

See discussions, stats, and author profiles for this publication at:
<https://www.researchgate.net/publication/232362313>

Absolute pKa determination for carboxylic acids using density functional theory and the polarizable continuum model

ARTICLE *in* CHEMICAL PHYSICS LETTERS · MAY 2003

Impact Factor: 1.9 · DOI: 10.1016/S0009-2614(03)00607-9

CITATIONS

126

READS

44

3 AUTHORS:



Gloria A A Saracino

Azienda Ospedaliera Niguarda Ca' Gra...

16 PUBLICATIONS 330 CITATIONS

SEE PROFILE



Roberto Improta

Italian National Research Council

148 PUBLICATIONS 4,770 CITATIONS

SEE PROFILE



Vincenzo Barone

Scuola Normale Superiore di Pisa

774 PUBLICATIONS 44,771 CITATIONS

SEE PROFILE



Absolute pK_a determination for carboxylic acids using density functional theory and the polarizable continuum model

Gloria A.A. Saracino, Roberto Improta, Vincenzo Barone *

Dipartimento di Chimica, Università Federico II, Complesso Universitario Monte S. Angelo, Via Cintia, 80126 Napoli, Italy

Received 7 February 2003; in final form 27 March 2003

Abstract

The absolute pK_a s of eight carboxylic acids have been computed using a Born–Haber cycle, the most recent experimental value for the proton solvation energy and completely *ab initio* structures, energies, and harmonic frequencies of acids and conjugated bases both in vacuo and in aqueous solution. The proton affinities computed by the PBE0 hybrid functional are in very good agreement with experimental values and with the results of the most sophisticated (and expensive) quantum mechanical models (G2, G3, and CBS). The Gibbs energies in aqueous solution have been computed by our last version of the polarizable continuum model, which takes into the proper account escaped charge effects in a very effective computational implementation. The encouraging results obtained for this training set and the linear scaling both of the electronic and solvation models pave the route for the evaluation of reliable pK values for the large systems of interest in biological or material science frameworks.

© 2003 Published by Elsevier Science B.V.

1. Introduction

The central role played by proton transfer reactions in chemistry, biology, and material science [1,2] makes the evaluation of reliable pK values a topic of continuing interest. Despite the general reliability of the different experimental procedures used with this aim [3], the situation is more involved when the molecule, or more generally the acid–base function of interest is embedded in a complex system such as a protein, a membrane or a micelle and also when its pK_a is very close to that

of the surrounding medium [4,5]. This explains why in the last decades many attempts have been made to develop effective strategies for the *ab initio* computation of accurate relative [6,7] and absolute [8–23] pK_a values. A detailed review of the results of all those studies is obviously outside the scope of this Letter. Schematically, the used approaches differ in: (i) the thermodynamic cycle which the theoretical procedure is based on; (ii) the level of theory used for the gas phase calculation; and (iii) the solvation model. In particular, use of the experimental value of the proton solvation energies allows to avoid not completely solved ambiguities connected with the H_3O^+ species [16,19,22,23]. At the same time the dimensions of the cavity including the solute in continuum solvent

* Corresponding author. Fax: +39-081-674-090.
E-mail address: baronev@unina.it (V. Barone).

models are very critical. Unfortunately, it is well known that isodensity models [21], in spite of their conceptual elegance, are not able to provide reliable solvation energies for charged species in aqueous solution [24]. One possible solution is the explicit consideration of water molecules belonging to the first solvation shell [22]. However, recent works have shown that our UAHF model [24], including a limited number of parameters, provides remarkable results when using bare solutes and G2 [25], G3 [26], or CBS [27] model chemistries for gas-phase computations [18–20]. Here we would like to analyse if similar results could be obtained using cheaper methods rooted into the density functional theory (DFT). At the same time, the recent introduction of an improved version of the polarizable continuum model (PCM) allowing an improved treatment of anionic species and more effective geometry optimisations [28] calls for a reinvestigation of previous results including also the effects of geometry reoptimisation and frequency evaluation in aqueous solution, which were neglected in most of the previous studies. In this Letter we will deal with the pK_a 's of prototype carboxylic acids. However, the essential linear scaling both of DFT and last generation PCM would pave the route for the application of our procedure to large systems of chemical and biological interest.

2. Computational methodology

For an acid species AH the pK_a , defined as minus logarithm of the dissociation constant of the reaction



is given by the well-known thermodynamic relation:

$$pK_a = \Delta G_{aq,AH} / 2.303RT \quad (2)$$

The Gibbs energy variation of the deprotonation reaction in aqueous solution is calculated in our procedure by adding a solvation contribution to the gas phase value as follows:

$$\Delta G_{aq,AH} = \Delta G_{gas,AH} + \Delta \Delta G_{solv,AH} \quad (3)$$

$$\Delta G_{gas,AH} = G_{gas,A^-} + G_{gas,H^+} - G_{gas,AH} \quad (4)$$

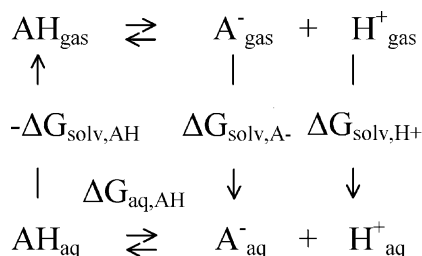
$$\Delta \Delta G_{solv,AH} = \Delta G_{solv,A^-} + \Delta G_{solv,H^+} - \Delta G_{solv,AH} \quad (5)$$

according to the thermodynamic cycle reported in Scheme 1.

The most recent experimental values have been taken for both G_{gas,H^+} (−6.28 kcal/mol [29]) and G_{solv,H^+} (−263.98 kcal/mol [30]). The latter value was obtained by Tissandier et al. using the cluster-pair-based-approximation without extra thermodynamic assumptions. This value is more negative than the lower end of the range of values based on measurements of the standard hydrogen electrode potentials [11], and is consistent with the theoretical one obtained from the experimental thermodynamic cycle of acetic acid [18].

For what concerns the calculation of the gas phase terms of cycle 1, our approach is based on the Density Functional Theory (DFT), which, because of its favourable scaling with the number of active electrons, opens the possibility to treat very large molecules. We have chosen, in particular, the hybrid PBE0 functional [31], which consistently provides satisfactory results for several structural and thermodynamic properties.

Solvation energies are computed by our most recent implementation of the Polarizable Continuum Model (PCM) [28], which takes into proper account escaped charge effects without quite arbitrary compensation procedures. Note that it is also necessary to add in the computations a term $\{-RT \ln (24.46)\}$ taking into account the transformation of the concentration units from atm in the gas phase to mol dm^{−3} in aqueous solution.



Scheme 1. Thermodynamic cycle.

The presence of anionic species makes the use of diffuse functions mandatory and for the sake of consistence we prefer to use the same basis set for all the computations. As a consequence we select the 6-31+G(d,p) basis set for geometry optimisations and evaluations of harmonic frequencies both in the gas phase and in aqueous solution.

All the calculations have been made by a development version of the Gaussian package [32].

3. Results and discussion

The procedure outlined above has been applied to the calculation of the pK_a for a training set of eight carboxylic acids: formic, acetic, fluoroacetic, chloroacetic, cyanoacetic, oxalic, propanoic and pivalic acid.

Starting our analysis from the gas phase contribution, several test computations showed that converged results can be obtained using the 6-311+G(2d,2p) basis set for single point energy evaluations at 6-31+G(d,p) geometries. Inspection of Table 1 shows that the ΔG_{gas} values obtained by this computational model are in very good agreement both with experimental values [18,33] and with the results of refined post-Hartree–Fock computational schemes, like the G2 [25], G3 [26], and CBS [27] models. These approaches provide thermodynamic data of chemical accuracy, but their scaling with the number of active electrons is prohibitive for the study of large molecules. Since

UAHF radii were optimised for HF computations, solvation Gibbs energies have been estimated by single point HF/6-31+G(d,p) calculations at geometries optimised in aqueous solution at the PBE0/6-31+G(d,p) level. Together with the details given above for gas phase computations, this completely defines our model, which will be referred to, in the following, as S03.

As a first step, we have compared the results obtained by the new PCM with those obtained by the CPCM version [34] used in several previous studies [10,18,20,22] using the standard UAHF radii.

As expected, both models provide comparable results for neutral molecules, whereas larger differences are obtained for anionic species, where escaped charge effects become more significant. Agreement between both sets of results can be restored using in the new PCM version $\gamma_{O^-} = 0.44$ (in place of the original value of 0.30) and this value has been used in all further computations. Note that γ_{O^-} rules the reduction of the oxygen radius as a function of its formal charge ($-1/2$ in carboxylates) and the use of a smaller cavity is coherent with an improved treatment of escaped charge effects. At the same time we refrained to perform other small adjustments of UAHF parameters related to more subtle differences between present computations (diffuse functions also for neutral species, geometry reoptimisation in solution) and those on which the parametrization was based in the attempt of fully testing the robustness of the proposed procedure.

Table 1

Comparison between gas phase basicities (kcal/mol) obtained by different computational methods and the corresponding experimental values

Acid	CBS	G2	G3	PBE0 ^a	Experiment ^b
HCOOH	338.1	337.9	339.4	338.4	338.5
CH ₃ COOH	341.6	340.5	341.4	341.7	341.4
FCH ₂ COOH	–	331.1	331.9	332.2	331.5
ClCH ₂ COOH	329.1	328.8	329.5	329.3	328.9
CNCH ₂ COOH	323.3	322.8	323.6	323.1	323.7
HOCCOOH	–	320.6	321.6	323.3	–
CH ₃ CH ₂ COOH	–	339.8	340.6	340.3	340.3
(CH ₃) ₃ CCOOH	337.3	338.6	339.3	339.5	337.8

^a 6-31+G(d,p)/6-311+G(2d,2p).

^b See [18,33].

A second point concerns the presence of both syn and anti conformers for each acid. For the gas phase computations there is no problem because, due to the formation of an intramolecular hydrogen bridge, syn conformers are much more stable than their anti counterparts (e.g., 4.0 and 5.2 kcal/mol for formic and acetic acid, respectively). The situation is more involved in aqueous solution because the energy difference between syn and anti conformers becomes so small (e.g., 0.4 and 0.8 kcal/mol for formic and acetic acid, respectively) to require, in principle, a Boltzmann averaging between the different forms. However, the pK_a difference is also quite small (0.29 and 0.60 for formic and acetic acid, respectively), so that the results of the complete statistical treatment never differ by more than 0.2 kcal/mol from the results obtained considering only the syn conformers. As a consequence we report in the tables only the results obtained for this latter kind of computations.

Table 2 shows that the pK_a 's calculated by the S03 procedure are close to their experimental counterparts [35], thus supporting the reliability of our procedure. In particular separate geometry optimisation in the gas phase and in aqueous solution not only ensures a more realistic and physically coherent description of the solvation phenomenon, but significantly improves the agreement between calculated and experimental pK_a 's (see Table 2). As it could be expected, geometry optimisation affects the solvation energy of charged species to a larger extent than that of the

neutral compounds. On the other hand the difference between harmonic frequencies computed in the gas phase and those in aqueous solution is generally negligible for the computation of non potential energy contributions to pK_a 's. This further improves the effectiveness of our computational model.

In summary, our results indicate that it is indeed possible to develop a computational procedure exhibiting all the features sketched in the introduction. The use of the PBE0/DFT model gives indeed results comparable to those obtained by the most refined post-Hartree–Fock models and, at the same time, it enables the application of the S03 procedure to very large systems. Calculating the free energy of each species always on geometries optimised (both in the gas phase and in aqueous solution) at the same level of theory further improves the physical soundness and the internal coherence of our procedure. Finally, the use of realistic and carefully tailored cavities allows to obtain accurate pK values without including explicit solvent molecules in the calculations [22]. On the one hand, it is true that charged species are likely to be strongly solvated in solution: thus explicit inclusion of water molecules belonging to the first solvation shell could improve the physical soundness and the accuracy of the computational procedure. On the other hand, a fully continuum approach does not require tedious geometry re-optimisations of the adducts between solute and a variable number of solvent molecules and allows to avoid the cumbersome and intrinsically dy-

Table 2

Solvation energies (kcal/mol) of acid/base pairs and corresponding pK_a obtained using geometries optimised in the gas phase or in aqueous solution are compared with experimental pK_a values (from [35])

Acids AH	Gas phase geometries			Aqueous solution geometries			pK_a^{exp}
	$\Delta G_{\text{solv,AH}}$	$\Delta G_{\text{solv,A}^-}$	pK_a	$\Delta G_{\text{solv,AH}}$	$\Delta G_{\text{solv,A}^-}$	pK_a	
HCOOH	-8.68	-76.35	4.93	-8.63	-77.07	4.37	3.77
CH ₃ COOH	-8.17	-78.36	5.29	-8.14	-79.23	4.62	4.76
FCH ₂ COOH	-11.26	-74.55	3.60	-11.24	-75.59	2.83	2.59
ClCH ₂ COOH	-10.89	-70.94	3.80	-10.87	-72.06	2.96	2.85
CNCH ₂ COOH	-15.00	-70.04	2.85	-14.9	-71.66	1.61	2.45
HOOC ⁺ COOH	-16.07	-72.55	1.81 ^a	-16.13	-74.12	0.70 ^a	1.23
CH ₃ CH ₂ COOH	-8.14	-76.85	5.57	-7.98	-77.92	4.67	4.87
(CH ₃) ₃ CCOOH	-6.85	-72.75	7.00	-6.77	-74.53	5.64	5.03

^a Including the correction for the presence of two equivalent ionisable groups.

dynamic problem of defining the exact number of solvent molecules that have to be included in the calculations. A last comment is in order concerning the choice of the thermodynamic cycle to be used in the computation of pK 's. It has been recently shown [23] that results based on the experimental values for the bare proton can be reconciled with those based on the H_3O^+/H_2O pair provided that a value of -110.2 kcal/mol is used for the solvation energy of H_3O^+ in place of the conventional value of -104.0 . In the present framework, this can be obtained simply modifying the γ_{O^+} value from -0.26 to -0.28 in the UAHF model, without changing none of the solvation energies of Table 2. As a consequence, the selection of a proper thermodynamic cycle becomes a simple matter of preference, since very close pK values are obtained for any reasonable choice.

References

- [1] C. Hansch, P.G. Sammes, J.B. Taylor (Eds.), *Comprehensive Medicinal Chemistry. Quantitative Drug Design*, vol. 4, Pergamon Press, Oxford, UK, 1990.
- [2] G. Shüürmann, in: G. Shüürmann, B. Markert (Eds.), *Ecotoxicology*, Wiley, New York, 1998, p. 665.
- [3] R.F. Cookson, *Chem. Rev.* 74 (1974) 5.
- [4] W. Reeve, C.M. Erikson, P.F. Aluotto, *Can. J. Chem.* 57 (1979) 2747.
- [5] F.G. Bordwell, *Acc. Chem. Res.* 21 (1988) 456.
- [6] W.L. Jorgensen, J.M. Briggs, *J. Am. Chem. Soc.* 111 (1989) 4190.
- [7] A.M. Toth, M.D. Liptak, D.L. Phillips, G.C. Shields, *J. Chem. Phys.* 114 (2001) 4595.
- [8] G. Shüürmann, *Quant. Struct. Act. Relat.* 15 (1996) 121.
- [9] G. Shüürmann, in: F. Chen, G. Shüürmann (Eds.), *Quantitative Structure–Activity Relationship in Environmental Sciences*, vol. VII, SETAC Press, Pensacola, FL, 1997.
- [10] G. Shüürmann, M. Cossi, V. Barone, J. Tomasi, *J. Phys. Chem. A* 102 (1998) 6706.
- [11] C. Lim, D. Bashford, M. Karplus, *J. Phys. Chem.* 95 (1991) 5610.
- [12] W.H. Richardson, C. Peng, D. Bashford, L. Noodleman, D.A. Case, *Int. J. Quantum Chem.* 61 (1997) 207.
- [13] J. Adzelm, C. Kölmel, A. Klamt, *J. Chem. Phys.* 103 (1995) 9312.
- [14] C.J. Cramer, D. Truhlar, *Science* 256 (1992) 213.
- [15] G. Shüürmann, *Chem. Phys. Lett.* 302 (1999) 471.
- [16] C.O. da Silva, E.C. da Silva, M.A.C. Nascimento, *J. Phys. Chem. A* 103 (1999) 11194.
- [17] C.O. da Silva, E.C. da Silva, M.A.C. Nascimento, *J. Phys. Chem. A* 104 (2000) 2402.
- [18] M.D. Liptak, G.C. Shields, *J. Am. Chem. Soc.* 123 (2001) 7314.
- [19] M.D. Liptak, G.C. Shields, *Int. J. Quantum Chem.* 85 (2001) 727.
- [20] M.D. Liptak, K.C. Gross, P.G. Seybold, S. Feldgus, G.C. Shields, *J. Am. Chem. Soc.* 124 (2002) 6421.
- [21] D.M. Chipman, *J. Phys. Chem. A* 106 (2002) 7413.
- [22] J.R. Pliego, J.M. Riveros, *J. Phys. Chem. A* 106 (2002) 7434.
- [23] J.R. Piego Jr., *Chem. Phys. Lett.* 367 (2003) 145.
- [24] V. Barone, M. Cossi, J. Tomasi, *J. Chem. Phys.* 107 (1997) 3210.
- [25] L.A. Curtiss, K. Raghavachari, G.W. Trucks, J.A. Pople, *J. Chem. Phys.* 94 (1991) 7221.
- [26] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople, *J. Chem. Phys.* 109 (1998) 7764.
- [27] J.W. Ochtersky, G.A. Petersson, J.A. Montgomery, *J. Chem. Phys.* 104 (1996) 2598.
- [28] M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Chem. Phys.* 117 (2002) 43.
- [29] I.A. Topol, G.J. Tawa, S.K. Burt, A.A. Rashin, *J. Chem. Phys.* 111 (1999) 10998.
- [30] M.D. Tissandier, K.A. Cowen, W.Y. Feng, E. Gundluach, M.H. Cohen, A.D. Earhart, J.V. Coe, T.R. Tuttle, *J. Phys. Chem. A* 102 (1998) 7787.
- [31] C. Adamo, V. Barone, *J. Chem. Phys.* 110 (1999) 6158.
- [32] M.J. Frisch, et al. G03, Gaussian Inc., Pittsburg PA, 2003.
- [33] J.E. Bartness, in: W.G. Mallard, P.J. Linstrom, Eds., *NIST Chemistry WebBook, NIST Standard Reference Database No. 69*, National Institute of Standards and Technology, Gaithersburg, MD, February 2000. Available from <<http://webbook.nist.gov>>.
- [34] V. Barone, M. Cossi, *J. Phys. Chem. A* 102 (1998) 1995.
- [35] E.P. Sargeant, B. Dempsey, *Ionisation Constants of Organic Acids in Aqueous Solution*; IUPAC No. 23, Pergamon Press, New York, 1979.