

# Calculated Gas-Phase Acidities Using Density Functional Theory: Is It Reliable?

Grant N. Merrill and Steven R. Kass\*

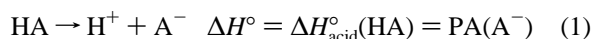
Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455

Received: May 29, 1996; In Final Form: July 30, 1996<sup>®</sup>

Proton affinities for a diverse set of 35 anions were calculated using 7 different levels of density functional theory (S-null, S-VWN, S-LYP, B-null, B-VWN, B-LYP, and Becke3-PW91). These combinations of functionals cover the range of those that are widely available. The basis-set dependence of these methods was explored using split-valence, double- vs triple-zeta, and correlation- vs noncorrelation-consistent bases (6-31+G(d), and 6-311++G(2df,2pd) vs aug-cc-pVDZ and aug-cc-pVTZ). The results are compared to modest and high-level ab initio acidities (MP2(fc)/6-31+G(d)//HF/6-31+G(d) and G2, respectively) as well as to experimental results.

## Introduction

One of the major contributions of computational chemistry is the accurate calculation of thermodynamic data. Gas-phase acidities (viz. proton affinities)



are of fundamental interest, and they provide a valuable means for evaluating the accuracy of different theoretical methods, since a wide range of experimental data are available for direct comparison.<sup>1</sup> Semiempirical calculations are computationally inexpensive, but they are of limited utility, since anion geometries are often poorly described.<sup>2</sup> This results in large errors in the computed proton affinities.<sup>3</sup> High-level ab initio methods such as the Gaussian and complete basis set schemes are remarkably accurate, typically providing acidities within 2 or 3 kcal/mol of the experimental values.<sup>4</sup> Unfortunately, these calculations are computationally demanding, and thus, their application is restricted to relatively small systems. At the present time, this usually means molecules with no more than about five or six non-hydrogenic (heavy) atoms.<sup>5</sup> More modest levels of theory enable larger systems to be examined and still can be quantitatively useful. For example, Saunders recently reported that the proton affinities of a wide variety of carbanions (16 in all) are well reproduced at the MP2(fc)/6-31+G(d)//HF/6-31+G(d) level owing to a fortuitous cancellation of errors resulting from basis-set deficiencies and a nonvariational treatment of correlation.<sup>6</sup>

Density functional theory (DFT) has undergone a renaissance within the chemical community. This is due largely to its favorable scaling behavior with respect to ab initio methods, and thus, it offers a means of carrying out calculations on large systems.<sup>7</sup> Specifically, Hartree–Fock theory formally scales as  $ON^4$ , where  $O$  is the number of occupied orbitals in the molecule and  $N$  is the number of basis functions used to describe the species.<sup>8</sup> With an efficient implementation applied to large systems, this cost can be reduced to  $ON^2$ . A MP2 calculation scales formally as  $ON^5$  with more sophisticated correlation treatments scaling even more poorly (e.g., MP3 ( $ON^6$ ), MP4, or QCISD(T) ( $ON^7$ )). DFT, on the other hand, scales formally as  $ON^3$ , and an efficient implementation can lower this dependence to  $ON^2$  or even  $ON$ ! Since density functional theory explicitly includes electron correlation, it offers great promise as a means for obtaining reliable and accurate thermodynamic

data. Consequently, we are very interested in finding out how reliable DFT is when it comes to calculating proton affinities.

In the process of carrying out this work, Smith and Radom reported very encouraging results using a mixed ab initio/DFT approach.<sup>9</sup> They used MP2/6-31G(d) geometries and HF/6-31G(d) vibrational frequencies in conjunction with single-point DFT energies obtained with a few gradient-corrected, nonlocal functionals. The best results were obtained with the “hybrid” functionals Becke3-P86 and Becke3-LYP (or more simply, B3-P86 and B3-LYP).<sup>10</sup>

In this investigation (1) all optimizations were performed *self-consistently* within the density functional framework as were the vibrational analyses, (2) a wide variety of functionals representative of those commonly available were examined, (3) a diverse set of 35 anions, including the entire list of compounds studied by Saunders and some problem cases for traditional ab initio methods, were explored, (4) the performance of the functionals as a function of the basis set quality (i.e., Pople vs Dunning and split-valence, double- vs triple-zeta) was investigated,<sup>11,12</sup> and (5) the performance of the MP2(fc)/6-31+G(d)//HF/6-31+G(d) (Saunders)<sup>6</sup> level of theory was examined for non-carbon acids.

## Experimental Section

All the computations reported in this work were carried out using Gaussian 92/DFT,<sup>13</sup> Gaussian 94,<sup>14</sup> or GAMESS<sup>15</sup> on UNIX-based workstations or Cray supercomputers at the Minnesota Supercomputer Institute.

**DFT Calculations.** Full geometry optimizations and vibrational analyses were performed using seven different levels of theory (S-null, S-VWN, S-LYP, B-null, B-VWN, B-LYP, and B3-PW91)<sup>16</sup> with the 6-31+G(d) and aug-cc-pVDZ basis sets.<sup>11,12</sup> The former basis set was chosen because it is commonly used and was employed by Saunders in his study, while the latter basis set is Dunning’s correlation-consistent counterpart. Single-point energy calculations were subsequently carried out with each functional using the 6-311++G(2df,2pd) or aug-cc-pVTZ basis set upon the 6-31+G(d) or aug-cc-pVDZ optimized structures, respectively.<sup>11,12</sup> A pruned, fine integration grid consisting of 75 radial shells and 302 angular points/shell for a total of approximately 7000 integration points/atom was employed in all these calculations. The resulting proton affinities given in Tables 1–3 incorporate the unscaled zero-point energy (ZPE), a finite-temperature correction (0–298 K), and the pressure–volume work term (+ $RT$  for eq 1).

<sup>®</sup> Abstract published in *Advance ACS Abstracts*, October 15, 1996.

**TABLE 1: Calculated Acidities of Neutral Acids HA (or Proton Affinities of A<sup>-</sup>) for Various Functionals and Pople-Type Basis Sets<sup>a,b</sup>**

HA	pt grp <sup>c</sup>	S-null		S-VWN		S-LYP		B-null		B-VWN		B-LYP		B3-PW91		expt <sup>d</sup>
		I	II	I	II	I	II	I	II	I	II	I	II	I	II	
CH <sub>4</sub>	<i>T<sub>d</sub></i> (C <sub>3v</sub> )	396.6	395.3	410.4	408.7	401.8	400.1	409.2	408.5	422.9	422.1	414.1	413.2	420.9	419.5	416.7 ± 0.7
C <sub>2</sub> H <sub>4</sub>	<i>D<sub>2h</sub></i> (C <sub>s</sub> )	388.5	388.4	400.7	400.7	393.0	392.9	399.7	400.5	411.9	412.8	404.0	405.1	410.1	410.4	409.4 ± 0.6
CH <sub>3</sub> F	<i>C<sub>3v</sub></i> (C <sub>3s</sub> )	390.6	390.7	403.0	402.9	395.5	395.3	398.8	399.8	411.1	412.0	404.0	404.7	411.9	412.4	409.0 ± 4.0 <sup>e</sup>
NH <sub>3</sub>	<i>C<sub>3v</sub></i> (C <sub>2v</sub> )	386.6	386.7	397.7	397.8	390.7	390.7	395.7	396.2	406.7	407.4	399.5	400.1	407.0	407.5	404.0 ± 0.4
c-CH <sub>2</sub> N <sub>2</sub>	<i>C<sub>2v</sub></i> (C <sub>1</sub> )	381.5	381.3	392.5	392.4	385.6	385.3	389.0	389.7	401.0	401.8	394.0	394.7	401.3	401.5	401 ± 3 <sup>f</sup>
C <sub>6</sub> H <sub>6</sub>	<i>D<sub>6h</sub></i> (C <sub>2v</sub> )	381.3		392.8		385.0		393.1		404.6		396.8		402.5		401.7 ± 0.5 <sup>g</sup>
H <sub>2</sub>	<i>D<sub>∞h</sub></i> ( <sup>1</sup> S <sub>0</sub> )	430.6	374.9	435.5	389.9	431.4	379.5	442.7	393.3	448.2	408.0	443.6	397.8	441.1	401.3	400.4 ± .002
CH <sub>3</sub> Cl	<i>C<sub>3v</sub></i> (C <sub>3s</sub> )	380.9	379.3	393.1	390.9	386.0	383.6	387.0	387.3	399.9	400.0	393.2	393.1	400.5	399.5	396.1 ± 3.1 <sup>e</sup>
CH <sub>3</sub> SCH <sub>3</sub>	<i>C<sub>2v</sub></i> (C <sub>s</sub> )	372.8	370.7	384.7	382.8	376.4	374.4	385.4	384.3	397.6	396.5	389.3	388.1	396.9	395.3	393.2 ± 2.1 <sup>e</sup>
CH <sub>3</sub> CH=CH <sub>2</sub>	<i>C<sub>s</sub></i> (C <sub>2v</sub> )	368.4	367.0	379.2	377.3	371.2	369.5	383.1	382.3	394.0	392.9	385.5	384.5	391.3	390.1	390.8 ± 2.1
H <sub>2</sub> O	<i>C<sub>2v</sub></i> (C <sub>∞v</sub> )	372.5	376.8	381.3	385.6	375.3	379.6	380.1	384.5	388.8	393.3	382.6	387.2	389.9	394.4	390.7 ± 0.1
CH <sub>3</sub> OH	<i>C<sub>s</sub></i> (C <sub>3v</sub> )	359.4	361.3	368.4	371.4	362.1	364.7	370.3	372.7	379.0	382.4	372.5	375.6	378.9	382.3	381.6 ± 0.7
C <sub>6</sub> H <sub>5</sub> CH <sub>3</sub>	<i>C<sub>1</sub></i> (C <sub>s</sub> ) <sup>h</sup>	362.1	360.7	371.4	369.4	356.8	354.9	377.1	376.3	386.5	385.4	379.0	378.1	384.4	383.2	382.3 ± 0.3 <sup>i</sup>
C <sub>2</sub> H <sub>2</sub>	<i>D<sub>∞h</sub></i> (C <sub>∞v</sub> )	358.7	361.6	369.7	372.4	362.3	365.2	368.9	372.5	380.3	383.5	372.5	375.9	377.5	380.2	378.0 ± 0.6
HOOH	<i>C<sub>2</sub></i> (C <sub>s</sub> )	359.3	364.0	367.2	371.9	361.7	366.5	365.3	370.0	373.1	378.0	367.6	372.5	374.4	379.5	375.9 ± 2.1 <sup>e</sup>
CH <sub>3</sub> SOCH <sub>3</sub>	<i>C<sub>s</sub></i> (C <sub>1</sub> )	358.2	357.4	367.5	366.7	359.7	358.8	372.8	373.0	382.5	382.5	374.4	374.3	379.3	379.0	373.5 ± 2.1 <sup>e</sup>
CH <sub>2</sub> N <sub>2</sub>	<i>C<sub>2v</sub></i> (C <sub>s</sub> )	353.0	353.8	361.5	362.3	354.7	355.4	365.7	367.2	374.5	376.0	367.3	368.7	373.8	374.9	373 ± 3 <sup>j</sup>
CH <sub>3</sub> CN	<i>C<sub>3v</sub></i> (C <sub>s</sub> ) <sup>k</sup>	353.4	352.9	362.1	361.4	354.9	354.1	368.4	368.4	377.1	376.9	369.3	369.1	374.0	373.6	372.9 ± 2.1
SiH <sub>4</sub>	<i>T<sub>d</sub></i> (C <sub>3v</sub> )	348.4	348.4	360.8	360.9	341.8	351.5	366.1	367.2	378.8	380.1	369.4	370.4	372.5	373.4	372.2 ± 2.0
CH <sub>3</sub> CO <sub>2</sub> CH <sub>3</sub>	<i>C<sub>1</sub></i> (C <sub>1</sub> )	354.8	354.0	363.0	362.1	355.8	354.8	368.4	368.3	376.9	376.7	369.4	369.2	375.0	374.8	371.9 ± 2.1 <sup>e</sup>
PH <sub>3</sub>	<i>C<sub>3v</sub></i> (C <sub>2v</sub> )	346.6	347.8	358.4	359.3	350.2	350.7	361.4	363.1	373.3	375.3	364.7	366.3	367.7	369.1	370.8 ± 2.0
HF	<i>C<sub>∞v</sub></i> ( <sup>1</sup> S <sub>0</sub> )	350.6	358.7	357.0	365.2	351.8	360.2	357.2	365.2	363.5	371.7	358.3	366.6	364.6	373.0	371.5 ± 0.2
CH <sub>3</sub> COCH <sub>3</sub>	<i>C<sub>2</sub></i> (C <sub>s</sub> )	350.6	350.0	359.1	358.4	352.1	351.2	364.7	364.7	373.2	373.1	365.5	365.3	370.9	370.7	369.1 ± 2.1 <sup>e</sup>
CH <sub>3</sub> CHO	<i>C<sub>s</sub></i> (C <sub>1</sub> )	347.0	346.0	355.2	354.3	348.0	347.0	361.5	361.0	369.9	369.6	362.1	361.7	367.5	367.3	365.8 ± 2.2
CH <sub>3</sub> SH	<i>C<sub>s</sub></i> (C <sub>3v</sub> )	337.9	341.7	347.6	351.3	340.4	344.4	350.1	354.7	359.8	363.4	352.5	356.3	355.9	359.7	356.9 ± 2.2
CH <sub>3</sub> NO <sub>2</sub>	<i>C<sub>1</sub></i> (C <sub>2v</sub> )	338.9	339.0	345.4	345.6	338.9	338.9	352.3	353.0	358.6	359.5	351.7	352.4	356.3	357.3	356.4 ± 2.2 <sup>e</sup>
c-C <sub>3</sub> H <sub>6</sub>	<i>C<sub>2v</sub></i> (D <sub>3h</sub> )	335.4	335.0	343.9	343.2	336.4	335.8	351.9	352.0	360.4	360.2	352.4	352.3	355.4	355.0	353.9 ± 2.2 <sup>e</sup>
H <sub>2</sub> NCN	<i>C<sub>s</sub></i> (C <sub>s</sub> )	330.9	332.2	338.6	339.9	332.6	333.9	341.5	343.4	349.1	351.0	342.8	344.7	347.4	349.1	352 ± 3 <sup>f</sup>
H <sub>2</sub> S	<i>C<sub>2v</sub></i> (C <sub>∞v</sub> )	331.2	335.2	341.2	345.0	334.0	338.0	343.7	347.5	353.7	357.5	346.2	350.1	349.3	353.5	351.1 ± 2.0
HCN	<i>C<sub>∞v</sub></i> (C <sub>2v</sub> )	334.3	336.4	343.5	345.5	336.7	338.7	344.9	347.5	354.2	356.7	347.2	349.8	350.6	352.8	351.4 ± 0.5
HCO <sub>2</sub> H	<i>C<sub>s</sub></i> (C <sub>2v</sub> )	324.8	329.0	331.7	336.2	326.0	330.5	334.6	339.1	341.3	346.0	335.5	340.2	340.3	345.2	345 ± 2
HCl	<i>C<sub>∞v</sub></i> ( <sup>1</sup> S <sub>0</sub> )	312.1	319.4	320.1	327.3	313.8	321.3	322.8	329.3	330.9	337.3	324.3	331.1	327.0	334.3	333.4 ± 0.03
unsigned mean error <sup>l</sup>		19.5	18.3	9.8	8.7	17.4	16.1	7.8	6.0	3.6	3.8	5.3	3.6	2.4	2.0	
mean deviation in error <sup>l,m</sup>		1.7	2.6	2.0	2.4	2.5	3.0	2.5	2.1	1.5	1.6	2.1	1.3	1.6	1.0	

<sup>a</sup> All values are in kcal/mol and have been zero-point energy (ZPE) and temperature corrected using unscaled density functional frequencies.

<sup>b</sup> Level of theory I: *X* (*X* = S-null, S-VWN, etc.)/6-31+G(d). Level of theory II: *X*/6-311++G(2df,2pd)/*X*/6-31+G(d). <sup>c</sup> The symmetry of the acid is given first, and the parenthetical value is for the conjugate base. <sup>d</sup> All experimental values from ref 19a unless otherwise noted. <sup>e</sup> See ref 1.

<sup>f</sup> See ref 19d. <sup>g</sup> See ref 19c. <sup>h</sup> The acid has *C<sub>s</sub>* symmetry with the following functionals: S-null, S-VWN, and S-LYP. <sup>i</sup> See ref 19b. <sup>j</sup> See ref 19e.

<sup>k</sup> The anion has *C<sub>2v</sub>* symmetry with the following functionals: S-null, S-VWN, and S-LYP. <sup>l</sup> Error excludes the acidity for H<sub>2</sub> (see ref 20 for justification). <sup>m</sup> The mean deviation error is defined as follows:  $\sum_i |x_i - |\bar{x}||/N$ .

The seven levels of theory used in this study cover the gamut of currently available functionals: local and nonlocal exchange (Slater's exchange functional (S-null) and Becke's gradient correction to Slater's exchange functional (B-null), respectively); local and nonlocal correlation (VWN and LYP, respectively); a "hybrid" functional involving DFT and HF exchange, Becke-3-PW91 (B3-PW91). The exchange–correlation energy,  $E_{xc}$ , associated with this latter functional is given in eq 2, where the parameters  $a_0 = 0.20$ ,  $a_x = 0.72$ , and  $a_c = 0.81$  were obtained by fitting 56 atomization energies, 42 ionization potentials, and 8 proton affinities to experimental values found in the G1 data set.<sup>4e,10</sup>

$$E_{xc} = E_{xc}^{S-VWN} + a_0(E_x^{HF} - E_x^{S-VWN}) + a_x \Delta E_x^{B-null} + a_c \Delta E_c^{PW91} \quad (2)$$

It is worth noting that within the current implementation found in the Gaussian series of programs (i.e., G92/DFT<sup>13</sup> and G94<sup>14</sup>) DFT calculations are more expensive than their HF counterparts but less so than MP2 computations. This is a result of the need to carry out a subsequent series of numerical integrations of the functional at each step in the SCF routine. Density functional wave functions and geometries are also often more difficult to converge than their HF counterparts.<sup>17</sup>

**Ab Initio Calculations.** Full geometry optimizations and vibrational frequency analyses were carried out at the Hartree–Fock (HF) level with the 6-31+G(d) basis set. Single-point energies were obtained at the MP2(fc)/6-31+G(d) level, and the resulting proton affinities were zero-point energy corrected. Unlike the DFT calculations, however, the ZPEs were scaled (0.9). In several cases we also carried out G2 calculations,<sup>4a</sup> which effectively correspond to the QCISD(T)/6-311+G(3df,-2p)//MP2(fu)/6-31G(d) level of theory and represent the highest level ab initio calculations employed in this study. The G2 prescription calls for a vibrational analysis at the HF/6-31G(d) level, but MP2 geometries are used for subsequent single-point calculations. For some species this is a potential pitfall, since the curvature of the potential energy surface can differ at correlated and uncorrelated levels.<sup>18</sup>

## Results and Discussion

Structures of 32 acids and their conjugate bases were optimized using 7 different combinations of functionals (S-null, S-VWN, S-LYP, B-null, B-VWN, B-LYP, and B3-PW91)<sup>16</sup> with the 6-31+G(d) basis set. The resulting proton affinities (PA) along with their literature values and the experimental uncertainties are given in Table 1.<sup>1,19</sup> The three functionals using Slater exchange, S-null, S-VWN, and S-LYP fare

**TABLE 2: Calculated Acidities of Neutral Acids HA (or Proton Affinities of A<sup>-</sup>) for Various Functionals and Dunning-Type Basis Sets<sup>a,b</sup>**

HA	pt gr <sup>c</sup>	S-null		S-VWN		S-LYP		B-null		B-VWN		B-LYP		B3-PW91		expt <sup>d</sup>
		III	IV	III	IV	III	IV	III	IV	III	IV	III	IV	III	IV	
CH <sub>4</sub>	<i>T<sub>d</sub></i> (C <sub>3v</sub> )	392.9	393.6	406.5	407.8	397.8	398.9	405.7	406.9	419.5	421.3	410.5	412.1	417.6	419	416.7 ± 0.7
C <sub>2</sub> H <sub>4</sub>	<i>D<sub>2h</sub></i> (C <sub>s</sub> )	385.9	386.9	398.6	400.2	390.7	392.0	397.7	399.4	410.3	412.4	402.2	404.2	408.4	410.1	409.4 ± 0.6
CH <sub>3</sub> F	<i>C<sub>3v</sub></i> (C <sub>3s</sub> )	387.5	388.8	400.0	401.8	392.3	393.8	396.2	398.6	408.6	411.4	401.2	403.7	409.6	411.9	409.0 ± 4.0 <sup>e</sup>
NH <sub>3</sub>	<i>C<sub>3v</sub></i> (C <sub>2v</sub> )	383.7	384.3	395.4	396.4	388.1	388.9	393.2	394.1	404.9	406.2	397.4	398.6	405.2	406.5	404.0 ± 0.4
c-CH <sub>2</sub> N <sub>2</sub>	<i>C<sub>2v</sub></i> (C <sub>1</sub> )	379.3	380.7	390.5	392.3	383.3	384.9	387.3	389.4	399.4	401.8	392.3	394.5	400.0	401.9	401 ± 3 <sup>f</sup>
C <sub>6</sub> H <sub>6</sub>	<i>D<sub>6h</sub></i> (C <sub>2v</sub> )	380.2		391.9		384.2		392.2		403.7		396.2		401.8		401.7 ± 0.5 <sup>g</sup>
H <sub>2</sub>	<i>D<sub>∞h</sub></i> (1S <sub>0</sub> )	371.5	373.0	387.0	388.8	376.4	378.1	388.9	391.6	404.0	407.2	393.5	396.5	398.4	400.5	400.4 ± .002
CH <sub>3</sub> Cl	<i>C<sub>3v</sub></i> (C <sub>3s</sub> )	376.8	378.6	388.6	390.5	381.4	383.2	384.1	386.8	396.6	399.5	389.7	392.4	396.9	399.3	396.1 ± 3.1 <sup>e</sup>
CH <sub>3</sub> SCH <sub>3</sub>	<i>C<sub>2v</sub></i> (C <sub>s</sub> )	369.7	370.3	382.4	383.0	374.2	374.8	382.9	384.2	395.5	397.0	387.0	388.3	395.1	395.8	393.2 ± 2.1 <sup>e</sup>
CH <sub>3</sub> CH=CH <sub>2</sub>	<i>C<sub>s</sub></i> (C <sub>2v</sub> )	365.9	376.0	376.5	387.1	368.5	378.9	380.7	391.7	391.2	402.7	383.0	394.2	389.1	388.3	390.8 ± 2.1
H <sub>2</sub> O	<i>C<sub>2v</sub></i> (C <sub>∞v</sub> )	372.5	373.4	382.5	383.7	375.8	376.8	380.0	381.1	389.7	391.0	383.2	384.4	391.3	392.7	390.7 ± 0.1
CH <sub>3</sub> OH	<i>C<sub>s</sub></i> (C <sub>3v</sub> )	359.1	360.5	369.4	370.7	362.5	363.7	370.0	371.4	380.0	381.7	373.1	374.6	380.4	381.9	381.6 ± 0.7
C <sub>2</sub> H <sub>2</sub>	<i>D<sub>∞h</sub></i> (C <sub>∞v</sub> )	358.9	361.4	370.1	372.7	362.8	365.4	368.9	372.1	380.1	383.5	372.8	376.0	377.7	380.5	378.0 ± 0.6
HOOH	<i>C<sub>2</sub></i> (C <sub>s</sub> )	360.8	361.7	369.4	370.5	363.6	364.6	366.6	367.9	375.2	376.7	369.5	370.8	377.1	378.7	375.9 ± 2.1 <sup>e</sup>
CH <sub>3</sub> SOCH <sub>3</sub>	<i>C<sub>s</sub></i> (C <sub>1</sub> )	356.0	357.4	365.8	367.2	358.0	359.2	370.6	372.9	380.4	382.7	372.2	374.4	377.7	379.2	373.5 ± 2.1 <sup>e</sup>
CH <sub>3</sub> N <sub>2</sub>	<i>C<sub>2v</sub></i> (C <sub>s</sub> )	351.9	353.4	360.7	362.4	353.9	355.4	364.8	366.9	373.7	376.0	366.3	368.5	373.0	374.9	373 ± 3 <sup>h</sup>
CH <sub>3</sub> CN	<i>C<sub>3v</sub></i> (C <sub>s</sub> ) <sup>i</sup>	352.0	353.0	360.5	361.7	353.2	354.3	367.0	368.4	375.5	377.1	367.7	369.3	372.7	373.8	372.9 ± 2.1
SiH <sub>4</sub>	<i>T<sub>d</sub></i> (C <sub>3v</sub> )	345.9	347.6	358.4	360.3	349.2	350.9	363.9	366.3	376.8	379.5	367.2	369.7	370.8	372.9	372.2 ± 2.0
CH <sub>3</sub> CO <sub>2</sub> CH <sub>3</sub>	<i>C<sub>1</sub></i> (C <sub>1</sub> )	352.8	353.9	361.0	362.2	353.6	354.7	366.6	368.1	375.2	376.7	367.6	369.2	373.7	374.8	371.9 ± 2.1 <sup>e</sup>
PH <sub>3</sub>	<i>C<sub>3v</sub></i> (C <sub>2v</sub> )	345.7	346.7	357.4	358.7	349.2	350.4	360.4	362.1	372.2	374.2	363.6	365.4	366.9	368.5	370.8 ± 2.0
HF	<i>C<sub>2v</sub></i> (1S <sub>0</sub> )	353.6	355.4	361.0	363.0	355.5	357.5	359.7	361.7	367.1	369.3	361.5	363.7	368.9	371.4	371.5 ± 0.2
CH <sub>3</sub> COCH <sub>3</sub>	<i>C<sub>2</sub></i> (C <sub>s</sub> )	348.9	349.9	357.5	358.5	350.3	351.2	363.1	364.6	371.6	373.1	363.8	365.3	369.8	370.9	369.1 ± 2.1 <sup>e</sup>
CH <sub>3</sub> CHO	<i>C<sub>s</sub></i> (C <sub>s</sub> )	344.8	345.9	353.2	354.3	346.0	347.0	359.4	360.9	367.8	369.4	359.9	361.5	366.2	367.4	365.8 ± 2.2
CH <sub>3</sub> SH	<i>C<sub>s</sub></i> (C <sub>3v</sub> )	338.9	340.7	348.5	350.5	341.5	343.5	350.7	352.8	360.4	362.8	353.2	355.5	356.9	359.0	356.9 ± 2.2
CH <sub>3</sub> NO <sub>2</sub>	<i>C<sub>1</sub></i> (C <sub>2v</sub> )	338.3	339.1	344.9	345.9	338.1	339.1	351.6	352.9	358.2	359.7	351.1	352.5	356.5	357.7	356.4 ± 2.2 <sup>e</sup>
c-C <sub>3</sub> H <sub>6</sub>	<i>C<sub>2v</sub></i> (D <sub>5h</sub> )	334.1	334.8	342.3	343.0	334.9	335.6	350.6	351.7	359.0	360.1	350.9	352.1	354.3	355.0	353.9 ± 2.2 <sup>e</sup>
H <sub>2</sub> NCN	<i>C<sub>s</sub></i> (C <sub>s</sub> )	330.7	332.0	338.6	340.0	332.5	333.8	341.7	343.3	349.4	351.1	343.0	344.6	347.7	349.3	352 ± 3 <sup>f</sup>
H <sub>2</sub> S	<i>C<sub>2v</sub></i> (C <sub>∞v</sub> )	332.8	334.4	342.7	344.4	335.6	337.3	344.6	346.6	354.6	356.8	347.2	349.3	350.7	354.0	351.1 ± 2.0
HCN	<i>C<sub>∞v</sub></i> (C <sub>∞v</sub> )	334.9	336.4	343.9	345.6	337.2	338.8	345.3	347.6	354.4	356.9	347.5	349.9	351.0	353.1	351.4 ± 0.5
HCO <sub>2</sub> H	<i>C<sub>s</sub></i> (C <sub>2v</sub> )	327.8	328.7	335.1	336.1	329.4	330.4	337.4	338.7	344.6	345.9	338.7	339.9	343.9	345.3	345 ± 2
HCl	<i>C<sub>∞v</sub></i> (1S <sub>0</sub> )	316.6	318.7	324.7	326.9	318.7	320.8	326.1	328.5	334.2	336.9	328.0	330.5	332.1	333.9	333.4 ± 0.03
unsigned mean error <sup>j</sup>		20.3	18.7	10.4	8.6	17.6	15.9	8.7	6.6	2.2	3.9	6.0	4.2	1.2	1.9	
mean deviation in error <sup>j,k</sup>		2.3	2.4	1.8	2.1	2.1	2.4	2.4	2.7	1.2	2.0	1.6	1.5	0.9	0.9	

<sup>a</sup> All values are in kcal/mol and have been zero-point energy (ZPE) and temperature corrected using unscaled density functional frequencies.<sup>b</sup> Level of theory III: X (X = S-null, S-VWN, etc.)/aug-cc-pVDZ. Level of theory IV: X/aug-cc-pVTZ//X/aug-cc-pVDZ. <sup>c</sup> The symmetry of the acid is given first, and the parenthetical value is for the conjugate base. <sup>d</sup> All experimental values from ref 19a unless otherwise noted. <sup>e</sup> See ref 1.<sup>f</sup> See ref 19d. <sup>g</sup> See ref 19c. <sup>h</sup> See ref 19e. <sup>i</sup> The anion has C<sub>2v</sub> symmetry with the following functionals: S-null, S-VWN, and S-LYP. <sup>j</sup> Error excludes the acidity for H<sub>2</sub> (see ref 20 for justification). <sup>k</sup> The mean deviation error is defined as follows:  $\sum_i |x_i - |\bar{x}||/N$ .

extremely poorly in that the average unsigned errors are 19.5, 9.8, and 17.4 kcal/mol, respectively.<sup>20</sup> The mean deviation errors, however, are relatively small (~2 kcal/mol), indicating that there is a systematic error in the computed proton affinities. This is also apparent in that the computed values are uniformly too small (i.e., too acidic). Excellent linear correlations were obtained for the S-null and S-VWN results by plotting the experimental vs calculated proton affinities (Figures 1 and 2). Although it is undesirable to have to scale the computed results using a linear correlation, if one does this, the average unsigned errors drop to 1.7 and 1.5 kcal/mol, respectively, and all the proton affinities are within 5 kcal/mol of the experimental values.<sup>21</sup> In fact, only three compounds deviate by 4 kcal/mol or more for the S-null/6-31+G(d) results (4.6 (PH<sub>3</sub>), 4.3 (SiH<sub>4</sub>), and 4.1 ((CH<sub>3</sub>)<sub>2</sub>SO) kcal/mol), and there are only two outlying species (4.6 (CH<sub>3</sub>Cl) and 4.1 (HF) kcal/mol) when the VWN local correlation functional is included.<sup>22</sup> The PA of methyl chloride also happens to have an experimental uncertainty of ±3.1 kcal/mol.<sup>1</sup> Interestingly, the results are not as satisfactory when the LYP nonlocal correlation functional is used,<sup>23</sup> and single-point energies with a significantly larger basis set (6-311++G(2df,2pd)) offer no improvement; the average unsigned error drops by about 1 kcal/mol, but the mean deviation error gets worse. The results also do not scale as well.

B-null, B-VWN, and B-LYP all make use of Becke's gradient correction to Slater's local exchange, and the average unsigned

errors in the proton affinities (7.8, 3.6, and 5.3 kcal/mol, respectively) with the 6-31+G(d) basis set are much reduced compared to those results using only Slater's exchange. All three methods, nevertheless, have a significant number of deviant points that differ from experimental results by 5 kcal/mol or more and thus are unreliable.<sup>24</sup> The B-null and B-LYP proton affinities are uniformly too small,<sup>25</sup> and the mean deviation errors are similar to those found with the functionals employing Slater exchange. The correlation between the experimental and calculated data is only modest, and the scaled acidities are not as accurate as the S-null and S-VWN results.<sup>26</sup> The B-VWN proton affinities are more randomly distributed, some being too large and others too small. This leads to a mean deviation error of only 1.5 kcal/mol, but fitting the results does little to improve them.<sup>27</sup> Single-point energies with the 6-311++G(2df,2pd) basis set, unlike the previous results with only local exchange, improve the B-null and B-LYP proton affinities. Although the agreement with experimental results is still unsatisfactory, the scaled results with all three functionals and the larger basis set (in particular, B-LYP) are now quite good; there are few outlying points and the average unsigned errors drop to 1.3–1.7 kcal/mol, which is better than or equal to the 1.7 kcal/mol average experimental uncertainty in the data.<sup>20,28</sup>

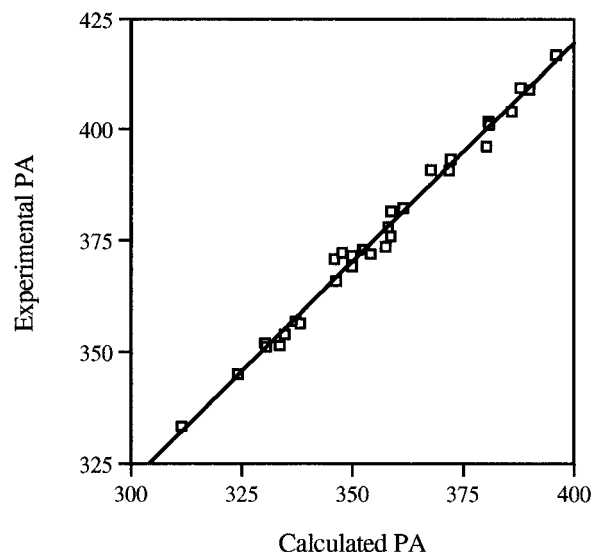
The final functional we examined, B3-PW91, is a "hybrid" functional that combines Hartree–Fock and density functional

**TABLE 3: Comparison of Calculated and Experimental Acidities of Neutral Acids HA (or Proton Affinities of A<sup>-</sup>)<sup>a,b</sup>**

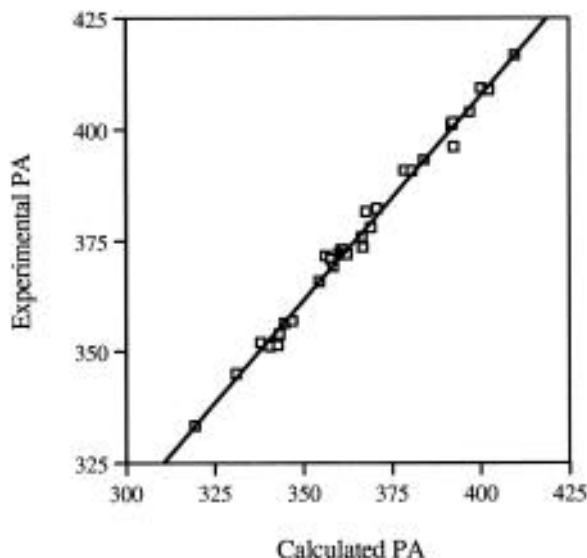
HA	pt grp <sup>c</sup>	MP2/6-31+G(d)/HF/6-31+G(d)	B-VWN III	B-LYP II	B3-PW91 III	G2	expt <sup>d</sup>
CH <sub>4</sub>	<i>T<sub>d</sub></i> ( <i>C<sub>3v</sub></i> )	418.2 <sup>e</sup>	419.5	413.2	417.6	418.5 <sup>f</sup>	416.7 ± 0.7
C <sub>2</sub> H <sub>4</sub>	<i>D<sub>2h</sub></i> ( <i>C<sub>s</sub></i> )	407.7 <sup>e</sup>	410.3	405.1	408.4	408.5 <sup>f</sup>	409.4 ± 0.6
CH <sub>3</sub> F	<i>C<sub>3v</sub></i> ( <i>C<sub>s</sub></i> )	408.7	408.6	404.7	409.6	411.0	409.0 ± 4.0 <sup>g</sup>
NH <sub>3</sub>	<i>C<sub>3v</sub></i> ( <i>C<sub>2v</sub></i> )	402.6	404.9	400.1	405.2	403.9 <sup>f</sup>	404.0 ± 0.4
<i>c</i> -CH <sub>2</sub> N <sub>2</sub>	<i>C<sub>2v</sub></i> ( <i>C<sub>1</sub></i> )	397.5	399.4	394.7	400.0	398.4	401 ± 3 <sup>i</sup>
C <sub>6</sub> H <sub>6</sub>	<i>D<sub>6h</sub></i> ( <i>C<sub>2v</sub></i> )	397.4 <sup>e</sup>	403.7		401.8		401.7 ± 0.5 <sup>j</sup>
H <sub>2</sub>	<i>D<sub>∞h</sub></i> ( <sup>1</sup> <i>S</i> <sub>0</sub> )	443.7	404.0	397.8	398.4	401.5 <sup>f</sup>	400.4 ± .002
CH <sub>3</sub> Cl	<i>C<sub>3v</sub></i> ( <i>C<sub>s</sub></i> )	402.6 <sup>e</sup>	396.6	393.1	396.9	397.9 <sup>f</sup>	396.1 ± 3.1 <sup>8</sup>
CH <sub>3</sub> SCH <sub>3</sub>	<i>C<sub>2v</sub></i> ( <i>C<sub>s</sub></i> )	395.9 <sup>e</sup>	395.5	388.1	395.1		393.2 ± 2.1 <sup>8</sup>
CH <sub>3</sub> CH=CH <sub>2</sub>	<i>C<sub>s</sub></i> ( <i>C<sub>2v</sub></i> )	390.3 <sup>e</sup>	391.2	384.5	389.1	391.6	390.8 ± 2.1
H <sub>2</sub> O	<i>C<sub>2v</sub></i> ( <i>C<sub>∞v</sub></i> )	383.0	389.7	387.2	391.3	389.8 <sup>f</sup>	390.7 ± 0.1
BCP (1)	<i>C<sub>2v</sub></i> ( <i>C<sub>s</sub></i> ) <sup>k</sup>	386.7	387.2	381.2	385.1		386 ± 3 <sup>i</sup>
CH <sub>3</sub> OH	<i>C<sub>s</sub></i> ( <i>C<sub>3v</sub></i> )	376.4	380.0	375.6	380.4	382.6 <sup>f</sup>	381.6 ± 0.7
C <sub>6</sub> H <sub>5</sub> CH <sub>3</sub>	<i>C<sub>1</sub></i> ( <i>C<sub>s</sub></i> )	385.1 <sup>e</sup>		378.1			382.3 ± 0.3 <sup>m</sup>
C <sub>2</sub> H <sub>2</sub>	<i>D<sub>∞h</sub></i> ( <i>C<sub>∞v</sub></i> )	374.3 <sup>e</sup>	380.1	375.9	377.7	377.4 <sup>f</sup>	378.0 ± 0.6
HOOH	<i>C<sub>2</sub></i> ( <i>C<sub>s</sub></i> )	367.3	375.2	372.5	377.1	376.4	375.9 ± 2.1 <sup>8</sup>
pyrazoline (2)	<i>C<sub>s</sub></i> ( <i>C<sub>1</sub></i> )	369.5	375.8	368.8	374.0		375.0 ± 3.0 <sup>n</sup>
pyridazine (3)	<i>C<sub>2</sub></i> ( <i>C<sub>1</sub></i> )	363.7	369.8	363.3	367.7		375.0 ± 3.0 <sup>n</sup>
CH <sub>3</sub> SOCH <sub>3</sub>	<i>C<sub>s</sub></i> ( <i>C<sub>1</sub></i> )	375.1 <sup>e</sup>	380.4	374.3	377.7	377.2	373.5 ± 2.1 <sup>8</sup>
CH <sub>2</sub> N <sub>2</sub>	<i>C<sub>2v</sub></i> ( <i>C<sub>s</sub></i> )	368.5	373.7	368.7	373.0	372.8 <sup>h</sup>	373 ± 3 <sup>o</sup>
CH <sub>3</sub> CN	<i>C<sub>3v</sub></i> ( <i>C<sub>s</sub></i> )	373.6 <sup>e</sup>	375.5	369.1	372.7		372.9 ± 2.1
SiH <sub>4</sub>	<i>T<sub>d</sub></i> ( <i>C<sub>3v</sub></i> )	374.2	376.8	370.4	370.8	373.8 <sup>f</sup>	372.2 ± 2.0
CH <sub>3</sub> CO <sub>2</sub> CH <sub>3</sub>	<i>C<sub>1</sub></i> ( <i>C<sub>1</sub></i> )	372.6 <sup>e</sup>	375.2	369.2	373.7		371.9 ± 2.1 <sup>8</sup>
PH <sub>3</sub>	<i>C<sub>3v</sub></i> ( <i>C<sub>2v</sub></i> )	366.0	372.2	366.3	366.9	368.1 <sup>f</sup>	370.8 ± 2.0
HF	<i>C<sub>∞v</sub></i> ( <sup>1</sup> <i>S</i> <sub>0</sub> )	358.2	367.1	366.6	368.9	370.7 <sup>f</sup>	371.5 ± 0.2
CH <sub>3</sub> COCH <sub>3</sub>	<i>C<sub>2</sub></i> ( <i>C<sub>s</sub></i> )	367.6 <sup>e</sup>	371.6	365.3	369.8		369.1 ± 2.1 <sup>8</sup>
CH <sub>3</sub> CHO	<i>C<sub>s</sub></i> ( <i>C<sub>s</sub></i> )	364.7 <sup>e</sup>	367.8	361.7	366.2	367.4 <sup>f</sup>	365.8 ± 2.2
CH <sub>3</sub> SH	<i>C<sub>s</sub></i> ( <i>C<sub>3v</sub></i> )	353.3	360.4	356.3	356.9	358.0 <sup>f</sup>	356.9 ± 2.2
CH <sub>3</sub> NO <sub>2</sub>	<i>C<sub>1</sub></i> ( <i>C<sub>2v</sub></i> )	353.9 <sup>e</sup>	358.2	352.4	356.5		356.4 ± 2.2 <sup>8</sup>
<i>c</i> -C <sub>5</sub> H <sub>6</sub>	<i>C<sub>2v</sub></i> ( <i>D<sub>5h</sub></i> )	348.6 <sup>e</sup>	359.0	352.3	354.3		353.9 ± 2.2 <sup>8</sup>
H <sub>2</sub> NCN	<i>C<sub>s</sub></i> ( <i>C<sub>s</sub></i> )	344.5	349.4	344.7	347.7	348.4 <sup>h</sup>	352 ± 3 <sup>i</sup>
H <sub>2</sub> S	<i>C<sub>2v</sub></i> ( <i>C<sub>∞v</sub></i> )	347.2	354.6	350.1	350.7	352.1 <sup>f</sup>	351.1 ± 2.0
HCN	<i>C<sub>∞v</sub></i> ( <i>C<sub>∞v</sub></i> )	345.9 <sup>e</sup>	354.4	349.8	351.0	350.4 <sup>f</sup>	351.4 ± 0.5
HCO <sub>2</sub> H	<i>C<sub>s</sub></i> ( <i>C<sub>2v</sub></i> )	335.0	344.6	340.2	343.9	343.5 <sup>f</sup>	345 ± 2
HCl	<i>C<sub>∞v</sub></i> ( <sup>1</sup> <i>S</i> <sub>0</sub> )	325.1	334.2	331.1	331.2	334.1 <sup>f</sup>	333.4 ± 0.03
unsigned mean error <sup>p</sup>		4.3(4.6)	2.2(2.0)	4.0(3.7)	1.4(1.4)	1.4	
mean deviation in error <sup>p,q</sup>		2.6(2.7)	1.3(1.3)	1.5(1.5)	1.1(1.2)	0.8	

<sup>a</sup> All values are in kcal/mol and have been zero-point energy and temperature corrected. Hartree–Fock frequencies and ZPE's were scaled by 0.8929 and 0.9135, respectively (Pople, J. A.; Scott, A. P.; Wong, M. W.; Radom, L. *Isr. J. Chem.* **1993**, 33, 345), while the corresponding density functional values were unscaled. <sup>b</sup> Level of theory II: *X* (*X* = B-VWN, B-LYP, B3-PW91)/6-311++G(2df,2pd)//*X*/6-31+G(d). Level of theory III: *X*/aug-cc-pVDZ//*X*/aug-cc-pVDZ. <sup>c</sup> The symmetry of the acid is given first, and the parenthetical value is for the conjugate base. <sup>d</sup> All experimental values from ref 19a unless otherwise noted. <sup>e</sup> See ref 6. <sup>f</sup> See ref 9. <sup>g</sup> See ref 1. <sup>h</sup> Russell, A. J.; Scott, A. P.; Radom, L. *Aust. J. Chem.* **1995**, 48, 401. <sup>i</sup> See ref 19d. <sup>j</sup> See ref 19c. <sup>k</sup> The acid has *C<sub>s</sub>* symmetry at the HF level. <sup>l</sup> Unpublished data. <sup>m</sup> See ref 19b. <sup>n</sup> Kroeker-Sachs, R. Ph.D. Thesis, University of Minnesota, 1993. <sup>o</sup> See ref 19e. <sup>p</sup> Parenthetical values are the errors resulting from the 23 acids for which G2 value are provided. Note that H<sub>2</sub> has been excluded from the error analysis in both cases (see ref 20 for the justification). <sup>q</sup> The mean deviation error is defined as follows:  $\sum_i^N |x_i - \bar{x}|/N$ .

exchange. As such, it is computationally more expensive than any of the purely DFT methods, but it still scales favorably relative to high-level ab initio calculations. It also gives the best results of the seven density functional levels of theory that were examined. When the 6-31+G(d) basis set is used, 3 compounds (HF, HCl, and (CH<sub>3</sub>)<sub>2</sub>SO) have errors greater than 5 kcal/mol and 11 out of the 32 acidities deviate from experimental results by at least 3 kcal/mol. Better results can be obtained by fitting the proton affinities with a linear regression ( $y(\text{expt}) = 0.911(\text{calc}) + 33.29$ ,  $r^2 = 0.988$ ); the average unsigned error drops from 2.4 to 1.8 kcal/mol and the number of outlying points ( $\geq 3$  kcal/mol from experiment) goes from 11 to 5 (6.1 (HF), 5.3 ((CH<sub>3</sub>)<sub>2</sub>SO), 3.2 (*c*-C<sub>5</sub>H<sub>6</sub>), 3.1 (CH<sub>3</sub>OH), and 3.0 (CH<sub>3</sub>CO<sub>2</sub>CH<sub>3</sub>) kcal/mol). However, if a larger, more flexible basis set is used to calculate the energies (i.e., B3-PW91/6-311++G(2df,2pd)//B3-PW91/6-31+G(d)), the resulting proton affinities are roughly comparable to high-level ab initio results. Six values deviate by 3 kcal/mol or more (5.5 ((CH<sub>3</sub>)<sub>2</sub>SO), 3.7 (H<sub>2</sub>O), 3.6 (HOOH), 3.5 (*c*-CH<sub>2</sub>N<sub>2</sub>), 3.4 (CH<sub>3</sub>Cl), and 3.4 (NH<sub>3</sub>) kcal/mol), and the average unsigned error, 2.0 kcal/mol, is now only slightly more than the 1.7 kcal/mol average experimental uncertainty.<sup>20</sup> Scaling the data ( $y(\text{expt}) = 0.969(\text{calc}) + 9.89$ ,  $r^2 = 0.994$ ) leads to a further improvement in



**Figure 1.** Experimental vs calculated S-null/6-31+G(d) proton affinities. All values are in kcal/mol, and the line ( $y(\text{expt}) = 0.985(\text{calc}) + 24.75$ ,  $r^2 = 0.990$ ) was obtained by a least squares analysis.

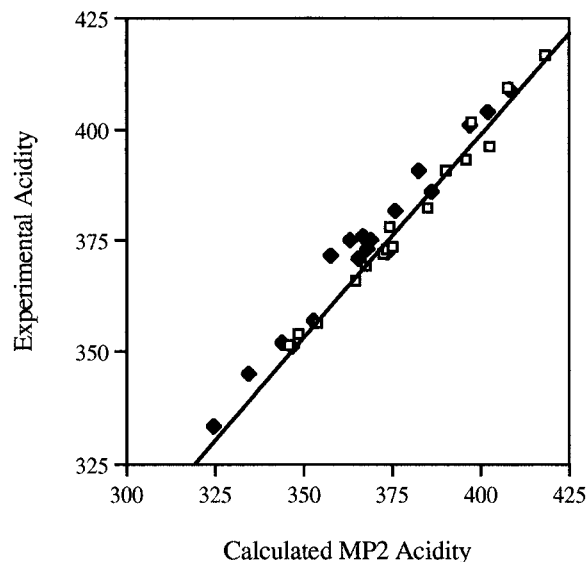


**Figure 2.** Experimental vs calculated S-VWN/6-31+G(d) proton affinities. All values are in kcal/mol, and the line ( $y(\text{expt}) = 0.925(\text{calc}) + 37.13$ ,  $r^2 = 0.991$ ) was obtained by a least squares analysis.

that the average error drops to 1.1 kcal/mol, and the worst errors (3.8 ( $\text{NH}_2\text{CN}$ ), 3.6 ( $(\text{CH}_3)_2\text{SO}$ ), and 3.3 ( $\text{PH}_3$ ) kcal/mol) are similar to the G2 results (3.6, 3.7, and 2.7 kcal/mol, respectively). The experimental uncertainties for these compounds also happen to be in the 2–3 kcal/mol range.

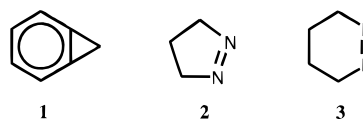
In order to examine basis-set dependence further, Dunning's correlation-consistent basis sets were used to recalculate all the proton affinities with each of the seven combinations of DFT functionals.<sup>12</sup> Fully optimized DFT/aug-cc-pVDZ structures served as the basis for single-point energies with the aug-cc-pVTZ bases (Table 2). The results obtained closely parallel those found with the Pople basis sets.<sup>29</sup> It is worth noting the following. (1) B-VWN provides reasonable proton affinities with the aug-cc-pVDZ basis set. The average unsigned error is 2.2 kcal/mol, the mean deviation error is 1.2 kcal/mol, but eight values differ from experiment by  $\geq 3$  kcal/mol (6.9 ( $(\text{CH}_3)_2\text{SO}$ ), 5.1 ( $c\text{-C}_5\text{H}_6$ ), 4.6 ( $\text{SiH}_4$ ), 4.4 ( $\text{HF}$ ), 3.5 ( $\text{CH}_3\text{SH}$ ), 3.5 ( $\text{H}_2\text{S}$ ), 3.3 ( $\text{CH}_3\text{CO}_2\text{CH}_3$ ), and 3.0 ( $\text{HCN}$ ) kcal/mol). (2) *B3-PW91/aug-cc-pVDZ proton affinities rival those obtained via high-level ab initio calculations.* The average unsigned error is only 1.2 kcal/mol, the mean deviation error is 0.9 kcal/mol, and there are only three outlying values  $\geq 3$  kcal/mol (4.3 ( $\text{NH}_2\text{CN}$ ), 4.2 ( $(\text{CH}_3)_2\text{SO}$ ), and 3.9 ( $\text{PH}_3$ ) kcal/mol).<sup>30</sup> (3) Proton affinities calculated with S-null, S-VWN, S-LYP, B-null, and B-LYP using the aug-cc-pVDZ basis set are linearly related to the experimental values, and this can be used to significantly reduce the computed errors.<sup>31</sup> This is especially the case for the B-LYP results, which upon scaling are as good as the B3-PW91 proton affinities (i.e., the average unsigned error drops from 6.0 to 1.3 kcal/mol and there are only three points that deviate by more than 3 kcal/mol (4.7 ( $(\text{CH}_3)_2\text{SO}$ ), 4.2 ( $\text{HF}$ ), and 3.7 ( $\text{NH}_2\text{CN}$ ) kcal/mol). (4) It may be possible to get comparable results between the aug-cc-pVDZ and 6-31+G(d) basis sets by adding a set of p-polarization functions to hydrogen in the latter case. Conversely, the cost of using the former basis set might be reduced without affecting the results by eliminating the diffuse functions of hydrogen.<sup>32</sup>

Saunders has recently shown that MP2(fc)/6-31+G(d) acidities using HF/6-31+G(d) geometries provide reliable results for a variety of carbon acids at a modest computational cost.<sup>6</sup> Since this is a very practical level of theory that can be applied to large numbers of compounds, we felt it was important to test this approach further with a bigger data set, which includes non-



**Figure 3.** Experimental vs calculated MP2/6-31+G(d)/HF/6-31+G(d) acidities from the data in Table 3. The line is from a linear least squares fit of Saunders' 16 carbon acids (see text for details).

carbon acids. Consequently, MP2 acidities for all the compounds in Table 1, which includes all 16 of Saunders' carbon acids, were examined along with three additional compounds of interest to us (**1**, **2**, and **3**, Table 3). The results are



surprisingly poor with an average unsigned error of 4.3 kcal/mol, a mean deviation of nearly 3 kcal/mol, and almost half of the values differ from experiment by at least 4 kcal/mol (more than a third are in error by  $\geq 5$  kcal/mol). A linear least squares fit of the experimental and computed proton affinities gives a relatively poor correlation ( $y(\text{expt}) = 0.882(\text{calc}) + 47.15$ ,  $r^2 = 0.969$ ), in contrast to the better results reported by Saunders for the 16 carbon acids ( $y(\text{expt}) = 0.913(\text{calc}) + 33.50$ ,  $r^2 = 0.981$ ).<sup>33</sup> This discrepancy is readily explained by looking at Figure 3 where it can be seen that all the additional compounds' acidities, except for silane and **1**, are above the line going through the original 16 acids. Moreover, this difference clearly illustrates one of the problems with scaling computed results: the scaled answer can change depending on what compounds are used to establish the linear fit. If one looks more closely at the MP2 acidities, it appears that hydrogen halides and oxy acids are described particularly poorly, and the overall results are less reliable and less consistent than the best DFT data.

Given the poor performance of the MP2 method, it seemed worthwhile to compare the experimental acidities in our data set to a well-established and highly accurate ab initio procedure. Therefore, 24 G2 acidities, including a few new values, were compiled in Table 3. All the results are within the target accuracy of  $\pm 2$  kcal/mol except for  $(\text{CH}_3)_2\text{SO}$ ,  $\text{NH}_2\text{CN}$ ,  $\text{PH}_3$ , and  $c\text{-CH}_2\text{N}_2$ , which differ from the experimental acidities by 3.7, 3.6, 2.7, and 2.6 kcal/mol, respectively. If one considers the experimental uncertainties for these compounds,  $\pm 2$  ( $(\text{CH}_3)_2\text{SO}$  and  $\text{PH}_3$ )<sup>1</sup> and  $\pm 3$  ( $\text{NH}_2\text{CN}$  and  $c\text{-CH}_2\text{N}_2$ )<sup>19d</sup> kcal/mol, then it is clear that the overall agreement between G2 theory and experiment is excellent. It also is clear that the best DFT results, B-VWN/aug-cc-pVDZ and B3-PW91/aug-cc-pVDZ, are nearly

as good as the G2 acidities, but the former calculations are computationally much less demanding.

## Conclusions

Slater's local exchange functional provides very poor proton affinities, even if a local (VWN) or nonlocal (LYP) correlation functional is included. The results, however, are systematically too low and can be improved dramatically by linearly fitting the data. In this case the best acidities were obtained using S-null/6-31+G(d) and S-VWN/6-31+G(d) geometries and energies.

Becke's gradient correction to Slater's exchange functional leads to better proton affinities, and reasonable agreement with experimental results can be obtained in most cases by including a local correlation functional (e.g., B-VWN/6-31+G(d)).<sup>34</sup> More accurate acidities can be obtained by scaling  $X/6-311++G(2df,2pd)$  energies, where  $X$  = B-null, B-VWN, or B-LYP. The best results were obtained in this case by using a nonlocal correlation functional (i.e., B-LYP).

Our best results were obtained using the B3-PW91 "hybrid" functional. In this case the proton affinities obtained with a flexible basis set (i.e., B3-PW91/6-311++G(2df,2pd))/B3-PW91/6-31+G(d) or, even better, B3-PW91/6-31+G(d) rival those computed via high-level (G2) *ab initio* calculations but at a fraction of the computational cost. Once again, the agreement with experimental results can be improved somewhat by linearly scaling the calculated acidities, but we do not recommend this.

Dunning's split-valence, double- and triple-zeta, correlation-consistent basis sets provide results similar to those of their Pople-type analogues. In some cases one is preferred over the other, but in general the biggest difference is between aug-cc-pVDZ and 6-31+G(d), with the former yielding somewhat closer agreement with experimental results.

*Ab initio* calculations at a modest level (MP2(fc)/6-31+G(d))/HF/6-31+G(d)) provide reasonable acidities for hydrocarbons owing to a fortuitous cancellation of errors: a relatively small basis set far from the Hartree-Fock limit and an incomplete accounting of electron correlation. However, the results are surprisingly poor when a wider range of compounds is considered. They also are not as reliable as the density functional results.

**Acknowledgment.** Support from the National Science Foundation, the donors of the Petroleum Research Foundation, as administered by the American Chemical Society, and the Minnesota Supercomputer Institute are gratefully acknowledged.

## References and Notes

- (1) Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G. *J. Phys. Chem. Ref. Data, Suppl. 1* **1988**, 17, or the slightly updated form available on a personal computer, *NIST Negative Ion Energetics Database*, Version 3.00; NIST Standard Reference Database, 19B; NIST: Gaithersburg, MD, 1993.
- (2) Unpublished results.
- (3) (a) Dewar, M. J. S.; Dieter, K. M. *J. Am. Chem. Soc.* **1986**, 108, 8075. (b) Kass, S. R. *J. Comput. Chem.* **1989**, 11, 94.
- (4) (a) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, 94, 7221. (b) Smith, B. J.; Radom, L. *J. Phys. Chem.* **1991**, 95, 10549. (c) Petersson, G. A.; Al-Laham, M. A. *J. Chem. Phys.* **1991**, 94, 6081. (d) Petersson, G. A.; Tensfeldt, T. G.; Montgomery, J. A., Jr. *J. Chem. Phys.* **1991**, 94, 6091. (e) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, 90, 5622. (f) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. *J. Chem. Phys.* **1988**, 89, 2193. (g) DeFrees, D. J.; McLean, A. D. *J. Comput. Chem.* **1986**, 7, 321. (h) Petersson, G. A.; Yee, A. K.; Bennett, A. *J. Chem. Phys.* **1985**, 83, 5105. (i) Petersson, G. A.; Braunstein, M. *J. Chem. Phys.* **1985**, 83, 5129.
- (5) Ochterski, J. W.; Petersson, G. A.; Wiberg, K. B. *J. Am. Chem. Soc.* **1995**, 117, 11299.
- (6) Saunders, W. H., Jr. *J. Phys. Org.* **1994**, 7, 268.
- (7) (a) Paar, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989. (b) Jones, R. O.; Gunnarsson, O. *Rev. Mod. Phys.* **1989**, 61, 689. (c) Dreizler, R. M.; Gross, E. K. U. *Density Functional Theory*; Springer: Berlin, 1990. (d) Ziegler, T. *Chem. Rev.* **1991**, 91, 651. (e) Labanowski, J. K.; Andzelm, J. W., Eds. *Density Functional Methods in Chemistry*; Springer-Verlag: New York, 1991. (f) Bartolotti, L. J.; Flurchick, K. *Rev. Comput. Chem.* **1996**, 7, 187. (g) St-Amant, A. *Rev. Comput. Chem.* **1996**, 7, 217.
- (8) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1992**, 97, 7846.
- (9) Smith, B. J.; Radom, L. *Chem. Phys. Lett.* **1995**, 245, 123.
- (10) Becke, A. D. *J. Chem. Phys.* **1993**, 98, 5648.
- (11) (a) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, 54, 724. (b) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, 56, 2257. (c) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, 28, 213. (d) Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, 27, 209. (e) Gordon, M. S. *Chem. Phys. Lett.* **1980**, 76, 163. (f) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, 72, 5639. (g) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, 4, 294. (h) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, 80, 3265.
- (12) (a) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, 90, 1007. (b) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, 96, 6796.
- (13) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Wong, M. W.; Foresman, J. R.; Robb, M. A.; Head-Gordon, M.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzales, C. A.; Martin, R. C.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *GAUSSIAN-92/DFT*; Gaussian Inc.: Pittsburgh, PA, 1993.
- (14) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, Revisions A, B, and C; Gaussian, Inc.: Pittsburgh, PA, 1995.
- (15) Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comput. Chem.* **1993**, 14, 1347.
- (16) (a) Slater, J. C. *Quantum Theory of Molecules and Solids, Vol. 4: The Self-Consistent Field for Molecules and Solids*; McGraw-Hill: New York, 1974. (b) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, 58, 1200. (c) Lee, C.; Yang, W.; Paar, R. G. *Phys. Rev. B* **1988**, 37, 785. (d) Becke, A. D. *Phys. Rev. A* **1988**, 38, 3098. (e) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, 45, 13244.
- (17) It is also worth noting that density functional theory often gives positive eigenvalues for the highest occupied molecular orbitals of anions. Although this is troubling, it does not necessarily diminish the utility of DFT for calculating proton affinities of anionic systems. Similar studies involving anions, specifically the calculation of electron affinities, give good results. See: (a) Galbraith, J. M.; Schaefer, H. F., III *J. Chem. Phys.* **1996**, 105, 862. (b) Gill, P. M. W.; Johnson, B. G.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1992**, 197, 499. (c) Ziegler, T.; Gutsev, G. L. *J. Comput. Chem.* **1992**, 13, 70.
- (18) Allyl anion within the  $C_{2v}$  point group is a minimum at the HF level but becomes a second-order transition structure when second-order Møller-Plesset perturbation theory is used. We have used the  $C_{2v}$  structure, nevertheless, in calculating the G2 acidity. In part, this was done because  $\Delta G_2(C_{2v} - C_2) = -3.5$  kcal/mol (i.e., the  $C_{2v}$  structure is more stable), where the ZPE's have been omitted, since we were unable to locate a  $C_2$  stationary point at the HF level.
- (19) (a) Berkowitz, J.; Ellison, G. B.; Gutman, D. *J. Phys. Chem.* **1994**, 98, 2744. (b) Davico, G. E.; Bierbaum, V. M.; DePuy, C. H.; Ellison, G. B. *J. Am. Chem. Soc.*, submitted for publication. (c) Davico, G. E.; Bierbaum, V. M.; DePuy, C. H.; Ellison, G. B.; Squires, R. R. *J. Am. Chem. Soc.* **1995**, 117, 2590. (d) Kroeker, R. L.; Kass, S. R. *J. Am. Chem. Soc.* **1990**, 112, 9024. (e) DePuy, C. H.; Van Doren, J. M.; Gronert, S.; Kass, S. R.; Motell, E. L.; Ellison, G. B.; Bierbaum, V. M. *J. Org. Chem.* **1989**, 54, 1846.
- (20) The acidity of molecular hydrogen has been omitted from all the statistical analyses in this paper because  $H^-$  is poorly described by most common basis sets.
- (21) The average experimental uncertainty for the acids in Tables 1 and 2 is 1.7 kcal/mol.
- (22) Two compounds are in error by 3–4 kcal/mol using the S-null results (3.8 ( $CH_3Cl$ ) and 3.2 ( $CH_2=CHCH_3$ ) kcal/mol), and three species deviate by this much for the S-VWN proton affinities (3.7 ( $CH_3OH$ ), 3.6 ( $(CH_3)_2SO$ ), and 3.5 ( $HCN$ )).

(23) A plot of experimental vs calculated proton affinities gives a modest linear correlation ( $y(\text{expt}) = 0.925(\text{calc}) + 44.24$ ,  $r^2 = 0.975$ ), but several points still have significant errors after scaling (i.e., 11.8 (SiH<sub>4</sub>), 8.0 (CH<sub>3</sub>C<sub>6</sub>H<sub>5</sub>), 5.2 (CH<sub>3</sub>Cl), and 4.3 (HCN) kcal/mol).

(24) B-null, B-VWN, and B-LYP have 24, 6, and 14 proton affinities, respectively, which differ from experimental values by  $\geq 5$  kcal/mol, and these numbers increase to 28, 11, and 20 if one sets the limit to  $\geq 4$  kcal/mol.

(25) The only exception to this is the B-LYP acidity for (CH<sub>3</sub>)<sub>2</sub>SO, which is 0.9 kcal/mol too big.

(26) A least squares fit affords  $y(\text{expt}) = 1.010(\text{calc}) + 4.13$ ,  $r^2 = 0.978$ , and  $y(\text{expt}) = 0.959(\text{calc}) + 20.48$ ,  $r^2 = 0.983$ , for B-null and B-LYP proton affinities, respectively. In the former case 3 acidities deviate by  $\geq 5$  kcal/mol, 5 by  $\geq 4$  kcal/mol, and 12 by  $\geq 3$  kcal/mol. The corresponding numbers for the latter functional are 2, 3, and 5, respectively.

(27) About one-fourth of the B-VWN proton affinities are too small, and three-fourths of the values are too large. The data can be described with the line  $y(\text{expt}) = 0.939(\text{calc}) + 20.90$ ,  $r^2 = 0.975$ , but the results are only slightly improved and there still are an unacceptable number of outlying points (i.e., 3 at  $\geq 5$  kcal/mol, 7 at  $\geq 4$  kcal/mol, and 10 at  $\geq 3$  kcal/mol).

(28) A direct comparison between theory and experiment reveals that the B-null, B-VWN, and B-LYP/6-311++G(2df,2pd) proton affinities have 23, 12, and 14 values that are in error by  $\geq 4$  kcal/mol (22 and 21 of the B-VWN and B-LYP values, respectively, are off by  $\geq 3$  kcal/mol). Linear correlations (B-null,  $y(\text{expt}) = 1.079(\text{calc}) - 23.01$ ,  $r^2 = 0.990$ ; B-VWN,  $y(\text{expt}) = 1.000(\text{calc}) - 3.74$ ,  $r^2 = 0.988$ ; B-LYP,  $y(\text{expt}) = 1.019(\text{calc}) - 3.51$ ,  $r^2 = 0.993$ ) lead to average unsigned errors of 1.6 (B-null), 1.7 (B-VWN), and 1.3 (B-LYP) kcal/mol, with no egregious deviations from experimental values. The largest errors after scaling are as follows. B-null: 6.0 ((CH<sub>3</sub>)<sub>2</sub>SO), 4.5 (NH<sub>2</sub>CN), and 3.5 (*c*-CH<sub>2</sub>N<sub>2</sub>) kcal/mol. All the other errors are  $< 3$  kcal/mol. B-VWN gives 5.3 ((CH<sub>3</sub>)<sub>2</sub>SO), 4.7 (NH<sub>2</sub>CN), 4.2 (SiH<sub>4</sub>), and 3.5 (HF) kcal/mol. B-LYP gives 4.4 ((CH<sub>3</sub>)<sub>2</sub>SO) and 4.3 (NH<sub>2</sub>CN) kcal/mol.

(29) The aug-cc-pVDZ basis set is only slightly larger for heavy (non-hydrogenic) atoms than the 6-31+G(d) basis set, whereas the aug-cc-pVTZ basis set is approximately 50% bigger than the 6-311++G(2df,2pd) basis set for non-hydrogenic atoms.

(30) It is likely that other parametrized functionals (e.g., B3-LYP and B3-P86) will also perform well.

(31) Linear fits of the computed proton affinities to the experimental data are as follows:  $y(\text{expt}) = 1.055(\text{calc}) + 0.83$ ,  $r^2 = 0.986$  (S-null/aug-cc-pVDZ or, more briefly, S-null/III);  $y(\text{expt}) = 1.048(\text{calc}) + 1.68$ ,  $r^2 = 0.983$  (S-null/aug-cc-pVTZ or, more briefly, S-null/IV);  $y(\text{expt}) = 0.985(\text{calc}) + 15.72$ ,  $r^2 = 0.990$  (S-VWN/III);  $y(\text{expt}) = 0.974(\text{calc}) + 18.23$ ,  $r^2 = 0.988$  (S-VWN/IV);  $y(\text{expt}) = 1.002(\text{calc}) + 16.87$ ,  $r^2 = 0.986$  (S-LYP/III);  $y(\text{expt}) = 0.993(\text{calc}) + 18.40$ ,  $r^2 = 0.984$  (S-LYP/IV);  $y(\text{expt}) = 1.085(\text{calc}) - 22.45$ ,  $r^2 = 0.988$  (B-null/III);  $y(\text{expt}) = 1.069(\text{calc}) - 19.03$ ,  $r^2 = 0.981$  (B-null/IV);  $y(\text{expt}) = 1.005(\text{calc}) - 3.24$ ,  $r^2 = 0.988$  (B-VWN/III);  $y(\text{expt}) = 0.987(\text{calc}) + 1.42$ ,  $r^2 = 0.981$  (B-VWN/IV);  $y(\text{expt}) = 1.026(\text{calc}) - 3.63$ ,  $r^2 = 0.993$  (B-LYP/III);  $y(\text{expt}) = 1.011(\text{calc}) - 0.14$ ,  $r^2 = 0.987$  (B-LYP/IV);  $y(\text{expt}) = 0.971(\text{calc}) + 10.97$ ,  $r^2 = 0.994$  (B3-PW91/III);  $y(\text{expt}) = 0.977(\text{calc}) + 7.36$ ,  $r^2 = 0.993$  (B3-PW91/IV). When the calculated acidities are scaled, the average unsigned errors are reduced to 2.0 (S-null/III), 2.2 (S-null/IV), 1.8 (S-VWN/III), 2.0 (S-VWN/IV), 2.1 (S-LYP/III), 2.4 (S-LYP/IV), 1.8 (B-null/III), 2.2 (B-null/IV), 1.8 (B-VWN/III), 2.2 (B-VWN/IV), 1.3 (B-LYP/III), 1.8 (B-LYP/IV), 1.1 (B3-PW91/III), and 1.2 (B3-PW91/IV) kcal/mol.

(32) Del Bene, J. E. *J. Phys. Chem.* **1993**, 97, 107.

(33) The given equation is slightly different from the one originally reported by Saunders (ref 6),  $y(\text{expt}) = 0.915(\text{calc}) + 32.54$ ,  $r^2 = 0.978$ , owing to small differences in the experimental data that was used.

(34) B-VWN calculations are also faster than those carried out using B-LYP.

JP961557X