Ab Initio Calculations of Absolute pK_a Values in Aqueous Solution I. Carboxylic Acids

Clarissa O. da Silva, Edilson C. da Silva, and Marco Antonio Chaer Nascimento*

*Departamento de Físico-Química, Instituto de Química, Universidade Federal do Rio de Janeiro, CT,
bloco A, sala 412, Cidade Universitária, Ilha do Fundão, Rio de Janeiro, RJ, Brazil CEP 21949−900*

Received: September 9, 1998; In Final Form: September 6, 1999

A thermodynamical cycle is proposed to calculate absolute pK_a values for a Brönsted acid in aqueous solution. The solvent (water) was represented by a dielectric using the polarizable continuum model (PCM), and the absolute pK_a values of some aliphatic carboxylic acids were computed. The results indicate that the proposed methodology seems to be capable of predicting reasonably good absolute pK_a values, although in some cases appreciable deviations are observed, which can be related to neglecting the molecular motion contributions (ΔG_{Mm}) to the solvation energy (ΔG_{solv}).

1. Introduction

According to the Brönsted definition, any compound which has a hydrogen atom is an acid, since it may be lost as a proton. Depending on the molecule, this process requires more or less energy, and in some cases, the process may be spontaneous. As proton transfer reactions are crucial in chemistry, it is important to quantify the tendency of the molecule to lose its hydrogen atom as an acidic proton. This is the role played by the quantity defined as pK_a.

Indeed, proton transfer is the way chosen by nature to link the extracellular (EC) and intracellular (IC) media. Information about the EC environment arrives at the cell interior throughout some chemical reactions, mainly proton transfer reactions. In pharmacokinetics, the pK_a of a certain drug is mandatory information in order to check its efficiency. Most drugs are weak acids or basis, present in solution as both the ionized and nonionized species. But only the nonionized species is able to cross the phospholipidic cellular membrane, because of its lipophilic character. Therefore, the transmembrane distribution of a weak electrolyte is usually determined by its pK_a, which is directly related to the ratio between the nonionized and ionized species, and the pH gradient across the membrane.[1]

The equilibrium of dissociation of a Brönsted acid depends on the interaction of the acid and its conjugate base with solvent molecules. Therefore, the pK_a value depends on the solvent medium were the measurements took place, and any reference to the pK_a value of a certain compound will be meaningful only if the solvent is specified. Experimentally, the most studied medium is water, which justifies our choice for the medium used in this paper. Although water is itself a Brönsted acid, the processes studied will not be affected by the solvent self-ionization because even the least acidic compound considered is still approximately 1×10^9 times more acidic than pure water.

The experimental pK_a values of several compounds, mainly organic acids in water,[2,3] are determined through very well-established methods.[4] On the other hand, from the theoretical side, the methodologies so far available furnish only relative pK_a values[5] or correlations between pK_a and other energy quantities in solution, using semiempirical[6−8] or ab initio methods.[9−11] More recently, Richardson et al.[9b] proposed a scheme to calculate pK_a values in solution. The Poisson−Boltzmann approach was used to compute the electrostatic contribution to ΔG_{solv} for the acid, and the nonelectrostatic contributions were estimated from experimental solvation energies of the alkane nearest in size and shape to the molecule in question. Also, the ΔG_{solv}^{el} for the base was computed making use of experimental pK_a values and experimental proton affinities. Thus, what is needed is a theoretical methodology that is able to predict absolute pK_a values. This would be important not only as an alternative route to obtain pK_a values but also, and most importantly, because in some situations the pK_a of a compound is so close to that of the solvent that its experimental determination becomes extremely difficult or almost impossible without the introduction of some artifice.[12,13]

In almost all of the previous works cited above, [9−11] the pK_a calculations in a vacuum were also performed and compared to the respective values in aqueous solution. The difference in pK_a values in both phases reflects the influence of the medium. It is important to stress that there are nowadays very well established and reliable methodologies to compute pK_a values in a vacuum.[14−16] In liquid phase, the main difficulties reside on how to take into account the solvent molecules and properly describe the modifications in the solute molecule, which arise from its interaction with the solvent.

The description of the ionization process in solution is the subject of this paper, where calculations of absolute pK_a values, in aqueous solution, are reported for the following organic carboxylic acids: formic, acetic, propanoic, butanoic, trimethyl acetic, fluoroacetic, and chloroacetic acids. Although alcohols and phenols are also under investigation, those compounds were chosen in order to be able to compare our results with the ones, for the same compounds, recently obtained using a methodology also based on a polarizable continuum model.

2. Theory

Following the conventional definition, a Brönsted acid is any compound that can loose a proton, and, in gas phase, the process is governed by the equilibrium constant K:

$$AH_g \overset{K}{\rightleftharpoons} A_g^- + H_g^+ \qquad (1)$$
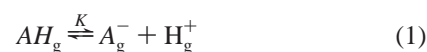
* Corresponding author fax, +55−21−2904746; e-mail, chaer@iq.ufrj.br

Carboxylic Acid Absolute pK_a Values

J. Phys. Chem. A, Vol. 103, No. 50, 1999 **11195**

The equilibrium dissociation constant is related to the variation in Gibbs standard free energy (ΔG_g^0) through the relationship

$$\Delta G_g^0 = -2.303RT\log K \text{ or } \Delta G_g^0 = 2.303RT\text{p}K_a{}^{gas} \quad (2)$$

The equivalent process, in aqueous solution, may be written as

$$A H_{aq} + H_2O_{aq} \overset{K'}{\rightleftharpoons} A_{aq}^- + HO_{3aq}^+ \quad (3)$$
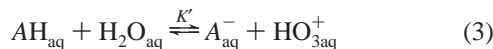
with

$$K' = \frac{[A_{aq}^-][H_3O_{aq}^+]}{[AH_{aq}][H_2O_{aq}]} \quad (4)$$

Since all of the H_3O^+ species in this model come from H^+ solvated by one water molecule ($[HO_{3aq}^+] = [H_{aq}^+]$), K is related to K' by the equation

$$K' = \frac{K}{[H_2O_{aq}]} \quad (5)$$

Thus, the relationship between the Gibbs standard free energy change (ΔG^0) and pK_a, in aqueous solution, becomes

$$\Delta G^0 = -2.303RT\log\left(\frac{K}{[H_2O_{aq}]}\right) \quad (6)$$

or

$$\Delta G^0 \text{ (kcal/mol)} = 1.36\ \text{p}K_a + 2.36 \quad (7)$$

if R and T are taken equal 1.98 kcal/mol.K and 298.15K, respectively. Equation 7 expresses a linear relationship between pK_a and ΔG^0, which should be used with caution.

There are many experimental measurements of the pK_a values, but if a comparison between experimental and theoretical values is intended, the reference state or initial conditions used in the experience must be mimicked or explicitly included in the theoretical model, in order to ensure a suitable comparison, since a thermodynamical methodology is adopted to obtain the ΔG^0 value. To make use of eq 7, it is necessary to calculate a thermodynamical quantity, the Gibbs standard free energy change (ΔG^0), related to the process described in ref 3. But, since ΔG^0 is a state property, it is completely determined by the initial and final states of the system and is independent of the path connecting them. In another words, it is almost always possible to propose a set of thermodynamical intermediate steps for the whole process, as long as the final and initial states are the same as the process being decomposed.

One of the many possible alternatives is the following Born–Haber thermodynamical cycle (TC), shown in Figure 1. From Figure 1, ΔG^0 is given by

$$\Delta G^0 = -\Delta G_{solv}(AH) + \Delta G_{vap}(H_2O) + \Delta G_{vac}^0 +$$
$$\Delta G_{solv}(A^-) + \Delta G_{solv}(H_3O^+) \quad (8)$$

where the $\Delta G_{vap}(H_2O)$ is the Gibbs free energy change related to the vaporization process and ΔG_{solv} are the solvation energy quantities.

According to Ben-Naim's[17,18] definition "the process of solvating *s* in *l* is the process of transferring *s* from a fixed position in the gas, into a fixed position in the liquid *l*, the process being carried out at a given temperature T and pressure
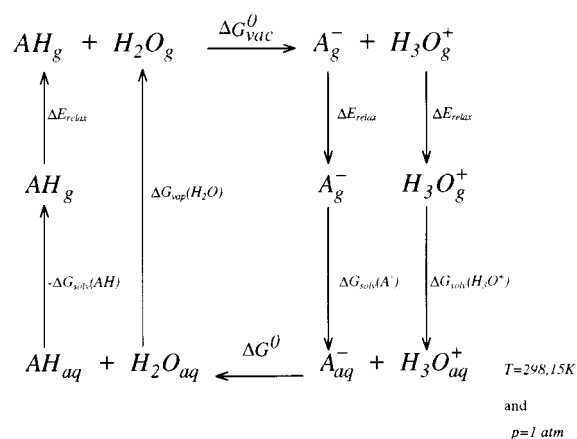

Figure 1. Proposed thermodynamical cycle (TC).

p". Before we proceed, it is important to comment on the differences in the definition of (ΔG_{solv}) based on the thermodynamical model[18] and the one used in the present work. In the TC chosen (Figure 1), the solute molecules are brought from the liquid to the gas phase, but keeping their respective solution phase optimized geometries. At those same geometries, the energy of the molecules, in gas phase, is computed and, in the last step, the molecules are allowed to relax in gas phase by performing a geometry optimization. ΔE_{relax} is the difference in energy when the molecular geometry is relaxed in gas phase. On the other hand, the solvation free energy (ΔG_{solv}) was originally defined[18] considering the molecules as hard spheres. Therefore, the values of solvation energy calculated from that approach cannot be directly compared to our ΔG_{solv} values, obtained by a quantum-mechanical model. Even if we had used the gas phase optimized geometry for both phases, the comparison would still not be appropriate because in our calculation the details of the molecular geometry of each species were taken into account. Besides that, we allowed for geometry relaxation, when going from the solution to the gas phase, which makes the comparison with the thermodynamical model even more improper. It is true that the original thermodynamical model also includes the possibility of releasing the constraint on the fixed position of the center of mass of the molecule, as discussed in ref 18, but apparently no solvation energies have been calculated using this extended model.

The way of treating the solvent (water) has to be consistent with the method used for all the other species in the TC. In principle one would have to optimize the geometry of the water molecule in water, bring it to the gas phase and finally let its geometry relax. Thus, at the end one would have computed not only the work to insert one water molecule in liquid water but, following the TC, also the work performed against all of the attractive forces exerted by the solvent molecules on a molecule of the same kind when removing it from the liquid phase. It would be more appropriate to identify the transformation involving the water molecule in the TC with the vaporization process, and not simply as a solvation. Thus, using $\Delta G_{solv}(H_2O)$ instead of $\Delta G_{vap}(H_2O)$ in the TC represents a source of error that leads to an overestimation of the pK_a values[11] as will be further discussed.

The choice of the H_3O^+ species to represent the solvated state of the proton deserves some comment. The structure of the hydrated proton is very important for understanding the properties of aqueous solutions of acids, including their pK_a values. The enthalpy of hydration of a proton is the largest among the singly charged atomic cations, and calculations show that much of this stabilization comes from adding the first water molecule.

Therefore, the convention of writing the proton in water as the species H_3O^+ is consistent with these values, but gives no indication of how the H_3O^+ ion is hydrated. Ab initio calculations[19] show that the hydrogen atoms in H_3O^+ carry substantial net positive charge to make three hydrogen bonds to other water molecules, forming the $H_9O_4^+$ species.[19] If another water molecule is added to $H_9O_4^+$ it adds preferentially to one of the external H_2O fragments rather than to the central H_3O^+. Thus, the first three water molecules complete the first solvation shell, and therefore one can look at the $H_9O_4^+$ species as the best representation of the hydrated proton. Structural and spectroscopic predictions making use of this model show excellent agreement with the experimental data.[19] Based on these studies, the choice of the H_3O^+ species to represent the hydrated proton in the proposed thermodynamic cycle is the best one consistent with the PCM model, if one considers that following this model the parametrization of the cavity already takes into account the solute first solvation shell. Thus, it would be inconsistent to use $H_5O_2^+$, $H_7O_3^+$, or even $H_9O_4^+$ to represent the hydrated proton because the extra water molecules, being part of the first solvation shell, would be counted twice.

Regarding the other molecules involved in the TC, their ΔG_{solv} energies are computed from the equation

$$\Delta G_{solv} = G_{solv} - G_g \qquad (9)$$

where G_{solv} is the total free energy of the system in solution, the proper reference states being considered, and G_g is the equivalent quantity in a vacuum. As stated in ref 20, it would follow Ben-Naim's[18] solvation energy definition if the gas phase optimized geometry was kept for both phases. The ΔG_{solv} energy was computed using the new version of the polarizable continuum model[21,22] (PCM) to describe the solvent and its interaction with the solute. The first version of this particular apparent surface charge (ASC) methodology was developed by Tomasi et al. in 1981,[23] and a representative review about the continuum solvation models available may be found in ref 20. A detailed explanation about the use of PCM will be presented in the next section.

ΔG_{vac}^0 is the standard Gibbs free energy change for the process in a vacuum, and is given by

$$\Delta G_{vac}^0 = \Delta H_{vac}^0 + T\Delta S_{vac}^0 \quad (T = 298.15 \text{ K}; p = 1 \text{ atm}) \quad (10)$$

As the entropic term at room temperature (ΔS_{vac}^0) is very small compared to the enthalpy variation (ΔH_{vac}^0) (for acetic acid, for instance, $\Delta S_{vac} = 0.64$ kcal/mol while $\Delta H_{vac} = 348.48$ kcal/mol, at room temperature), it can be neglected, and

$$\Delta G_{vac}^0 \approx \Delta H_{vac}^0 \quad (T = 298.15 \text{ K}; p = 1 \text{ atm}) \quad (11)$$

where

$$\Delta H_{vac}^0 = \Delta E_{vac}^0 + \Delta pV \quad (T = 298.15 \text{ K}; p = 1 \text{ atm}) \quad (12)$$

Assuming ideal behavior

$$\Delta H_{vac}^0 = \Delta E_{vac}^0 + \Delta nRT \quad (T = 298.15 \text{ K}; p \approx 0 \text{ atm}) \quad (13)$$

but as $\Delta n = 0$, the second term on the right-hand side of eq 13 disappears and

$$\Delta E_{vac}^0 = \Delta E_{vac,0 \text{ K}} + \Delta ZPE + \Delta E_{298.15 \text{ K}}^{rot} + \\ \Delta E_{298.15 \text{ K}}^{vib} + \Delta E_{298.15 \text{ K}}^{trans} \quad (14)$$

when the system is brought from $T = 0$ K to $T = 298.15$ K. It may also be assumed that the rotational, translational, and vibrational corrections, when the system is brought from $T = 0$ K to $T = 298.15$ K, are very similar for products and reactants. So, their contributions can be neglected. The change in internal energy (ΔE_{vac}^0) becomes

$$\Delta E_{vac}^0 \approx \Delta E_{vac,0 \text{ K}} + \Delta ZPE \qquad (15)$$

where $\Delta E_{vac,0 \text{ K}}$ is the variation in the internal energy at $T = 0$ K and ΔZPE is the difference between the total zero point energy correction for the products and that of the reactants.

Finally, as a continuum dielectric is employed to describe the solvent, it is not possible to compute its energetic contribution to the whole system explicitly. Otherwise, it would not be necessary to split the process into many intermediate steps. The construction of a TC, with a number of intermediate steps, while increasing the number of terms to be calculated in the expression of ΔG^0, makes possible the absolute determination of pK_a values.

3. Computational Details

According to the TC shown in Figure 1, we first optimized the geometry of each molecule in solution phase. The molecules were then transferred to the gas phase and their energies, at the respective solution phase optimized geometries, were computed. Finally, a geometry optimization in gas phase was performed, for each molecule, to obtain the relaxation energy.

All calculations have been performed using a 6-31+G** basis set, at the Hartree–Fock (HF) level. The gas phase calculations were performed using standard procedures, and, for the zero point energy corrections, the frequencies were scaled by a factor 0.9181, as suggested by Scott and Radom[24] which is suitable to the basis set and level of calculation employed.

For the solution phase calculations we used the PCM model[20,21] as implemented in the *Gaussian 94*[25] package. This new version contains very efficient subroutines for geometry optimization.[25–27] The initial geometries inputed to the PCM program were obtained from molecular mechanics calculations using the Dreiding II force field.[29]

To take full advantage of the proposed thermodynamic cycle to obtain pK_a values, one should consider electronic correlation effects, which are known to be important in the calculations of gas-phase acidity,[14–16] and presumably equally important in the solution phase. In the proposed TC, geometry relaxation is allowed for all of the species in both phases. However, with the PCM program we cannot perform geometry optimization in the solution phase including correlation effects, but just single point calculations. Thus, to use the TC consistently, the correlation effects in both phases were neglected. On the other hand, since as we follow the cycle, the two deprotonation processes being considered are in opposite directions (see Figure 1 and eq 8), the correlation effects, in the different phases, may hopefully compensate each other. Another possibility would be to optimize the solution-phase geometries at the HF level and use them to compute the correlation corrections. However, this procedure is known to lead to inconsistent results in both the gas[30,31] and solution[32,33] phases.

The geometry of each molecule was first optimized in solution phase using the PCM model. Subsequently, each molecule was transferred to the gas phase and its energy computed at the same geometry of the solution phase. As previously discussed, the solvation energy which would be computed from the energy difference between these two calculations would be the closest related to the thermodynamical model. Finally, the geometry
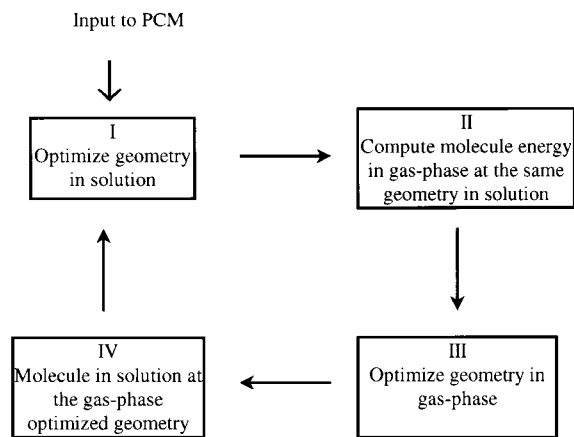
Carboxylic Acid Absolute pK_a Values

J. Phys. Chem. A, Vol. 103, No. 50, 1999 **11197**

Input to PCM

Figure 2. Geometry optimization scheme.

of each molecule was optimized in the gas phase and its relaxation energy determined.

To check the algorithm consistency, after geometry relaxation in gas phase, the molecules were put back into the solution phase for geometry reoptimization (see Figure 2).

Surprisingly, for some molecules the final geometries and energies differed appreciably from the ones obtained at the beginning of the cycle, i.e., after the first geometry optimization in solution. At first we believed that this discrepancy could be due to the charge renormalization procedure employed[34] in the geometry optimization. In the PCM model, when optimizing geometry, it is not presently possible to use the most efficient charge renormalization procedure. To check this possibility we performed the same sequence of calculations using the integral equation formalism (IEF)[35] for optimizing the geometries in solution. Both models allow the same kind of charge renormalization corrections, in geometry optimization procedures, but the corrections associated to the electric field operators (PCM) are larger than the ones related to the electric potential (IEF), and the former may be properly obtained if the best charge renormalization procedure is employed. In fact, with the IEF method, all of the discrepancies observed when reoptimizing the geometries in solution (IV → I, in Figure 2) disappeared. Thus, for all cases that presented such discrepancies, we first optimized the geometry in solution using the IEF model and then performed a single point calculation using the PCM model and the most efficient charge renormalization procedure. In principle these results would confirm our belief that the observed discrepancy had to do with the charge renormalization procedure. However, as pointed out by Mennucci,[36] the same result should be obtained using either the PCM or IEF formalism, and therefore the observed discrepancy was most certainly caused by problems in the version of D-PCM which we have used.

The ΔG_{solv} energy can be partitioned into five different contributions:

$$\Delta G_{solv} = \Delta G_{eletr} + \Delta G_{cav} + \Delta G_{rep} + \Delta G_{disp} + \Delta G_{Mm} \quad (16)$$

In the PCM version used, just the electrostatic (ΔG_{eletr}) component is solved self-consistently together with the solute wave function. The cavitation (ΔG_{cav}), repulsion (ΔG_{rep}), and dispersion (ΔG_{disp}) energy terms are calculated using semiempirical formulations. The self-consistent field determination of the repulsion and dispersion contributions is presently being implemented in the PCM model.[37] As the molecular motion contributions (ΔG_{Mm}) are normally very small, they are generally neglected.

Among the possible different definitions of cavity, the united atom topological model[38] (UATM) was chosen as it traditionally furnishes solvation energies very close to the experimental values. As shown in Figure 3, in this molecular shaped cavity model,[39] interlocking spheres are centered in the atoms nuclei, one sphere per atom, except for hydrogen atoms, which are inserted in a sphere centered on the atom to which it is bound.

4. Results and Discussion

The pK_a values obtained using the approach proposed in this work are shown in Table 1, together with the results derived from the methodology presented in ref 11 and with the available experimental data.

Figure 3 presents an alternative way of visualizing the thermodynamical cycle, proposed in this work, to calculate pK_a values in solution. It shows the molecules in both phases as well as the cavity model used. The geometries shown in the Figure 3 correspond to the optimized ones in each phase, for the formic acid.

A comparison with the results presented in ref 11 is particularly interesting because it reveals how differences in the methodology can drastically affect the pK_a values. It is important to mention that although pK_a values have not been explicitly given, the paper by Schüürmann et al.[11] contains all the data needed to compute them. A simple manipulation of eqs 11 and 13 of that paper allows us to write (using their notation)

$$\Delta G_{aq}^{SCF} = \Delta G^{SCF} + \Delta E_{aq}^{SCF} - \Delta E^{SCF}$$

ΔE_{aq}^{SCF} being equivalent to ΔG^0 defined in eq 7 of this paper. Thus, using the values of ΔG^{SCF}, ΔE_{aq}^{SCF}, and ΔE^{SCF} listed by those authors, the pK_a values shown in the second column of Table 1 have been computed. If we now consider our thermodynamical cycle (Figure 1) with $\Delta G_{solv}(H_2O)$ instead of ΔG_{vap}-(H_2O), the final expression for ΔG^0 (eq 8) would be equivalent to the one used by Schüürmann et al.[11] Thus, one should expect pK_a values, calculated with our TC modified with $\Delta G_{solv}(H_2O)$, similar to the ones obtained with the data and methodologies presented in ref 11. Comparing the pK_a values shown in columns 2 [pK_a(a)] and 4 [pK_a(b)] of Table 1, one sees that, except for the trimethylacetic acid, the pK_a values differ at most by 0.8 pK_a units. Such differences may be attributed to the procedure used for the cavity tessellation to compute the ΔG_{elect} term. While the number of initial tesserae on each sphere was set equal 196 in ref 11, we have used 60. Another possibility for the observed deviations could be related to differences in the entropic and other thermodynamical corrections. However, since the authors[11] did not indicate how those corrections have been made, that possibility could not be verified.

Independently of the sources of deviation, the results obtained using either the methodology and data from ref 11 or our data with the modified TC, show large deviations from the available experimental pK_a values. On the other hand, comparing the last two columns of Table 1 one sees that, except for the formic and halogenated acids, the present results, obtained from the TC shown in Figure 1, are in reasonably good agreement with the experiments. In fact, to our knowledge, these are the best absolute pK_a values reported in the literature. Even for the three compounds showing the largest deviations, the results obtained using the proposed TC figure among the ones exhibiting the closest agreement with the experimental data. Despite the large deviations, the theoretical values for those compounds are in the correct order of pK_a values and their relative pK_a values are very similar to the experimental ones. It is also worth noticing
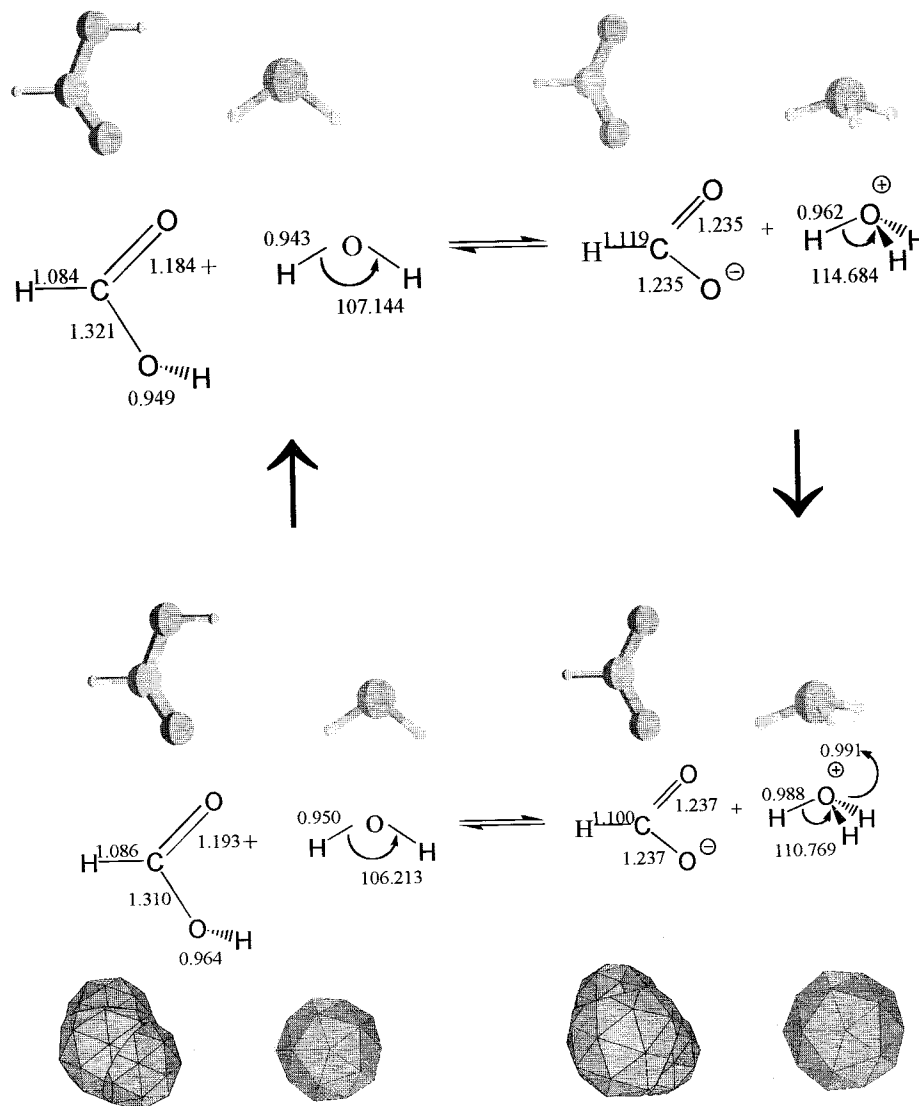
Figure 3. Thermodynamical cycle for formic acid.

TABLE 1: Comparison among Different pK_a Values[a]

compound	ΔG_{aq}^{SCF} [b]	pK_a [b]	$\Delta G^{0'}$ [c]	pK_a' [c]	ΔG^0 [d]	pK_a [d]	p$K_{a_{exp}}$ [e]
HCOOH	11.70	6.87	12.54	7.49	6.33	2.92	3.77
CH$_3$COOH	13.47	8.17	14.28	8.77	8.23	4.32	4.76
CH$_3$CH$_2$COOH	14.88	9.21	14.60	9.00	8.84	4.76	4.88
CH$_3$(CH$_2$)$_2$COOH			14.29	8.77	8.60	4.59	4.82−4.95 for CH$_3$(CH$_2$)$_n$COOH (n=2 to 7)
(CH$_3$)$_3$CCOOH	16.17	10.15	14.06	8.60	9.74	5.42	5.05
FCH$_2$COOH	9.51	5.25	9.95	5.58	3.91	1.14	2.66
ClCH$_2$COOH	9.41	5.18	10.51	5.99	4.50	1.57	2.86

[a] All ΔG values are in kcal/mol. [b] Reference 11. [c] Calculations using solvation and gas-phase energy values, according to Schüürmann's methodology: ΔG_{solv}(H$_2$O) computed instead of ΔG_{vap}(H$_2$O) and no geometry optimization in solution-phase. The TΔS term has been also computed. [d] Present work. [e] Reference 3.

that those compounds are exactly the ones presenting the lowest pK_a values. Therefore, the theoretical pK_a values for those compounds will be much more affected by small errors in the calculation of ΔG_{solv}.

From our experience with other classes of compounds (aliphatic alcohols, for instance), the smaller member of the class (methanol) always exhibits the largest deviation from the experimental pK_a value. This could be related to the fact that we did not consider the ΔG_{Mm} energy term when computing the ΔG_{solv} (eq 16). Neglecting this term is completely equivalent

to assuming that the neutral system (*A*H) and its conjugate base (*A*$^-$) would exhibit a similar motion in the solution medium. This assumption is certainly more drastic for the smaller anionic system since they present stronger electric fields at the surface of the cavity. Changes in the volume of the cavity could also be responsible for the larger deviations presented by the smaller member of a class of compounds. This effect is indeed larger for the formic acid, as expected, although the corresponding change in ΔG_{cav} is still very small to be responsible for the observed deviation in its pK_a value. As the size of the molecule

Carboxylic Acid Absolute pK_a Values

J. Phys. Chem. A, Vol. 103, No. 50, 1999 **11199**

TABLE 2: Quantities Employed in the pK_a Calculation Following the Approach Proposed in This Work

compound	$\Delta G_{solv} + \Delta E_{relax}{}^a$	$\Delta G_{solv}{}^b$	$E_{vac, 0 K}{}^c$	ZPE^d
HCOOH	−8.29	−6.58^e	−188.77751	21.27
		−5.5^f		
HCOO$^-$	−77.42		−188.20969	12.75
CH$_3$COOH	−7.86	−6.69^e	−227.82841	38.26
CH$_3$COO$^-$	−79.09	−82.2^g	−227.25421	29.73
		−77^h		
CH$_3$CH$_2$COOH	−8.14	−6.86^e	−266.86653	56.04
CH$_3$CH$_2$COO$^-$	−78.23		−266.29284	47.30
CH$_3$(CH$_2$)$_2$COOH	−7.33		−305.90473	73.51
CH$_3$(CH$_2$)$_2$COO$^-$	−77.45		−305.33121	64.68
(CH$_3$)$_3$CCOOH	−7.02		−344.94052	90.32
(CH$_3$)$_3$CCOO$^-$	−74.54		−344.36948	81.58
FCH$_2$COOH	−11.01		−326.67310	34.34
FCH$_2$COO$^-$	−76.30		−326.11567	26.09
ClCH$_2$COOH	−10.42		−686.71863	33.30
ClCH$_2$COO$^-$	−72.54		−686.16527	25.01
H$_2$O	−6.77	−6.3^f	−76.03123	13.32
H$_3$O$^+$	−106.61	−104^h	−76.31105	21.13

^{a} Solvation and relaxation energy (kcal/mol) calculated in the present work. ^{b} Solvation energies. ^{c} Total energy (hartree) for the system in gas phase, at $T = 0$ K. ^{d} Scaled zero point correction energy (kcal/mol). ^{e} Reference 40. ^{f} Reference 17 or 18. ^{g} Reference 9. ^{h} Reference 41.

increases the assumption that the neutral system (AH) and its conjugate base (A^-) would exhibit a similar motion in the solution medium becomes more acceptable. It is possible that the deviation exhibited by the halogenated acids is also related to neglecting the ΔG_{Mm} term (mainly vibrational and rotational motions), in this case not because of the size effect, but due to the interaction between the negative charge (A^-), mainly localized at the COO$^-$ group, and the dipole moment of the C−X (X=F, Cl) bond. The introduction of the water molecules into the solute cavity is under investigation, since it may reduce the errors in both cases.

Comparing the approach used in the present work with the one adopted in ref 11, it is clear that the much better pK_a values obtained with our TC resulted from the way water was treated (as discussed in the previous section) and from the consideration of geometry relaxation in the solution phase. Mainly for charged systems, geometry optimization is an important step when dealing with solvation.[25] Table 2 shows all of the quantities needed to obtain pK_a values according to the TC in Figure 1. From Table 2 one shows that although the differences in the ΔG_{solv} values, with and without considering geometry relaxation, are not very large, they are large enough to introduce significant change in the pK_a values.

5. Conclusions

The model proposed in this work seems to be capable of predicting reasonably good absolute pK_a values. The deviations observed for the formic acid can be attributed to neglecting of the ΔG_{Mm} term, when computing ΔG_{solv}, and is consistent with the deviation found for the smaller member of other classes of compounds being investigated. For the halogenated acids the deviations can be also attributed to neglecting the ΔG_{Mm} term, but further investigation is needed.

The relative success of the model must certainly be related with the proper way of treating water in the thermodynamical cycle and also to the consideration of geometry relaxation in solution phase.

Although the compounds investigated in this paper comprise a small range of pK_a values, the methodology is completely generic and may be applied to the other classes of compounds.

References and Notes

(1) *The Pharmacological Basis of Therapeutics*, 8th ed.; Guilman, A. G., Rall, T. W., Nies, A. S., Taylor, P., Eds; Pergamon Press: New York, 1990.

(2) Christen, H. R.; Vögtle, F. *Organische Chemie-Von den Grundlagen zur Forschung*; Otto Salle, Ed.; Verlag: Frankfurt, Germany, 1988; Vol. 1, p 419.

(3) March, J. *Advanced Organic Chemistry − Reactions, Mechanisms and Structure*, 4th ed.; Wiley-Interscience: New York, 1992; Chapter 8.

(4) Cookson, R. F. *Chem. Rev.* **1974**, *74*, 5.

(5) Jorgensen, W. L.; Briggs, J. M. *J. Am. Chem. Soc.* **1989**, *111*, 4190.

(6) Schüürmann, G. *Quant. Struct.-Act. Relat.* **1996**, *15*, 121.

(7) Schüürmann, G. *Quantitative Structure−Activity Relationships in Environmental Sciences*; Chen, F., Schüürmann, G., Eds.; SETAC Press: Pensacola, FL, 1997; Vol VII, p 225.

(8) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213.

(9) (a) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610. (b) Richardson, W. H.; Peng, C.; Bashford, D.; Noodleman, L.; Case, D. A. *Int. J. Quantum Chem.* **1997**, *61*, 207. (c) Li, G. S.; Ruiz-Lópes, M. F.; Maigret, B. *J. Phys. Chem A* **1997**, *101*, 7885.

(10) Andzelm, J.; Kölmel, C.; Klamt, A. *J. Chem. Phys.* **1995**, *103*, 9312.

(11) Schüürmann, G.; Cossi, M.; Barone, V.; Tomasi, J. *J. Phys. Chem. A* **1998**, *102*, 6706.

(12) Reeve, W.; Erikson, C. M.; Aluotto, P. F. *Can. J. Chem.* **1979**, *57*, 2747.

(13) Bordwell, F. G. *Acc. Chem. Res.* **1988**, *21*, 456.

(14) Motta Netto, J. D.; Nascimento, M. A. C. *J. Phys. Chem.* **1996**, *100*, 15105.

(15) Siggel, M. R. F.; Thomas, D. T.; Saethre, L. J. *J. Am. Chem. Soc.* **1988**, *110*, 91.

(16) Gao, J.; Garner, D. S.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1986**, *108*, 4784.

(17) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016.

(18) Ben-Naim, A. *Solvation Thermodynamics*; Plenum Press: New York, 1987; Chapter 1.

(19) (a) Newton, M. D.; Ehrenson, S. *J. Am. Chem. Soc.* **1971**, *93*, 4971. (b) Newton, M. D. *J. Chem. Phys.* **1997**, *67*, 5535.

(20) Tomasi, J.; Persico, M. *Chem Rev.* **1994**, *94*, 2027.

(21) Cammi, R.; Tomasi, J. *J. Comput. Chem.* **1995**, *16*, 1449.

(22) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett.* **1996**, *255*, 327.

(23) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.

(24) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.

(25) *Gaussian 94* (Revision E.2) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; 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. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc.: Pittsburgh, PA, 1995.

(26) Cammi, R.; Tomasi, J. *J. Chem. Phys.* **1994**, *101*, 3888.

(27) Cossi, M.; Mennucci, B.; Cammi, R. *J. Comput. Chem.* **1996**, *17*, 57.

(28) Barone, V.; Cossi, M.; Tomasi, J. *J. Comput. Chem.* **1998**, *19*, 404.

(29) POLYGRAF v 3.2.1, Molecular Simulation Inc., 1992.

(30) Curtis, L. A.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 1293.

(31) Bauschlicher, C. W.; Partridge, H. *J. Chem. Phys.* **1995**, *103*, 1788.

(32) Adamo, C.; Cossi, M.; Barone, V. *J. Comput. Chem.* **1997**, *18*, 1993.

(33) Barone, V.; Arnaud, R. *J. Chem. Phys.* **1997**, *106*, 8727.

(34) Cossi, M.; Mennucci, B.; Pitarch, J.; Tomasi, J. *J. Comput. Chem.* **1998**, *19*, 833.

(35) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032.

(36) Mennucci, B., personal communication.

(37) Amovilli, C.; Mennucci, B. *J. Phys. Chem. B.* **1997**, *101*, 1051.

(38) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210.

(39) The molecular cavities in Figure 3 were obtained using the program GEOMVIEW v1.6.1 (Software Development Group, Geometry Center, 1300 South 2nd St, Suite 500, Minneapolis, MN 55454, USA). The interface GEOMVIEW/GAUSSIAN94 was written by C. Pommeli.

(40) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Sol. Chem.* **1981**, *10*, 563.

(41) Pearson, R. G. *J. Am. Chem. Soc.* **1986**, *108*, 6109.