Theoretical Calculation of pKa Using the Cluster −Continuum Model

2 AUTHORS, INCLUDING:

Josefredo R Pliego
Federal University of São João del-Rei

76 PUBLICATIONS **1,728** CITATIONS

Theoretical Calculation of pK_a Using the Cluster−Continuum Model

Josefredo R. Pliego, Jr. and José M. Riveros*

Instituto de Química, Universidade de São Paulo, Caixa Postal 26077, São Paulo, Brazil, CEP 05513-970

Received: April 10, 2002

The pK_a's of 17 species from −10 to 50 were calculated using the ab initio MP2/6-311+G(2df,2p) level of theory and inclusion of solvent effects by the cluster−continuum model, a hybrid approach that combines gas-phase clustering by explicit solvent molecules and solvation of the cluster by the dielectric continuum. In addition, the pure continuum methods SM5.42R and PCM were also used for comparison purposes. Species such as alcohols, carboxylic acids, phenol, acetaldehyde and its hydrate, thiols, hydrochloric acid, amines, and ethane were included. Our results show that the cluster−continuum model yields much better agreement with experiment than do the above-mentioned pure continuum methods, with a rms error of 2.2 pK_a units as opposed to 7 pK_a units for the SM5.42R and PCM methods. The good performance of the cluster−continuum model can be attributed to the introduction of strong and specific solute−solvent interactions with the molecules in the first solvation shell of ions. This feature decreases the dielectric continuum contribution to the difference in the solvation free energy between ions, making the method less susceptible to error because of the continuum contribution to solvation. Because the method is not based on extensive parametrizations and it is shown to fare well for several functional groups, the present results suggest that this method could be used as a general approach for predicting reliable pK_a values.

Introduction

The ionization of Brønsted acids and bases in aqueous solution is a well-known chemical phenomenon that plays a central role in much of chemistry and biochemistry. Acidity is generally quantified by the equilibrium constant K_a of the prototype reaction 1 or more conveniently by the pK_a of the equilibrium constant.

$$HA(aq) \rightleftharpoons H^+(aq) + OH^-(aq) \qquad (1)$$

Acidities of a wide variety of organic and inorganic compounds have been studied over the years, and pK_a's have been measured by different experimental methods.[1] Yet, the accurate experimental determination of pK_a is not a trivial task in many cases and may in fact pose a significant experimental challenge. This is so for species such as reaction intermediates, very strong acids, and very weak acids. Nevertheless, a priori knowledge of pK_a's can be crucially important in situations such as those encountered in the development of biologically active molecules. Thus, it is of considerable interest to develop reliable theoretical protocols that can be used to calculate this property and complement experimental techniques. This is in general a very ambitious and difficult objective to achieve, considering the fact that the ± 2 kcal mol^{-1} benchmark criterion usually adopted for theoretical calculations of thermochemical parameters, the so-called chemical accuracy, amounts to an error of 1.47 pK units!

Many reports describing empirical or ab initio approaches to the problem of calculating pK_a can be found in the literature. For example, Jorgensen et al.[2,3] pioneered the use of ab initio methods coupled with free-energy perturbation to include the effect of the solvent. The pK_a's of very weak acids such as CH$_3$-CN, CH$_3$NH$_2$, and CH$_3$CH$_3$ were calculated using intermolecular potentials capable of reproducing the ion−water interaction. For CH$_3$CN, a pK_a of 28.4 was calculated, which is in excellent agreement with the most recent experimental and reliable value of 28.9.[4] For the other species, comparison is more difficult because of uncertainties in the experimental data. Subsequent theoretical studies adopted mainly continuum models to represent the solvent and to include explicitly the effect of the solvent in the calculation. This is illustrated in the study of Karplus and co-workers[5] in which a continuum dielectric method based on the solution of the Poisson equation was applied to calculate absolute pK_a values for some organic acids. This approach led to relatively poor agreement with experimental data and deviations approaching 8 pK_a units. Tunon et al.[6] obtained much better agreement between theory and experiment by using the polarizable continuum model (PCM) in the calculation of protonation free energies of amines in aqueous solution, with a deviation of $\sim$2.5 kcal mol^{-1}. Agreement is even better when relative values are compared. Subsequently, an adjusted radius was introduced for oxygen in the PCM model to account for the acidity of alcohols.[7] These calculations, when applied to methanol, yield a deprotonation solution free energy that is identical to the experimental data. However, the error increases for other alcohols and reaches 2.8 kcal mol^{-1} for t-BuOH. The apparently good results obtained with continuum models can therefore be attributed to the use of an adequate atomic radius and to the fact that the test molecules have the same functional group. This latter feature contributes to a cancellation of errors and to marked improvement in the final results.

Following some of these initial efforts, a number of reports have appeared in the literature in the last 7 years tackling the problem of calculating pK_a's.[8−24] Some of these can be briefly summarized. For example, Richardson et al.[8] have used density functional theory coupled to a dielectric continuum model to calculate the pK_a's of different organic species. Although

* Corresponding author. E-mail: jmrnigra@iq.usp.br. Phone/Fax: 55-11-3091-3888.

reasonable results were obtained, they recognized that ion solvation represents a difficult problem for continuum models. In fact, the calculations reported in ref 8 point out that to obtain reliable values of pK_a an accurate procedure is needed to calculate solvation free energies in addition to high-level ab initio methods for the gas-phase contribution. By comparison, Schuurmann et al.[9] concluded that the main source of error in the calculation of the pK_a's of a series of carboxylic acids stemmed from inaccurate gas-phase free energies. Shields et al.[10,11] addressed this problem by using high-level ab initio theory coupled with the CPCM continuum model, and this method proved to yield accurate pK_a's (within $^1/_2$ pK_a unit) for several carboxylic acids. A somewhat different and very difficult problem was studied by Li et al.[12] They focused on the problem of calculating the pK_a's of the hydrated metallic cations Mn^{2+}, Mn^{3+}, Fe^{2+}, and Fe^{3+}. Because of the nature of these systems, a combination of discrete and continuum methods had to be used to represent the solvent. For the highly charged cations, the deficiency of the dielectric continuum is evident, and the authors concluded that for these cases the second solvation shell must be explicitly included.

Continuum models that use a fixed atomic radius for each atom are clearly unable to predict accurate solvation free energies of ions in aqueous solution.[25] To overcome this problem, Tomasi et al.[26] proposed a new parametrization of the PCM model in which atoms in different functional groups have different atomic radii. This approach, named PCM-UAHF, led to considerable improvement over the PCM model. However, this method suffers serious drawbacks when it is used to compute activation free energies because transition states do not necessarily have well-defined functional groups. Furthermore, its good performance is limited to the restricted functional groups used in the parametrization, and there is no certainty that it would work for other systems. Thus, general theoretical methods that are more reliable than pure continuum models are needed to calculate the solvation free energy of ions. In short, it is not possible to predict reliable pK_a's unless the corresponding solvation free energies (ΔG^*_{solv}) can be accurately calculated.

Liquid simulation coupled with free-energy perturbation[27,28] is an accurate approach to the calculation of ΔG^*_{solv}. However, it requires knowledge of intermolecular potentials including many-body effects[29,30] and the use of various perturbation steps and thus becomes a less practical approach. Recently, we have proposed a cluster−continuum model to be used in calculating the solvation free energy of ions.[25] This is a hybrid method in which the ion interacts explicitly with solvent molecules to form a cluster, and the resulting super molecule is allowed to interact with the dielectric continuum. We have shown that this approach is superior to pure continuum models such as the PCM[31] and SM5.42R[32] applied to the bare gas-phase ions and is capable of predicting solvation free energies with a homogeneous deviation from experimental data. Inclusion of explicit solvent molecules eliminates a key deficiency of continuum models when dealing with ions, namely, the strong and specific solute− solvent interactions. Our method has also been applied to the theoretical characterization of the free-energy profile of the OH^- + methyl formate reaction in aqueous solution and resulted in calculated reaction and activation free energies that are in excellent agreement with experimental data.[33] This is a particularly important result because reactions involving hydroxide ion in aqueous solution are considered to be very difficult to describe theoretically.[34] In the present work, we have extended the use of the cluster−continuum method to the challenging problem of calculating the pK_a's of different organic species. Our results show that this approach works very well and suggests a very general method for calculating pK_a's that is not based on any kind of extensive parametrization procedure.

Calculation of pK_a

The theoretical calculation of pK_a can be carried out by following the direct definition shown in eq 1 or by using the proton-transfer reaction (eq 2):

$$HA + H_2O \rightarrow A^- + H_3O^+ \qquad (2)$$

Both reactions involve the formation of charged species starting from neutral molecules, a difficult problem for theoretical methods. By comparison, reactions that conserve the number of charged species are more suitable for accurate calculations of changes in solvation free energies. Thus, we have opted to use a proton-transfer reaction between the acid and the hydroxide anion for our prototype process, as shown in eq 3,

$$HA + OH^- \rightarrow A^- + H_2O \qquad \Delta G^*_{sol} \qquad (3)$$

that leads to the following equilibrium in eq 4:

$$\frac{[H_2O][A^-]}{[HA][OH^-]} = e^{-\Delta G^*_{sol}/RT} \qquad (4)$$

Using the autoionization equilibrium for water,

$$K_w = [H^+][OH^-] = 1.0 \times 10^{-14} \qquad (5)$$

and taking $[H_2O] = 55.5$ mol L^{-1}, we can arrive at our final expression for the pK_a:

$$pK_a(HA) = \frac{\Delta G^*_{sol}}{(2.303)RT} + 15.74 \qquad (6)$$

The reaction free energy in aqueous solution (ΔG^*_{sol}) can then be calculated through eq 7:

$$\Delta G^*_{sol} = \Delta G^*_g + \Delta G^*_{solv}(A^-) + \Delta G^*_{solv}(H_2O) - \Delta G^*_{solv}(OH^-) - \Delta G^*_{solv}(HA) \quad (7)$$

where ΔG^*_g is the gas-phase reaction free energy (1 mol L^{-1} standard state) relative to eq 3 and $\Delta G^*_{solv}(X)$ is the solvation free energy of species X according to the Ben-Naim definition.[35] By combining eqs 6 and 7, the pK_a is then obtained for the HA acid of eq 3.

Similarly, the reaction shown in eq 8 was used for the ionization of BH^+ acids,

$$BH^+ + H_2O \rightarrow B + H_3O^+ \qquad \Delta G^*_{sol} \qquad (8)$$
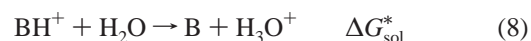
with the pK_a calculated according to eqs 9 and 10:

$$pK_a(BH^+) = \frac{\Delta G^*_{sol}}{(2.303)RT} - 1.74 \qquad (9)$$

$$\Delta G^*_{sol} = \Delta G^*_g + \Delta G^*_{solv}(H_3O^+) + \Delta G^*_{solv}(B) - \Delta G^*_{solv}(BH^+) - \Delta G^*_{solv}(H_2O) \quad (10)$$

The use of these equations in calculating the pK_a of any HA or BH^+ acid requires the computation of the gas-phase reaction

free energy by high-level ab initio methods. The solvation contribution can be calculated by several different procedures: (i) Monte Carlo free-energy perturbation, (ii) continuum models, or (iii) a hybrid discrete-continuum approach such as the cluster−continuum model.[25] In the last case, it is advantageous to consider the ion that is explicitly solvated as being directly involved in the proton-transfer reaction. Thus, in the cluster−continuum model, the calculation of the pK_a of an acid HA makes use of the reaction given by eq 11:

$$HA + OH^-(H_2O)_3 \rightarrow A^-(H_2O)_n + (4-n)H_2O \qquad \Delta G^*_{sol}$$
(11)
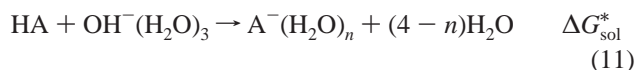
The idea of this method is to represent the ion as a cluster formed by the ion and an optimal number of solvent molecules and then to solvate the cluster using a dielectric continuum. This optimal number is chosen in such a way as to achieve maximum calculated stability for the ion in solution. For the hydroxide ion, the optimum number of water molecule corresponds to three, as previously demonstrated.[25] For many other monovalent ions, the ideal number of solvent molecules varies between two and three. On the right side of eq 11, the A^- species is again explicitly solvated with water molecules whereas the solvation of the neutral HA is calculated by a dielectric continuum. Thus, on the basis of eq 11, the pK_a of HA can be obtained by eq 12:

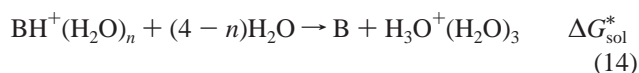$$pK_a(HA) = \frac{\Delta G^*_{sol}}{(2.303)RT} + 15.74 + (3-n)\log[H_2O] \qquad (12)$$

and

$$\Delta G^*_{sol} = \Delta G^*_g + \Delta G^*_{solv}(A^-(H_2O)_n) + (4-n)\Delta G^*_{solv}(H_2O) -$$
$$\Delta G^*_{solv}(OH^-(H_2O)_3) - \Delta G^*_{solv}(HA) \qquad (13)$$

As in eq 7, ΔG^*_g is the gas-phase reaction free energy (1 mol L^{-1} standard state) relative to eq 11 whereas $\Delta G^*_{solv}(X)$ is the solvation free energy of species X calculated by a continuum model.

For BH^+ acids, we make use of the reaction given by eq 14:

$$BH^+(H_2O)_n + (4-n)H_2O \rightarrow B + H_3O^+(H_2O)_3 \qquad \Delta G^*_{sol}$$
(14)

The calculation of pK_a is then obtained from eq 15

$$pK_a(BH^+) = \frac{\Delta G^*_{sol}}{(2.303)RT} - (4-n)\log[H_2O] \qquad (15)$$

where the solution free energy is calculated by

$$\Delta G^*_{sol} = \Delta G^*_g + \Delta G^*_{solv}(H_3O^+(H_2O)_3) + \Delta G^*_{solv}(B) -$$
$$\Delta G^*_{solv}(BH^+(H_2O)_n) - (4-n)\Delta G^*_{solv}(H_2O) \qquad (16)$$

In our present work, eqs 11−16 were used in the calculation of pK_a using the cluster−continuum model, and following the procedure outlined in our original report,[25] the continuum contribution was calculated by the isodensity polarizable continuum model (IPCM).[36] In addition, for comparison, we have tested the performance of pure continuum models such as the PCM[31] and the SM5.42R.[32] In these cases, eq 3 and eqs

6−10 were used in calculating the pK_a's that refer to the bare gas-phase ion solvated by a continuum.

Ab Initio Calculations

The structures of neutral molecules, ions, and ionic clusters were obtained by full optimization at the HF/6-31+G(d,p) level of theory, and the nature of the stationary points (minimum) was confirmed by harmonic frequency calculations. Single-point calculations at the MP2/6-311+G(2df,2p) level were made for all optimized species to obtain reliable proton-transfer energies. The data generated by ab initio methods were then used to calculate the gas-phase thermodynamic properties by standard statistical mechanics methods.[37]

The solvation of neutral species and ionic clusters by the bulk solvent was calculated by the continuum IPCM method using an isodensity of 0.0004 in conjunction with the MP2/6-31+G-(d,p) wave function and a dielectric constant of 78.0. These calculations were applied in the computation of pK_a using the cluster−continuum model, as explained in the previous section. Furthermore, the pure continuum models PCM and SM5.42R were also used in conjunction with the bare ions and neutrals to calculate pK_a. For the PCM method, we have used the atomic radii internally stored in Gamess and the BEM routines including correction for the escape charges. The HF/6-31+G(d,p) wave functions were used in the PCM calculations. For the solvation model of Truhlar et al.,[32] we have made use of the SM5.42R/HF/6-31G(d) method. The gas-phase ab initio calculations and the continuum IPCM method were performed using the Gaussian 94 program[38] whereas the calculations using PCM and SM5.42R were made using the Gamess[39] and Gamesol[40] programs, respectively.

Results and Discussion

The pK_a values (from −10 to 50) of 17 species including different functional groups were calculated in our study. The calculated thermodynamic data for the proton-transfer reactions are listed in Tables 1 and 2. Results for pK_a's calculated by the cluster−continuum model, the PCM method, and the SM5.42R method are listed in Table 3 along with the respective experimental values. The performance of each theoretical model is displayed in Figures 1−3 as graphs correlating theoretical and experimental pK_a values. The much higher accuracy of the cluster−continuum model is easily observed and leads to a root-mean-squared error (rms error) of 2.2 pK_a units whereas the SM5.42R and PCM methods have rms errors in the vicinity of 7 pK_a units. The largest error of the cluster−continuum method occurs for $CH_3NH_3^+$ and amounts to 4.6 pK_a units. By comparison, the SM5.42R and PCM methods reach their largest error of 14 pK_a units with NH_4^+ and 15 pK_a units with CH_3-NH_3^+. For CH_3CH_3 and CH_3NH_2, no reliable experimental pK_a's are available. However, Jorgensen and Briggs[3] have estimated the pK_a's of CH_3CH_3 and CH_3NH_2 using free-energy perturbation and have obtained values of 52 ± 2 and 33.3, respectively, which are in excellent agreement with the values of 50.6 and 32.2 that are predicted by the cluster−continuum model. However, the SM5.42R method predicts 53.0 and 42.8 for these two compounds whereas the PCM method predicts 47.8 and 32.6, respectively.

As mentioned above, accurate calculations of pK_a's also require good-quality gas-phase proton-transfer energies besides accurate calculations of solvation free energies. Indeed, an important source of error in the cluster−continuum model is the inaccuracy of the MP2/6-311+G(2df,2p) method for car-boxylic acids, as shown in Table 1. The error for the proton-

TABLE 1: Calculated Free Energies of the Gas-Phase and Aqueous Solution Proton-Transfer Reactions from HA to OH⁻ and from BH⁺ to H₂O[a]

HA/BH⁺	MP2/6-311+G(2df,2p)	ΔG^*_g (theor)	ΔG^*_g (exptl)[b]	$(\Delta\Delta G_{solv})^c$ (SM5.42R)	$(\Delta\Delta G_{solv})^c$ (PCM)	$(\Delta G^*_{sol})^d$ (SM5.42R)	$(\Delta G^*_{sol})^d$ (PCM)
CH₃OH	−6.80	−8.53	−8.9	21.71	10.68	13.18	2.15
EtOH	−10.45	−12.69	−12.4	26.60	13.76	13.91	1.07
HCOOH	−47.70	−49.93	−45.8	34.32	21.48	−15.61	−28.45
CH₃COOH	−43.86	−45.98	−43.0	35.76	20.04	−10.22	−25.94
HCL	−60.69	−55.62	−56.0	26.78	15.76	−28.84	−39.86
PhOH	−42.14	−43.14	−41.8	44.83	27.81	1.69	−15.33
H₂S	−41.06	−38.38	−39.3	20.08	17.38	−18.30	−21.00
CH₃SH	−34.88	−32.76	−33.5	25.27	19.24	−7.49	−13.52
NH₃	14.97	12.46	12.8	18.66	4.63	31.12	17.09
CH₃NH₂	14.35	11.10		25.86	11.92	36.96	23.02
CH₃CHO	−24.58	−25.27	−25.1	34.05	22.35	8.76	−2.92
CH₃C(OH)₂	−24.81	−26.97		34.39	20.70	7.42	−6.27
CH₃CH₃	31.51	27.26	27.3	23.58	16.45	50.84	43.71
NH₄⁺	41.09	38.76	38.0	−4.48	−4.16	34.28	34.60
CH₃NH₃⁺	51.92	49.56	48.9	−15.04	−11.55	34.52	38.01
CH₃OH₂⁺	15.20	15.48	15.5	−12.23	−11.01	3.25	4.47
EtOH₂⁺	20.06	20.59	20.3	−17.98	−15.84	2.61	4.75

[a] Equations 3 and 8. Units of kcal mol⁻¹. [b] Experimental data taken from ref 41. [c] $\Delta\Delta G_{solv}$ represents the free-energy contribution calculated by the continuum method. [d] See eqs 7 and 10.

TABLE 2: Calculated Free Energies of the Gas-Phase and Aqueous Solution Proton-Transfer Reactions Shown in Reactions 11 and 14[a]

HA/BH⁺	n[b]	MP2/6-311+G(2df,2p)	ΔG^*_g	$(\Delta\Delta G_{solv})^c$	$(\Delta G^*_{sol})^d$
CH₃OH	3	−2.05	−1.69	2.08	0.40
EtOH	3	−3.31	−2.72	3.18	0.46
HCOOH	2	−11.47	−19.38	−3.03	−22.42
CH₃COOH	2	−8.29	−16.85	−3.50	−20.35
HCl[e]	2	−18.31	−26.58	−9.90	−36.48
PhOH	3	−20.16	−18.01	6.35	−11.66
H₂S	2	0.85	−7.99	−8.09	−16.08
CH₃SH	2	5.86	−3.62	−5.96	−9.58
NH₃	2	36.92	27.70	−8.16	19.54
CH₃NH₂	2	36.43	26.97	−6.87	20.09
CH₃CHO	3	−5.05	−2.28	5.29	3.01
CH₃C(OH)₂	2	6.88	−1.75	−2.03	−3.79
CH₃CH₃	2	62.1	51.24	−6.03	45.22
NH₄⁺	3	15.72	16.45	1.95	18.39
CH₃NH₄⁺	2	8.43	16.90	8.73	25.64
CH₃OH₂⁺	2	−12.43	−5.10	7.50	2.40
EtOH₂⁺	2	−10.56	−3.33	2.42	−0.91

[a] Units of kcal mol⁻¹. [b] Number of water molecules in the cluster. [c] $\Delta\Delta G_{solv}$ represents the free-energy contribution calculated by the IPCM method. [d] See eqs 13 and 16. [e] Includes an anharmonic correction.

transfer reaction between the carboxylic acids and the hydroxide ion amounts to 4 kcal mol⁻¹ for HCOOH and 3 kcal mol⁻¹ for

CH₃COOH. If we use the experimental gas-phase proton-transfer reaction values listed in Table 1, the pK_a's predicted by the cluster−continuum model would be 4.1 and 4.8 for HCOOH and CH₃COOH, respectively. These corrected values are in excellent agreement with the experimental data, with deviations of less than 0.4 pK_a units.

The SM5.42R and PCM models apparently fare poorly in describing the solvation of ions with the charge localized on oxygen atoms, and this leads to obvious consequences in the calculation of pK_a. This is clearly noticeable in Table 3 where the SM5.42R method predicts erroneous pK_a's for the respective acids, with the exception of HCOOH. The PCM method works better but predicts erroneous pK_a's for HCOOH, CH₃COOH, and CH₃CHO. The same poor performance of the continuum models occurs with H₂S and CH₃SH. However, our cluster− continuum approach works very well for these systems. The largest error occurs for phenol, but this is partly due to inaccurate calculation of the gas-phase energetics. This is verified by the fact that a pK_a of 8.2 is obtained by using the experimental gas-phase proton-transfer free energy. Likewise, systems with charge located on nitrogen are not well described by the continuum models. The very weak base NH₃ has an extrapolated pK_a of 33, and our cluster−continuum model predicts a value

TABLE 3: Comparison between Calculated and Experimental pK_a Values

HA/BH⁺	experimental[a]	cluster−continuum model	SM5.42R	PCM
CH₃OH	15.5	16.03	25.40	17.32
EtOH	15.9	16.08	25.93	16.52
HCOOH	3.75	1.03	4.30	−5.11
CH₃COOH	4.76	2.57	8.25	−3.27
HCl	−6.1	−9.25	−5.40	−13.47
PhOH	9.99	7.19	16.98	4.50
H₂S	7.05	5.69	2.33	0.35
CH₃SH	10.33	10.46	10.25	5.83
NH₃	33	31.82	38.55	28.27
CH₃NH₂		32.21	42.83	32.61
CH₃CHO	16.73	17.95	22.17	13.60
CH₃C(OH)₂	13.31	14.71	21.18	11.14
CH₃CH₃		50.63	53.00	47.78
NH₄⁺	9.25	11.73	23.38	23.61
CH₃NH₃⁺	10.66	15.30	23.56	26.11
CH₃OH₂⁺	−2.05	−1.73	0.64	1.53
EtOH₂⁺	−1.94	−4.16	0.17	1.74
rms error[b]		2.2	7.2	7.3

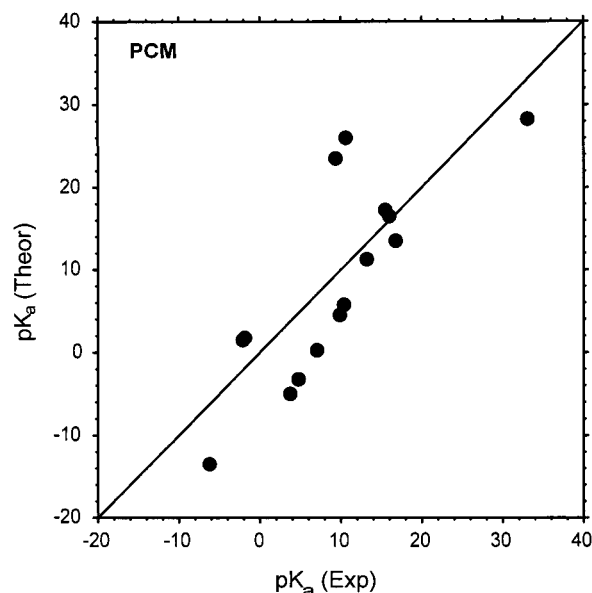[a] Values taken from ref 42. [b] Root-mean-square error.
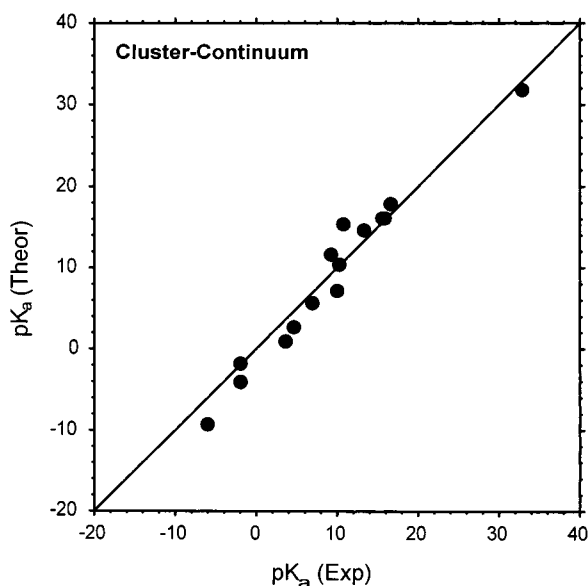
Figure 1. Correlation between experimental pK_a values and calculated values obtained from the cluster−continuum method.
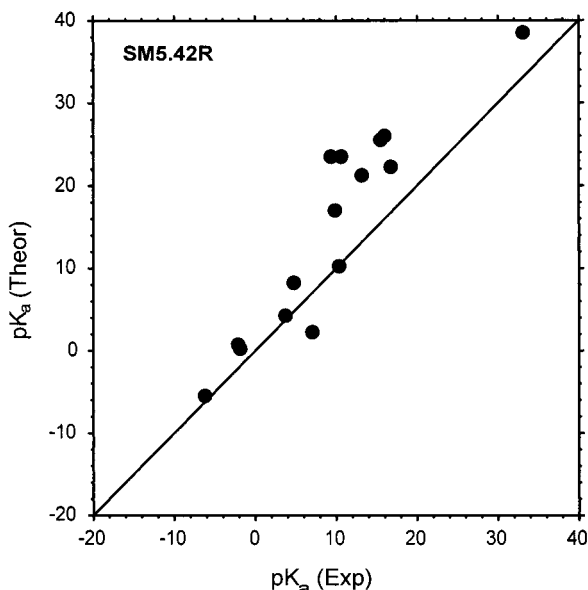
Figure 2. Correlation between experimental pK_a values and calculated values obtained from the SM5.42R method.

of 31.8. By comparison, the continuum SM5.42R and PCM models display high deviations from experimental data in opposite directions, with pK_a's of 38.6 and 28.3, respectively. Again, the stability of our hybrid model becomes very noticeable.

The fact that the cluster−continuum method outperforms pure continuum models can be attributed to the explicit introduction of strong and specific solute−solvent interactions in the first solvation shell around the ions. This results in a relatively small contribution of the dielectric continuum to the variation of the solvation free energy ($\Delta\Delta G_{\text{solv}}$) in the cluster−continuum model when compared with the much larger effect in the SM5.42R and PCM methods, as shown in Tables 1 and 2. Whereas the SM5.42R and PCM methods display $\Delta\Delta G_{\text{solv}}$'s as high 45 and 28 kcal mol^{-1}, respectively, the $\Delta\Delta G_{\text{solv}}$ contribution in the cluster−continuum model does not exceed 10 kcal mol^{-1}. As a consequence, the differential solvation in the hybrid model is mainly introduced by the explicit solvent molecules resulting in more accurate solution-phase reaction free energies.

Figure 3. Correlation between experimental pK_a values and calculated values obtained from the PCM method.

The failure of the cluster−continuum model in predicting an accurate pK_a for $CH_3NH_3^+$ can be traced to the continuum part of the calculation, specifically, the IPCM method. For cations, the isodensity surface used in this method yields a smaller cavity than required and overestimates the dielectric continuum contribution to the solvation free energy. This could be corrected by using fixed atomic cavities such as those used in the PCM method. Thus, application of this hybrid method in conjunction with the PCM rather than the IPCM would probably lead to more reliable differences in solvation free energies. Likewise, higher-level ab initio calculations (MP4 or CCSD(T) level) should help in improving the gas-phase energetics calculation and yield a more accurate approach to the calculation of pK_a's in aqueous solution.

It is important to emphasize that the advantage of our hybrid approach is that it does not require the extensive parametrization that is used in many of the pure continuum models. This becomes unnecessary in the cluster−continuum method because the solute−near-solvent molecule interactions are explicitly considered. Furthermore, the performance of our method has been tested for several functional groups, and the model has been shown to be very stable and to yield very creditable pK_a values. Thus, we believe that this method can be extended to a wide variety of systems and should be able to predict reliable pK_a's. We also find that pure continuum models are not reliable when applied to different types of systems, and even the highly parametrized PCM-UAHF[26] is probably accurate only for systems containing the functional groups used in the parametrization.

Conclusions

The cluster−continuum model and the pure continuum SM5.42R and PCM methods were applied to the pK_a calculations of several organic species having different functional groups. Our calculations show that the hybrid model is superior to the pure continuum methods, having an rms error of 2.2 pK_a units as opposed to 7 pK_a units for the SM5.42R and PCM methods. These results show that inclusion of the explicit solvent molecules leads to considerable improvement over pure continuum models, a fact that can be attributed to the introduction of specific solute−solvent interactions.

References and Notes

(1) Albert, A.; Serjeant, E. P. *The Determination of Ionization Constants*; Chapman and Hall: New York, 1984.

(2) Jorgensen, W. L.; Briggs, J. M.; Gao, J. *J. Am. Chem. Soc.* **1987**, *109*, 6857.

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

(4) Richard, J. P.; Williams, G.; Gao, J. *J. Am. Chem. Soc.* **1999**, *121*, 715.

(5) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610.

(6) Tunon, I.; Silla, E.; Tomasi, J. *J. Phys. Chem.* **1992**, *96*, 9043.

(7) Tunon, I.; Silla, E.; Pascual-Ahuir, J. *J. Am. Chem. Soc.* **1993**, *115*, 2226.

(8) Richardson, W. H.; Peng, C.; Bashford, D.; Noodleman, L.; Case, D. A. *Int. J. Quantum Chem.* **1997**, *61*, 207.

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

(10) Toth, A. M.; Liptak, M. D.; Phillips, D. L.; Shields, G. C. *J. Chem. Phys.* **2001**, *114*, 4595.

(11) Liptak, M. D.; Shields, G. C. *J. Am. Chem. Soc.* **2001**, *123*, 7314.

(12) Li, J.; Fisher, L.; Chen, J. L.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1996**, *35*, 4694.

(13) Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F. *Chem. Phys. Lett.* **1995**, *240*, 199.

(14) Gao, J.; Li, N.; Freindorf, M. *J. Am. Chem. Soc.* **1996**, *118*, 4912.

(15) Kallies, B.; Mitzner, R. *J. Phys. Chem. B* **1997**, *101*, 2959.

(16) Topol, I. A.; Tawa, G. J.; Burt, S. K.; Rashin, A. A. *J. Phys. Chem. A* **1997**, *101*, 10075.

(17) Shapley, W. A.; Bacskay, G. B.; Warr, G. G. *J. Phys. Chem. B* **1998**, *102*, 1938.

(18) da Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. *Int. J. Quantum Chem.* **1999**, *74*, 417.

(19) da Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. *J. Phys. Chem. A* **1999**, *103*, 11194.

(20) Wiberg, K. B.; Clifford, S.; Jorgensen, W. L.; Frisch, M. J. *J. Phys. Chem. A* **2000**, *104*, 7625.

(21) da Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. *J. Phys. Chem. A* **2000**, *104*, 2402.

(22) Liptak, M. D.; Shields, G. C. *Int. J. Quantum Chem.* **2001**, *85*, 727.

(23) Jang, Y. H.; Sowers, L. C.; Cagin, T.; Goddard, W. A., III. *J. Phys. Chem. A* **2001**, *105*, 274.

(24) Klicic, J. J.; Friesner, R. A.; Liu, S.; Guida, W. C. *J. Phys. Chem. A* **2002**, *106*, 1327.

(25) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2001**, *105*, 7241.

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

(27) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050.

(28) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395.

(29) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. B* **2000**, *104*, 5155.

(30) Pliego, J. R., Jr.; Riveros, J. M. *J. Chem. Phys.* **2000**, *112*, 4045.

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

(32) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1999**, *103*, 9.

(33) Pliego, J. R., Jr.; Riveros, J. M. *Chem.−Eur. J.* **2002**, *8*, 1945.

(34) Pitarch, J.; Ruiz-Lopez, M. F.; Silla, E.; Pascual-Ahuir, J. L.; Tunon, I. *J. Am. Chem. Soc.* **1998**, *120*, 2146.

(35) Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 792.

(36) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. *J. Phys. Chem.* **1996**, *100*, 16098.

(37) Hill, T. L. *An Introduction to Statistical Thermodynamics*; Addison-Wesley Pub. Co.: Reading, MA, 1960.

(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, revision D.2; Gaussian, Inc.: Pittsburgh, PA, 1995.

(39) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comput. Chem.* **1993**, *14*, 1347.

(40) Li, J.; Zhu, T.; Hawkins, G. D.; Chuang, Y.-Y.; Fast, P. L.; Liotard, D. A.; Rinaldi, D.; Cramer, C. J.; Truhlar, D. G. *GAMESOL*, version 2.2.4; University of Minnesota: Minneapolis, MN, 1999.

(41) NIST Chemistry WebBook. http://webbook.nist.gov; NIST Standard Reference Database Number 69; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, July 2001 (accessed).

(42) Pliego, J. R., Jr.; Riveros, J. M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622.