J. Am. Chem. Soc. 1988, 110, 91-96

Ab Initio Calculation of Brønsted Acidities

Michele R. F. Siggel,[1a-c] **T. Darrah Thomas,***[1b] **and Leif J. Saethre**[1a]

Contribution from the Department of Chemistry, Institute of Mathematical and Physical Sciences, University of Tromsø, N-9001 Tromsø, Norway, and the Department of Chemistry, Oregon State University, Corvallis, Oregon 97331. *Received May 4, 1987*

Abstract: Ab initio calculations that have a basis set capable of near-Hartree–Fock accuracy and an adequate treatment of correlation give gas-phase acidities that are in excellent agreement with experiment. Acidities (ΔH°_{298}) for methane and formic acid calculated at the MP4(SDTQ)/6-311++G(2d,p) level are within 0.09 eV (2 kcal/mol) of the experimental values. Acidities have also been calculated at lower ab initio levels (3-21+G and 6-31+G) for a range of compounds, including hydrocarbons, alcohols, and carboxylic acids. These calculations, as well as others from the literature that lack either an adequate basis set or suitable treatment of electron correlation, give acidities that differ systematically from the experimental values. In particular, the theoretical results overestimate the acidity differences between compounds. However, the theoretical acidities show good linear correlations with the experimental values, and these correlations can be used in conjunction with theory to give estimates of experimental acidities. This procedure is applied to isopropenyl and allyl alcohols to give estimates of their acidities and of the keto–enol energy difference in acetone.

The Brønsted acidity, which is the ability of a molecule to give up a proton, is a quantity of fundamental chemical interest. Considerable effort has been devoted to the measurement of acidities (both in solution and in the gas phase), to qualitative interpretation of the relative strengths of acids, and to theoretical calculations of acidities.

Substituents may affect acidities either by establishing a charge distribution in the neutral molecule that favors proton removal or by stabilization of the anion through charge delocalization after the removal of the proton. The traditional view of organic acids is that their relative strengths are largely determined by effects that stabilize the anion, such as resonance delocalization. However, recent experimental and theoretical results presented by Siggel and Thomas[2] call this view into question and show, for instance, that the greater acidity of acetic acid relative to isopropyl alcohol is almost entirely due to the effects of the charge distribution in the neutral molecule.

To help establish a firm basis for understanding acidity we have calculated gas-phase acidities for a number of molecules using ab initio theory. One goal has been to find out what level of calculation is necessary to provide reasonably accurate values of both absolute and relative acidities. Since such calculations for large and chemically interesting molecules are prohibitively expensive at present, a second goal has been to determine whether lower level calculations can be calibrated against known experimental acidities to provide a relatively simple way to predict acidities of substances for which it is difficult or impossible to measure the acidity.

The acidity may be expressed as either the equilibrium constant, the pK value, or the free-energy change, ΔG°_T, for the reaction

$$\mathrm{RH} \rightarrow \mathrm{R}^- + \mathrm{H}^+ \qquad (1)$$

In considering gas-phase acidities, it has been common, however, to quote ΔH°_T for this reaction. The principal difference between the magnitude of ΔH°_T and that of ΔG°_T arises because of the entropy of the hydrogen ion,[3] which cancels out in a comparison of relative acidities. From a theoretical point of view, one most easily calculates the difference in energy, ΔE°_{eq}, between the energy minimum for RH and that for R$^-$. Corrections ranging between 0.2 and 0.4 eV (5 and 10 kcal/mol) for the zero-point energy give ΔE°_0 (and ΔH°_0). Further corrections of about 0.1 eV (2 kcal/mol) give a theoretical value of ΔH°_T for comparison with the experimental value.

The results of a variety of ab initio calculations of acidities have been reported.[4a,b,5] These are usually presented along with the experimental results (both with and without the corrections needed to get ΔH°_T from ΔE°_{eq}, or vice versa). However, these presentations have usually been concerned primarily with the details of the calculations and have lacked significant discussion of the degree of agreement between theory and experiment.

A useful way to consider the agreement between experiment and theory is to calculate the linear regression between the theoretical and experimental quantities and to calculate the mean and root-mean-square differences between the experimental and theoretical quantities. The degree of agreement between the two is then reflected in the slope and intercept of the correlation line, the mean difference, the root-mean-square difference, and the correlation coefficient, r^2. The last two depend on the scatter of the points about either a reference line of unit slope and zero intercept or the correlation line. A slope different from unity implies systematic deviations between experiment and theory and that the theoretical values of relative acidities will be systematically either too high or too low. A mean difference that is not zero reflects an overall bias in the absolute values of the calculated acidity.

Previous ab Initio Calculations of Acidities and Proton Affinities

In Table I we have summarized the results of comparing theoretical and experimental acidities from some of the studies that have been reported. In each study a single basis set has been used to calculate the acidities of a range of compounds. Included here are the slopes of the lines that describe a linear regression between experiment and theory, the correlation coefficient, r^2, and the mean and root-mean-square differences between experiment and theory.

The most striking feature of these results is that, although most calculations show a good linear relationship between experiment and theory, the slope of the line is invariably significantly greater than unity, except for the most sophisticated calculations (8, 9 in Table I). In particular, all but one[6] of the calculations at the

---

[1] (a) University of Tromsø. (b) Oregon State University. (c) Present address: Department of Chemistry, University of California, Berkeley, CA 94720.

[2] Siggel, M. R.; Thomas, T. D. *J. Am. Chem. Soc.* **1986**, *108*, 4360.

[3] Bartmess, J. E.; Scott, J. A.; McIver, R. T., Jr. *J. Am. Chem. Soc.* **1979**, *101*, 6046.

[4] (a) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley and Sons: New York, 1986; (a) p 314, (b) p 313, (c) p 315, (d) p 311, (e) p 312, (f) p 260.

[5] (a) Gordon, M. S.; Davis, L. P.; Burggraf, L. W.; Damrauer, R. *J. Am. Chem. Soc.* **1986**, *108*, 7889. (b) Lee, T. J.; Schaefer, H. F., III *J. Chem. Phys.* **1985**, *83*, 1784. (c) Chandrasekhar, J.; Andrade, J. G.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1981**, *103*, 5609. (d) Gao, J.; Garner, D. S.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1986**, *108*, 4784.

[6] The slope of 1.07 (entry 5b of Table I) results only if HCN is omitted from the correlation. Lee and Schaefer (ref 5b) have noted that the HCN theoretical value appears to be anomalously low.

0002-7863/88/1510-0091$01.50/0

Table I. Summary of Theoretical Calculations of Acidities and Proton Affinities

	basis set[a]	slope[b]	correl[b] coeff r^2	mean[c] deviation	root-mean[c] square deviation	ref and comment
		Acidities				
1	RHF/3-21+G//3-21+G	1.42	0.943	0.11	0.51	*d, e*
2	RHF/6-31+G*//3-21G	1.25	0.993	0.03	0.29	*f, g*
3	RHF/6-31++G(d,p)//6-31G*	1.16	0.980	0.10	0.27	*h, i*
4	RHF/6-311++G(3df,2pd)//6-31G*	1.12	0.985	0.13	0.24	*h, i*
5a	near-Hartree–Fock	1.18	0.993	0.23	0.30	*j, k*
5b		1.07	0.993	0.32	0.33	*j, k, l*
6a	MP2/4-31+G//4-31+G	1.19	0.943	−0.39	0.55	*m, g*
6b		1.16	0.994	−0.19	0.30	*m, g, n*
6c		1.15	0.972	0.20	0.30	*m, e, n*
7	MP2/6-31+G*//6-31G*	1.10	0.964	−0.05	0.18	*o, e*
8	MP4(SDTQ)/6-31++G(d,p)//6-31G*	1.03	0.988	0.04	0.14	*h, i*
9	MP4(SDTQ)/6-311++G(3df,2pd)//6-31G*	0.992	0.999	−0.04	0.05	*h, i, p*
		Proton Affinities				
10a	MP4/6-31G**//6-31G*	0.98	0.985	0.25	0.32	*q, g*
10b		1.00	0.984	0.51	0.55	*q, e*
11	RHF/6-31G*//6-31G*	1.04	0.985	0.26	0.33	*q, g*
12	RHF/3-21G//3-21G	1.05	0.891	0.03	0.48	*r, g*

[a] The nomenclature for identifying Gaussian basis sets is described in ref 4, pp 68–87. We have used the notation given by the authors reporting the results. [b] Slope and correlation coefficient for least-squares fit of the theoretical to the experimental results. [c] Mean and root-mean-square deviation, theory minus experiment (eV). [d] Reference 4a; 13 assorted molecules. [e] Comparison between $\Delta E°_{eq}$(calcd) and $\Delta H°_{298}$(exptl). [f] Reference 4b; 10 assorted molecules. [g] Comparison between $\Delta E°_{eq}$(calcd) and $\Delta E°_{eq}$(exptl). [h] Reference 5a; 9 binary hydrides. [i] Comparison between $\Delta H°_0$(calcd) and $\Delta H°_{298}$(exptl). See ref 8. [j] Reference 5b; 4 binary hydrides plus HCN. [k] Comparison between $\Delta H°_{298}$(calcd) and $\Delta H°_{298}$(exptl). [l] Omitting results for HCN. [m] Reference 5c. 6a: 7 assorted molecules. 6b: 5 molecules. 6c: 12 molecules. [n] Omitting results for HF and H_2O. [o] References 5d; 11 assorted molecules. [p] Reference 8. [q] Reference 4c. 10a and 11: 10 assorted molecules. 10b: 12 molecules. [r] Reference 4d; 11 assorted molecules.

RHF level give slopes that are 10% or more greater than 1.0. This is true even for those calculations with large enough basis sets that they may be expected to approach Hartree–Fock accuracy. Furthermore, inclusion of electron correlation in the calculations does not solve this problem except for the most extensive basis sets that have been studied (8, 9).

The failure to achieve unit slope in these correlations does not depend on whether the experimental values of $\Delta H°_{298}$ are corrected for zero-point energy and to 0 K. Calculations 6b and 6c show the effects of making the comparison with and without the corrections. The slopes are nearly the same and the only significant difference is the mean deviation of theory from experiment, which is 0.39 eV (9.0 kcal/mol) more positive without the corrections than with.

If there is a sufficiently flexible basis set and adequate treatment of correlation, then there is excellent agreement between theory and experiment. The acidities of nine hydrides have been calculated by Gordon, Davis, Burggraf, and Damrauer[5a] using 6-31++G(d,p) and 6-311++G(3df,2pd) basis sets[7] (3, 4, 8, 9 in Table I). They have considered the effects of electron correlation through MP4(SDTQ). For the highest level calculations (9) the mean difference between theory and experiment is 0.04 eV (1 kcal/mol) and the root-mean-square deviation is 0.05 eV. The linear regression line between the experimental and theoretical results has a slope of 0.992 ($r^2 = 0.999$). Thus, at this level there is excellent agreement between theory and experiment, with no systematic deviations.[8] Their results at the MP4(SDTQ)/6-31++G(d,p) level (8) are slightly inferior to those with the more extensive basis set. At the SCF level (3, 4) the agreement is worse, and there is a systematic deviation between theory and experiment.

From these results it is apparent that calculation of accurate gas-phase acidities requires a high-quality basis set and the inclusion of electron correlation. If either of these is missing, there is a systematic deviation between calculation and experiment, with the calculation tending to overestimate relative acidities. (It is also necessary to apply corrections for zero-point energy and the change between 0 K and the temperature appropriate to the experiment.)

Although calculated acidities deviate in a systematic way from the measured values, except with calculations done at a very high level, the situation is not so bad for the calculation of proton affinities of neutral molecules[4cd] (10–12 in Table I). In this case, the calculations do not involve the same difficulties that are encountered in calculating the total energies of anions, which require a diffuse basis set.

Proton affinities calculated at the MP4/6-31G**//6-31G* level[4c] (10a of Table I) show a good linear correlation with experimental values corrected to 0 K. The root-mean-square difference between experiment and theory is 0.32 eV. The only systematic deviation between theory and experiment is that the theoretical values average 0.25 eV higher than the experimental ones. Thus, the discrepancies between theory and experiment for relative proton affinities calculated at this level are quite small and even the absolute values are reasonably good.

Even without including the effects of electron correlation, proton affinities calculated with the 6-31G** basis set give similar agreement between theory and experiment (11). The average difference between experiment and theory is 0.26 eV and the root-mean-square difference is 0.33 eV. These are insignificantly different from those obtained at the MP4 level. The linear correlation between the RHF results and the experimental proton affinities (corrected to 0 K) has a slope of 1.04, which is considerably closer to unity than is found for most of the correlations for acidity. Moreover, the correction of the values of $\Delta H°_{298}$ to $\Delta E°_0$ is not essential in establishing this slope; the correlation between the MP4 calculations and the measured values of $\Delta H°$ (10b), uncorrected, has a slope of 1.00. This correction is, however, important in calculating the absolute values of the proton affinity. Without it, the average difference between experiment and theory is 0.51 eV.

The choice of basis set affects the overall quality of the agreement between experiment and theory, but it does not show any systematic effects. Thus, proton affinities calculated at the 3-21G//3-21G level (12) correlate linearly with the experimental values, corrected to 0 K, with a slope of 1.05 (essentially the same as found for the more extensive basis set mentioned above) but

---

(7) The nomenclature for identifying Gaussian basis sets is discussed in ref 4, pp 68–87.

(8) This excellent agreement may be either better than indicated or partly fortuitous because Gordon et al. have neglected the corrections between $\Delta H°_0$ and $\Delta H°_{298}$, which are of the order of 0.06 eV. This is approximately equal to the mean deviation between theory and experiment and is in the right direction. In addition, they have ignored the observation by Hehre, Radom, Schleyer, and Pople (ref 4f) that the zero-point-energies calculated from GAUSSIAN 82 are too high by about 10%.

(9) See ref 8 of Chandrasekhar et al. (ref 5c).

Table II. Energies Calculated at the 6-311++G(2d,p)//6-311++G(2d,p) Level[a]

	methane	methyl anion	formic acid	formate ion
RHF	-40.21041	-39.51968	-188.83147	-188.26233
MP2	-40.38826	-39.71053	-189.39935	-188.84473
MP3	-40.40717	-39.72420	-189.39625	-188.83229
MP4(DQ)	-40.40875	-39.72509	-189.39824	-188.83519
MP4(SDQ)	-40.40978	-39.72678	-189.40656	-188.84548
MP4(SDTQ)	-40.41453	-39.73420	-189.43206	-188.87501
$\Delta E°_{eq}$(RHF)	18.79 (433.4)		15.49 (357.1)	
$\Delta E°_{eq}$(SDTQ)	18.51 (426.9)		15.16 (349.6)	
corrections				
zero-point energy	-0.43		-0.34	
translation	0.04		0.04	
rotation	0.00		0.00	
vibration	0.00		0.00	
P-V	0.03		0.03	
total	-0.36		-0.28	
$\Delta H°_{298}$(SDTQ)	18.15 (418.5)		14.88 (343.1)	
$\Delta H°_{298}$(exptl)	18.07 (416.6)		14.97 (345.2)	

[a] Geometries optimized at the RHF level. Total energies in hartrees, $\Delta E°_{eq}$ in eV and kcal/mol in parentheses.

with a value of r^2 of 0.891, indicating considerable scatter to the points. The mean difference between experiment and theory is less than 1 kcal/mol; the root-mean-square difference is, however, high—0.48 eV (11 kcal/mol).

In summary, although proton affinities for neutral species can be calculated reasonably well with a modest effort, the proton affinities of anions (acidity of the neutral species) cannot be handled so easily. All but the most sophisticated calculations show a tendency to overestimate relative acidities (i.e., the slope of the line correlating calculated with measured acidities is always significantly greater than unity).

High-Level Calculations for Methane and Formic Acid

The results reported by Gordon et al.[5a] show that it is possible to calculate acidities that are in good agreement with experiment. Their investigation was, however, limited to simple hydrides with only one heavy atom. To provide information on more chemically interesting systems, we have made calculations with an extensive basis set and corrections for electron correlation on formic acid and methane. These represent substances near the high and low end of the acidity scale. Formic acid is more representative of organic acids than are the acids that have so far been treated with high-level calculations.

These high-level calculations were done with Gaussian 82[10] and a VAX 8600 computer at the University of Tromsø. Geometry optimizations for methane and formic acid and their anions were carried out at the RHF/6-311++G(2d,p) level.[11] With use of this geometry and basis set, total energies were calculated at the RHF level and with inclusion of electron correlation through MP4(SDTQ). Results of these calculations are summarized in Table II. We see that most of the improvement in the calculated total energies comes from including the MP2 correction for electron correlation. This correction is greater for the anion than for the neutral molecule; hence RHF calculations tend to give acidities that are too high. The calculated acidities for methane are within about 0.1 eV (2 kcal/mol) of those calculated by others using basis sets that approach Hartree–Fock accuracy.[12]

For comparison with experimental acidities, we have calculated zero-point energies for the relevant species at the 3-21+G//3-21+G level and, in keeping with an observation made by Hehre, Radom, Schleyer, and Pople,[4f] have reduced the resulting energies by 10% to bring them into closer agreement with experimental values. The vibrational contribution to $H_{298}-H_0$, taken from the Gaussian 82 output, is almost negligible, but it has also been reduced by 10%. The rotational, translational, and PV contributions to this energy difference were calculated classically. The values for each of these corrections are listed in Table II, where we also give both the theoretically calculated and experimental values of $\Delta H°_{298}$. The theoretical result for formic acid is just within the experimental uncertainty (0.09 eV) of the experimental value. For methane, the disagreement between experiment and theory of 0.08 eV is about twice the experimental uncertainty.

Lower Level ab Initio Calculations for a Wider Range of Molecules

Calculations of the sort described above are prohibitively expensive to apply to a series of complex molecules. For instance, the calculations needed to determine the acidity of formic acid required 4.5 days of cpu time on a VAX 8600. As a result, the only high-level calculations that have been reported for series of molecules have been for hydrides with only one or two heavy atoms. It is, therefore, of interest to see whether ab initio calculations with smaller basis sets can provide useful information about acidities.

To explore this possibility, we have calculated acidities for a number of molecules with two different basis sets (3-21+G and 6-31+G) and with treatment of electron correlation at various levels. These basis sets include diffuse functions, which are essential for the correct calculation of the energies of anions, and, hence, of acidities.[4e,9] The compounds are listed in Table III and include 4 hydrocarbons, 4 alcohols, 3 carboxylic acids, and acetone; these were chosen to provide several different types of compounds and, within each type, a range of acidities.

The calculations were done with Gaussian 82. Some were done on the VAX 750 at Oregon State University and others on the VAX 8600 at Tromsø. For all calculations geometries were optimized at the RHF level with the 3-21+G basis set.[11,14] Acidities were calculated for all of these compounds with this basis set. For all except monofluoroacetic acid and 2,2-difluoroethanol acidities were also calculated at the 6-31+G//3-21+G level. Calculations were done at the MP2/6-31+G//3-21+G level for all except the two fluoro compounds mentioned and cyclopentadiene. The acidities, ΔE_{eq}, resulting from these calculations are summarized in Table III. Also shown in this table are the linear regression parameters obtained by fitting the calculated acidities to the experimental values of $\Delta H°_{298}$ by least squares. (No corrections for differences in temperature or for zero-point energies have been made, since the calculation of these is very expensive for the large molecules considered here.)

We have also done calculations for all of these compounds except cyclopentadiene and acetone at the MP4/3-21+G//3-

---

(10) Binkley, J. S.; Frisch, M. J.; DeFrees, D. J.; Rahgavachari, K.; Whiteside, R. A.; Schlegel, H. B.; Fluder, E. M.; Pople, J. A. Department of Chemistry, Carnegie-Mellon University, Pittsburgh, PA.

(11) Results of the geometry optimizations are available as supplementary material.

(12) (a) Kollmar, H. *J. Am. Chem. Soc.* **1978**, *100*, 2665. (b) Keil, F.; Ahlrichs, R. *J. Am. Chem. Soc.* **1976**, *98*, 4787.

(13) (a) Bartmess, J. E.; McIver, R. T., Jr. *Gas Phase Ion Chemistry*; Bowers, M. T., Ed.; Academic: New York, 1979; Vol. 2, p 87. (b) DePuy, C. H.; Bierbaum, V. M.; Damrauer, R. *J. Am. Chem. Soc.* **1984**, *106*, 4051.

(14) Clark, T.; Chandrasekhar, J.; Spitznagel, G.; Schleyer, P. v. R. *J. Comput. Chem.* **1984**, *4*, 294.

Table III. Calculated and Experimental Acidities (eV)

	compd	exptl[a]	RHF/3-21+G//3-21+G	RHF/6-31+G//3-21+G	MP2/6-31+G//3-21+G
1	monofluoroacetic acid	14.64	14.58		
2	formic acid	14.97	14.93	15.07	14.67
3	acetic acid	15.11	15.12	15.27	14.90
4	cyclopentadiene	15.44	15.85	15.88	
5	2,2-difluoroethanol	15.92	15.98		
6	acetone	15.99	16.30	16.47	16.19
7	isopropyl alcohol	16.22	16.83	16.88	16.30
8	ethanol	16.31	16.88	17.01	16.48
9	methanol	16.44	16.94	17.08	16.62
10	ethene	17.61[b]	18.23	18.35	18.08
11	methane	18.07	18.72	18.80	18.44
12	ethane	18.26[b]	18.95	19.03	18.64
			Linear Regression Parameters		
	r^2		0.988	0.991	0.995
	intercept		-3.008	-2.320	-3.167
	slope		1.2074	1.1740	1.2004
	mean dev[c]		0 362	0.541	0.150
	rms dev[c]		0.457	0.588	0.289
	rms dev[d]		0.150	0.122	0.099

[a] Except as noted, from ref 13a. [b] Reference 13b. [c] Theory minus experiment. [d] Deviation between theoretical values and the regression line.

Table IV. Linear Correlation Parameters for 3-21+G//3-21+G Calculations

level	intercept	slope	r^2
RHF	-3.0080	1.2074	0.9883
MP2	-4.1853	1.2537	0.9952
MP3	-3.2798	1.2116	0.9941
MP4(DQ)	-3.5130	1.2242	0.9944
MP4(SDQ)	-3.8042	1.2374	0.9950
VAR1	-2.8218	1.1915	0.9930
CISD,4	-2.8500	1.1907	0.9947

21+G level. The parameters describing the linear correlations between theory and experiment are summarized in Table IV. Although the linear correlations are excellent, their slopes are all approximately 1.2.

A comparison of calculated and theoretical acidities for the RHF/3-21+G//3-21+G calculations is shown in Figure 1, where we have plotted the theoretical acidities versus the experimental values.[13] Also shown is the linear regression line (solid line) and a reference line of unit slope and zero intercept (dashed line). It is apparent that there are systematic differences between theory and experiment both in absolute value and in slope. The average difference, theory minus experiment, is 0.36 eV (8.3 kcal/mol) and the root-mean-square difference is 0.46 eV (10.5 kcal/mol). The slope, which is greater than unity, is consistent with the slopes for other calculations, discussed in the introduction. It is, however, apparent that the scatter of the data around the linear regression line is small; the root-mean-square deviation of the points from the line is 0.15 eV (3.5 kcal/mol). Thus, although the calculations reproduce the experimental values rather poorly, there is a very good linear relationship between the theoretical values and the experimental ones, even for the 3-21+G basis set. As can be seen from the parameters given in Tables III and IV, there are also good correlations for the higher level calculations. However, inclusion of corrections for electron correlation at either the 3-21+G or 6-31+G level has only a small effect on the slope of the correlation line.

Inspection of the results given in Table III for the 6-31+G //3-21+G calculations shows an improvement of the linear correlation with the more extensive basis set and with the inclusion of electron correlation. As the quality of the calculation is improved the values of r^2 approach closer to unity and the values of the root-mean-square deviation between the calculated values and the linear regression line approach closer to zero. The absolute acidities calculated at the RHF/6-31+G level are, however, on the average in worse agreement with experiment than those from the RHF/3-21+G calculations. For the 6-31+G basis set the mean difference (theory minus experiment) is 0.54 eV, compared with 0.37 eV for 3-21+G. For the MP2/6-31+G calculations this is reduced to 0.15 eV. This low number does not, however, indicate that this method reliably predicts acidity. The slope of 1.20 and intercept of −3.17 for the linear correlation implies that the theoretical calculations will by systematically low for acidities less than 15.8 eV and high for acidities higher than this.

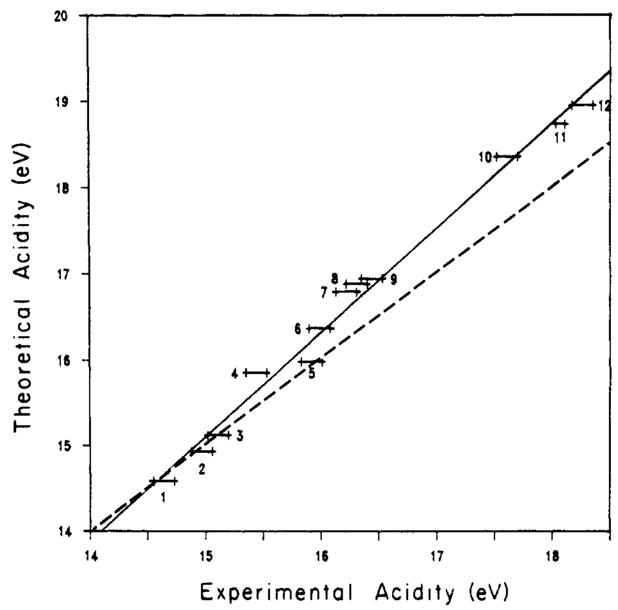

Figure 1. Theoretical acidities (RHF/3-21+G//3-21+G) plotted versus experimental values of ΔH°_{298}. The solid line shows a least-squares fit of a straight line to the data. The dashed line has unit slope and zero intercept. See Table III for names of compounds.

Prediction of Unknown Acidities

It is sometimes useful to be able to predict acidities where they either have not been or cannot be measured. Examples are allyl alcohol and isopropenyl alcohol, which have been of interest to us in our effects to understand some of the factors that affect acidity.[2] Although the first is a stable compound, its acidity has not been reported. The second is too unstable for its acidity to be readily measured. It is, however, reasonably straightforward to calculate the theoretical acidity for these substances. As can be seen from the foregoing discussion, the direct use of such calculations has the disadvantage that they do not reproduce the acidity with sufficient accuracy, except with the highest level calculations. It is, however, possible to use the linear correlations between theoretical and measured acidities to predict the experimental acidity from the theoretical values with reasonable accuracy. From the root-mean-square deviations of the calculated points from the correlation lines (0.09 to 0.15 eV) and from the slopes of the correlation lines (1.2) we conclude that estimates of the experimental acidity based on theoretical calculations and these correlation lines should be within about 0.12 to 0.13 eV (3 kcal/mol) of the correct value. This is only slightly worse than the uncertainty in many experimental values, typically 2 kcal/mol[13] (0.1 eV).

These estimates of uncertainties are calculated from the deviations of the points from the lines. The implicit assumption is that these are random deviations. That is, we assume that there is a systematic deviation between experiment and theory that depends only on the acidity and not on the details of the molecular structure. One cannot, however, fail to notice that the points with low acidities are all carboxylic acids and those with high acidities are all hydrocarbons, with the alcohols in between. It is, therefore, necessary to ask whether the compounds with high acidity are sufficiently different in structure from those with low acidity that the theory might work with different success for one kind of the structure than for the other. If this is the case, then each of these groups might have its own correlation line and the overall correlation line might result from such structural differences.

The results obtained by Gordon et al.[5a] provide some evidence that the deviations are related to the acidity rather than to the structure. They have calculated acidities for nine binary hydrides, each with a single central atom and with acidities spanning a wide range (100 kcal/mol). The analyses of their results (3 and 4 of Table I), however, show the same kinds of deviations as do those for results on compounds with more varied structures. Although there may be structural effects causing particular classes of compounds to have calculated acidities that differ systematically from the measured values, there is a definite trend that seems to be associated with the magnitude of the acidity alone.

We can also subdivide the data shown in Figure 1 and consider the groups separately. This procedure has the disadvantage of small numbers of points that are closely spaced. However, from such an analysis we find a slope of 1.13 to fit the data for carboxylic acids, 1.10 to fit those for the hydrocarbons, and 1.95 for the alcohols. For each type of structure there is a systematic deviation, although for two of the classes it is less than for the overall group.

How does the possibility for such structural dependence affect the use of the regression line to predict unknown acidities from theoretical values? Each structural type spans a fairly narrow range of acidities. Even if each should have its own correlation line with slope different from that of the main correlation line, this will probably not deviate from the main line in the region where one would expect to find appropriate compounds. Thus, we might expect reasonably good predictions from the main correlation line even if there are deviations that result from structure as well as from overall acidity. However, we should also note that our correlation line has been established for only a few types of molecules, containing carbon and hydrogen, and might not be applicable to ones that are quite different in type or structure. For example, we note that the calculated acidity for nitric acid[4a] falls 0.7 eV below the regression line shown in Figure 1.

Two examples of using these linear correlations are shown in Table V, where we illustrate their application to estimating the experimental acidities of isopropenyl and allyl alcohol from theoretical calculations. The first column of numbers in this table shows the acidities, ΔE°_{eq}, for isopropenyl alcohol calculated by using different basis sets and degrees of electron correlation. The second column of numbers gives the values of ΔH°_{298} inferred from these theoretical results and the linear correlation parameters given in Tables III and IV. We see that although there is a range of 0.57 eV (13 kcal/mol) in the theoretical values, there is a range of only 0.16 eV (4 kcal/mol) in the inferred values of ΔH°_{298}. Averaging these values gives an estimated value of ΔH°_{298} of 15.64 eV (360.6 kcal/mol) with all values falling within 0.08 eV of this.[15] This is to be compared with an estimate made by Bartmess[16] (on the basis of the systematics of keto–enol energy differences) of 15.65 eV (361 kcal/mol). There is excellent agreement between our value, based on the theoretical calculations and the linear correlations between ΔE°_{eq}(theo) and ΔH°_{298}(exptl), and this estimate. A similar treatment for allyl alcohol, using the numbers shown in the last two columns of Table V, gives a value for ΔH°_{298} of 16.24 eV (374.5 kcal/mol) with all values within 0.06 eV of this. From a comparison of the experimental acidities of propanol, propylamine, and allylamine, Bartmess[16] has estimated a value for ΔH°_{298} of 16.22 ± 0.09 eV (374 ± 2 kcal/mol), in excellent agreement with the value we have derived from theory.

Table V. Acidity of Isopropenyl and Allyl Alcohol from Theory and Linear Correlation (eV)

	isopropenyl		allyl	
method[a]	theor acidity	acidity from correl	theor acidity	acidity from correl
RHF/3-21+G	15.90	15.66	16.65	16.28
MP2/3-21+G	15.43	15.65	16.15	16.22
MP3/3-21+G	15.75	15.71	16.41	16.25
MP4(DQ)/3-21+G	15.72	15.71	16.37	16.24
MP4(SDQ)/3-21+G	15.62	15.70	16.27	16.22
VAR1/3-21+G	15.90	15.72	16.59	16.29
CISD,4/3-21+G	15.86	15.72	16.54	16.28
RHF/6-31+G	16.00	15.60	16.77	16.26
MP2/6-31+G	15.52	15.56	16.26	16.18
weighted av[b]		15.64		16.24

[a] Geometry for all calculations optimized at the 3-21+G level. [b] Weighted average was calculated by first averaging the MP3, MP4, VAR1, and CISD values into a single value, which was then averaged with the others.

Since both acetone and isopropenyl alcohol form the same anion when a proton is removed, the difference between the acidities of these two compounds, 0.35 eV (8 kcal/mol), is equal to the energy difference between acetone and isopropenyl alcohol, which are keto and enol isomers of one another. This value is to be compared with an experimental value of 0.61 eV (14 kcal/mol) reported by Holmes and Lossing[17] and a theoretical value (6-31G*//3-21G) of 0.73 eV (16.8 kcal/mol).[18] Our direct calculations of the keto–enol energy difference give 0.40 eV (9 kcal/mol) (RHF/3-21+G//3-21+G and RHF/6-31+G//3-21+G) and 0.47 eV (11 kcal/mol) (MP2/6-31+G//3-21+G). However, the analysis described above indicates that these theoretical relative acidities (and, hence, the energy differences) are too large by about 20%. Therefore, a high-level calculation might be expected to give a keto–enol energy difference of 0.3 to 0.4 eV (7 to 9 kcal/mol), which is in agreement with our estimate of 0.35 eV (8 kcal/mol) but lower than the reported experimental value of 0.6 eV (14 kcal/mol).[17]

Conclusions

High-level calculations in which the basis set is capable of yielding near-Hartree–Fock accuracy and which include extensive corrections for correlation are capable of producing acidities that are in excellent agreement with experimental values. For methane and formic acid calculated at the MP4/6-311++G(2d,p) level, the calculated acidities are within 0.09 eV of the experimental values.

Calculations that lack either Hartree-Fock accuracy or adequate treatment of electron correlation give acidities that differ in a systematic way from the measured values. In particular, such calculations all overestimate relative acidities.

Results have been presented for calculations at the 3-21+G and 6-31+G levels for a variety of compounds, including hydrocarbons,

---

(15) In applying this procedure, one would in general select a single method of calculation (presumably the best that the resources allow) rather than average the results from several different levels.

(16) Bartmess, J. E., private communication.

(17) Holmes, J. L.; Lossing, F. P. *J. Am. Chem. Soc.* **1982**, *104*, 2648.

(18) Heinrich, N.; Koch, W.; Frenking, G.; Schwarz, H. *J. Am. Chem. Soc.* **1986**, *108*, 593.

alcohols, carboxylic acids, and acetone. These all show the systematic effects mentioned above. In particular, the linear correlation line between the calculated and experimental acidities has a slope close to 1.2 for all of the calculations.

Because of these systematic errors, lower level ab initio calculations of the sort considered here cannot by themselves give reliable predictions for the acidity. Typical calculated values can deviate from the experimental ones by 0.3 eV and, for different levels of calculation for the same compound, may range over as much as 0.6 eV. However, the correlation coefficients for the linear correlation between experiment and theory are very close to unity. Therefore, the calculated acidities fall very close to the regression lines, and these lines can be used together with calculated numbers to give predicted experimental values close to the true ones. Even though the calculated acidities for a given compound range over 0.6 eV, the values of acidity estimated from these theoretical results and the linear correlations agree with one another within 0.14 eV (3 kcal/mol). These linear correlations can, therefore, be used with the theoretical acidities to make predictions of experimental acidities that are probably within 0.13 eV (3 kcal/mol) of the correct values.

Acknowledgment. This work was supported in part by the National Science Foundation and the Norwegian Council for Scientific and Industrial Research. We are indebted to John Bartmess for providing us with his estimates of the acidities of allyl and isopropenyl alcohol and to Thomas Hellman Morton for a suggestion that led ultimately to this work.

Registry No. Fluoroacetic acid, 144-49-0; fluoroacetate anion, 513-62-2; formic acid, 64-18-6; formate anion, 71-47-6; acetic acid, 64-19-7; acetate anion, 71-50-1; cyclopentadiene, 542-92-7; cyclopentadiene anion, 12127-83-2; 2,2-difluoroethanol, 359-13-7; 2,2-difluoroethanol anion, 104745-47-3; acetone, 67-64-1; acetone anion, 24262-31-5; isopropyl alcohol, 67-63-0; isopropyl alcohol anion, 15520-32-8; ethanol, 64-17-5; ethanol anion, 16331-64-9; methanol, 67-56-1; methanol anion, 3315-60-4; ethene, 74-85-1; ethene anion, 25012-81-1; methane, 74-82-8; methane anion, 15194-58-8; ethane, 74-84-0; ethane anion, 25013-41-6; allyl alcohol, 107-18-6; allyl alcohol anion, 71695-00-6; isopropenyl alcohol, 29456-04-0.

Supplementary Material Available: Tables of optimized geometries and total energies for all of the neutral molecules and their anions (4 pages). Ordering information is given on any current masthead page.

Valency Correlation Diagrams

Prabha Siddarth and M. S. Gopinathan*

Contribution from the Department of Chemistry, Indian Institute of Technology, Madras 600 036, India.
Received April 27, 1987

Abstract: It is shown that the recently proposed molecular orbital (MO) valency serves as a quantitative ordinate for the Walsh-type correlation diagrams. The essential requirements for a molecular orbital quantity to be a successful ordinate for such correlation diagrams are formulated. Universal correlation diagrams are drawn with MO valency as ordinate for AH_2, AH_3, HAB, and AB_2 classes of molecules and shown to be remarkably similar to the original Walsh diagrams. With the notable exception of core and lone pair orbitals, MO eigenvalues also lead to acceptable correlation diagrams. The question whether the sum of the ordinate quantities leads to the correct bond angle for a given molecule is also examined. It is shown that MO valency predicts the bond angles well whereas the MO eigenvalue often fails in this respect. A simple valency method is proposed to predict bond angles of electronically excited or ionized states of a given molecule by using just its ground-state wave function. This method is shown to be capable of predicting bond angles for excited and ionized states remarkably well.

I. Introduction

Walsh diagrams[1,2] are plots of certain "molecular orbital energies" as a function of some geometry parameter, usually a bond angle. These diagrams have been useful in predicting the geometry of the ground, excited, or ionized states of molecules merely from the number of valence electrons. However, such predictions are only qualitative since Walsh did not precisely define the ordinate "molecular orbital energy" and the Walsh diagrams were based on qualitative arguments regarding atomic overlaps and their relation to molecular orbital (MO) binding energies. This fact has prompted numerous investigators to search for a molecular orbital quantity in the SCF theory that could be used as the ordinate. These attempts have been described in a review by Buenker and Peyerimhoff.[3] Quantities such as MO eigenvalues[4] and their variants[5-7] and MO forces[8] have been tried, but these attempts have been only partially successful. We have recently introduced the idea of a valency for each molecular orbital[9] and have shown that MO valency as the ordinate reproduces the Walsh-type diagrams quite accurately for a few molecules.

In this paper, we first set down the essential properties that a MO quantity should have to qualify as an ordinate for the qualitative Walsh diagrams (Section II). We then examine how well MO valency as well as the usual ordinate, MO eigenvalue, satisfy these criteria (Section III). The quantitative predictions of the equilibrium bond angles for the ground state of molecules by both mo valency and mo eigenvalue are discussed in Sec. IV. In Sec. V, a simple method, which utilizes only the ground state molecular orbitals, is proposed and demonstrated to be remarkably successful for the prediction of bond angles of excited states.

II. Criteria for the Walsh Ordinate

Walsh diagrams are meant to be used only for qualitative predictions of the shape of molecules. The prediction of the exact geometry of a given molecule is not intended. However, there have been some studies in the literature[3,4,10] to see if the sum of the ordinates, normally the eigenvalues, gives the same equilibrium bond angle as does the total energy. While it is desirable to have this property for certain purposes, as will be discussed later, it is not an essential requirement for a Walsh ordinate.

The essential criteria for a MO quantity to serve as a Walsh ordinate may be formulated as follows:

---

(1) Walsh, A. D. *J. Chem. Soc.* **1953**, (a) 2260, (b) 2266, (c) 2288, (d) 2296, (e) 2301, (f) 2306, (g) 2321.

(2) Mulliken, R. S. *Rev. Mod. Phys.* **1942**, *14*, 204.

(3) Buenker, R. J.; Peyerimhoff, S. D. *Chem. Rev.* **1974**, *74*, 127.

(4) Peyerimhoff, S. D.; Buenker, R. J.; Allen, L. C. *J. Chem. Phys.* **1966**, *45*, 734.

(5) Coulson, C. A.; Neilson, A. H. *Discuss. Faraday Soc.* **1963**, *35*, 71, 217.

(6) Stenkamp, L. Z.; Davidson, E. R. *Theor. Chim. Acta* **1973**, *30*, 283.

(7) Mehrotra, P. K.; Hoffmann, R. *Theor. Chim. Acta* **1978**, *48*, 301.

(8) Coulson, C. A.; Deb, B. M. *Int. J. Quantum Chem.* **1971**, *5*, 411.

(9) Gopinathan, M. S.; Siddarth, P.; Ravimohan, C. *Theor. Chim. Acta* **1986**, *70*, 303.

(10) Allen, L. C.; Russell, J. D. *J. Chem. Phys.* **1967**, *46*, 1029.