JOC Article

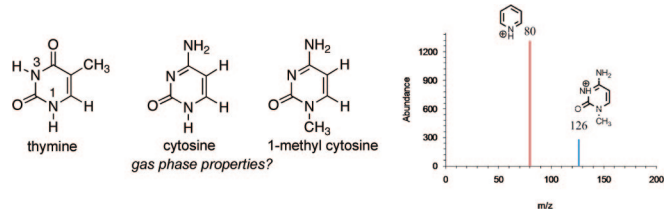
Gas-Phase Thermochemical Properties of Pyrimidine Nucleobases

Min Liu, Tingting Li, F. Sedinam Amegayibor, Daisy S. Cardoso, Yunlin Fu, and Jeehiun K. Lee*

Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854

jeehiun@rci.rutgers.edu

Received August 14, 2008



The gas-phase acidity and proton affinity of thymine, cytosine, and 1-methyl cytosine have been examined using both theoretical (B3LYP/6-31+G*) and experimental (bracketing, Cooks kinetic) methods. This paper represents a comprehensive examination of multiple acidic sites of thymine and cytosine and of the acidity and proton affinity of thymine, cytosine, and 1-methyl cytosine. Thymine exists as the most stable "canonical" tautomer in the gas phase, with a ΔH_{acid} of 335 ± 4 kcal mol^{-1} (ΔG_{acid} = 328 ± 4 kcal mol^{-1}) for the more acidic N1−H. The acidity of the less acidic N3−H site has not, heretofore, been measured; we bracket a ΔH_{acid} value of 346 ± 3 kcal mol^{-1} (ΔG_{acid} = 339 ± 3 kcal mol^{-1}). The proton affinity (PA = ΔH) of thymine is measured to be 211 ± 3 kcal mol^{-1} (GB = ΔG = 203 ± 3 kcal mol^{-1}). Cytosine is known to have several stable tautomers in the gas phase in contrast to in solution, where the canonical tautomer predominates. Using bracketing methods in an FTMS, we measure a ΔH_{acid} for the more acidic site of 342 ± 3 kcal mol^{-1} (ΔG_{acid} = 335 ± 3 kcal mol^{-1}). The ΔH_{acid} of the less acidic site, previously unknown, is 352 ± 4 kcal mol^{-1} (345 ± 4 kcal mol^{-1}). The proton affinity is 228 ± 3 kcal mol^{-1} (GB = 220 ± 3 kcal mol^{-1}). Comparison of these values to calculations indicates that we most likely have a mixture of the canonical tautomer and two enol tautomers and possibly an imine tautomer under our conditions in the gas phase. We also measure the acidity and proton affinity of cytosine using the extended Cooks kinetic method. We form the proton-bound dimers via electrospray of an aqueous solution, which favors cytosine in the canonical form. The acidity of cytosine using this method is ΔH_{acid} = 343 ± 3 kcal mol^{-1}, PA = 227 ± 3 kcal mol^{-1}. We also examined 1-methyl cytosine, which has fewer accessible tautomers than cytosine. We measure a ΔH_{acid} of 349 ± 3 kcal mol^{-1} (ΔG_{acid} = 342 ± 3 kcal mol^{-1}) and a PA of 230 ± 3 kcal mol^{-1} (GB = 223 ± 3 kcal mol^{-1}). Our ultimate goal is to understand the intrinsic reactivity of nucleobases; gas-phase acidic and basic properties are of interest for chemical reasons and also possibly for biological purposes because biological media can be quite nonpolar.

Introduction

The intrinsic gas-phase acidic and basic properties of nucleobases are of interest for purely chemical reasons but also could be of importance for biological reasons because biological environs can be relatively nonpolar in nature.[1,2] Hydrogen bonding modulates recognition of DNA and RNA bases, and the interaction energy between two complementary nucleobases that are held together by NH−O and NH−N hydrogen bonds is dependent on the intrinsic basicity of the acceptor atoms as well as on the acidity of the NH donor groups.[3,4]

Knowledge of and comparison of gas-phase acidities and proton affinities to solution values will yield valuable information on intrinsic nucleobase reactivity and the role of solvent

(1) Simonson, T.; Brooks, C. L. *J. Am. Chem. Soc.* **1996**, *118*, 8452–8458.
(2) Jordan, F.; Li, H.; Brown, A. *Biochemistry* **1999**, *38*, 6369–6373.
(3) Nguyen, M. T.; Chandra, A. K.; Zeegers-Huyskens, T. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 1277–1280.
(4) Chandra, A. K.; Nguyen, M. T.; Uchimaru, T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **1999**, *103*, 8853–8860.

10.1021/jo801822s CCC: $40.75
Published on Web 10/31/2008

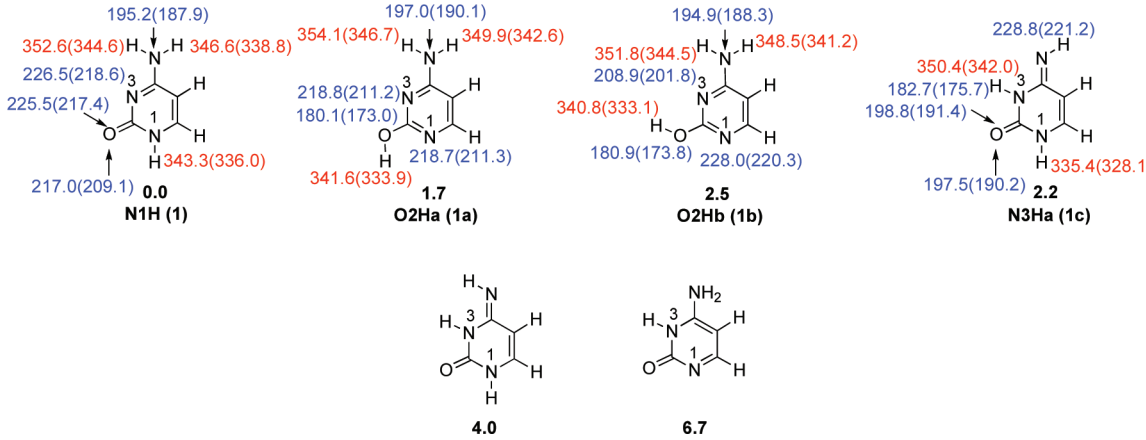


FIGURE 1. Relative enthalpies (ΔH in kcal mol^{-1}) of six possible tautomers of cytosine and the acidities (red values: ΔH_{acid} with ΔG_{acid} in parentheses, all values in kcal mol^{-1}) and proton affinities (blue values: PA with GB in parentheses, all values in kcal mol^{-1}) of the four most stable tautomers calculated at B3LYP/6-31+G* (298 K).

in affecting base reactivity.[5−13] In essence, gas-phase experiments can provide the link between calculations and condensed-phase data.

In previous work, we have reported the gas-phase thermochemical properties of uracil and adenine as well as several damaged purine nucleobases.[5−13] Damaged DNA bases differ in structure and properties from normal nucleobases and, therefore, intervene with gene replication and expression, leading to cell death, aging, and carcinogenesis.[14−17] Our studies are motivated by understanding the mechanisms by which mutated bases are cleaved, focusing on the glycosylase enzymes that excise damaged bases.[5,6,8,9,11,13,18] The first step toward understanding how normal bases differ from damaged bases is to characterize the naturally occurring normal compounds. The pyrimidine nucleobases have previously been the subject of some study both theoretically and experimentally; to our knowledge, there have been two experimental studies of the gas-phase proton affinity of the most basic site and one on the most acidic site of cytosine and thymine.[4,19−31] Herein, we provide a comprehensive examination of the gas-phase thermochemical properties of the naturally occurring pyrimidine bases, thymine and cytosine, as well as of the 1-methyl derivative of cytosine. We measure *multiple* acidities of more acidic and less acidic sites not heretofore accomplished as well as the proton affinity of thymine and cytosine. We also provide new data for 1-methyl cytosine, which unlike cytosine, does not have several possible stable tautomeric states in the gas phase.

Results

Cytosine. (i) Calculations: Cytosine Tautomers. Cytosine, as with all nucleobases, has several possible tautomeric forms (six most stable in Figure 1; see Supporting Information for higher-energy tautomers).[4,24−26,28−41] At B3LYP/6-31+G*, we find that the canonical tautomer, N1H (**1**), is the most stable, but that three other tautomers (O2H enols, **1a** and **1b**, and N3H imine, **1c**) are very close in energy to the canonical form. The next most stable tautomers are 4.0 and 6.7 kcal mol^{-1} less stable than the canonical; the remaining tautomers are all quite higher in energy.

(5) Kurinovich, M. A.; Lee, J. K. *J. Am. Chem. Soc.* **2000**, *122*, 6258–6262.
(6) Kurinovich, M. A.; Lee, J. K. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 985–995.
(7) Kurinovich, M. A.; Phillips, L. M.; Sharma, S.; Lee, J. K. *Chem. Commun. (Cambridge, U.K.)* **2002**, 2354–2355.
(8) Lee, J. K. *Int. J. Mass Spectrom.* **2005**, *240*, 261–272.
(9) Liu, M.; Xu, M.; Lee, J. K. *J. Org. Chem.* **2008**, *73*, 5907–5914.
(10) Pan, S.; Sun, X.; Lee, J. K. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1383–1395.
(11) Sharma, S.; Lee, J. K. *J. Org. Chem.* **2002**, *67*, 8360–8365.
(12) Sharma, S.; Lee, J. K. *J. Org. Chem.* **2004**, *69*, 7018–7025.
(13) Sun, X.; Lee, J. K. *J. Org. Chem.* **2007**, *72*, 6548–6555.
(14) O'Brien, P. J.; Ellenberger, T. *J. Biol. Chem.* **2004**, *279*, 9750–9757.
(15) Speina, E.; Kierzek, A. M.; Tudek, B. *Mutat. Res.* **2003**, *531*, 205–217.
(16) Stivers, J. T.; Jiang, Y. L. *Chem. Rev.* **2003**, *103*, 2729–2759.
(17) Berti, P. J.; McCann, J. A. B. *Chem. Rev.* **2006**, *106*, 506–555, and references therein.
(18) Bennett, M. T.; Rodgers, M. T.; Hebert, A. S.; Ruslander, L. E.; Eisele, L.; Drohat, A. C. *J. Am. Chem. Soc.* **2006**, *128*, 12510–12519.
(19) Meot-Ner (Mautner) , M. *J. Am. Chem. Soc.* **1979**, *101*, 2396–2403.
(20) Greco, F.; Liguori, A.; Sindona, G.; Uccella, N. *J. Am. Chem. Soc.* **1990**, *112*, 9092–9096.
(21) Chen, E. C. M.; Chen, E. S. *J. Phys. Chem. B* **2000**, *104*, 7835–7844.
(22) Chen, E. C. M.; Herder, C.; Chen, E. S. *J. Mol. Struct.* **2006**, *798*, 126–133, and references therein.
(23) *NIST Chemistry WebBook* [Online]; Linstrom, P. J., Mallard, W. G., Eds.; NIST Standard Reference Database Number 69; National Institute of Standards and Technology: Gaithersburg, MD, 2005; http://webbook.nist.gov.
(24) Huang, Y.; Kenttamaa, H. *J. Phys. Chem. A* **2003**, *107*, 4893–4897, and references therein.
(25) Podolyan, Y.; Gorb, L.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 7346–7352, and references therein.
(26) Yang, Z.; Rodgers, M. T. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2749–2757, and references therein.
(27) Yao, C.; Cuadrado-Peinado, M. L.; Polášek, M.; Turecek, F. *J. Mass Spectrom.* **2005**, *40*, 1417–1428, and references therein.
(28) Yao, C.; Turecek, F.; Polce, M. J.; Wesdemiotis, C. *Int. J. Mass Spectrom.* **2007**, *265*, 106–123.
(29) Turecek, F.; Yao, C. *J. Phys. Chem. A* **2003**, *107*, 9221–9231.
(30) Kobayashi, R. *J. Phys. Chem. A* **1998**, *102*, 10813–10817.
(31) Wolken, J. K.; Yao, C.; Turecek, F.; Polce, M. J.; Wesdemiotis, C. *Int. J. Mass Spectrom.* **2007**, *267*, 30–42.
(32) Trygubenko, S. A.; Bogdan, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Sponer, J.; Slavicek, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4192–4203.
(33) Barker, D. L.; Marsh, R. E. *Acta Crystallogr.* **1964**, *17*, 1581–1587.
(34) McClure, R. J.; Craven, B. M. *Acta. Crystallogr.* **1973**, *B29*, 1234–1238.
(35) Weber, H. P.; Craven, B. M.; McMullan, R. K. *Acta. Crystallogr.* **1980**, *B36*, 645–649.
(36) Ueda, T.; Fox, J. J. *J. Am. Chem. Soc.* **1963**, *85*, 4024–4028.
(37) Dreyfus, M.; Bensaude, O.; Dodin, G.; Dubois, J. E. *J. Am. Chem. Soc.* **1976**, *98*, 6338–6349.
(38) Sambrano, J. R.; de Souza, A. R.; Queralt, J. J.; Andrés, J. *Chem. Phys. Lett.* **2000**, *317*, 437–443.
(39) Nir, E.; Müller, M.; Grace, L. I.; de Vries, M. S. *Chem. Phys. Lett.* **2002**, *355*, 59–64.
(40) Szczesniak, M.; Szczepaniak, K.; Kwiatkowski, J. S.; KuBulat, K.; Person, W. B. *J. Am. Chem. Soc.* **1988**, *110*, 8319–8330.
(41) Chandra, A. K.; Michalska, D.; Wysokinsky, R.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2004**, *108*, 9593–9600, and references therein.

TABLE 1. Summary of Results for Acidity Bracketing of the More Acidic Site of Cytosine

reference compound	ΔH_{acid}[a] (kcal mol^{-1})	ΔG_{acid}[a] (kcal mol^{-1})	proton transfer[b] ref. acid	proton transfer[b] conj. base
acetic acid	348.1 ± 2.2	341.1 ± 2.0	−	+
formic acid	345.3 ± 2.2	338.3 ± 2.0	−	+
2,4-pentadione	343.8 ± 2.1	336.7 ± 2.0	−	+
ethoxyacetic acid	342.0 ± 2.2	335.0 ± 2.0	+	+
3-chloropropanoic acid	340.8 ± 2.7	333.8 ± 2.0	+	−
trifluoro-*m*-cresol	339.3 ± 2.1	332.4 ± 2.0	+	−
methyl cyanoacetate	340.8 ± 0.6	334.5	+	−
2-chloropropanoic acid	337.0 ± 2.1	330.4 ± 2.0	+	−
difluoroacetic acid	331.0 ± 2.2	323.8 ± 2.0	+	−
1,1,1-trifluoro-2,4-pentadione	328.3 ± 2.9	322.0 ± 2.0	+	−

[a] Acidities are in kcal mol^{-1}.[23] [b] "+" indicates the occurrence of proton transfer, and "−" indicates the absence of proton transfer.

(ii) Calculations: Cytosine Acidity. The calculated values for the acidity of the four most stable tautomers of cytosine are shown in red in Figure 1.[42] The ΔH_{acid} of the most acidic site for the canonical and enol tautomers appears to be around 341−343 kcal mol^{-1} (ΔG_{acid} = 333−336 kcal mol^{-1}), whereas for the imine tautomer, **1c**, the most acidic site is calculated to be closer to $\Delta H_{acid} \sim$ 335 kcal mol^{-1} ($\Delta G_{acid} \sim$ 328 kcal mol^{-1}).

(iii) Calculations: Cytosine Proton Affinity. The calculated values for the proton affinity (PA, ΔH) and gas-phase basicity (GB, ΔG) of the four most stable tautomers of cytosine are shown in blue in Figure 1. The PA/GB of the most basic site of cytosine varies by tautomer. For the canonical tautomer, there are two sites that are close in basicity: O2 (PA = 225.5 kcal mol^{-1}; GB = 217.4 kcal mol^{-1}) and N3 (PA = 226.5 kcal mol^{-1}; GB = 218.6 kcal mol^{-1}). The most basic sites for the enol **1b** and the imine **1c** have PAs similar to that of the canonical (**1b**, PA = 228.0 kcal mol^{-1}, GB = 220.3 kcal mol^{-1}; and **1c**, PA = 228.8 kcal mol^{-1}, GB = 221.2 kcal mol^{-1}). In contrast, the most basic site of the enol tautomer, **1a**, has a calculated PA of 219 kcal mol^{-1} (GB = 211 kcal mol^{-1}) corresponding to either the N1 or the N3 site.

(iv) Experiments: Cytosine Acidity. The bracketing results for the more acidic site of cytosine are summarized in Table 1[23] (we can measure multiple acidic sites on a molecule; see Experimental Section). "Ref. acid" refers to the reaction of deprotonated cytosine with the neutral reference acid. "Conj. base" refers to the reaction of the conjugate base of the reference acid with neutral cytosine. We find that deprotonated cytosine can deprotonate neutral ethoxyacetic acid (ΔH_{acid} = 342.0 kcal mol^{-1}; ΔG_{acid} = 335.0 kcal mol^{-1}) and more acidic compounds but cannot deprotonate neutral 2,4-pentadione (ΔH_{acid} = 343.8 kcal mol^{-1}; ΔG_{acid} = 336.7 kcal mol^{-1}) or less acidic compounds. In the opposite direction, ethoxyacetate deprotonates neutral cytosine as do more basic bases. Less basic bases do not deprotonate neutral cytosine. We are, therefore, able to bracket the acidity of the more acidic site of cytosine as ΔH_{acid} = 342 ± 3 kcal mol^{-1} (ΔG_{acid} = 335 ± 3 kcal mol^{-1}).[5−9,11−13]

Using methodology developed in our laboratory, we also measured the less acidic site of cytosine (Table 2).[5−9,11−13] We find that deprotonated cytosine under "less acidic" conditions (Experimental Section) cannot deprotonate acids as acidic as and less acidic than 4-(trifluoromethyl)aniline (ΔH_{acid} = 353.3 kcal mol^{-1}; ΔG_{acid} = 346.0 kcal mol^{-1}) but can deprotonate acids as acidic as and more acidic than *p*-cresol (ΔH_{acid} = 350.3 kcal mol^{-1}; ΔG_{acid} = 343.4 kcal mol^{-1}). We therefore bracket the less acidic site of cytosine as ΔH_{acid} = 352 ± 4 kcal mol^{-1} (ΔG_{acid} = 345 ± 4 kcal mol^{-1}).

(42) If more than one value is listed for an atom, the arrows show the site of protonation (for example, the O2 of cytosine can be protonated on the N1 side and on the N3 side).

TABLE 2. Summary of Results for Acidity Bracketing of Less Acidic Site of Cytosine

reference compound	ΔH_{acid}[a] (kcal mol^{-1})	ΔG_{acid}[a] (kcal mol^{-1})	proton transfer[b]
pyrrole	359.6 ± 2.9	351.8 ± 2.0	−
chloroacetonitrile	357.7 ± 2.2	350.0 ± 2.0	−
1-propane thiol	354.2 ± 2.2	347.9 ± 2.0	−
2-propane thiol	353.4 ± 2.2	347.1 ± 2.0	−
4-(trifluoromethyl)aniline	353.3 ± 2.1	346.0 ± 2.0	−
p-cresol	350.3 ± 2.1	343.4 ± 2.0	+
m-cresol	349.6 ± 2.1	342.7 ± 2.0	+
acetic acid	348.1 ± 2.2	341.1 ± 2.0	+
formic acid	345.3 ± 2.2	338.3 ± 2.0	+
2,4-pentadione	343.8 ± 2.1	336.7 ± 2.0	+
methyl cyanoacetate	340.8 ± 0.6	334.5	+

[a] Acidities are in kcal mol^{-1}.[23] [b] "+" indicates the occurrence of proton transfer, and "−" indicates the absence of proton transfer.

TABLE 3. Summary of Results for PA Bracketing of the More Basic Site of Cytosine

reference compound	PA[a] (kcal mol^{-1})	GB[a] (kcal mol^{-1})	proton transfer[b] ref. base	proton transfer[b] conj. acid
1-methyl piperidine	232.1 ± 2.0	224.7 ± 2.0	+	−
1-methyl pyrrolidine	230.8 ± 2.0	223.4 ± 2.0	+	−
piperidine	228.0 ± 2.0	220.0 ± 2.0	+	+
pyrrolidine	226.6 ± 2.0	218.8 ± 2.0	−	+
3-picoline	225.5 ± 2.0	217.9 ± 2.0	−	+
pyridine	222.3 ± 2.0	214.7 ± 2.0	−	+
N-methyl propanamide	220.0 ± 2.0	212.6 ± 2.0	−	+

[a] Values are in kcal mol^{-1}.[23] [b] "+" indicates the occurrence of proton transfer, and "−" indicates the absence of proton transfer.

(v) Experiments: Cytosine Proton Affinity. The bracketing results for the proton affinity of the most basic site of cytosine are shown in Table 3. "Ref. base" refers to the reaction of the neutral reference base with protonated cytosine. "Conj. acid" refers to the reaction of the protonated reference base with neutral cytosine. We find that the reaction between cytosine and piperidine (PA = 228.0 kcal mol^{-1}; GB = 220.0 kcal mol^{-1}) proceeds in both directions (cytosine deprotonates protonated piperidine and piperidine deprotonates protonated cytosine); however, with 1-methyl pyrrolidine (PA = 230.8 kcal mol^{-1}; GB = 223.4 kcal mol^{-1}) the reaction with protonated cytosine does occur but the opposite reaction does not. With pyrrolidine (PA = 226.6 kcal mol^{-1}; GB = 218.8 kcal mol^{-1}), cytosine can deprotonate protonated pyrrolidine, but the opposite reaction does not proceed. These data point to a PA for cytosine as 228 ± 3 kcal mol^{-1} (ΔG = gas-phase basicity, GB = 220 ± 3 kcal mol^{-1}).

1-Methyl Cytosine. (i) Calculations: 1-Methyl Cytosine Tautomers. As with cytosine, 1-methyl cytosine has more than one tautomer that is low-lying in energy, but fewer than with cytosine (five most stable tautomers in Figure 2, all optimized tautomers in Supporting Information).[27,43] The canonical tautomer is calculated to be 2.9 kcal mol^{-1}, which is more stable than the next nearest (imine) tautomer; the other imine tautomer lies 4.6 kcal mol^{-1} higher in energy than the canonical structure.

(43) Szczesniak, M.; Leszczynski, J.; Person, W. B. *J. Am. Chem. Soc.* **1992**, *114*, 2731–2733.

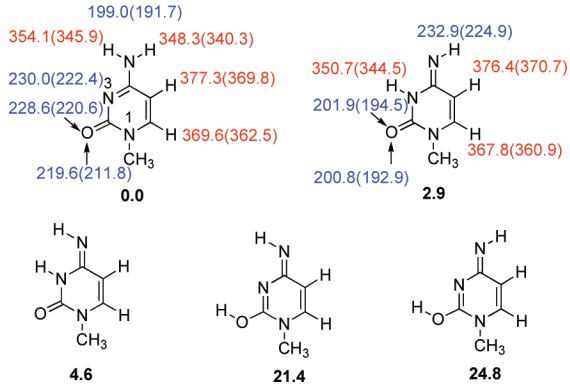


FIGURE 2. Relative enthalpies (ΔH in kcal mol^{-1}) of five possible tautomers of 1-methyl cytosine and acidities (red values: ΔH_{acid} with ΔG_{acid} in parentheses, all values in kcal mol^{-1}) and proton affinities (blue values: PA with GB in parentheses, all values in kcal mol^{-1}) of the two most stable tautomers calculated at B3LYP/6-31+G* (298 K).

TABLE 4. Summary of Results for Acidity Bracketing of the More Acidic Site of 1-Methyl Cytosine

reference compound	ΔH_{acid}[a] (kcal mol^{-1})	ΔG_{acid}[a] (kcal mol^{-1})	proton transfer[b] ref. acid	proton transfer[b] conj. base
acetone	369.1 ± 2.1	361.9 ± 2.0	−	+
pyrrole	359.6 ± 2.9	351.8 ± 2.0	−	+
4-(trifluoromethyl)aniline	353.3 ± 2.1	346.0 ± 2.0	−	+
m-cresol	349.6 ± 2.1	342.7 ± 2.0	−	+
acetic acid	348.1 ± 2.2	341.1 ± 2.0	+	−
butanoic acid	346.5 ± 2.2	339.5 ± 2.0	+	−
formic acid	345.3 ± 2.2	338.3 ± 2.0	+	−
2,4-pentadione	343.8 ± 2.1	336.7 ± 2.0	+	−
trifluoro-*m*-cresol	339.3 ± 2.1	332.4 ± 2.0	+	−

[a] Acidities are in kcal mol^{-1}.[23] [b] "+" indicates the occurrence of proton transfer, and "−" indicates the absence of proton transfer.

(ii) Calculations: 1-Methyl Cytosine Acidity. The acidities of the two lowest-lying tautomers of 1-methyl cytosine are shown in Figure 2. The acidity of both tautomers is similar (canonical, ΔH_{acid} = 348.3 kcal mol^{-1}, ΔG_{acid} = 340.3 kcal mol^{-1}; higher-energy tautomer, ΔH_{acid} = 350.7 kcal mol^{-1}, ΔG_{acid} = 344.5 kcal mol^{-1}).

(iii) Calculations: 1-Methyl Cytosine Proton Affinity. The proton affinities of 1-methyl cytosine are also in Figure 2.[27] There are two sites on the canonical tautomer that are comparable in PA: O2 (PA = 228.6 kcal mol^{-1}; GB = 220.6 kcal mol^{-1}) and N3 (PA = 230.0 kcal mol^{-1}; GB = 222.4 kcal mol^{-1}). The imine tautomer is slightly more basic; the calculated PA value is 232.9 kcal mol^{-1} (GB = 224.9 kcal mol^{-1}).

(iv) Experiments: 1-Methyl Cytosine Acidity. The acidity bracketing results for 1-methyl cytosine are shown in Table 4. Deprotonated 1-methyl cytosine can deprotonate acetic acid (ΔH_{acid} = 348.1 kcal mol^{-1}; ΔG_{acid} = 341.1 kcal mol^{-1}) and more acidic acids but cannot deprotonate *m*-cresol (ΔH_{acid} = 349.6 kcal mol^{-1}; ΔG_{acid} = 342.7 kcal mol^{-1}) or acids that are less acidic. Likewise, *m*-cresolate can deprotonate 1-methyl cytosine but acetate cannot. We, therefore, bracket a ΔH_{acid} of 1-methyl cytosine to be 349 ± 3 kcal mol^{-1} (ΔG_{acid} = 342 ± 3 kcal mol^{-1}).

(v) Experiments: Proton Affinity of 1-Methyl Cytosine. The proton affinity bracketing results for 1-methyl cytosine are summarized in Table 5. 1-Methyl pyrrolidine (PA = 230.8 kcal mol^{-1}; GB = 223.4 kcal mol^{-1}) is basic enough to deprotonate protonated 1-methyl cytosine, but piperidine (PA = 228.0 kcal mol^{-1}; GB = 220.0 kcal mol^{-1}) is not. In the opposite direction, 1-methyl cytosine can deprotonate protonated piperidine but cannot deprotonate protonated 1-methyl pyrrolidine. The reaction with 2,4-lutidine (PA = 230.1 kcal mol^{-1}; GB = 222.5 kcal mol^{-1}) proceeds in both directions. We therefore bracket the PA of 1-methyl cytosine to be 230 ± 3 kcal mol^{-1} (GB = 223 ± 3 kcal mol^{-1}).

TABLE 5. Summary of Results for PA Bracketing of the More Basic Site of 1-Methyl Cytosine

reference compound	PA[a] (kcal mol^{-1})	GB[a] (kcal mol^{-1})	proton transfer[b] ref. base	proton transfer[b] conj. acid
triethylamine	234.7 ± 2.0	227.0 ± 2.0	+	−
1-methyl pyrrolidine	230.8 ± 2.0	223.4 ± 2.0	+	−
2,4-lutidine	230.1 ± 2.0	222.5 ± 2.0	+	+
piperidine	228.0 ± 2.0	220.0 ± 2.0	−	+
4-picoline	226.4 ± 2.0	218.8 ± 2.0	−	+
pyridine	222.0 ± 2.0	214.7 ± 2.0	−	+
2,4-pentadione	208.8 ± 2.0	200.0 ± 2.0	−	+

[a] Values are in kcal mol^{-1}.[23] [b] "+" indicates the occurrence of proton transfer, and "−" indicates the absence of proton transfer.

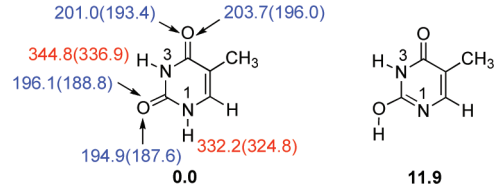


FIGURE 3. Relative enthalpies (ΔH in kcal mol^{-1}) of the two most stable thymine tautomers and acidities (red values: ΔH_{acid} with ΔG_{acid} in parentheses, all values in kcal mol^{-1}) and proton affinities (blue values: PA with GB in parentheses, all values in kcal mol^{-1}) of the canonical thymine tautomer calculated at B3LYP/6-31+G* (298 K).

Thymine. (i) Calculations: Thymine Tautomers. The canonical structure of thymine is the most stable by far; it is calculated to be about 12 kcal mol^{-1} more stable than the next nearest tautomer (Figure 3, higher-energy tautomers in Supporting Information).

(ii) Calculations: Thymine Acidity. The more acidic site of thymine has a calculated ΔH_{acid} of 332.2 kcal mol^{-1} (ΔG_{acid} = 324.8 kcal mol^{-1}) at N1 (Figure 3, in red).

(iii) Calculations: Thymine Proton Affinity. The more basic site of thymine has a calculated PA of 203.7 kcal mol^{-1} (GB = 196.0 kcal mol^{-1}) at O4 (Figure 3, in blue).

(iv) Experiments: Thymine Acidity. The acidity bracketing results for the more acidic site of thymine are shown in Table 6. We find that deprotonated thymine cannot deprotonate 2-bromopropionic acid (ΔH_{acid} = 336.8 kcal mol^{-1}; ΔG_{acid} = 329.8 kcal mol^{-1}) but that the opposite reaction does occur. Deprotonated thymine does react with pyruvic acid (ΔH_{acid} = 333.5 kcal mol^{-1}; ΔG_{acid} = 326.5 kcal mol^{-1}), but pyruvate does not deprotonate thymine. We, therefore, bracket the acidity of thymine to be ΔH_{acid} = 335 ± 4 kcal mol^{-1} (ΔG_{acid} = 328 ± 4 kcal mol^{-1}).

We have also bracketed the less acidic site of thymine (Table 7).[23] Deprotonated thymine under less acidic conditions does not deprotonate butyric acid (ΔH_{acid} = 346.5 kcal mol^{-1}; ΔG_{acid} = 339.5 kcal mol^{-1}) but does deprotonate isovaleric acid (ΔH_{acid} = 345.5 kcal mol^{-1}; ΔG_{acid} = 338.5 kcal mol^{-1}). We, therefore, bracket the less acidic site of thymine to be ΔH_{acid} = 346 ± 3 kcal mol^{-1} (ΔG_{acid} = 339 ± 3 kcal mol^{-1}).

TABLE 6. Summary of Results for Acidity Bracketing of the More Acidic Site of Thymine

reference compound	ΔH_{acid}[a] (kcal mol^{-1})	ΔG_{acid}[a] (kcal mol^{-1})	proton transfer[b] ref. acid	proton transfer[b] conj. base
2,4 pentadione	343.8 ± 2.1	336.7 ± 2.0	−	+
trifluoro-*m*-cresol	339.3 ± 2.1	332.4 ± 2.0	−	+
2-bromopropionic acid	336.8 ± 2.1	329.8 ± 2.0	−	+
pyruvic acid	333.5 ± 2.9	326.5 ± 2.8	+	−
per-fluoro-*tert*-butanol	331.6 ± 2.2	324.0 ± 2.0	+	−
difluoroacetic acid	331.0 ± 2.2	323.8 ± 2.0	+	−
1,1,1-trifluoro-2,4-pentadione	328.3 ± 2.9	322.0 ± 2.0	+	−

[a] Acidities are in kcal mol^{-1}.[23] [b] "+" indicates the occurrence of proton transfer, and "−" indicates the absence of proton transfer.

TABLE 7. Summary of Results for Acidity Bracketing of the Less Acidic Site of Thymine

reference compound	ΔH_{acid}[a] (kcal mol^{-1})	ΔG_{acid}[a] (kcal mol^{-1})	proton transfer[b]
pentane thiol	352.5 ± 2.3	346.2 ± 2.5	−
m-cresol	349.6 ± 2.1	342.7 ± 2.0	−
acetic acid	348.1 ± 2.2	341.1 ± 2.0	−
butyric acid	346.5 ± 2.2	339.5 ± 2.0	−
isovaleric acid	345.5 ± 2.1	338.5 ± 2.0	+
formic acid	345.3 ± 2.2	338.3 ± 2.0	+
2,4-pentadione	343.8 ± 2.1	336.7 ± 2.0	+
methyl cyanoacetate	340.8 ± 0.6	334.5	+
trifluoro-*m*-cresol	339.3 ± 2.1	332.4 ± 2.0	+

[a] Acidities are in kcal mol^{-1}.[23] [b] "+" indicates the occurrence of proton transfer, and "−" indicates the absence of proton transfer.

TABLE 8. Summary of Results for PA Bracketing of the More Basic Site of Thymine

reference compound	PA[a] (kcal mol^{-1})	GB[a] (kcal mol^{-1})	proton transfer[b] ref. base	proton transfer[b] conj. acid
2-chloropyridine	215.3 ± 2.0	208.0 ± 2.0	+	−
o-toluidine	212.9 ± 2.0	205.3 ± 2.0	+	−
pyrimidine	211.7 ± 2.0	204.5 ± 2.0	+	−
pyrrole	209.2 ± 2.0	201.7 ± 2.0	−	+
2,4-pentadione	208.8 ± 2.0	200.0 ± 2.0	−	+
m-chloro-aniline	207.5 ± 2.0	199.9 ± 2.0	−	+
methyl styrene	206.0 ± 2.0	199.0 ± 2.0	−	+
diethyl sulfide	204.8 ± 2.0	197.7 ± 2.0	−	+
4-methyl-cyclohexanone	201.9 ± 2.0	194.3 ± 2.0	−	+
cyclohexanone	201.0 ± 2.0	193.9 ± 2.0	−	+
2-butanone	197.7 ± 2.0	190.1 ± 2.0	−	+
acetone	194.0 ± 2.0	186.9 ± 2.0	−	+

[a] Values are in kcal mol^{-1}.[23] [b] "+" indicates the occurrence of proton transfer, and "−" indicates the absence of proton transfer.

(v) Experiments: Thymine Proton Affinity. The proton affinity results for thymine are summarized in Table 8. We find that pyrimidine (PA = 211.7 kcal mol^{-1}; GB = 204.5 kcal mol^{-1}) does deprotonate protonated thymine but the opposite reaction does not occur. Pyrrole, however (PA = 209.2 kcal mol^{-1}; GB = 201.7 kcal mol^{-1}), is not basic enough to deprotonate protonated thymine; thymine does deprotonate protonated pyrrole. The PA of thymine is, therefore, bracketed to be 211 ± 3 kcal mol^{-1} (GB = 203 ± 3 kcal mol^{-1}).

Discussion

Cytosine. We have measured the most acidic site of cytosine to have a ΔH_{acid} = 342 ± 3 kcal mol^{-1} (ΔG_{acid} = 335 ± 3 kcal mol^{-1}). This value is in agreement with that previously measured by Chen and Chen (340 ± 2 kcal mol^{-1}) using a variety of methods: derivation from DMSO pK_a values, electron impact, and negative chemical ionization mass spectrometry.[21,22] We have also measured the less acidic site to have a ΔH_{acid} = 352 ± 4 kcal mol^{-1}. The proton affinity is bracketed to be 228 ± 3 kcal mol^{-1}. This PA value is also in agreement with previous measurements (227.0 kcal mol^{-1}).[19,20,23,44,45]

In aqueous solution and the solid state, the canonical tautomer of cytosine (**1**) is the predominant form.[32−38] One of the issues with the exploration of cytosine in the gas phase, however, is what tautomer(s) might be present. Our calculations (Figure 1) and others indicate that there are four tautomers (**1**, **1a**, **1b**, and **1c**, Figure 1) that are close in stability.[4,24−26,30−32,41] Turecek, Wesdemiotis, and co-workers have done a superb job of assessing the energy orderings of cytosine versus calculational level and conclude that while the relative energies are very sensitive to the type of basis set used (with B3LYP calculations, for example, consistently favoring **1**), it is clear that **1**, **1a**, and **1b** are all very close in energy and are likely to coexist in the gas phase with possibly a small amount of **1c**.[31] Experiments in the gas phase are consistent with calculations insofar as mixtures of tautomers are found. Resonance-enhanced multiphoton ionization (REMPI) experiments indicated that **1**, **1a**, and possibly **1b** coexist in the gas phase.[39] This same mixture plus **1c** was observed in IR matrix isolation studies.[40] Molecular beam microwave (MW) spectroscopy studies showed the presence of **1**, **1a**, and **1c**.[46]

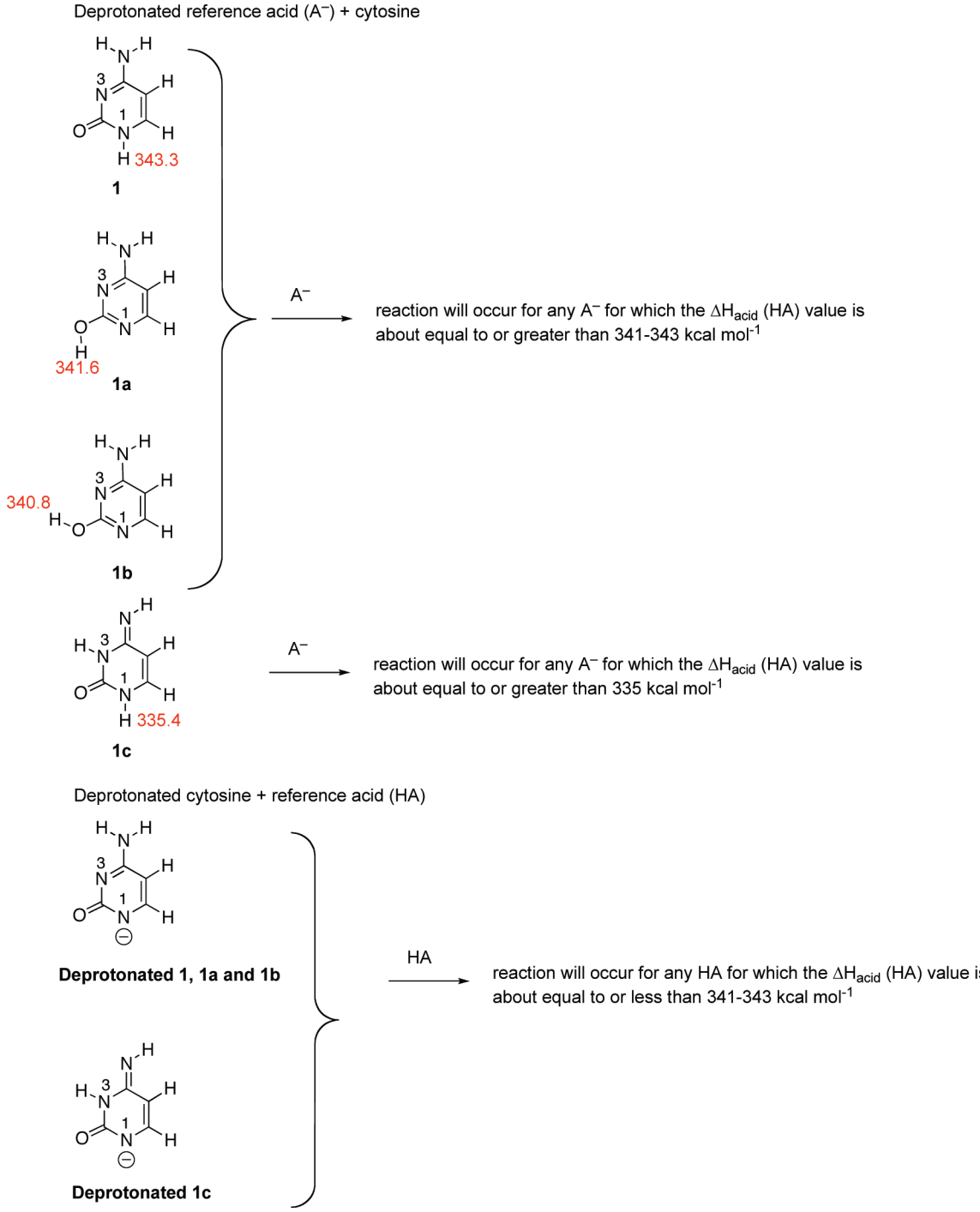
Can we use our bracketed acidity and proton affinity values and compare them to the calculated values (Figure 1) to discern which tautomers are present under our gas-phase conditions? If we presume that we have a mixture of **1**, **1a**, **1b**, and **1c**, and we conduct the more acidic gas-phase bracketing experiment, we would expect the reactivity depicted in Scheme 1. The bracketing reaction is conducted in two directions: deprotonated reference acid (A^-) plus cytosine and deprotonated cytosine plus reference acid HA. In the former, tautomers **1**, **1a**, and **1b** are indistinguishable; any A^- that is basic enough, for example, a ΔH_{acid} value for HA of at least around 341−343 kcal mol^{-1}, corresponding to the most acidic site of all three tautomers, will deprotonate each. The experiment cannot differentiate among those three tautomers. Tautomer **1c** is different; its predicted acidity is ΔH_{acid} ∼ 335 kcal mol^{-1}. If **1c** is present, an anion, A^-, whose ΔH_{acid} (HA) is equal to or greater than 335 kcal mol^{-1} should result in proton transfer. Therefore, in this direction, when a series of conjugate bases of reference acids, A^-, is introduced, if an A^- with a ΔH_{acid} value (for the conjugate acid HA) lower than ∼342 deprotonates cytosine, then **1c** is present. Interestingly, the reaction in the opposite direction is not as informative (Scheme 1). Deprotonation of **1**, **1a**, or **1b** results in the same anion. Reaction with HA could protonate N1 or O2 to produce any of those three tautomers; for all three, the reaction should proceed if ΔH_{acid} (HA) is equal to or less than about 342 kcal mol^{-1}. If **1c** is also present, the reactivity will not change; the reaction turns "on", or results in a "+",

(44) Ucella and co-workers (ref 20) used the kinetic method to obtain a cytosine PA of 225.9 kcal mol^{-1}. Mautner (ref 19) conducted high-pressure mass spectrometry equilibrium measurements, obtaining a cytosine PA of 224.9 kcal mol^{-1}. NIST subsequently evaluated these values, updating for changes in the reference acid and base scale, to report an evaluated PA of 227.0 kcal mol^{-1} (ref 23).

(45) Wilson, M. S.; McCloskey, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 3436–3444.

(46) Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. *J. Am. Chem. Soc.* **1989**, *111*, 2308–2310.

SCHEME 1



for any HA with a ΔH_{acid} value of 342 kcal mol^{-1} or lower, including an HA with an acidity of 335 kcal mol^{-1}.

This complicated situation is well summarized in a hypothetical acidity bracketing table (Table 9). In essence, if multiple tautomers with differing acidities are present in appreciable amounts, one will not see a clean crossover point in an acidity bracketing table. Instead, there will be a large section where there are two "+" indicators, spanning the acidities of the different tautomers (in this hypothetical example, 335−343 kcal mol^{-1}).

We do not see such a pattern (Table 1), which would imply that under our conditions, the imine tautomer **1c** ($\Delta H_{acid(calc)}$ = 335 kcal mol^{-1}) is not present. However, it is also possible that the imine tautomer is simply present in very small amounts (calculation by Turecek, Wesdemiotis, and co-workers estimates 5% at 473 K, the temperature of a typical ion source), and we, therefore, do not see its contribution to reactivity.[31] Thus, we cannot rule out the presence of **1c**.

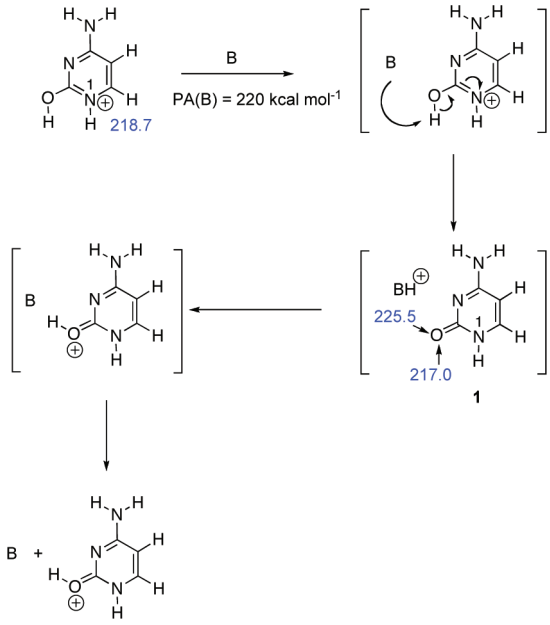
We can also compare our less acidic site acidity measurement with calculations. From previous experiments, we have found that when measuring the acidity of a less acidic site, we measure the least acidic ion present; thus, any tautomer with a proton that has an acidity near 352 kcal mol^{-1} could contribute to the measurement of that value.[5−7,9,13] Therefore, the less acidic measurement could be attributable to any of the four most stable tautomers: **1**, **1a**, **1b**, and **1c**. For PA, the measured proton

TABLE 9. Expected Acidity Bracketing Table if Cytosine Tautomers 1, 1a, 1b, and 1c Are Present

ΔH_{acid}[a] (kcal mol^{-1}) of ref acid	proton transfer[b]	
	ref. acid	conj. base
347	−	+
345	−	+
343	+	+
341	+	+
339	+	+
337	+	+
335	+	+
333	+	−
331	+	−

[a] Acidities are in kcal mol^{-1}.[23] [b] "+" indicates the occurrence of proton transfer, and "−" indicates the absence of proton transfer.

SCHEME 2



affinity of cytosine is 228 kcal mol^{-1}. This could correspond to tautomers **1** (probably both the O2 and N3 sites, which will be indistinguishable), **1b**, and/or **1c**. Tautomer **1a** has a much lower calculated PA of ~219 kcal mol^{-1}. Because protonated cytosine under our conditions does not react with reference bases whose PAs are 226.6 kcal mol^{-1} and lower, it is possible that **1a** is not present. However, once **1a** is protonated, a series of proton transfers could occur that simply would produce an isomeric protonated ion, which would appear as a lack of reaction (Scheme 2 with initial N1-protonated ion of **1a**). This is simply a limitation of the mass spectrometric experiment.

Thus, under our FTMS bracketing conditions, wherein cytosine is introduced via a heated solids probe, our results are consistent with a cytosine mixture of **1**, **1a**, **1b**, and possibly **1c**. This is also consistent with the previous gas-phase results.[39,40,46]

We followed up the bracketing experiments with extended Cooks kinetic method measurements of the acidity and proton affinity of cytosine. In this experiment, a proton-bound dimer of cytosine and a reference acid or base is formed in solution and vaporized via electrospray. Details are in the Experimental Section, but the point of interest for discussion is that the cytosine is electrosprayed from aqueous solution as opposed to sublimed from the solid state; most likely the canonical form of cytosine is, therefore, the predominant reactant.[27,37] The extended Cooks kinetic method experiments yield a ΔH_{acid} of 343 ± 3 kcal mol^{-1} and a PA of 227 ± 3 kcal mol^{-1}.[47−52] These data are in agreement with the calculated values for the most acidic and most basic site of the canonical cytosine tautomer **1** as well as, of course, **1b** and **1c**, though given the aqueous conditions, we consider the presence of those tautomers unlikely.

1-Methyl Cytosine. We synthesized and examined this derivative of cytosine because methylation of the 1-position decreases the number of energetically accessible tautomers. We find that at B3LYP/6-31+G* the canonical tautomer is the most stable by about 3 kcal mol^{-1} (Figure 2). This value is in agreement with calculations conducted at B3-MP2/6-311++G**// B3LYP/6-31+G** by Turecek and co-workers.[27] These same authors also predict that at 298 K the canonical tautomer will predominate in the gas phase (98.3% equilibrium fraction); at 473 K the canonical tautomer still predominates (92%). We would, therefore, expect to see prevalent reactivity of the canonical tautomer of 1-methyl cytosine. We measure a ΔH_{acid} = 349 ± 3 kcal mol^{-1} (ΔG_{acid} = 342 ± 3 kcal mol^{-1}) and a PA of 230 ± 3 kcal mol^{-1} (GB = 223 ± 3 kcal mol^{-1}). These values are in agreement with the calculated acidity and basicity values for the most acidic and most basic site of the canonical tautomer and provide benchmarking data that this level of calculation is reasonable for the acidity and proton affinity of cytosine and derivatives.

Thymine. The canonical tautomer of thymine is by far the most stable; by our calculations, the next nearest tautomer is about 12 kcal mol^{-1} less stable (Figure 3).[4,24,25] The acidity of thymine at the most acidic site, N1, is calculated to be ΔH_{acid} = 332.2 kcal mol^{-1} (ΔG_{acid} = 324.8 kcal mol^{-1}). We measure the acidity to be ΔH_{acid} = 335 ± 4 kcal mol^{-1} (ΔG_{acid} = 328 ± 4 kcal mol^{-1}), which is also in agreement with a previously measured gas-phase value (ΔH_{acid} = 333 ± 2 kcal mol^{-1}, using different methods).[21,22] The next acidity brackets to 346 ± 3 kcal mol^{-1} (ΔH_{acid}, ΔG_{acid} = 339 ± 3 kcal mol^{-1}), which is in agreement with the calculated value for N3−H acidity of 344.8 kcal mol^{-1} (ΔH_{acid}, ΔG_{acid} = 336.9 kcal mol^{-1}). The proton affinity measurement yields a value of PA = 211 ± 3 kcal mol^{-1} (GB = 203 ± 3 kcal mol^{-1}), which agrees with previous measurements (PA = 210.5 kcal mol^{-1}).[19,20,23,45,53] The value of 211 kcal mol^{-1} is somewhat surprising in light of the

(47) We calculated the $\Delta\Delta S$ values for the cytosine acidity and the PA for the extended Cooks kinetic method experiments, using the Armentrout method (ref 48). For the acidity, the $\Delta\Delta S$ value is 3.2 cal K^{-1} mol^{-1}; for the PA, the value is 2.2 cal K^{-1} mol^{-1}. It has been noted that the $\Delta\Delta S$ value is related to the accuracy of the PA value obtained by the extended Cooks kinetic method (ref 51). Ideally, the actual $\Delta\Delta S$ value should be less than or equal to about 5 cal K^{-1} mol^{-1}; otherwise, the extended kinetic method may underestimate the PA. Another caveat is that the $\Delta\Delta S$ value obtained from the extended method is often itself underestimated. There is, therefore, a possibility that the values obtained via our extended Cooks kinetic method experiment are underestimated; however, given that they are in agreement with the bracketing results, we are inclined to believe that the extended kinetic method value is not too low.

(48) Armentrout, P. B. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 371–379.
(49) Cheng, X.; Wu, Z.; Fenselau, C. *J. Am. Chem. Soc.* **1993**, *115*, 4844–4848.
(50) Nold, M. J.; Cerda, B. A.; Wesdemiotis, C. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 1–8.
(51) Wesdemiotis, C. *J. Mass Spectrom.* **2004**, *39*, 998–1003.
(52) Williams, T. I.; Denault, J. W.; Cooks, R. G. *Int. J. Mass Spectrom.* **2001**, *210/211*, 133–146.
(53) Ucella and co-workers (ref 20) used the Cooks kinetic method to obtain a thymine PA of 209.0 kcal mol^{-1}. Mautner (ref 19) conducted high-pressure mass spectrometry equilibrium measurements, obtaining a thymine PA of 210.9 kcal mol^{-1}. NIST subsequently evaluated these values, updating for changes in the reference acid and base scale, to report an evaluated PA of 210.5 ± 2.0 kcal mol^{-1} (ref 23).

calculated value of 203.7 kcal mol^{-1}. We are not certain why the calculated and experimental values are quite different; we saw a similar result when we examined uracil, which only differs from thymine by the lack of the methyl group on C5.[7,25,54] Higher-level calculations do bring the calculated values closer to the experimental; we find that at B3LYP/6-311++G**, the PA of thymine is calculated to be 206.6 kcal mol^{-1}.[4,55]

Conclusions

We have calculated and measured the acidity and proton affinity of cytosine, 1-methyl cytosine, and thymine to probe the intrinsic reactivity of these pyrimidine nucleobases. We are interested in particular in how damaged bases differ from normal bases. DNA is inevitably damaged by environmental mutagens as well as chemotherapeutics; such mutations are linked to carcinogenesis and aging.[15−17] Our laboratory studies the mechanisms by which enzymes, primarily glycosylases, might cleave damaged bases from DNA, thereby protecting our genome.[5,6,8,9,11,13,18] Previous results from our laboratory have shown that the properties of normal versus damaged bases lend insight into the mechanisms by which the damaged bases are cleaved.[5,6,8,9,11,13,18] The first step toward understanding how normal bases differ from damaged bases is to characterize the naturally occurring normal compounds, which motivates the study herein.

Cytosine has two measurable acidic sites, measured using the acidity bracketing method: ΔH_{acid} = 342 ± 3 and 352 ± 4 kcal mol^{-1} (ΔG_{acid} = 335 ± 3 and 345 ± 4 kcal mol^{-1}). The PA of cytosine brackets to be 228 ± 3 kcal mol^{-1} (GB = 220 ± 3 kcal mol^{-1}). Comparison of these values to theoretical data indicates that under our conditions, we probably have a mixture of the canonical tautomer, **1**, the enol tautomers, **1a** and **1b**, and possibly the imine tautomer, **1c**. We also measured the acidity of the most acidic site and the proton affinity of the most basic site of cytosine using the extended Cooks kinetic method; in these experiments, protonated dimers with cytosine are electrosprayed from aqueous solution and are, therefore, more likely to consist predominantly of the canonical tautomer. Using the Cooks method, we measure the ΔH_{acid} to be 343 ± 3 kcal mol^{-1} and the PA to be 227 ± 3 kcal mol^{-1}. We also examined the properties of 1-methyl cytosine, which is predicted to be predominantly the canonical tautomer in the gas phase and, therefore, allows benchmarking of the calculational level. We measure the ΔH_{acid} to be 349 ± 3 kcal mol^{-1} (corresponding to exocyclic NH_2, ΔG_{acid} = 342 ± 3 kcal mol^{-1}) and the PA to be 230 ± 3 kcal mol^{-1} (corresponding to O2 and/or N3, GB = 223 ± 3 kcal mol^{-1}); both these values are in agreement with the calculation. Last, we examined the pyrimidine nucleobase thymine. Thymine exists as the canonical tautomer in the gas phase and has a ΔH_{acid} of 335 ± 4 kcal mol^{-1} (ΔG_{acid} = 328 ± 4 kcal mol^{-1}) for the more acidic N1 site and a ΔH_{acid} of 346 ± 3 kcal mol^{-1} (ΔG_{acid} = 339 ± 3 kcal mol^{-1}) for the less acidic N3 site. The PA brackets to be 211 ± 3 kcal mol^{-1} (corresponding to O4, GB = 203 ± 3 kcal mol^{-1}). Thus, it is much easier to deprotonate neutral thymine in the gas phase than it is to deprotonate neutral cytosine, which is also true in solution (pK_a of thymine = 9.9, pK_a of cytosine = 12.2).[56] For protonation of the neutral bases, cytosine is more basic than thymine in the gas phase; this is also true in solution (pK_a of protonated thymine = 0, pK_a of protonated cytosine = 4.45).[56] Studies are currently underway on adenine and guanine to measure the thermochemical properties, which will allow for a full comparison of all five RNA−DNA nucleobases in the gas phase and in solution.

Experimental Section

All chemicals except 1-methyl cytosine are commercially available and were used as received. 1-Methyl cytosine was synthesized by following the literature procedure.[57]

Acidity and proton affinity experiments were conducted using a Fourier transform ion cyclotron resonance mass spectrometer (FTMS) with a dual cell setup, which has been described previously.[5,6,8] In our FTMS, two adjoining 1-in. cubic cells are positioned collinearly with the magnetic field produced by a 3.3 T superconducting magnet. The pressure of the dual cell is pumped down to less than 1 × 10^{-9} Torr. The solid nucleobases are introduced into the cells via a heatable solids probe. Ions are generated via reaction with H_3O^+ or OH^- ions. Ions can be transferred from one cell to the second cell via a 2-mm hole in the center of the central trapping plate. Transferred ions are cooled by a pulse of argon that raises the cell pressure to 10^{-5} Torr.[58] Experiments are conducted at ambient temperature.

Acidity, proton affinity, and gas-phase basicity are assessed using bracketing experiments in the FTMS, which have been described previously.[5,6,8] Briefly, for acidity bracketing, hydroxide ions are generated first by pulsing water into the FTMS cell and sending an electron beam (8 eV, 6 μA, beam time 0.5 s) through the center of the cell. The hydroxide ions deprotonate neutral molecules, M (either nucleobases or reference bases), to yield the $[M - H]^-$ ions. The $[M - H]^-$ ion is allowed to react with the neutral nucleobase or reference base. The same procedure is used for bracketing proton affinity, where hydronium ions (20 eV, 6 μA, beam time 0.2 s) are used for protonation. The occurrence of proton transfer is regarded as evidence that the reaction is exothermic ("+" in tables). Because we can measure multiple acidic and basic sites on a molecule (vida infra), we refer to these bracketing experiments as "more acidic" or "more basic" conditions. Charged species are all either monodeprotonated in acidity studies or monoprotonated in PA/GB studies.

We have recently developed an FTMS method for the bracketing of less acidic and less basic sites in molecules that have multiple acidic and basic sites; the experimental procedure and limitations have been described previously.[5,6,11,12] Briefly, in this setup, nucleobase ions produced after reaction of the corresponding neutral nucleobase with hydroxide ions are immediately removed from the first cell and transferred into the second cell. Reference acids are then injected into the second cell and allowed to react with the nucleobase ions. The first reaction cell is rich in neutral nucleobase concentration, and over time neutral-catalyzed isomerization leads to survival of only the most acidic ions. Transferring ions into the second cell immediately after their generation allows us to carry out the reaction between reference acids and nucleobase ions in the absence of neutral nucleobase. The same procedure can be applied to the bracketing of less basic sites as well.[7] In our experience, we usually measure two sites: the most acidic or basic site present and then a second site, which would be the least acidic or basic site present. The reasons for our measuring only two sites, and not more, have been previously discussed.[5−9,11−13] We refer to the conditions under which we run this experiment as "less

(54) Wolken, J. K.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1065–1071.

(55) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **1998**, *102*, 6010–6016.

(56) *Handbook of Biochemistry and Molecular Biology: Physical and Chemical Data*; Fasman, G. D., Ed.; CRC Press, Inc.: Cleveland, OH, 1976; Vol. 1.

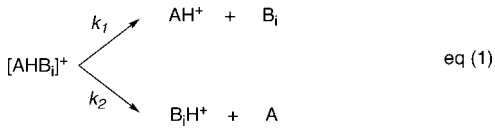
(57) Papoulis, A.; Al-Abed, Y.; Bucala, R. *Biochemistry* **1995**, *34*, 648–655.

(58) Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325–1337, and references therein.

acidic" or "less basic" conditions. Again, charged species are all either monodeprotonated or monoprotonated.[5−9,11−13]

In our experiments, we have pseudo-first-order conditions, where the amount of the neutral substrate is in excess relative to the reactant ions. Reading the pressure from an ion gauge is often unreliable because of both the gauge's remote location and the varying sensitivity for different substrates.[12,59] We, therefore, "back out" the neutral pressure from a control reaction; this procedure has been described previously by us.[12] Briefly, we obtain the pseudo-first-order rate constant for the reaction of hydroxide and the neutral substrate. Because hydroxide is very basic, we assume this reaction proceeds at the theoretical collision rate.[11,12,60,61] We then can use the calculated collisional rate constant to back out the neutral pressure.

We also used the Cooks kinetic method in a quadrupole ion trap (LCQ) mass spectrometer[62−66] to measure the proton affinity and acidity of cytosine. The procedure for conducting these experiments in our laboratory has been described previously.[9,13] Briefly, for PA experiments, this method involves the formation of a proton-bound complex, or dimer, of the unknown (in our case, cytosine) and a reference base of known proton affinity (eq 1, where A is cytosine and B_i is a series of reference bases). Collision-induced dissociation (CID) of this dimer leads to the formation of either the protonated unknown or the protonated reference base. The ratio of these two protonated products yields the relative proton affinities of the two compounds of interest, assuming that the dissociation has no reverse activation energy barrier and that the dissociation transition structure is late and, therefore, indicative of the stability of the two protonated products. Both these assumptions are generally true for proton-bound systems.[66−68] The same type of experiment can be done for acidity measurements (proton-bound dimer of deprotonated analyte and reference acid).



$$\ln(k_1/k_2) = [(\mathrm{PA(A)}/RT_{\mathrm{eff}}) - \Delta(\Delta S)/R] - \mathrm{PA(B_i)}/(RT_{\mathrm{eff}}) \quad (2)$$

$$\ln(k_1/k_2) = \ln([\mathrm{AH^+}]/[\mathrm{B_iH^+}]) \quad (3)$$

$$\mathrm{GB^{app}(A)}/RT_{\mathrm{eff}} = \mathrm{PA(A)}/(RT_{\mathrm{eff}}) - \Delta(\Delta S)/R \quad (4)$$

$$y_{01}' = [(\mathrm{PA(A)} - \mathrm{PA_{avg}})/RT_{\mathrm{eff}}] - \Delta(\Delta S)/R \quad (5)$$

Data were analyzed using the extended Cooks kinetic method.[48−50,52,69] This method has been well-described and involves acquiring ion abundance ratios at different collision energies and, therefore, different effective temperatures, vide infra, which allows for deconvolution of the enthalpic and entropic contributions. The data were worked up using a method developed by Armentrout, which is related to a method developed by Fenselau and Wesdemiotis.[48−50] Equations 2−5 summarize the data analysis. T_{eff} is the effective temperature of the dissociating proton-bound complex in kelvin. The term $\Delta(\Delta S)$ is the difference in the ΔS associated with the two channels in eq 1. In the Fenselau−Wesdemiotis method, a plot of $\ln(k_1/k_2)$, which is equal to $\ln([\mathrm{AH^+}]/[\mathrm{B_iH^+}])$ (eq 3), versus $\mathrm{PA(B_i)}$ yields the T_{eff} from the slope (eq 2) and the $\mathrm{GB^{app}(A)}$ from the intercept (eqs 2, 4). Plotting eq 4 at different values of T_{eff} yields the proton affinity and $\Delta(\Delta S)$ for the Cooks measurement. However, Armentrout noted that a more statistically rigorous analysis would involve using eqs 1−4, but instead of plotting $\ln(k_1/k_2)$ versus $\mathrm{PA(B_i)}$ (eq 2), one would plot $\ln(k_1/k_2)$ versus $[\mathrm{PA(B_i)} - \mathrm{PA_{avg}}]$, where $\mathrm{PA_{avg}}$ is equal to the average PA of the reference bases used. This would result in a new *y*-intercept of y_{01}' (eq 5). Plotting the y_{01}' values versus $1/RT_{\mathrm{eff}}$ (eq 5) will yield $\Delta(\Delta S)/R$ (from the *y*-intercept) and PA(A) (from the slope, which equals $\mathrm{PA(A)} - \mathrm{PA_{avg}}$).

Proton-bound complex ions are generated by electrospray (ESI).[70] For each experiment, a solution of the nucleobase and reference base or acid is prepared (10^{-3} to 10^{-4} M solutions in a 20% methanol aqueous solution). The typical flow rate is 25 μL/min. An electrospray needle voltage of ~4500 V was used. The proton-bound complex ions were isolated and then dissociated by applying collision-induced dissociation (CID); the complexes were activated for about 30 ms. A total of 40 scans was averaged for the product ions.

For the Cooks kinetic measurement of cytosine acidity, the following reference acids were used: trifluoro-*m*-cresol ($\Delta H_{\mathrm{acid}} = 339.3 \pm 2.1$ kcal mol^{-1}), methoxyacetic acid ($\Delta H_{\mathrm{acid}} = 341.9 \pm 2.1$ kcal mol^{-1}), ethoxyacetic acid ($\Delta H_{\mathrm{acid}} = 342.0 \pm 2.2$ kcal mol^{-1}), and 2-chlorophenol ($\Delta H_{\mathrm{acid}} = 343.4 \pm 2.3$ kcal mol^{-1}). For the cytosine PA measurement, we used 1-methyl pyrrolidine (PA = 230.8 ± 2.0 kcal mol^{-1}), piperidine (PA = 228.0 ± 2.0 kcal mol^{-1}), pyrrolidine (PA = 226.6 ± 2.0 kcal mol^{-1}), 3-picoline (PA = 225.5 ± 2.0 kcal mol^{-1}), pyridine (PA = 222.3 ± 2.0 kcal mol^{-1}), and 1-octanamine (PA = 222.0 ± 2.0 kcal mol^{-1}).

The B3LYP method and the 6-31+G* basis set as implemented in Gaussian03 were used for all the gas-phase calculations.[71−76] This level has been shown previously to be reasonably accurate for gas-phase acidity and proton affinity calculations of nucleobases.[5−9,11−13] All the geometries are fully optimized, and frequencies are calculated; no scaling factor is applied. Reported values herein are at 298 K.[77]

Acknowledgment. We gratefully acknowledge the support of NSF, the Alfred P. Sloan Foundation, and the National Center for Supercomputer Applications.

Supporting Information Available: Cartesian coordinates for all calculated species, including higher energy tautomers, and full citations for references with greater than 16 authors. This material is available free of charge via the Internet at http://pubs.acs.org.

JO801822S

(59) Bartmess, J. E.; Georgiadis, R. M. *Vacuum* **1983**, *33*, 149–153.
(60) Su, T.; Chesnavich, W. J. *J. Chem. Phys.* **1982**, *76*, 5183–5185.
(61) Chesnavich, W. J.; Su, T.; Bowers, M. T. *J. Chem. Phys.* **1980**, *72*, 2641–2655.
(62) Cooks, R. G.; Kruger, T. L. *J. Am. Chem. Soc.* **1977**, *99*, 1279–1281.
(63) McLuckey, S. A.; Cameron, D.; Cooks, R. G. *J. Am. Chem. Soc.* **1981**, *103*, 1313–1317.
(64) McLuckey, S. A.; Cooks, R. G.; Fulford, J. E. *Int. J. Mass Spectrom. Ion Processes* **1983**, *52*, 165–174.
(65) Brodbelt-Lustig, J. S.; Cooks, R. G. *Talanta* **1989**, *36*, 255–260.
(66) Green-Church, K. B.; Limbach, P. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 24–32.
(67) Ervin, K. M. *Chem. Rev.* **2001**, *101*, 391–444, and references therein.
(68) Gronert, S.; Feng, W. Y.; Chew, F.; Wu, W. *Int. J. Mass Spectrom.* **2000**, *196*, 251–258.
(69) Cerda, B. A.; Wesdemiotis, C. *J. Am. Chem. Soc.* **1996**, *118*, 11884–11892.
(70) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64–71.
(71) Frisch, R. C.; et al. Gaussian03; Gaussian, Inc.: Wallingford, CT, 2004.
(72) Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974–12980.
(73) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(74) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
(75) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
(76) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
(77) To calculate the ΔH_{298}, we account for both the translational energy of the proton (3/2 *RT*) and the work term associated with the dissociation of one molecule into two (*RT*). For the ΔG_{298} values, we also account for the ΔS_{298} of the proton (experimental value = 26 eu, ref 23).