text). A direct determination of k_{chem} would require stopped-flow experiments with substrates containing a fluorophore that reports on base-flipping (i.e., 2-aminopurine),[40] which is beyond the scope of this work.
(40) Stivers, J. T.; Pankiewicz, K. W.; Watanabe, K. A. *Biochemistry* **1999**, *38*, 952−963.
(41) Waters, T. R.; Swann, P. F. *Mutat. Res.* **2000**, *462*, 137−147.

Table 1. Kinetic Parameters[a] for hTDG

substrate	k_{max} (min^{-1}) [b]	relative to GU	relative to GT	relative to GC	k_{non} (min^{-1}) [c]	rate enhancement [f]
G·**U**	2.6 ± 0.3	(1)	11.8		2.4×10^{-9}	$10^{9.0}$
G·**T**	0.22 ± 0.04	0.08	(1)		7.5×10^{-10}	$10^{8.5}$
G·**FU**	202 ± 16	78	918		2.9×10^{-8}	$10^{9.9}$
G·**ClU**	126 ± 16	48	572		5.0×10^{-8}	$10^{9.4}$
G·**BrU**	11.6 ± 1.0	4.5	53		4.2×10^{-8} [d]	$10^{8.4}$
G·**IU**	0.06 ± 0.006	0.023	0.27		9.5×10^{-9}	$10^{6.8}$
G·**hoU**	2.2 ± 0.2	0.85	10			
G·**hmU**	1.9 ± 0.2	0.73	8.6			
G·**C**	$1.2 \times 10^{-5} \pm 0.5 \times 10^{-5}$	$10^{-5.3}$	$10^{-4.3}$	(1)	5×10^{-12} [e]	$10^{6.4}$
G·**FC**	0.035 ± 0.004			$10^{3.5}$		
G·**BrC**	0.008 ± 0.0007			$10^{2.8}$		
G·**hoC**	0.010 ± 0.001			$10^{2.9}$		

[a] Rate constants for the maximal activity of hTDG against various nucleobases (bold) paired with guanine, determined from single turnover experiments under saturating enzyme conditions. [b] The k_{max} values are the average of at least three determinations. [c] Rate constants for the nonenzymatic 2′-deoxynucleoside hydrolyses (k_{non}) at 22 °C are from Arrhenius plots constructed from data reported for higher temperatures.[22,23,42] [d] For BdU, the k_{non} shown represents the average of values determined using the data of Shapiro and Kang (5.3×10^{-8} min^{-1})[22] and the data of Van Schepdael et al. (3.0×10^{-8} min^{-1}).[42] [e] Rate constants for the nonenzymatic dC hydrolyses were previously determined at pH < 6 and temperatures ≥75 °C. Thus, k_{non} for dC at 22 °C, pH 7.4, was calculated here by extrapolating the reported rate versus pH curve (data at 95 °C) to pH 7.4, and converting this rate constant to one at 22 °C using a ratio of rates for 22 and 95 °C from an Arrhenius plot constructed from data at high temperautures (≥75 °C) and pH ≈ 1.5.[23] [f] Rate enhancement is k_{max}/k_{non}.

obtained for G·T19 are constant (within experimental error) for hTDG concentrations ranging from 1 to 10 μM (Figure 4C). Similar results are observed for the U·G19 substrate over the same range of concentrations of hTDG (data not shown). For all the other substrates examined here, data were collected at 5 and 10 μM concentrations of hTDG and gave equivalent k_{max} values (within experimental error). The results of these kinetic experiments are summarized in Table 1.

Theoretical Results. Theoretical structures for the neutral and N1 deprotonated nucleobases were calculated as described in the Theoretical Calculations section. The calculated deprotonation enthalpies and free energies are summarized in the Supporting Information (Table S1). Independent zero-point energy, basis set superposition error, and thermal corrections are provided for all nucleobases listed. In previous work, we also calculated the N1 and N3 acidities of uracil at the CBS-Q level of theory.[43] Those calculations suggested that the MP2-(full)/6-311+G(2d,2p)//MP2(full)/6-31G* level of theory used here somewhat overestimates the acidities, but showed that the relative acidities are accurately reproduced. Thus, the trends in the MP2 acidities should be good descriptors of the influence of the 5-substituent on the acidity of the nucleobase. Our calculated deprotonation enthalpy for uracil N1 (Table S1, Supporting Information) is in good agreement with previously reported experimental[44,45] and calculated[44,46−48] values. Likewise, good absolute and excellent relative agreement between our calculations and previous studies is obtained for N1 acidities of thymine,[47,49,50] cytosine,[47,51] 5-flurouracil (FU),[50,52] 5-chlorouracil (ClU),[52] and 5-hydroxyuracil (hoU).[50] It was also previously established that the use of ECPs for the halogen substituent produced only a modest effect on the calculated acidities of BrU,[43] and thus, it is expected that the calculated acidity of IU is also reasonably accurate.

The calculated molar volumes of the isolated nucleobases are summarized in Table 2 along with other steric and electrostatic properties. The calculated electrostatic potential maps are shown in Figure 5.

Structure−activity correlations. Previous studies indicate that nonenzymatic pyrimidine hydrolysis reactions are dissociative,[22,23] where the rate depends on the leaving ability of the nucleobase. This suggested that we could test the role of *N*-glycosidic bond stability in the specificity of hTDG by determining its maximal activity against substrates with varying leaving-group ability. Such structure−activity correlations often require substrates with significantly different steric and electrostatic properties and are therefore not feasible for many enzymes, including probably most DNA glycosylases. However, hTDG is a good candidate for this approach, because it has a relatively accommodating and nonspecific active site.[20,21] The substrates examined here differ significantly in leaving-group ability, where the pK_a^{N1} for a 5-substituted uracil or cytosine base depends on the electronic effect (σ_m) of the C5 substituent. The C5 substituent can also affect the steric and electrostatic properties of a nucleobase, which could potentially impact its interaction with the active site and its rate of excision by hTDG. These properties are given in Table 2.

hTDG Activity against 5-Substituted Uracils. We found that 5-halogen substituents dramatically alter the excision rate of uracil by hTDG. Indeed, the activity against G·FU pairs, $k_{max} = 202$ min^{-1} (Figure 6A), is *78 times faster* than for G·U pairs (Table 1). Thus, the small increase in size of FU over U (Table 2, Figure 5) does not appear to significantly impact its access to the hTDG active site. The propensity of FU to form electrostatic interactions is probably similar to that of U, because the change in polarity due to the 5-F substitution is relatively

(42) Van Schepdael, A.; Ossembe, N.; Herdewijn, P.; Roets, E.; Hoogmartens, J. *J. Pharm. Biomed. Anal.* **1993**, *11*, 345−351.
(43) Yang, Z.; Rodgers, M. T. *J. Am. Chem. Soc.* **2004**, *126*, 16217−16226.
(44) Kurinovich, M. A.; Lee, J. K. *J. Am. Chem. Soc.* **2000**, *122*, 6258−6262.
(45) Miller, T. M.; Arnold, S. T.; Viggiano, A. A.; Miller, A. E. S. *J. Phys. Chem. A* **2004**, *108*, 3439−3446.
(46) Nguyen, M. T.; Chandra, A. K.; Zeegers-Huyskens, T. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 1277−1280.
(47) Huang, Y. Q.; Kenttamaa, H. *J. Phys. Chem. A* **2003**, *107*, 4893−4897.
(48) Di Laudo, M.; Whittleton, S. R.; Wetmore, S. D. *J. Phys. Chem. A* **2003**, *107*, 10406−10413.
(49) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **1998**, *102*, 6010−6016.
(50) Whittleton, S. R.; Hunter, K. C.; Wetmore, S. D. *J. Phys. Chem. A* **2004**, *108*, 7709−7718.
(51) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Mol. Struct.* **2000**, *519*, 1−11.
(52) Chandra, A. K.; Uchimaru, T.; Zeegers-Huyskens, T. *J. Mol. Struct.* **2002**, *605*, 213−220.
(53) Nakanishi, K.; Suzuki, N.; Yamazaki, F. *Bull. Chem. Soc. Jpn.* **1961**, *34*, 53−57.
(54) Berens, K.; Shugar, D. *Acta Biochim. Pol.* **1963**, *10*, 25−48.
(55) Wempen, I.; Fox, J. J. *J. Am. Chem. Soc.* **1964**, *86*, 2474−2477.

Table 2. Parameters for the 5-Substituted Uracil and Cytosine Bases

base	C5 substituent	pK_a^{N1} [a]	electronic substituent constant (σ_m)[b]	hydrophobic substituent constant (π)[c]	projection along C5-X axis (Å)[d]	volume of base (Å^3)
U	H	9.76	(0)	(0)	2.28	127.6
T	CH_3	10.19	−0.07	0.56		150.5
FU	F	8.43	0.34	0.14	2.81	132.7
ClU	Cl	8.14	0.37	0.71	3.49	149.9
BrU	Br	8.24	0.39	0.86	3.80	155.4
IU	I	8.44	0.35	1.12	4.01	166.7
hoU	OH	9.34	0.12	−0.67		137.3
hmU	CH_2OH	9.82	0.00	−1.03		160.2
C	H	12.20	(0)	(0)	2.28	133.7
FC	F	10.87	0.34	0.14	2.83	138.6
BrC	Br	10.33	0.39	0.86	3.81	160.7
hoC	OH	11.66	0.12	−0.67		143.7

[a] The observed ionization constants (K_a) for uracil and 5-substituted uracils are composites of the microscopic ionization constants for the N1 (K_a^{N1}) and N3 (K_a^{N3}) sites. The (N1) pK_a values (pK_a^{N1}) for U and T were calculated previously[22] using the relative populations of the N1$^-$ (N1 deprotonated) and N3$^-$ monoanions (determined spectrophotometrically)[53] and the equations: $K_a = K_a^{N1} + K_a^{N3}$, and $K_T = K_a^{N1}/K_a^{N3}$, where K_T is the tautomeric equilibrium constant. Using these equations, we calculated pK_a^{N1} values for FU, ClU, BrU, and IU using the macroscopic pK_a values[54] and the population of N1$^-$ and N3$^-$ monoanions in aqueous solution.[55,56] Previously, Shapiro and Kang calculated pK_a^{N1} = 8.49 for BrU using a population of 36% N1$^-$ (64% N3$^-$).[22] However, the correct populations are 64% N1$^-$ and 36% N3$^-$,[55] from which one obtains pK_a^{N1} = 8.24 for BrU. We calculated the macroscopic pK_a for hoU and hmU from a linear fit of pK_a versus σ_m[57] for U, T, and the 5-halouracils, and this calculated pK_a was used to determine the pK_a^{N1} using the above equations and assuming a population of 50% for N1$^-$ (the actual population values have not been reported to our knowledge). The pK_a^{N1} values for C, 5FC, and 5BrC were reported,[58] and that of hoC was calculated from a linear fit of pK_a^{N1} versus σ_m for C, FC, and BrC. [b] The electronic substituent (Hammett) constant for the meta position (σ_m) gives the total electronic effect, where positive values indicate electron-withdrawing, and negative values, electron-donating substituents.[57] [c] The hydrophobic substituent constant (π) gives the hydrophobic character of a substituent as determined from partitioning studies of substituted benzenes into the octanol−water solvent system,[57] where increasingly positive π values indicate greater hydrophobicity. [d] The "projection along the C5-X bond axis" is the sum of the reported van der Waals radius of the substituent[59] and the C5-X bond length, obtained from the structures of the nucleobases optimized at the MP2(full)/6-31G* level of theory. Values are not given for nucleobases with multi-atom C5-substituents (T, hmU, hoU, and hoC).

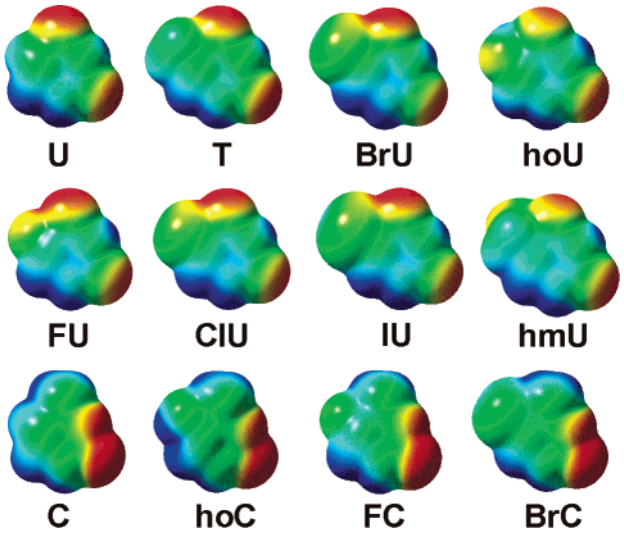

Figure 5. Electrostatic potential maps of the 5-substituted uracil and cytosine bases used in this work. The range of electrostatic potential varies between −31 kcal/mol (red) to +31 kcal/mol (blue), on an isosurface of 0.002 electrons/Å^3; of the total SCF electron density.

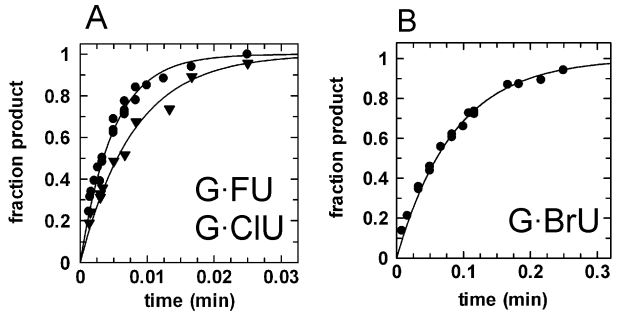

Figure 6. Representative data for the activity of hTDG against 5-halouracils. (A) hTDG rapidly excises FU, k_{max} = 202 min^{-1} (●), and ClU, k_{max} = 126 min^{-1} (▼). (B) The removal of BrU is significantly slower, k_{max} = 11.6 min^{-1}, likely due to steric effects with the active site (see text).

small (π, Table 2), and it has been shown that C−F groups in aromatic systems are poor H-bond acceptors[60] and poor halogen bond donors.[61] Indeed, the enthalpies of base pairing for isolated A·U and A·FU pairs were calculated to be respectively 12.2 and 12.6 kcal/mol at the MP2(full)/6-311+G(2d,2p)//B3LYP/6-31G* level of theory.[43] Thus, the greater hTDG activity against FU can be largely attributed to the superior leaving ability of FU over that of U ($\Delta pK_a^{N1} = -1.33$), and is consistent with the higher intrinsic reactivity of FdU compared to that of dU (k_{non}, Table 1).

We also discovered that hTDG is highly active against G·ClU pairs, k_{max} = 126 min^{-1} (Figure 6A), corresponding to a 48-fold and 572-fold increase over G·U and G·T pairs, respectively (Table 1).[62] This is remarkable, considering that ClU is much bulkier than U and is nearly isosteric with T (Figure 5, Table 2). The H-bonding properties of ClU and T are probably not substantially different, because the 5-Cl and 5-CH_3 substituents have a similar hydrophobic effect (π, Table 2), and previous studies have shown that the hydrophobicity of uracil was increased to the same extent by halogen (5-Br) and methyl (5-CH_3) substituents.[65,66] Moreover, the enthalpies of base pairing for the isolated A·T and A·ClU pairs were calculated to be respectively 12.1 and 12.8 kcal/mol at the MP2(full)/6-

(56) Wierzchowski, K. L.; Litonska, E.; Shugar, D. *J. Am. Chem. Soc.* **1965**, *87*, 4621−4629.
(57) Hansch, C.; Leo, A.; Unger, S. H.; Kim, K. H.; Nikaitani, D.; Lien, E. J. *J. Med. Chem.* **1973**, *16*, 1207−1216.
(58) Wempen, I.; Fox, J. J. *J. Med. Chem.* **1963**, *122*, 688−693.
(59) Bondi, A. *J. Am. Chem. Soc.* **1964**, *68*, 441−451.
(60) Kool, E. T. *Acc. Chem. Res.* **2002**, *35*, 936−943.
(61) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16789−16794.

311+G(2d,2p)//B3LYP/6-31G* level of theory.[43,67] Thus, the 572-fold greater activity against ClU can be most readily explained by the superior leaving ability of ClU compared to that of T ($\Delta pK_a^{N1} = -2.05$), and is consistent with the higher intrinsic reactivity (k_{non}) of CldU over that of dT (Table 1).

In contrast with FU and ClU, the limits of the hTDG active site appear to be tested by BrU and IU. The activity against G·BrU pairs, $k_{max} = 11.6\ min^{-1}$ (Figure 6B), is significantly lower than for G·FU and G·ClU pairs, although still 53 times greater than for G·T pairs (Table 1). Given that BrU and ClU are nearly equivalent in leaving ability and have similar inherent reactivity (k_{non}), the 11-fold lower hTDG activity against BrU indicates that its access to the active site is limited, presumably by steric effects (see "projection along C5-X axis", Table 2). Consistent with this, hTDG activity against the bulky IU is sharply reduced, $k_{max} = 0.06\ min^{-1}$ (data not shown), and is about 4-fold *slower* than for G·T pairs. The leaving group ability of IU is similar to that of the other 5-halouracils, suggesting that its low reactivity is due to steric hindrance in the active site (Table 2). Indeed, IU is removed $10^{-3.5}$ times more slowly than FU, even though FU and IU have the same pK_a^{N1} values, and FdU and IdU exhibit similar nonenzymatic rates (Table 1).

We also examined the activity against 5-hydroxyuracil (hoU) and 5-hydroxymethyluracil (hmU), which are similar to U in leaving ability but differ significantly in their steric and electrostatic properties (Table 2, Figure 5). The activity against G·hmU pairs, $k_{max} = 1.9\ s^{-1}$, and G·hoU pairs, $k_{max} = 2.2\ s^{-1}$ (data not shown) is about the same as for G·U pairs.[68] This finding is consistent with indications from the results above that hTDG activity depends on glycosidic bond stability.

Cytosine Analogues with Improved Leaving Ability Are Substrates of hTDG. The strong specificity of hTDG for G·T over A·T pairs[9] likely arises from H-bond interactions that are compatible with the Watson−Crick base-pairing groups of guanine but not adenine.[19] Given its specificity for removing bases that are paired with guanine, hTDG must employ a stringent mechanism to avoid the excision of cytosine from the million-fold excess of G·C pairs over G·T mispairs. Accordingly, hTDG activity against G·C pairs has not been previously observed.[13,18] Nevertheless, under the saturating enzyme conditions used in this work, we were able to measure an exceedingly weak activity against G·C pairs, $k_{obs} = 1.2 \times 10^{-5}\ min^{-1}$ (Figure S1, Supporting Information). Given the strong hTDG activity against FU and ClU (Table 1), we suspected that improving the leaving ability of cytosine ($pK_a^{N1} = 12.2$) could substantially increase its rate of excision by hTDG. Indeed, hTDG exhibits significant activity against 5-fluorocytosine (FC), $k_{max} = 0.035\ min^{-1}$, 5-bromocytosine (BrC), $k_{max} = 0.008\ min^{-1}$, and 5-hydroxycytosine (hoC), $k_{max} = 0.010\ min^{-1}$ (Figure 7). Remarkably, these cytosine analogues are removed $10^{2.8}$- to $10^{3.5}$-fold faster than cytosine (Table 1), consistent with their increased N1 acidities (Table 2). Although the leaving ability of BrC is better than that of FC and nearly the same as that of T, BrC is excised 4-fold and 28-fold slower than FC and T, respectively, suggesting that the access of BrC to the active site is limited, probably by steric effects (Table 2, Figure 5). The observation that hTDG excises cytosine, albeit very slowly, and readily removes cytosine analogues that are *bulkier than cytosine* indicates that cytosine can flip into the hTDG active site.

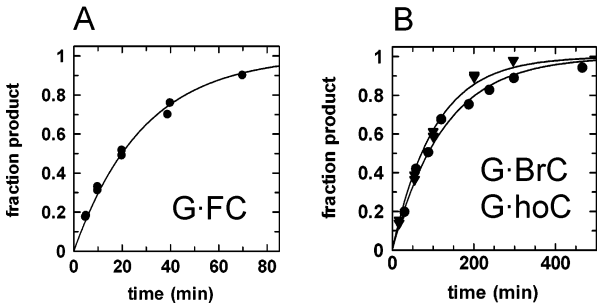

Figure 7. hTDG excises cytosine analogues with improved leaving-group ability. Shown are representative data from single turnover experiments for (A) the removal of FC from G·FC pairs, $k_{max} = 0.035\ min^{-1}$, and (B) the removal of BrC from G·BrC pairs, $k_{max} = 0.008\ min^{-1}$ (●), and hoC from G·hoC pairs, $k_{max} = 0.010\ min^{-1}$ (▼).

As an important control, we sought to determine whether the halogen substitutions perturbed the stability of G·U and G·C pairs in the duplex substrates, because previous studies indicate that a target base which is more prone to flipping out of the duplex may be more rapidly removed by a DNA glycosylase.[8,70] This could lead to an increase in k_{max} unrelated to glycosidic bond stability, i.e., due to a change in K_{flip} rather than k_{chem} (Figure 3). To address this question, we conducted melting studies on the DNA substrates. We find that the 5-halogen substitutions do not alter the stability of the G·U or G·C base pairs in the substrates examined here (Supporting Information).

Linear Free Energy Relationships for *N*-Glycosidic Bond Hydrolysis. Taken together, our observations indicate a strong dependence of hTDG activity on the leaving ability of the target base, suggesting that a quantitative analysis would be informative. Indeed, a plot of $\log(k_{max})$ versus leaving-group pK_a^{N1} reveals a Brønsted-type linear free energy relationship (LFER) with a large negative slope of $\beta_{lg} = -1.6 \pm 0.2$ (Figure 8). The large negative slope of the LFER indicates that negative charge accumulates on the departing base in the transition state of the hTDG reaction and that the *N*-glycosidic bond is largely or perhaps fully ruptured such that the rate of C−N bond

(62) Halogen substitutions lower the pK_a for both the N1 and N3 sites of uracil, raising the possibility that the 5-halouracils could be partially ionized at pH 7.5, which could lead to decreased k_{max} values because the ionized base is a poor substrate. However, previous studies reported that for a G·BrU pair in a 7 bp duplex, BrU ionizes with $pK_a^{N3} = 8.6$ and is neutral at pH 7.5.[63] Similarly, for a G·FU pair in a 7 bp duplex, FU ionizes with $pK_a^{N3} = 8.3$ and is predominantly neutral at pH 7.5.[64] Nevertheless, we repeated the single turnover experiments at pH 7.0 and found that the k_{max} values are the same within error (not shown). As an important control, k_{max} is the same for G·U pairs at pH 7.0 and pH 7.5, where uracil is known to be neutral ($pK_a^{N3} = 9.7$).

(63) Sowers, L. C.; Goodman, M. F.; Eritja, R.; Kaplan, B.; Fazakerley, G. V. *J. Mol. Biol.* **1989**, *205*, 437−447.

(64) Sowers, L.; Eritja, R.; Kaplan, B.; Goodman, M.; Fazakerly, G. *J. Biol. Chem.* **1988**, *263*, 14794−14801.

(65) Shih, P.; Pedersen, L. G.; Gibbs, P. R.; Wolfenden, R. *J. Mol. Biol.* **1998**, *280*, 421−430.

(66) Giesen, D. J.; Chambers, C. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, *101*, 5084−5088.

(67) Yang, Z.; Rodgers, M. T. *Int. J. Mass Spectrom.* **2005**, *241*, 225−242.

(68) Ionization of the hydroxyl group of hoU, hmU, and hoC must be considered, because this would significantly increase pK_a^{N1} and therefore lower k_{max}. This is not a concern for the G·hmU substrate, because a $pK_a^{OH} = 9.33$ was reported for 5-hm-dU.[69] The ionization of hoC in DNA can be predicted from the $pK_a^{OH} = 8.5$ reported for 5-hydroxy-dC-5′-monophosphate, and a similar value is likely for hoU in DNA.[69] Nevertheless, for the hoU and hoC substrates, we collected the single turnover data at pH 6.5 to ensure that the k_{max} values were not impacted by 5-OH ionization, and found the same k_{max} values within experimental error.

(69) La Francois, C. J.; Jang, Y. H.; Cagin, T.; Goddard, W. A.; Sowers, L. C. *Chem. Res. Toxicol.* **2000**, *13*, 462−470.

(70) Krosky, D. J.; Schwarz, F. P.; Stivers, J. T. *Biochemistry* **2004**, *43*, 4188−4195.

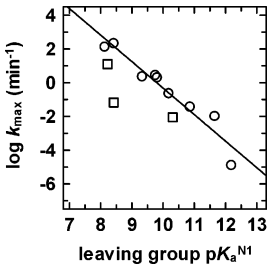

Figure 8. Brønsted-type linear free energy relationship (LFER) for the hTDG reaction. The LFER has a good correlation coefficient ($r = 0.96$) and a large negative slope of $\beta_{lg} = -1.6 \pm 0.2$ (Brønsted coefficient). The LFER includes data for U, T, FU, ClU, hoU, hmU, FC, hoC, and C (○) and covers a range in leaving-group pK_a^{N1} of >4 log units. A determination of the LFER using only the 5-substituted uracils gives essentially the same result, $\beta_{lg} = -1.4 \pm 0.2$, $r = 0.96$. Data for BrU, IU, and BrC are shown (□) but were not included in the LFER because the k_{max} values for these bases suggests limited access to the active site (see text).

cleavage is highly sensitive to the influence of the 5-substituent, where electron-withdrawing substituents (e.g., Cl) increase the rate by stabilizing the transition state, and electron-donating groups (e.g., CH_3) destabilize the transition state and slow the reaction. Thus, our findings suggest the hTDG-catalyzed reaction follows a highly dissociative or perhaps a stepwise mechanism, consistent with previous work showing that nearly all enzymatic and nonenzymatic *O*- and *N*-glycoside hydrolyses proceed through a highly dissociative A_ND_N (S_N2) or a stepwise $D_N{*}A_N$ (S_N1) mechanism.[24]

The Brønsted-type LFER shows that hTDG activity (k_{max}) is highly correlated with the N1 acidity of the target base in *aqueous solution*. However, structural studies indicate that the hTDG active site is hydrophobic,[20,21] and nucleobase acidity varies with the polarity of the environment.[44,56] Indeed, the N1 and N3 acidities of uracil differ by 14 kcal/mol in the gas phase,[44,46,56] whereas they are nearly equivalent in aqueous solution.[22] Thus, as indicated by previous studies, gas-phase acidities can be informative regarding the leaving ability of a nucleobase within a hydrophobic active site, and may be relevant to the mechanism of substrate discrimination.[44,71] We therefore examined the dependence of hTDG activity on N1 acidity in a nonpolar environment by calculating the gas-phase N1 acidities of the nucleobases at 298 K ($\Delta G_{acid,g}$, Table S1, Supporting Information). (For comparison with other studies, the more commonly reported deprotonation enthalpies at 0 and 298 K are also given). A plot of $(2.3RT)\log(k_{max})$ versus $G_{acid,g}$ gives a LFER with a good correlation coefficient ($r = 0.94$) and a slope of $m = -0.42 \pm 0.06$ (Figure S2, Supporting Information). The strong correlation of $\log(k_{max})$ with the nucleobase N1 acidities in both the gas phase ($\Delta G_{acid,g}$) and in aqueous solution (pK_a^{N1}) is explained by the finding that $\Delta G_{acid,g}$ is highly correlated with pK_a^{N1} ($r = 0.98$, Figure S2, Supporting Information), which is perhaps not surprising for acids of similar structure. The slope of this correlation ($m = 3.7 \pm 0.2$) is similar to that observed previously for substituted carboxylic acids,[72,73] and indicates that differences in N1 acidities are nearly 4-fold greater in the gas phase than in aqueous solution. The enhanced acidity in the gas phase as compared to aqueous solution is relevant because it predicts that pK_a^{N1} differences should be enhanced in a hydrophobic active site, leading to greater differences in k_{max} between various substrates than would be expected for an aqueous environment. As discussed below, our findings indicate that a hydrophobic active site increases the specificity of hTDG for G·T mispairs relative to G·C pairs.

Discussion

Implications of the Brønsted-Type LFER for the Mechanism of hTDG. We found that the hTDG reaction is remarkably amenable to Brønsted-type LFER analysis (Figure 8), to our knowledge, the first such study for a DNA glycosylase. Indeed, this approach is probably not feasible for most DNA glycosylases, because a selective active site may not accommodate the perturbations in substrate size and/or polarity that can accompany changes in leaving-group ability. Even the closely related eMUG is probably too restrictive for such studies, because its activity is exceedingly slow for uracil analogues with modestly sized 5-substituents such as T (5-CH_3) and hmU (5-CH_2OH).[74] In contrast, T, hmU, and ClU are well fit by the LFERs for hTDG (Figure 8 and Figure S2, Supporting Information). Moreover, the Brønsted LFER for hTDG includes nine substrates, covers a large range in pK_a^{N1} (>4 log units), and exhibits a good correlation coefficient ($r = 0.96$), indicating that the $\beta_{lg} = -1.6 \pm 0.2$ accurately reflects the dependence of k_{max} on the leaving ability of the target base (pK_a^{N1}).

It is informative to compare the results for hTDG with the nonenzymatic reaction. A plot of $\log(k_{non})$ versus pK_a^{N1} for the spontaneous hydrolysis of dU, dT, and 5-Br-dU at pH 6.5 and 75 °C has a slope of $\beta_{lg} = -0.86 \pm 0.03$,[22,75] and the same value ($\beta_{lg} = -0.86 \pm 0.05$) is obtained using k_{non} values extrapolated to 22 °C for dU, dT, 5-F-dU, 5-Cl-dU, and 5-Br-dU (Figure S3, Supporting Information). These β_{lg} values, together with the observed low and positive activation entropies,[22] are consistent with a highly dissociative transition state for the nonenzymatic reaction,[22,23,76] although it was concluded from computational studies that the mechanism is concerted (A_ND_N) rather than stepwise ($D_N{*}A_N$).[77] The more negative $\beta_{lg} = -1.6$ observed for hTDG indicates that the transition state (TS) of the enzymatic reaction is more sensitive to the development of negative charge on N1 of the leaving-group base than is the TS of the nonenzymatic reaction. Thus, the $\beta_{lg} = -1.6$ may reflect a more dissociative TS for hTDG compared to the nonenzymatic reaction. The difference in β_{lg} may also be attributable to the hydrophobic active site of hTDG.[20,21] As noted previously,[78,79]a more negative β_{lg} can be expected for a reaction occurring in an environment that is less polar than aqueous solution, where charge development in the TS of the latter is stabilized to a greater extent by solvent. Nevertheless, the large negative $\beta_{lg} = -1.6$ implies that the hTDG reaction

(71) Sharma, S.; Lee, J. K. *J. Org. Chem.* **2002**, *67*, 8360−8365.
(72) Wilson, B.; Georgiadis, R.; Bartmess, J. E. *J. Am. Chem. Soc.* **1991**, *113*, 1762−1766.
(73) Kumar, G. A.; McAllister, M. A. *J. Am. Chem. Soc.* **1998**, *120*, 3159−3165.
(74) O'Neill, R. J.; Vorob'eva, O. V.; Shahbakhti, H.; Zmuda, E.; Bhagwat, A. S.; Baldwin, G. S. *J. Biol. Chem.* **2003**.
(75) Shapiro and Kang measured the non-enzymatic rates (k_{non}) for the hydrolysis of dU, dT, and 5-Br-dU and showed that $\log(k_{non})$ versus pK_a^{N1} was highly linear.[22] Although they did not report a slope, one obtains $\beta_{lg} = -0.99 \pm 0.01$ using their data, which included p$K_a^{N1} = 8.49$ for BrU, a value they calculated based on 36% N1 deprotonated species ($N1^-$) and 64% N3 deprotonated species ($N3^-$), citing Wempen and Fox.[55] However, Wempen and Fox actually report the opposite: 64% $N1^-$ and 36% $N3^-$ for the BrU monoanion,[55] from which one obtains p$K_a^{N1} = 8.24$ for BrU and $\beta_{lg} = -0.86 \pm 0.03$ for the non-enzymatic reaction (Supporting Information).
(76) Stivers, J. T.; Jiang, Y. L. *Chem. Rev.* **2003**, *103*, 2729−2759.
(77) Dinner, A. R.; Blackburn, G. M.; Karplus, M. *Nature* **2001**, *413*, 752−755.

proceeds through a highly dissociative, perhaps stepwise mechanism. Accordingly, hTDG does not appear to have an active-site group that could serve as a general base catalyst, which would likely be needed to activate (deprotonate) the water nucleophile in a more associative reaction. Thus, our findings are consistent with previous work showing that nearly all enzymatic and nonenzymatic *O*- and *N*-glycoside hydrolyses follow a highly dissociative A_ND_N (S_N2) or stepwise $D_N{*}A_N$ (S_N1) mechanism.[24]

It is of interest to consider how hTDG could stabilize a highly dissociative transition state. The precedent for enzymatic *N*-glycosidic bond hydrolysis in DNA is the reaction catalyzed by uracil DNA glycosylase (UDG), which proceeds through a stepwise mechanism involving an oxacarbenium cation−uracil anion intermediate.[77,80−82] For UDG, two conserved residues each contribute about 5 kcal/mol of TS stabilization; an Asp side chain serves to stabilize the oxacarbenium cation and activate (or position) the water nucleophile, and a His side chain stabilizes the uracil anion leaving group via formation of a strong hydrogen bond.[80,83−85] Although these catalytic groups are absent in hTDG, the enzyme still provides over 12 kcal/mol of TS stabilization for a dU substrate (Table 2) as compared to 17 kcal/mol for UDG.[83] Structural studies of eMUG suggest that it and hTDG may provide some stabilization to an anionic nucleobase leaving group via H-bond interactions involving backbone amides and an active-site water molecule.[20] To stabilize the oxacarbenium ion, hTDG may use tactics employed by UDG, including electrostatic interactions with the anionic leaving group and with the negatively charged DNA backbone.[76,77,82,86]

Role of a Hydrophobic Active Site. Our findings indicate that the hydrophobic active site of hTDG[20,21] enhances the inherent differences in substrate reactivity, thereby increasing the specificity against cytosine. The potential impact of a hydrophobic active site is indicated by the highly linear plot of $\Delta G_{acid,g}$ versus $(2.3RT)pK_a^{N1}$ (Figure S2, Supporting Information), which shows that the N1 acidity differences are 4-fold greater in the gas phase than in aqueous solution. This relationship indicates that differences in nucleobase leaving ability (pK_a^{N1}) should be enhanced in the hydrophobic hTDG active site, thereby leading to larger differences in k_{max}. Consistent with this prediction, the Brønsted LFER for hTDG (Figure 8) shows that the dependence of $\log(k_{max})$ on pK_a^{N1} is steeper ($\beta_{lg} = -1.6$) than would be expected for an active-site environment that mimics aqueous solution ($\beta_{lg} = -0.86$). Thus, the hydrophobic active site of hTDG increases its specificity for G·T over G·C pairs by enhancing the difference in the inherent reactivity between dT and dC. Indeed, hTDG exhibits a $10^{4.3}$-fold difference in k_{max} for G·T relative to G·C pairs, much larger than the $10^{2.2}$-fold difference in the nonenzymatic rates (k_{non}) for dT and dC (Table 1). The observed $10^{4.3}$-fold k_{max} difference also exceeds the $10^{3.2}$-fold difference that would be predicted by the Brønsted LFER ($\Delta pK_a^{N1} = 2.0$, $\beta_{lg} = -1.6$), as indicated by the relatively poor fit of cytosine to the LFER (Figure 8). The origin of this additional specificity against C is presently unknown, although, as discussed below, our findings indicate that cytosine can flip into the active site. Given that hTDG is a promiscuous enzyme, the enhanced (20,000-fold) difference in hTDG activity for G·T versus G·C pairs, as compared to the 150-fold difference in intrinsic reactivity between dT and dC, may be essential for avoiding the danger of either excessive activity against normal G·C pairs or insufficient activity against G·T mispairs.

Our finding that a hydrophobic active site contributes to hTDG specificity by enhancing the differences in nucleobase N1 acidity is reminiscent of previous studies which found that the difference in uracil N1 and N3 acidities is much greater in the gas phase than in aqueous solution, suggesting this difference in acidities may be observed in a hydrophobic active site.[44,46] Although a hydrophobic active site may be expected to slow the reaction because charge development becomes less favorable, it has been shown that the N1 site of uracil is as acidic as HCl in the gas phase, and it was suggested that N1-deprotonated U may be a relatively good leaving group in a nonpolar active site.[44,87] A similar argument can be made for thymine, which is merely 1.5 kcal/mol less acidic than U (Table S1, Supporting Information). Moreover, computational studies indicate that the gas-phase N1 acidities of U, T, FU, and hoU can be significantly increased by H-bond interactions,[48,50] and structural studies indicate that such interactions are provided by backbone amides and/or a water molecule in the eMUG and hTDG active sites.[20]

It is of interest to compare our findings with recent studies of another promiscuous DNA glycosylase, AlkA, for which the base excision rate was also found to depend on substrate reactivity.[7] These studies showed that AlkA exhibits the same rate enhancement (k_{max}/k_{non}) for several damaged and normal purine bases.[7] In contrast, we find that hTDG exhibits *different* rate enhancements for the various pyrimidine substrates examined here, where the rate enhancement decreases with diminishing leaving ability of the nucleobase (Table 2). This observation can be largely attributed to the hydrophobic active site of hTDG, which enhances the inherent differences in substrate reactivity.

Role of Base-Flipping in the Specificity against Cytosine. It was previously suggested that the specificity of eMUG and hTDG against cytosine may be attributable to the greater stability of G·C pairs over that of G·U (and G·T) mispairs, such that the enzymes are unable to flip cytosine into the active site.[19] However, our observation that hTDG removes cytosine (albeit very slowly) and readily excises the *larger* cytosine analogues, FC, BrC, and hoC, indicates that cytosine has access to the active site. Nevertheless, given that the equilibrium constant for the spontaneous opening of G·C base pairs is 3−4 orders of magnitude smaller than for G·T mispairs,[76] it could be argued that greater G·C stability decreases the equilibrium for *enzymatic* base flipping (K_{flip}, Figure 3), thereby decreasing the lifetime of cytosine in the active site. However, our observations suggest otherwise. If K_{flip} was diminished for C, FC, BrC, and hoC compared to that for the 5-xU bases, then the k_{max} values would be lower (because $k_{max} = k_{chem}(K_{flip}/1 + K_{flip})$) and would not

(78) Jencks, W. P. *Cold Spring Harbor Symp. Quant. Biol.* **1972**, *36*, 1−11.
(79) Hollfelder, F.; Herschlag, D. *Biochemistry* **1995**, *34*, 12255−12264.
(80) Werner, R. M.; Stivers, J. T. *Biochemistry* **2000**, *39*, 14054−14064.
(81) Drohat, A. C.; Stivers, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 1840−1841.
(82) Jiang, Y. L.; Drohat, A. C.; Ichikawa, Y.; Stivers, J. T. *J. Biol. Chem.* **2002**, *277*, 15385−15392.
(83) Drohat, A. C.; Jagadeesh, J.; Ferguson, E.; Stivers, J. T. *Biochemistry* **1999**, *38*, 11866−11875.
(84) Drohat, A. C.; Stivers, J. T. *Biochemistry* **2000**, *39*, 11865−11875.
(85) Dong, J.; Drohat, A. C.; Stivers, J. T.; Pankiewicz, K. W.; Carey, P. R. *Biochemistry* **2000**, *39*, 13241−13250.
(86) Jiang, Y. L.; Ichikawa, Y.; Song, F.; Stivers, J. T. *Biochemistry* **2003**, *42*, 1922−1929.
(87) Kurinovich, M. A.; Lee, J. K. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 985−995.

provide a good fit to the LFERs (Figure 8 and Figure S2 of the Supporting Information). On the contrary, the LFERs provide no indication that the k_{max} values are *uniformly lower* for C, FC, BrC, and hoC relative to that for the uracil analogues. Thus, our findings indicate that cytosine can flip into the hTDG active site, but the enzyme simply lacks the necessary catalytic power to remove cytosine at a significant rate. In marked contrast, uracil DNA glycosylase (UDG) employs steric hindrance and selective electrostatic interactions to preclude cytosine from docking in its active site.[2] Moreover, UDG possesses enough catalytic power to remove cytosine, as a UDG variant (N123D) designed to form hydrogen bonds with cytosine can remove cytosine at a significant rate ($k_{cat} = 0.02$ min^{-1}).[88,89] Thus, by excluding cytosine from its active site, UDG can attain high catalytic power while avoiding the danger of aberrant cytosine excision.

Mechanism of hTDG-Catalyzed Excision of C and 5-xC Analogues. Our results indicate that the mechanism employed by hTDG in the excision of C and 5-xC analogues differs from that of the nonenzymatic and UDG-catalyzed reactions. The nonenzymatic dC and 5BrdC hydrolyses are acid catalyzed through N3 protonation.[23] The bromine substituent increases the acidity of the N1 and N3 sites, lowering the concentration of the N3-protonated species.[58] Thus, dC is hydrolyzed as rapidly as 5BrdC (at 95 °C, pH 5) even though BrC has a much lower pK_a^{N1} than C (Δp$K_a^{N1} = 1.9$).[23] Likewise, studies of a UDG variant (N123D) designed to remove cytosine concluded that the reaction involves protonation of cytosine N3 prior to glycosidic bond cleavage.[88] In contrast, our findings indicate that the hTDG-catalyzed excision of C, FC, BrC, and hoC is not acid catalyzed. Indeed, hTDG removes FC $10^{3.5}$-fold *faster* than C even though the population of protonated FC is expected to be 60-fold *lower* than protonated C (Δp$K_a^{N3} = 1.8$).[58] Similarly, hTDG removes BrC and hoC much more rapidly than C, despite the lower population of protonated BrC and hoC compared to protonated C. Consistent with our findings, hTDG does not appear to have an active-site side chain that could protonate cytosine N3.[20,21] Thus, we conclude that the excision of C, FC, BrC, and hoC by hTDG is not acid catalyzed, and the base departs as the monoanion rather than the neutral species, such that the rate depends largely on pK_a^{N1}.

Biological Relevance of hTDG Activity against ClU and BrU. The observation of strong hTDG activity against ClU and BrU may have biological implications, because these lesions arise in DNA due to oxidative processes associated with inflammation.[90,91] ClU and BrU can be incorporated opposite G[92] to give lesions that are mutagenic, genotoxic, and cytotoxic,[93,94] consistent with the correlation between chronic inflammation and carcinogenesis.[95,96] Our findings raise the possibility that hTDG provides some protection against ClU and BrU lesions, and previous studies indicate that such protection is not offered by other pyrimidine-specific enzymes. Indeed, human UDG and SMUG1 (single-stranded specific monofunctional uracil−DNA glycosylase) are inactive against ClU and BrU,[90,97,98] due likely to their restrictive active sites. Human MBD4 (methyl binding domain IV) reportedly removes ClU and U with similar efficiency and is much less effective against BrU than U.[99] The activity of hTDG will likely be lower for A·BrU and A·ClU pairs, given the reported $10^{2.4}$-fold decrease in hTDG activity for A·U versus G·U pairs.[9] Although the activity will also be somewhat lower for G·ClU and G·BrU lesions located outside of CpG sites, the activity may still be biologically relevant given the potent activity against ClU and BrU observed here (Table 2). The activity of hTDG against ClU, BrU, and other lesions located in DNA contexts other than CpG sites is under investigation.

Acknowledgment. We thank Dr. Yu Lin Jiang, East Tennessee State University, for converting 5-chloro-2′-deoxyuridine to its corresponding phosphoramidite, Dr. Primo Schär, University of Basel, for generously providing an expression plasmid for hTDG, and the reviewers for their helpful suggestions. This work was supported by grants from the NIH (GM 72711) to A.C.D and the NSF (CHE-0518262) to M.T.R, and by the University of Maryland Greenebaum Cancer Center.

Supporting Information Available: A figure showing the hTDG activity against G·C base pairs; experimental methods and results for melting temperature studies on the DNA substrates; a figure showing a plot of $(2.3RT)\log(k_{max})$ versus nucleobase N1 acidity in the gas phase ($\Delta G_{acid,g}$) and a plot of $\Delta G_{acid,g}$ versus$(2.3RT)$pK_a^{N1}; figures showing Brønsted-type LFERs for nonenzymatic pyrimidine hydrolyses at 75 and at 22 °C; and the complete reference for Frisch, M. J. et al., Gaussian, Inc.: Pittsburgh, PA, 2003 (ref 30). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0634829

(88) Kwon, K.; Jiang, Y. L.; Stivers, J. T. *Chem. Biol.* **2003**, *10*, 351−359.
(89) Kavli, B.; Slupphaug, G.; Mol, C. D.; Arvai, A. S.; Peterson, S. B.; Tainer, J. A.; Krokan, H. E. *Embo. J.* **1996**, *15*, 3442−3447.
(90) Jiang, Q.; Blount, B. C.; Ames, B. N. *J. Biol. Chem.* **2003**, *278*, 32834−32840.
(91) Henderson, J. P.; Byun, J.; Takeshita, J.; Heinecke, J. W. *J. Biol. Chem.* **2003**, *278*, 23522−23528.
(92) Yu, H.; Eritja, R.; Bloom, L.; Goodman, M. *J. Biol. Chem.* **1993**, *268*, 15935−15943.
(93) Morris, S. M. *Mutat. Res.* **1991**, *258*, 161−188.
(94) Morris, S. M. *Mutat. Res.* **1993**, *297*, 39−51.
(95) Ames, B. N.; Gold, L. S.; Willett, W. C. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5258−5265.
(96) Ohshima, H.; Tatemichi, M.; Sawa, T. *Arch. Biochem. Biophys.* **2003**, *417*, 3−11.
(97) Brandon, M. L.; Mi, L.-J.; Chaung, W.; Teebor, G.; Boorstein, R. J. *Mutat. Res.* **2000**, *459*, 161−169.
(98) Kubareva, E. A.; Volkov, E. M.; Vinogradova, N. L.; Kanevsky, I. A.; Oretskaya, T. S.; Kuznetsova, S. A.; Brevnov, M. G.; Gromova, E. S.; Nevinsky, G. A.; Shabarova, Z. A. *Gene.* **1995**, *157*, 167−171.
(99) Valinluck, V.; Liu, P.; Kang, J. I., Jr; Burdzy, A.; Sowers, L. C. *Nucleic Acids Res.* **2005**, *33*, 3057−3064.