Supporting Information

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[‡]*

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

E-mail:adroh001@umaryland.edu

[*] Corresponding author, Tel.: (410) 706-8118, E-mail: adroh001@umaryland.edu
[‡] Department of Biochemistry and Molecular Biology and Greenebaum Cancer Center, University of Maryland School of Medicine. [†] Department of Chemistry, Wayne State University
[§] Wadsworth Center, New York State Department of Health.

Activity of hTDG against normal G·C pairs. Some DNA glycosylases having broad specificity and a permissive active site exhibit weak activity against normal bases in DNA.[1,2] Although hTDG activity against G·C pairs has not previously been reported, we were able to detect weak activity for the removal of cytosine from G·C base pairs by hTDG, k_{max} = 1.2 x 10^{-5} min^{-1}, using high enzyme and substrate concentrations (Figure S1).

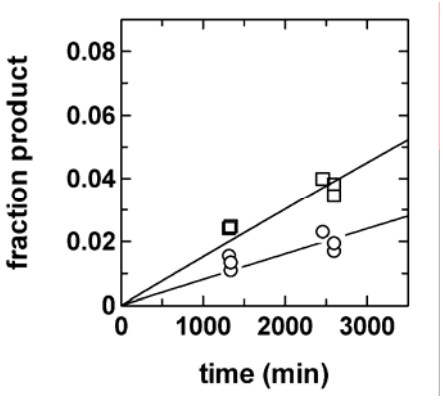

Figure S1. The removal of cytosine from a C·G pair by hTDG (10 μM) is exceedingly slow but measurable, giving an average rate constant of k_{max} = 1.2 x 10^{-5} min^{-1}. The activity against the G·C19 substrate (500 nM) involves the removal of cytosine from each of the two G·C pairs comprising the central CpG site, i.e. 5'-CpG/5'-CpG (Figure 4A). The removal of cytosine from the "target" strand occurs with a rate constant of k_{max} = 0.8 x 10^{-5} min^{-1} (○), and the rate constant

for excision of cytosine from the complementary strand is k_{max} = 1.5 x 10^{-5} min^{-1} (□), giving an average of k_{max} = 1.2 x 10^{-5} ± 0.5 min^{-1}. The G·C19 substrate (indeed all substrates examined here) contains only one CpG site, thus activity against other G·C base pairs, if it occurs, is too weak to detect.

Methods for the DNA thermal melting experiments. The duplex DNA samples consisted of 1 μM concentrations of the target and complementary strands in 0.01 M sodium phosphate pH 7.4, 0.1 M NaCl. For each melting temperature (T_m) determination, a 1.0 ml sample of duplex DNA was prepared from the single strand stocks of the constituent oligonucleotides, heated to 85 °C, and cooled slowly (> 3 hrs) to 20 ºC. The melting experiments were performed using tightly capped quartz cells with a 1 cm path-length and a Beckman DU 640 spectrophotometer equipped with a peltier device. Absorbance at 260 nm was monitored from 20 °C to 87 °C and in increments of 1 °C, with 0.5 °C increments in the transition region (T_m ± ~5 °C). Samples were held at each point for two minutes or until the absorbance remained constant. Samples were weighed before and after each experiment, and were found to differ by <0.1%. The absorption curves displayed sharp transitions, with 80% of the absorbance change occurring over a 14 ºC window. T_m values corresponding to the temperature of mid-transition (50% dissociated) were determined as previously described[3] and represent the average of at least three independent samples, with an uncertainty of ± 0.5 °C.

The stability of G·U and G·C base pairs is not significantly altered by the 5-halogen substitutions. We thought it important to determine whether the halogen substituents perturbed the stability of G·U and G·C pairs in the duplex substrates, because previous studies have indicated that if a substrate base is rendered more prone to flipping out of the DNA duplex, it may be more rapidly removed by a DNA glycosylase.[2,4] To address this, we determined the melting temperature (T_m) for the various DNA substrates. For the G·xU19 duplexes, the T_m values are the same within error: G·U19, 57.3 ºC; G·T19, 58.1 ºC; G·FU19, 57.3 ºC; G·ClU19, 57.7 ºC; G·BrU19, 57.4 ºC; and G·IU19, 57.5 ºC (data not shown). The G·xC duplexes also have very similar T_m values: G·C, 63.9 ºC; G·FC, 65.0 ºC; and G·BrC, 65.2 ºC. As expected, the T_m values for the G·xC duplexes are higher than the for G·xU duplexes, by about ~7 ºC. The large difference in T_m for the two nearly identical groups (G·5x<u>U</u> versus G·5x<u>C</u>) indicates that if base pair stability were significantly perturbed by the 5-halogen substitutions, we would see an effect on T_m, which we did not. Our results are consistent with previous findings that 15 bp duplexes with a central G·T, G·BrU, or G·IU pair exhibit nearly equivalent T_m values, and that G·BrC and G·IC duplexes have similar T_m values, albeit 4 ºC higher than a G·C duplex.[5] NMR studies have shown that heptamer duplexes containing hmU·G pairs form canonical B-type DNA, and that the hmU·G base pair adopts a structure that is essentially the same as a G·T mispair.[6] Similarly, the more limited spectroscopic data reported for DNA containing hoU·G pairs adopts non-distorted B-type DNA structure, and hoU forms H bond interactions with G or A partners.[6] Taken together, our observations and previous findings indicate that the changes we observed in hTDG activity (k_{max}) arising from the 5-substituents of U and C are not due to an increased propensity of these bases to flip out of the DNA duplex.

Table S1. Calculated Gas-Phase N1 Deprotonation Enthalpies and Acidities at 0 and 298 K (kcal/mol)[a]

Species	Theory[a]				
	D_e[b]	D_0[b,c]	$D_{0,BSSE}$[b,c,d]	$\Delta H_{acid,g}$[b,c,d]	$\Delta G_{acid,g}$[b,c,d]
uracil	334.0	331.8	329.4	330.7	324.6
5-methyluracil	341.7	333.4	331.0	332.3	326.1
5-hydroxyuracil	339.6	331.5	329.1	330.3	324.2
5-hydroxymethyluracil	335.2	327.2	324.8	326.0	320.0
5-hydroxymethyluracil[e]	339.1	331.0	328.2	329.5	323.2
5-fluorouracil	335.1	327.0	324.6	325.9	319.8
5-chlorouracil	333.0	325.0	322.5	323.8	317.7
5-bromouracil	332.4	324.3	321.9	323.2	317.0
5-iodouracil	332.4	324.4	321.9	323.2	317.1
cytosine	352.5	343.9	341.5	342.9	336.6
5-hydroxycytosine	352.2	343.5	341.1	342.1	336.2
5-fluorocytosine	347.8	339.4	337.0	338.4	332.17
5-bromocytosine	343.9	335.5	333.0	334.5	328.1

[a]MP2(full)/6-311+G(2d,2p)//MP2(full)/6-31G*. [b]Calculated enthalpies and free energies for N1 deprotonation. [c]Also includes ZPE corrections. [d]Also Includes BSSE corrections. [e]Values for excited conformation in which the hydroxymethyl substituent does not hydrogen bond to O4.

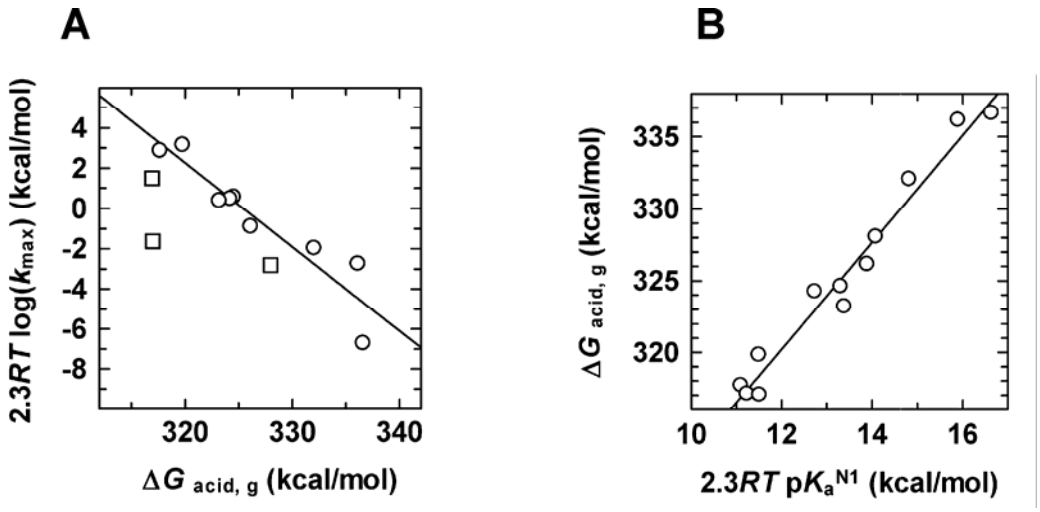

Figure S2. (**A**) LFER for the dependence of (2.3*RT*)log(k_{max}) on nucleobase N1-H acidity in the gas phase at 25 ºC ($\Delta G_{acid,g}$). The LFER includes data for U, T, FU, ClU, hoU, hmU, FC, hoC, and C (○), and has a slope of m = -0.42 ± 0.06 and a good correlation coefficient (r = 0.94). Data for BrU, IU, and BrC are shown (□) but not included in the LFER because the low k_{max} values obtained for these bases suggests limited access to the active site (see main text). (**B**) A plot of $\Delta G_{acid,g}$ versus (2.3*RT*)pK_a^{N1} is highly linear (r = 0.98), demonstrating that N1 acidity in aqueous solution is strongly correlated with the gas-phase acidity. The slope of m = 3.7 ± 0.2 indicates that the N1 acidity differences are nearly four-fold greater in the gas phase than aqueous solution.

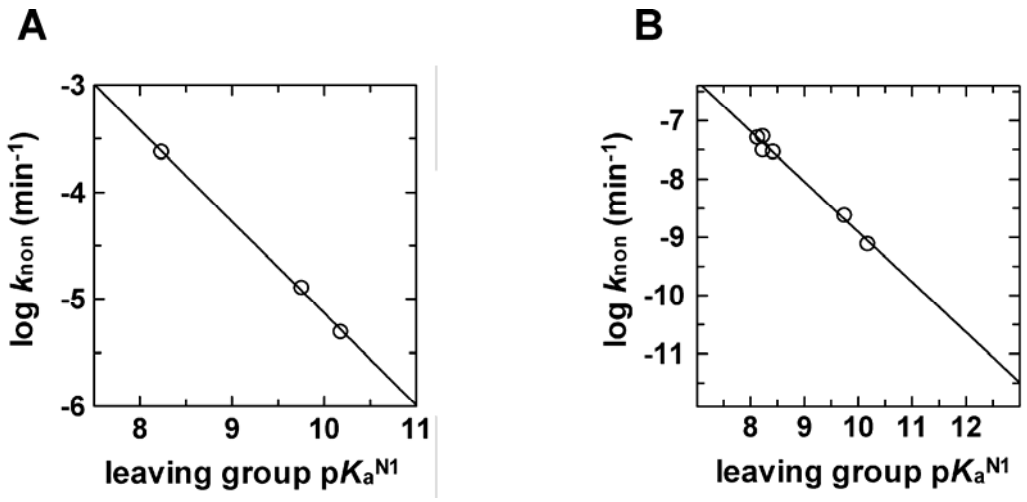

Figure S3. Brønsted-type LFERs for the non-enzymatic hydrolysis of 5-substituted 2'-deoxyuridines. (**A**) A slope of β_{lg} = -0.86 ± 0.03 (r = 1.00) is obtained for the hydrolysis of dU, dT, and 5-Br-dU at pH 6.5 and 75 ºC using the data reported by Shapiro and Kang,[7] but with pK_a^{N1} = 8.24 for BrU. Shapiro and Kang[7] used pK_a^{N1} = 8.49 for BrU, which is not correct (see footnote 75 of main text), giving a β_{lg} = -0.99 ± 0.01 (not shown). (B) A slope of β_{lg} = -0.86 ± 0.05 (r = 0.994) is obtained using k_{non} values extrapolated to 22 ºC for dU, dT, 5-F-dU, 5-Cl-dU, and 5-Br-dU (Table 1) and the pK_a^{N1} values for constituent bases from Table 2.[7,8]

Additional references. The complete reference for Frisch, M. J., et al., Gaussian, Inc.: Pittsburgh, PA, 2003. is given below.[9]

References

(1) Berdal, K. G.; Johansen, R. F.; Seeberg, E. *Embo. J.* **1998**, *17*, 363-367.

(2) O'Brien, P. J.; Ellenberger, T. *J. Biol. Chem.* **2004**, 9750-9757.

(3) Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515-537.

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

(5) Ferrer, E.; Wiersma, M.; Kazimierczak, B.; Muller, C. W.; Eritja, R. *Bioconjugate Chemistry* **1997**, *8*, 757-761.

(6) Lukin, M.; de Los Santos, C. *Chem Rev* **2006**, *106*, 607-686.

(7) Shapiro, R.; Kang, S. *Biochemistry* **1969**, *8*, 1806-1810.

(8) Van Schepdael, A.; Ossembe, N.; Herdewijn, P.; Roets, E.; Hoogmartens, J. *J. Pharm. Biomed. Anal.* **1993**, *11*, 345-351.

(9) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.;

Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, K. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; Gaussian, Inc.: Pittsburgh, PA, 2003.