Pergamon

Tetrahedron 57 (2001) 5923–5930

TETRAHEDRON

# The acidity of carbon acids in aqueous solutions: correlations with theoretical descriptors

Marcos Caroli Rezende

*Facultad de Química y Biología, Universidad de Santiago, Casilla 40, Correo 33, Santiago, Chile*

Received 1 October 2000; accepted 14 May 2001

**Abstract**—The acidity of 32 carbon acids in water, with p$K_a$ values ranging from 25.6 to $-6.2$, was correlated with a two-parameter equation, involving the theoretical descriptors $\Delta E$ and $\sum\Delta q_X$, calculated with the HF/3-21G//HF/3-21G and the B3LYP/6-31G(d)//HF/3-21G methods, with regression coefficients $r \geq 0.95$. The acid deprotonation energy ($\Delta E$) was associated with its intrinsic, or gas-phase acidity, whereas the total charge variation at the hard centers ($\sum\Delta q_X$) was a measure of the contributions of aqueous solvation to its acid behavior in water.

## 1. Introduction

The behavior of organic acids in the gas-phase often departs considerably from their acidity in aqueous solution. As a consequence, calculations in the gas-phase cannot be expected to yield reliable predictions concerning the acidity of organic compounds in water. In spite of this generally accepted view, reports of good correlations between theoretical gas-phase descriptors and the aqueous acidity of series of compounds are found in the literature. Examples include a series of substituted acetic acids,[1] of alkyl-substituted alcohols,[2] of *p*-substituted phenols[3] and a study on the hydroxyl acidity in zeolites.[4] The good correlation obtained between the p$K_a$ values of a wide range of organic acids and their calculated deprotonation enthalpies has been used to estimate the acidity of hydrofullerenes.[5] Reasonable correlations between theoretical descriptors and the gas-phase acid–base equilibria of organic compounds have been reported for amines,[6,7] for alcohols and thiols,[8] and for haloacetic acids.[9] The p$K_a$ values in water of fifteen organic acids (alcohols, thiols and carboxylic acids) have been calculated with variable success, employing the polarizable continuum method to account for solvent effects.[10] A good correlation was obtained between the acidity, calculated with the 3-21G(*) method, and the p$K_a$ values of a series of 37 nitrogen bases, after incorporation of solvent correction factors.[11]

Recently, various theoretical descriptors were investigated, in the search of correlations with the gas-phase acidity of compounds of the form $CH_3Z$.[12] With the modest HF/3-21G method, a reasonable correlation ($r = 0.961$) between the gas-phase acidity of the compounds and their calculated deprotonation enthalpies was obtained. Interestingly, the use of larger basis sets and polarization (6-31G(d)) or diffuse functions (6-31+G(3df,2p)) did not improve the correlations, in agreement with similar observations made by other authors.[7,11]

In the present paper, this study is extended to include the acid behavior in water of a larger set of compounds, exhibiting greater structural variations. The p$K_a$ values in water of 32 different carbon acids of the form CHWYZ, where W, Y and/or Z are electron-withdrawing groups were correlated with theoretical descriptors reflecting charge and energy variations upon deprotonation. This simple model was tested with a Hartree–Fock method employing a modest basis set (3-21G) and a hybrid DFT method employing a larger basis set (B3LYP/6-31G(d)). Unlike most theoretical studies dealing with acid–base equilibria in water, the model does not incorporate any explicit corrections for solvent effects, yielding reasonable correlations with readily obtained descriptors.

## 2. Experimental

Calculations were performed with a GAUSSIAN 94w package.[13] For each compound, the structures of the acid and of its conjugate base were fully optimized at the HF/3-21G level. The obtained heats of formation and atomic charges were then used to calculate energy and charge variations upon deprotonation. Single-point calculations were also performed on the optimized structures employing the hybrid DFT B3LYP method and the 6-31G(d) basis set.

---

*Keywords*: carbon acids; aqueous p$K_a$ values; HF and DFT calculations.

PII: S0040-4020(01)00563-4

**Table 1.** p$K_a$ Values of carbon acids CHWYZ in water

| Compound number | W | Y | Z | p$K_a$ Value[a] | Ref. |
|---|---|---|---|---|---|
| **1** | H | H | $CO_2Et$ | 25.6 | 25 |
| **2** | H | H | CN | (25)[b] | 21 |
| **3** | H | H | $SO_2Me$ | (23)[b] | 21 |
| **4** | H | H | COSEt | 21 | 25 |
| **5** | H | H | COMe | 19.3 | 25 |
| **6** | H | H | COPh | 18.3 | 25 |
| **7** | H | Cl | COMe | 16.5 | 21 |
| **8** | Cl | Cl | COMe | 15 | 21 |
| **9** | H | $SO_2Me$ | $SO_2Me$ | 14 | 21 |
| **10** | H | $CO_2Et$ | $CO_2Et$ | 13.3 | 21 |
| **11** | F | F | $NO_2$ | 12.4[c] | 14 |
| **12** | H | CN | CN | 11.2 | 21 |
| **13** | Me | COMe | COMe | 11 | 21 |
| **14** | H | COMe | $CO_2Me$ | 10.72[d] | 14 |
| **15** | H | H | $NO_2$ | 10.24 | 14 |
| **16** | Cl | F | $NO_2$ | 10.14 | 14 |
| **17** | H | COMe | $SO_2Me$ | 10.03[e] | 14 |
| **18** | H | COMe | COMe | 9.03 | 14 |
| **19** | H | Me | $NO_2$ | 8.6 | 21 |
| **20** | H | Cl | $NO_2$ | 7.2 | 14 |
| **21** | Cl | Cl | $NO_2$ | 5.99 | 14 |
| **22** | COMe | COMe | COMe | 5.85 | 21 |
| **23** | H | $CONH_2$ | $NO_2$ | 5.18 | 14 |
| **24** | H | COMe | $NO_2$ | 5.1 | 21 |
| **25** | Cl | $NO_2$ | $NO_2$ | 3.8 | 14 |
| **26** | H | $NO_2$ | $NO_2$ | 3.63[e] | 14 |
| **27** | $CH_2CN$ | $NO_2$ | $NO_2$ | 2.27 | 14 |
| **28** | $CONH_2$ | $NO_2$ | $NO_2$ | 1.30[e] | 14 |
| **29** | $NO_2$ | $NO_2$ | $NO_2$ | 0.06 | 14 |
| **30** | $CO_2Me$ | CN | CN | −2.8 | 14 |
| **31** | CN | CN | CN | −5.1 | 14 |
| **32** | CN | $NO_2$ | $NO_2$ | −6.2 | 14 |

[a] At 25°C, unless stated otherwise.
[b] Value estimated by extrapolation.
[c] No temperature given.
[d] At 20.8°C.
[e] At 20°C.

## 3. Results and discussion

Table 1 lists the series of carbon acids CHWYZ studied in this work, with their corresponding p$K_a$ values in water, in order of increasing acidity. Whenever possible, values at 25°C were used. The W, Y and/or Z groups comprised, in addition to alkyl and aryl groups, the following substituents: CN, $NO_2$, $COCH_3$, $CO_2CH_3$, $CO_2Et$, COSEt, $CONH_2$, $SO_2CH_3$, F, Cl and Br. Acidity values for aldehydes were not included, as they applied in general to the corresponding hydrates[14]

### 3.1. Definition of model and descriptors

The following assumptions, based on a simple model, guided the present search of theoretical descriptors for the aqueous acidity of the studied compounds:

1. The intrinsic, or gas-phase acidity of a carbon acid reflects the relative stability of its conjugate base vis-a-vis its protonated form. Two factors contribute to the stabilization of the conjugate base. The first is charge delocalization through resonance, whereby the negative charge is shared among the anionic carbon and the atoms conjugated with it. The second source of stabilization is due to field effects, whereby charge is shared with polarizable groups in the molecule.
2. Solvation by water may alter significantly the intrinsic acidity of the carbon acid. Anion solvation takes place through hydrogen-bonding between water and negatively charged atoms in the molecule. These comprise the anionic carbon and all atoms conjugated with it. The charge variations at these centers may thus be used as a measure of the degree of solvation of the acid and its conjugate base by water.

To describe the intrinsic, or gas-phase acidity of the CH acids, the deprotonation energy $\Delta E$ was adopted as a descriptor. This parameter is defined as the difference between the heats of formation of the conjugate base $R^-$ and the acid RH.

$$\Delta E = \Delta H_{R_-} - \Delta H_{RH} \qquad (1)$$

The adoption of this descriptor was suggested by the good correlation obtained in a previous study, between the gas-phase $\Delta G^\circ$ and $\Delta E$ values calculated for nine acids $CH_3Z$.[12] In the present study, which included polysubstituted acids, the choice of this descriptor was validated by equally good correlations. Table 2 lists the gas-phase $\Delta G^\circ$ values of 13 carbon acids,[15,16] most of them included in Table 1, together with the corresponding $\Delta E$ values calculated by the B3LYP/6-31G(d)//HF/3-21G method. The good correlation obtained ($r = 0.985$) is illustrated in Fig. 1. A similar plot,

**Table 2.** Gas-phase acidities of carbon acids $CH_2ZY$

| Compound number | Z | Y | $\Delta E$, a.u.[a,b] | $\Delta G^\circ$, kcal mol$^{-1}$[b] | Ref. |
|---|---|---|---|---|---|
| **33** | H | SOMe | 0.6466 | 366.4 | 15 |
| **2** | H | CN | 0.6335 | 365.2 | 15 |
| **34** | H | $CO_2Me$ | 0.6312 | 364.3 | 16 |
| **7** | H | COMe | 0.6262 | 361.9 | 15 |
| **35** | H | COH | 0.6231 | 359.7 | 16 |
| **3** | H | $SO_2Me$ | 0.6142 | 358.2 | 15 |
| **6** | H | COPh | 0.6138 | 354.5 | 15 |
| **15** | H | $NO_2$ | 0.6103 | 352.0 | 16 |
| **36** | $CH_3$ | $NO_2$ | 0.6090 | 351.7 | 16 |
| **10** | $CO_2Et$ | $CO_2Et$ | 0.5718 | 342.3 | 15 |
| **18** | COMe | COMe | 0.5641 | 336.7 | 15 |
| **37** | CN | $CO_2Et$ | 0.5661 | 333.6 | 15 |
| **12** | CN | CN | 0.5552 | 329.5 | 15 |

[a] Calculated by the B3LYP/6-31G(d)//HF/3-21G method.
[b] Conversion factors: 1 atomic unit=2625.5 kJ mol$^{-1}$, 1 kcal mol$^{-1}$=4.184 kJ mol$^{-1}$.

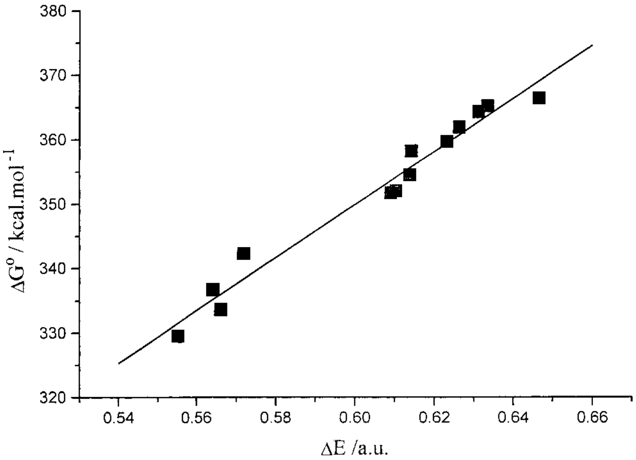

**Figure 1.** Correlation ($r = 0.985$, s.d. $= 2.3$ kcal mol$^{-1}$) between the gas-phase acidities of carbon acids $CH_2ZY$ and the calculated (B3LYP/HF/6-31G(d)// HF/3-21G) deprotonation energies $\Delta E$, according to the data of Table 2.

with $\Delta E$ values calculated by the HF/3-21G//HF/3-21G method, yielded a poorer, though still significant correlation ($r = 0.946$).

The fact that a good correlation was found between theoretical enthalpies and experimental free-energies in the gas phase suggests that, for the series under study, the entropic term in the acid–base equilibrium should be nearly constant. This was indeed found to be the case for the gas-phase acidity of a set of 27 carbon acids,[16] where the entropic contribution ($T\Delta S^\circ$) to $\Delta G^\circ$ at 298 K amounted to a nearly constant value of 6–7 kcal mol$^{-1}$. Similar observations were made by other authors[17,18] for a series of weak acids in DMSO.

A second type of descriptors was adopted to account for solvation effects by water. These descriptors were based on local charge variations upon deprotonation, since aqueous solvation arises from electrostatic attractions between partially charged atoms and water molecules. All carbon acids listed in Table 1 form resonance-stabilized anions upon deprotonation. Charge is therefore shared among the soft carbon and conjugated hard oxygen or nitrogen atoms. By separating the charge variations at the soft carbon center from the total variations at the hard centers, the relative contribution of these sites to the anion stabilization in solution may be estimated.

The charge descriptor at the soft carbon center was therefore defined as

$$\Delta q_C = q_C^{R-} - q_C^{RH} \qquad (2)$$

where $q$ is the charge at the ionizable carbon, and the superscripts $R^-$ and RH refer to the anion $^-CWYZ$ and the acid HCWYZ, respectively.

In an analogous way, the charge descriptor at the conjugated hard centers was defined as

$$\sum \Delta q_X = \sum \Delta q_X^{R-} - \sum \Delta q_X^{RH} \qquad (3)$$

where the sum extended over all hard centers X conjugated with the ionizable carbon (two oxygen atoms in each $NO_2$ group, one nitrogen atom in each –CN group, and one oxygen atom in each carbonyl group of a ketone, an ester or an amide). The two oxygen atoms of the sulfonyl group were also included in this set, though a direct conjugation of these centers with an adjacent anionic carbon should be precluded by geometrical considerations. An analysis of the optimized geometries of dimethylsulfone and its anion showed that, in spite of non-coplanarity between the p-orbital of the anionic carbon atom and the π S–O bonds, some degree of conjugation existed in the molecule. This was suggested by a shortening of the C–S bond and an elongation of the S–O bonds in the anion $^-CH_2SO_2CH_3$, when compared with the dimethylsulfone molecule. A similar conclusion had been reached by Bordwell et al.[19] when comparing the effect of the $PhSO_2$– and other substituents upon the acidity of carbon acids in DMSO.

### 3.2. Comparison of methods

Since single-point calculations on the optimized structures were performed with the HF/3-21G and the hybrid DFT B3LYP/6-31G(d) methods, a comparison of the two sets of results seemed of interest. The DFT method, which takes into account correlation energy contributions and utilizes a larger basis set, should be in principle superior to the rather modest HF/3-21G method. As an illustration of this point, for a set of 125 calculations involving thermochemical data (atomization energies, ionization potentials, proton and electron affinities), the B3LYP/6-31G(d)//HF/3-21G(d) method fared much better than the HF/3-21G(d)// HF/3-21G(d) method, with a smaller standard deviation (9.4 kcal mol$^{-1}$) than the latter (50.1 kcal mol$^{-1}$).[20] Single-point calculations with the B3LYP functional on geometries optimized at a lower level of theory seemed validated by the fact that, for the above set, the computed energies were surprisingly insensitive to the geometry optimization level, yielding results which were just as good with 3-21G(d) as with the 6-31G(d) levels.[20] Also, previous experience[3,6,7,12] had suggested that, for the particular case of acid–base equilibria, methods involving small basis sets (STO-3G, 3-21G, 3-21g(d)) are often reasonably accurate and that the use of HF methods with larger basis sets might

**Table 3.** Descriptor values and results of regression analysis, calculated with the HF/3-21G//HF/3-21G method

| Compound number | | | | $pK_a$ Value | |
|---|---|---|---|---|---|
| | $\Delta E$, a.u.[a] | $\sum\Delta q_X$ | $\Delta q_C$ | Calculated[b] | Experimental |
| **1** | 0.6324 | −0.206 | 0.026 | 21.8 | 25.6 |
| **2** | 0.6455 | −0.259 | −0.017 | 21.4 | (25) |
| **3** | 0.6186 | −0.200 | −0.005 | 20.6 | (23) |
| **4** | 0.5983 | −0.154 | 0.075 | 20.1 | 21 |
| **5** | 0.6376 | −0.266 | 0.059 | 20.3 | 19.3 |
| **6** | 0.6222 | −0.240 | 0.071 | 19.7 | 18.3 |
| **7** | 0.6035 | −0.257 | 0.111 | 17.3 | 16.5 |
| **8** | 0.5895 | −0.225 | 0.070 | 16.9 | 15 |
| **9** | 0.5533 | −0.279 | −0.183 | 11.7 | 14 |
| **10** | 0.5721 | −0.284 | 0.093 | 13.3 | 13.3 |
| **11** | 0.5776 | −0.411 | 0.179 | 9.8 | 12.4 |
| **12** | 0.5691 | −0.428 | 0.111 | 8.4 | 11.2 |
| **13** | 0.5895 | −0.306 | 0.154 | 14.3 | 11 |
| **14** | 0.5734 | −0.320 | 0.122 | 12.3 | 10.72 |
| **15** | 0.5925 | −0.400 | 0.166 | 11.6 | 10.24 |
| **16** | 0.5640 | −0.381 | 0.220 | 9.4 | 10.14 |
| **17** | 0.5661 | −0.330 | −0.043 | 11.3 | 10.03 |
| **18** | 0.5748 | −0.354 | 0.128 | 11.4 | 9.03 |
| **19** | 0.5922 | −0.392 | 0.265 | 11.8 | 8.6 |
| **20** | 0.5559 | −0.364 | 0.230 | 9.2 | 7.2 |
| **21** | 0.5515 | −0.349 | 0.254 | 9.2 | 5.99 |
| **22** | 0.5386 | −0.470 | 0.162 | 4.1 | 5.85 |
| **23** | 0.5403 | −0.427 | 0.271 | 5.6 | 5.18 |
| **24** | 0.5465 | −0.444 | 0.197 | 5.7 | 5.1 |
| **25** | 0.5120 | −0.446 | 0.358 | 2.3 | 3.8 |
| **26** | 0.5145 | −0.506 | 0.277 | 0.6 | 3.63 |
| **27** | 0.5066 | −0.449 | 0.464 | 1.6 | 2.27 |
| **28** | 0.5035 | −0.510 | 0.452 | −0.6 | 1.30 |
| **29** | 0.4886 | −0.563 | 0.471 | −3.8 | 0.06 |
| **30** | 0.5164 | −0.483 | 0.249 | 1.5 | −2.8 |
| **31** | 0.5077 | −0.576 | 0.316 | −2.3 | −5.1 |
| **32** | 0.4771 | −0.589 | 0.469 | −5.8 | −6.2 |

[a] Conversion factor: 1 atomic unit=2625.5 kJ mol$^{-1}$.

[b] From Eq. (3).

not necessarily lead to improvements in the correlations with the experimental data.

The calculated descriptors $\Delta E$, $\Delta q_C$, and $\sum\Delta q_X$ employing the two methods are listed in Tables 3 and 4. The tables also include the experimental and the calculated $pK_a$ values obtained by a regression analysis of the theoretical descriptors.

Although the individual calculated $pK_a$ values differed for the two methods, the statistical trends were rather similar. Of the three employed descriptors, only $\Delta E$ and $\sum\Delta q_X$ proved statistically significant. The regression equations, with corresponding standard deviations and correlation coefficients for each method, are given in Eqs. (4) and (5).

Method HF/3-21G//HF/3-21G:

$$pK_a^{calc} = 94.84\Delta E + 29.91\sum\Delta q_X - 3.05\Delta q_C - 32.09 \quad (4)$$

s.d. $= 2.5$ $pK_a$ units

correlation coefficient $r = 0.958$

Method B3LYP/6-31G(d)//HF/3-21G:

$$pK_a^{calc} = 117.15\Delta E + 19.92\sum\Delta q_X - 5.02\Delta q_C - 48.02 \quad (5)$$

s.d. $= 2.7$ $pK_a$ units

correlation coefficient $r = 0.951$

The suppression of the $\Delta q_C$ descriptor led to simpler Eqs. (6) and (7), which depended only on $\Delta E$ and $\sum\Delta q_X$, with the same correlation coefficients of the three-parameter equations and slightly better standard deviations. Eqs. (6) and (7) were therefore employed for the calculation of the theoretical $pK_a$ values listed in Tables 2 and 3.

Method HF/3-21G//HF/3-21G:

$$pK_a^{calc} = 98.09\Delta E + 32.11\sum\Delta q_X - 33.65 \quad (6)$$

s.d. $= 2.4$ $pK_a$ units

$r = 0.958$

correlation between $pK_a^{exp}$ and $\Delta E = 0.931$

correlation between $pK_a^{exp}$ and $\sum\Delta q_X = 0.920$

**Table 4.** Descriptor values and results of regression analysis, calculated with the B3LYP/6-31G(d)//HF/3-21G method

| Compound number | $\Delta E$, a.u. | $\sum\Delta q_X$ | $\Delta q_C$ | p$K_a$ Value Calculated[a] | p$K_a$ Value Experimental |
|---|---|---|---|---|---|
| **1** | 0.6291 | −0.216 | −0.016 | 21.7 | 25.6 |
| **2** | 0.6335 | −0.223 | −0.128 | 22.1 | (25) |
| **3** | 0.6142 | −0.220 | −0.018 | 19.7 | (23) |
| **4** | 0.6091 | −0.135 | 0.025 | 21.0 | 21 |
| **5** | 0.6262 | −0.285 | 0.028 | 19.8 | 19.3 |
| **6** | 0.6138 | −0.249 | 0.011 | 19.0 | 18.3 |
| **7** | 0.5988 | −0.273 | 0.012 | 16.6 | 16.5 |
| **8** | 0.5859 | −0.238 | 0.011 | 15.8 | 15 |
| **9** | 0.5510 | −0.316 | −0.023 | 9.7 | 14 |
| **10** | 0.5718 | −0.294 | −0.111 | 12.8 | 13.3 |
| **11** | 0.6232 | −0.561 | 0.009 | 13.3 | 12.4 |
| **12** | 0.5552 | −0.356 | −0.025 | 9.3 | 11.2 |
| **13** | 0.5790 | −0.292 | 0.226 | 13.7 | 11 |
| **14** | 0.5678 | −0.335 | 0.130 | 11.4 | 10.72 |
| **15** | 0.6103 | −0.525 | 0.074 | 12.5 | 10.24 |
| **16** | 0.5981 | −0.526 | 0.111 | 10.9 | 10.14 |
| **17** | 0.5652 | −0.341 | 0.032 | 10.9 | 10.03 |
| **18** | 0.5641 | −0.373 | 0.140 | 10.1 | 9.03 |
| **19** | 0.6090 | −0.521 | 0.152 | 12.4 | 8.6 |
| **20** | 0.5862 | −0.491 | 0.101 | 10.2 | 7.2 |
| **21** | 0.5731 | −0.476 | 0.148 | 8.9 | 5.99 |
| **22** | 0.5376 | −0.434 | 0.273 | 5.4 | 5.85 |
| **23** | 0.5575 | −0.537 | 0.187 | 5.6 | 5.18 |
| **24** | 0.5507 | −0.529 | 0.153 | 5.0 | 5.1 |
| **25** | 0.5318 | −0.598 | 0.273 | 1.1 | 3.8 |
| **26** | 0.5309 | −0.670 | 0.189 | −0.6 | 3.63 |
| **27** | 0.5212 | −0.608 | 0.362 | −0.5 | 2.27 |
| **28** | 0.5269 | −0.623 | 0.372 | −0.1 | 1.30 |
| **29** | 0.5110 | −0.678 | 0.384 | −3.2 | 0.06 |
| **30** | 0.5111 | −0.408 | 0.232 | 2.7 | −2.8 |
| **31** | 0.4939 | −0.462 | 0.175 | −0.7 | −5.1 |
| **32** | 0.4946 | −0.708 | 0.396 | −6.0 | −6.2 |

[a] From Eq. (4).

Method B3LYP/6-31G(d)//HF/3-21G:

$$\mathrm{p}K_a^{\mathrm{calc}} = 124.96\Delta E + 22.11\sum\Delta q_X - 52.16 \qquad (7)$$

s.d. = 2.6 p$K_a$ units

$r = 0.950$

correlation between $\mathrm{p}K_a^{\mathrm{exp}}$ and $\Delta E = 0.887$

correlation between $\mathrm{p}K_a^{\mathrm{exp}}$ and $\sum\Delta q_X = 0.809$

Plots of the calculated vs experimental p$K_a$ values for the two methods are shown in Figs. 2 and 3.

Before analyzing the results, it must be born in mind that

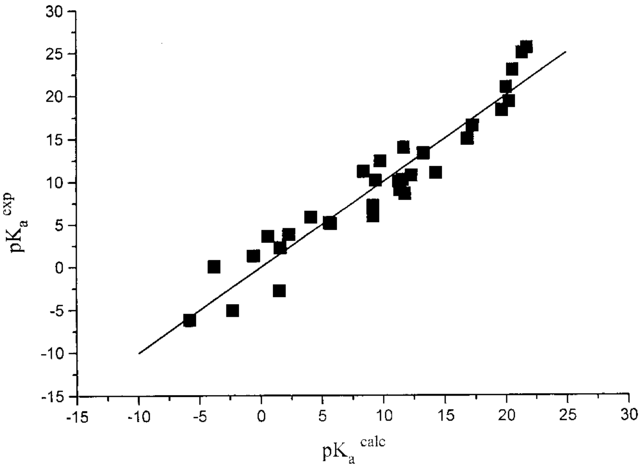

**Figure 2.** Correlation ($r = 0.958$) between the experimental vs calculated (HF/3-21G//HF/3-21G) p$K_a$ values of acids (**1**)–(**32**).

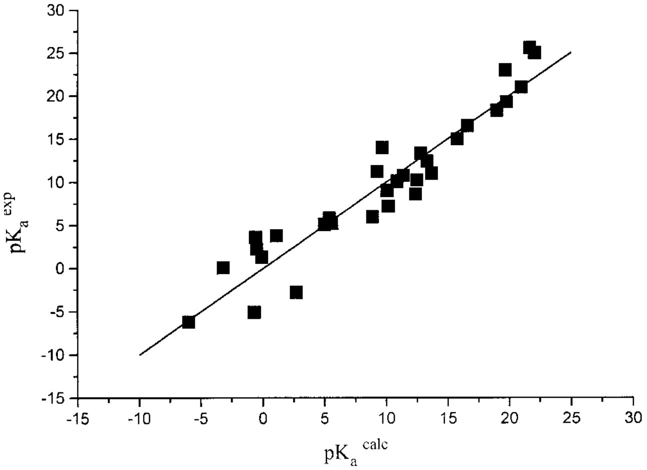

**Figure 3.** Correlation ($r = 0.950$) between the experimental vs. calculated (B3LYP/HF/6-31G(d)//HF/3-21G) p$K_a$ values of acids (**1**)–(**32**).

some of the p$K_a$'s of Table 1 are estimated values and that, for quite a few compounds, the quoted experimental value refers to the gross acid constant, uncorrected for enol content.[21] As shown by Kresge and Meng[22] in a comparative study of an aceto- and a thionoaceto-phenone, the acidity of the keto and enol forms of CH acids may be very different, with variable contributions to the gross p$K_a$ value in water. In addition, significant discrepancies in p$K_a$ values are found in literature for some compounds. For example, Bell[23] reports a value of 19.2 for the p$K_a$ of acetophenone. Another kinetic approach[24] places this p$K_a$ about three units lower (ca. 16). The value of 18.3, adopted in the present work, was taken from Richard and Amyes,[25] quoting data from Keefe and Kresge.[26]

Another example is the acidity of ethyl acetate in water. Pearson and Dillon[21] estimated a p$K_a$ value of 24.5 for this compound. This value had been adopted without discussion in the literature for nearly 30 years, until it was recently shown that the correct value should be about one p$K_a$ unit higher (25.6).[25] The same authors[21] reported a p$K_a$ value of 14 for $CH_2(SO_2CH_3)_2$, while Taft and Bordwell[15] found it to be more acidic (p$K_a$=12.7) by about one p$K_a$ unit. A substantial difference is found between the value reported for tricyanomethane by Pearson and Dillon[21] (estimated p$K_a$=0) and by Boyd[27] (ca. −5).

These examples show that the average deviations obtained in the above regressions are quite reasonable, in view of the uncertainty of some of the data from the literature. This uncertainty may also explain the relative lack of sensitivity of the obtained correlations to the level of theory employed in the calculations.

### 3.3. The interpretation of shifts in acidity

The two descriptors may be used in the interpretation of the shifts in acidity of a series of acids, when transferred from the gas phase to the aqueous solution.

The deprotonation energy $\Delta E$ is a measure of the gas-phase acidity of the compound. In general, acidity should increase with a decrease of the deprotonation energy, as illustrated in Fig. 1.

The $\sum \Delta q_X$ charge descriptor measures the total increase in negative charge at the conjugated hard centers of the conjugate base, and may be related with the ease of solvation of this species by hydrogen-bonds with the water molecules. In general, the larger the absolute value of this descriptor, the stronger is the base stabilization by solvation, and the stronger the acid in aqueous solution. These trends are apparent in Tables 3 and 4. The behavior of the other charge descriptor, $\Delta q_C$, is noteworthy. In most cases, there is a *decrease* of negative charge at the anionic carbon atom upon deprotonation. This is equivalent to saying that, in aqueous solution, this soft center tends to be *desolvated* with the acid dissociation. This contrasting behavior with the conjugated hard centers of the molecule indicates that hydrogen-bonding with the latter dominates all aspects of stability of the anion in water.

The energy $\Delta E$ and the charge parameter $\sum \Delta q_X$ are not entirely independent. A correlation $r = 0.87$ was obtained when the former values from Table 3 were plotted against the $\sum \Delta q_X$ values. This follows from the fact that, since the descriptor $\Delta E$ measures the ease of charge delocalization upon deprotonation, the greater this delocalization, the greater should be the amount of charge allocated to the conjugated hard centers. However, an important factor, besides resonance, in delocalizing charge in the gas-phase, is the polarizability of adjacent groups. In many cases, a more polarizable group may render a molecule more acidic in the gas-phase (smaller $\Delta E$ value), reducing, at the same time, charge delocalization at the conjugated hard centers. The result is a decrease of the $\sum \Delta q_X$ value, and a poorer anion solvation in aqueous solution. These opposing effects tend to cancel out, at least partially, in solution, with the result that two analogous acids with rather different acidities in the gas-phase, may exhibit similar acid behavior in aqueous solution. An example of this behavior is the pair acetophenone/acetone (**6**/**5**). The former is more acidic than the latter in the gas-phase by ca. 5 p$K_a$ units, estimated by the corresponding $\Delta G°$ values.[15] This result reflects the greater polarizability of the phenyl group, compared with the methyl substituent, and is in agreement with the corresponding calculated $\Delta E$ values, as mentioned above. In solution, however, the conjugate base of acetone (**5**)

($\sum\Delta q_X$=−0.266 [−0.285])[†] is better solvated than that of acetophenone (**6**) ($\sum\Delta q_X$=−0.240 [−0.249]) and this reduces the acidity gap between the two to about 1 p$K_a$ unit.

The same trend is observed for the series of ketones $CHX_2COCH_3$, and the series of nitroalkanes $CHX_2NO_2$ where X=H or Cl. Replacement of a hydrogen atom by a more polarizable, electronegative chlorine atom reduces the $\Delta E$ values and increases the intrinsic acidity of the members of the two series.

For the ketones, $\Delta E$ values decrease as follows: 0.6376 [0.6262] a.u. (acetone, **5**)>0.6035 [0.5988] a.u. (chloroacetone, **7**)>0.5895 [0.5859] a.u. (dichloroacetone, **8**). For the nitroalkanes, $\Delta E$ values are: 0.5925 [0.6103] a.u. (nitromethane, **15**)>0.5559 [0.5862] a.u. (chloronitromethane, **20**)>0.5515 [0.5731] a.u. (dichloronitromethane, **21**). In solution, however, the opposite trend is observed, as the absolute value of $\sum\Delta q_X$ decreases (less solvation of the conjugate base) with the increased replacement of H by Cl atoms. For the ketone series, $\sum\Delta q_X$ decreases as follows: 0.266 [0.285] (acetone, **5**)>0.257 [0.273] (chloroacetone, **7**)>0.225 [0.238] (dichloroacetone, **8**). For the nitroalkane series, $\sum\Delta q_X$ values are: 0.400 [0.525] (nitromethane, **15**)>0.364 [0.491] (chloronitromethane, **20**)>0.349 [0.476] (dichloronitromethane, **21**). From the good linear correlation between $\Delta G^\circ$ and $\Delta E$ values of Fig. 1, and with the $\Delta E$ values of Table 3, a difference of 16.5 kcal $mol^{-1}$ may be estimated between the $\Delta G^\circ$ values of acetone and dichloroacetone. At 298 K, this corresponds to a difference of ca. 12 p$K_a$ units in the gas-phase acidity of the two compounds. In water, this difference is reduced to 4.3 p$K_a$ units, in agreement with the opposing effects discussed above, which lead to a reduction of the acidity gap between the two compounds in aqueous solution.

When compounds of the form $CH_mZ_n$ are compared, a saturation effect is observed in their acidities, as the number of electron-withdrawing substituents Z increases. This effect has long been noticed.[21] In terms of resonance between group Z and the anionic carbon of the conjugate base $^-CZ_3$, it has been described[28] as "the extent to which depletion of the negative charge at a given site in a carbanion by resonance causes an attenuation in the stabilizing (or destabilizing) effect of a group when substituted into that ion." This attenuation is especially apparent for two series of acids of Tables 3 and 4, with Z=$NO_2$ and COMe. Values for the intrinsic acidity ($\Delta E$) of these compounds decrease, in an attenuated form, with the increasing substitution by Z. For Z=$NO_2$, $\Delta E$ values are: $CH_3NO_2$ (**15**), 0.5925 [0.6103] ≫ $CH_2(NO_2)_2$ (**26**), 0.5145 [0.5309]>$CH(NO_2)_3$ (**29**), 0.4886 [0.5110]. For Z=COMe, the following series is obtained: $CH_3COMe$ (**5**), 0.6376 [0.6262] ≫ $CH_2(COMe)_2$ (**18**), 0.5748 [0.5641]>$CH(COMe)_3$ (**22**), 0.5386 [0.5376]. For Z=CN, however, this attenuation is less evident: $CH_3CN$ (**2**), 0.6455 [0.6335] ≫ $CH_2(CN)_2$ (**12**), 0.5691 [0.5552] ≫ $CH(CN)_3$ (**31**), 0.5077 [0.4939]. This difference in behavior reflects the increasing repulsion, in polysubstituted systems, of the $NO_2$ and COMe groups, when stabilizing the anionic conjugate bases. This repulsion, which is not observed with the much smaller CN substituent, prevents full coplanarity and conjugation of the former groups in crowded systems with the anionic carbon atom. Thus, in polysubstituted acids, the nitro group is increasingly less effective than the cyano substituent in stabilizing the corresponding conjugate bases. This theoretical prediction is born out by experiment: nitromethane (p$K_a$=10.2) is much more acidic than acetonitrile (estimated p$K_a$ 25), whereas trinitromethane (p$K_a$=0.1) is a weaker acid than tricyanomethane (p$K_a$=−5.1).

In conclusion, the acidity of mono- and poly-substituted carbon acids in aqueous solution may be described with the aid of two readily available theoretical descriptors, $\Delta E$ and $\sum\Delta q_x$, without the need of explicit corrections arising from solvent effects.

The deprotonation energy $\Delta E$ is associated with the intrinsic, or gas-phase acidity of the compound, and is defined as the difference between the heats of formation of the acid anion and its protonated form. A good correlation ($r = 0.985$) between experimental $\Delta G^\circ$ values of thirteen carbon acids[15,16] and $\Delta E$ values calculated with the B3LYP/6-31G(d)//3-21G was obtained.

The charge descriptor $\sum\Delta q_x$ is defined as the total difference, between the anion and the neutral acid, of the charges at the hard centers conjugated with the ionizable carbon. This descriptor provides a measure of the contributions of aqueous solvation to the acid behavior in water.

A plot of the experimental vs calculated p$K_a$ values obtained by regression analysis of the data for thirty-two acids based on this model yielded a correlation coefficient $r = 0.958$ with a standard deviation of 2.4 p$K_a$ units, when a 3-21G//3-21G method was employed. When the hybrid DFT method B3LYP/6-31G(d)//3-21G was used, the correlation coefficient was $r = 0.950$ with a standard deviation of 2.6 p$K_a$ units. Thus, the use of a higher level of calculation did not lead to any improvement in the correlations with the experimental data in solution. This might be due to the uncertainties in the experimental data, which are often gross values that overlook keto–enol equilibria.

The two descriptors may be used to interpret various trends in acidity, observed upon transfer of the studied compounds from the gas-phase to aqueous solutions.

## Acknowledgements

The author is grateful to USACH-DICYT for supporting this work.

## References

1. De Proft, F.; Amira, S.; Choho, K.; Geerlings, P. *J. Phys. Chem.* **1994**, *98*, 5227–5233.
2. De Proft, F.; Langenaeker, W.; Geerlings, P. *Tetrahedron* **1995**, *51*, 4021–4032.

---

[†] Here, and in the following discussion, numbers quoted first refer to the values obtained by the 3-21G method of Table 3, numbers between square brackets refer to calculated values by the B3LYP/6-31G(d) method of Table 4.

3. Méndez, F.; Romero, M. L.; De Proft, F.; Geerlings, P. *J. Org. Chem.* **1998**, *63*, 5774–5778.
4. Tielens, F.; Langenaeker, W.; Geerlings, P. *Theochem* **2000**, *496*, 153–162.
5. Choho, K.; Van Lier, G.; Van de Woude, G.; Geerlings, P. *J. Chem. Soc. Perkin 2* **1996**, 1723–1732.
6. Yang, W.; Mortier, W. J. *J. Am. Chem. Soc.* **1986**, *108*, 5708–5711.
7. Contreras, R. R.; Fuentealba, P.; Galván, M.; Pérez, P. *Chem. Phys. Lett.* **1999**, *304*, 405–413.
8. Pérez, P.; Simón-Manso, Y.; Aizman, A.; Fuentealba, P.; Contreras, R. *J. Am. Chem. Soc.* **2000**, *122*, 4756–4762.
9. Pérez, P.; Toro-Labbé, A.; Contreras, R. *J. Phys. Chem. A* **2000**, *104*, 5882–5887.
10. Silva, C. O.; Da Silva, E. C.; Nascimento, M. A. C. *J. Phys. Chem. A* **2000**, *104*, 2402–2409.
11. D'Souza, F.; Zandler, M. E.; Deviprasad, G. R.; Kutner, W. *J. Phys. Chem. A* **2000**, *104*, 6887–6893.
12. Rezende, M. C. *J. Brazil. Chem. Soc.* **2001**, *12*, 73–80.
13. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, Jr., J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, J. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J.B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. GAUSSIAN 94W; Gaussian Inc., Pittsburgh PA, 1994.
14. Serjeant, E. P.; Dempsey, B. *Ionization Constants of Organic Acids in Aqueous Solution*, *IUPAC Chemical Data Series*; Pergamon Press: New York, 1979.
15. Taft, R. W.; Bordwell, F. G. *Acc. Chem. Res.* **1988**, *21*, 463–469.
16. Bartness, J. E.; Scott, J. A.; McIver, R. T. *J. Amer. Chem. Soc.* **1979**, *101*, 6049.
17. Bordwell, F. G.; Cornforth, F. J. *J. Org. Chem.* **1978**, *43*, 1763–1768.
18. Arnett, E. M.; Johnson, D. E.; Small, L. E.; Oancea, D. *Faraday Symp. Chem. Soc.* **1975**, *10*, 20.
19. Bordwell, F. G.; Der Puy, M. V.; Vanier, N. R. *J. Org. Chem.* **1976**, *41*, 1883–1885.
20. Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*, 2nd ed.; Gaussian Inc., Pittsburgh PA, 1995, pp. 144–150.
21. Pearson, R. G.; Dillon, R. L. *J. Am. Chem. Soc.* **1953**, *75*, 2439–2443.
22. Kresge, A. J.; Meng, Q. *J. Am. Chem. Soc.* **1998**, *120*, 11,830–11,831.
23. Bell, R. P. *Trans. Faraday Soc.* **1943**, *39*, 253.
24. Novak, M.; Loudon, G. M. *J. Am. Chem. Soc.* **1976**, *98*, 3591.
25. Amyes, T. L.; Richard, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 3129–3141.
26. Keefe, J. R.; Kresge, A. J. In *The Chemistry of Enols*; Rappoport, Z., Ed.; Wiley: Chichester, 1990.
27. Boyd, R. H. *J. Phys. Chem.* **1963**, *67*, 737.
28. Bordwell, F. G.; Bares, J. E.; Bartness, J. E.; McCollum, G. J.; Puy, M. V. D.; Vanier, N. R.; Matthews, W. S. *J. Org. Chem.* **1977**, *42*, 321–325.