Theor Chem Acc (2013) 132:1310
DOI 10.1007/s00214-012-1310-z

REGULAR ARTICLE

Isodesmic reaction for pK_a calculations of common organic molecules

Sebastián Sastre · Rodrigo Casasnovas · Francisco Muñoz · Juan Frau

Received: 18 September 2012 / Accepted: 28 November 2012 / Published online: 13 December 2012

Abstract Three quantum chemistry methods (B3LYP, M05-2X and CBS-4B3*) have been used, in combination with SMD and CPCM continuum solvent models, to calculate the aqueous pK_a values of common organic compounds (aliphatic alcohols, carboxylic acids, amines, phenols, benzoic acids and pyridines) by using an isodesmic reaction. Good precision is found for all the studied functional groups, resulting mean absolute deviations of 0.5–1 pK_a units (equivalent to the best results obtained with thermodynamic cycles). It is worthy to note that no explicit water molecules were needed with the isodesmic reaction. In addition, the quality of the results is not strongly dependent on the combination of quantum chemistry method, solvent model and reference species. Therefore, the isodesmic reaction could be successfully used when dealing with gas-phase unstable species, with species that undergo large conformational changes between gas-phase and solution-phase or other difficult cases for the thermodynamic cycles.

Keywords pK_a calculation · Isodesmic reaction · Continuum solvent model

Published as part of the special collection of articles derived from the 8th Congress on Electronic Structure: Principles and Applications (ESPA 2012).

Electronic supplementary material The online version of this article (doi:10.1007/s00214-012-1310-z) contains supplementary material, which is available to authorized users.

S. Sastre · R. Casasnovas · F. Muñoz · J. Frau (✉)
Institut d'Investigació en Ciències de la Salut (IUNICS), Departament de Química, Universitat de les Illes Balears, 07122 Palma de Mallorca, Spain
e-mail: juan.frau@uib.es

1 Introduction

Great efforts have been devoted during the last years to develop computational strategies for the accurate prediction of pK_a values [1].

The traditional approach in theoretical pK_a calculations is based on the use of thermodynamic cycles that combine gas-phase deprotonation free energies (ΔG_{gas}) and the solvation energies (ΔG_{solv}) of the involved species. The use of such cycles is due to the fact that it is not rigorously correct to obtain the free energy of deprotonation in solution (ΔG_{soln}) by calculating the free energies of the involved species according to the rigid rotor-harmonic approximation in the continuum solvent, as discussed previously in detail [2, 3]. In fact, the rigorous calculation would require much more costly simulation methods in which the solvent is explicitly considered, together with a free energy calculation procedure.

The use of thermodynamic cycles shows problems for those species that are gas-phase unstable or undergoes large conformational changes between gas-phase and solution-phase. However, it is possible to calculate the approximate free energies of deprotonation in solution by using the isodesmic reaction, since gas-phase energies are not required for the pK_a calculation (Scheme 1). Additionally, this reaction scheme, also known as relative pK_a calculation or proton exchange reaction, benefits from the

$$AH^{q}_{(soln)} + B^{m-1}_{(soln)} \xrightarrow{\Delta G_{(soln)}} A^{q-1}_{(soln)} + BH^{m}_{(soln)}$$

Scheme 1 Isodesmic reaction employed for the pKa calculation of AH^q. ΔG_{soln} stands for the free energy in solution of the acid–base reaction. The global charge of the acid species and its conjugate base are represented by q and $q - 1$, respectively. The charges of the reference species and its conjugate base are represented by m and $m - 1$, respectively

absence of the proton free energy of solvation, which is a potential source of error in the calculated free energies (2).

The pK_a of the acid AH can be calculated from Eqs. 1–3 and the experimental pK_a value of the reference species BH

$$\Delta G_{soln} = G_{soln}(BH^m) + G_{soln}(A^{q-1}) - G_{soln}(AH^q) - G_{soln}(B^{m-1}) \quad (1)$$

$$G_{soln} = E_{soln} + G_{nes} + \Delta G_{corr\_soln} \quad (2)$$

$$pK_a(AH^q) = \frac{\Delta G_{soln}}{2,303RT} + pK_a(BH^m) \quad (3)$$

As previously noted, it is not possible to calculate the exact free energies in solution because the partition functions of the involved species (AH^q, BH^m, A^{q-1} and B^{m-1}) are unknown. Our group [3] recently proposed to calculate approximate free energies in solution under the assumption that gas-phase vibrational partition functions of the solutes as an approximation for the solution-phase and, secondly, that rotational and translational contributions to the free energies corresponding to conjugate acid and base species are of similar magnitude so, according to Eq. 1, they mostly cancel out. According to Eq. 2, the approximate free energies of each species in solution (G_{soln}) are calculated as the sum of the potential energy of the solute (E_{soln}) which includes the electric response of the continuum solvent, all the non-electrostatic contributions to the solute–solvent interaction (namely dispersion, repulsion and cavitation energies) which are all included in the term G_{nes}, and the contribution to the free energies from the vibrational motion of the nuclei at 298 K (ΔG_{corr_soln}).

It should be noted that both solvation energies of neutral species and gas-phase free energies (of neutral and charged species) can be determined with an accuracy of 1 kcal/mol [1], however, solvation energies of charged species calculated with continuum solvent models show, at best, average errors of 4 kcal/mol [4, 5]. According to the isodesmic reaction scheme, such errors should not be present in the pK_a calculation as solvation energies are not required. In any case, a higher accuracy in the pK_a calculation is expected for the isodesmic reaction since a good cancellation of errors is expected if both the reference species (BH^m) and the acid species (AH^q) present the same electric charge and similar structure.

So far the isodesmic reaction has mainly been used for the calculation of enthalpies of formation [6] but, concerning pK_a calculations, it has been less used than the thermodynamic cycles. Li et al. [7] reported that relative instead of absolute pK_a values are predicted with higher precision when combining ab initio methods and continuum solvent models for the study of methylimidazoles in aqueous solution. A recent combined experimental and theoretical study performed by Ruiz-López et al. [8] also shows that remarkably low errors are obtained when using the isodesmic reaction for diprotic species. Govender and Cukrowski [9, 10] have employed this scheme, together with PCM solvent model and UA0 cavities for the calculation of the successive dissociation constants of nitrilopropanoic and nitrilotriacetic acids, obtaining errors of 0.2–3 pK_a units.

Our group [3] has recently employed the isodesmic reaction (Scheme 1) to calculate the pK_a of substituted pyridines and carbon acids. The results obtained for pyridines were equivalent to those of thermodynamic cycles [11] but no explicit water molecules were needed to increase the accuracy [3]. Concerning carbon acids, the mean absolute deviation of 2 pK_a units was also equivalent to the best results obtained with thermodynamic cycles [12, 13] with the advantage that the electric charge of the reference species was not determinant for a good precision [3]. We also reported the pK_a values of amino acids with very low errors (i.e., 0.22 and 0.19 pK_a units for the dissociation of carboxylic and amino groups, respectively) which are one of the paradigmatic difficult cases for thermodynamic cycles as the major tautomers in solution-phase and gas-phase are different [3].

The main objective in this work is to evaluate the robustness of the isodesmic reaction in the calculation of pK_a values of common acid–base organic functionalities and to compare such results with those calculated with thermodynamic cycles previously reported in the literature. Two different continuum solvent models were employed (namely CPCM [14–16] and SMD [17, 18]) in combination with three quantum chemistry methods, two common DFT functionals (B3LYP [19] and M05-2X [20, 21]) and the composite method CBS-4B3* [3, 22].

2 Methodology

The pK_a values of aliphatic alcohols, carboxylic acids, aliphatic amines, benzoic acids, phenols and pyridines calculated with the isodesmic reaction (Scheme 1) according to Eqs. 1–3 are displayed in Table 1.

The criterion followed to choose the reference species BH was the similitude of chemical structure with the studied species AH so; for alcohols, carboxylic acids and amines, the reference species are, respectively, ethanol, acetic acid and ethylamonium ion, while phenol, benzoic acid and pyridinium ion were the references for phenols, benzoic acids and pyridines, respectively.

Three quantum chemistry methods, specifically B3LYP [19], M05-2X [20, 21] and CBS-4B3* [3, 22], have been used to calculate the free energies of Eq. 2. In all cases, the geometries were optimized and characterized as energy minima by the absence of imaginary frequencies according

Table 1 Pyridines, aliphatic alcohols, carboxylic acids, amines, phenols and benzoic acids studied in this work

Pyridines	Alcohols	Carboxylic acids	Amines	Phenols	Benzoic acids
2-Methylpyridine	Methanol	Chloroacetic	Methylamine	*p*-Cyanophenol	*o*-Chlorobenzoic
3-Methylpyridine	2-Chloroethanol	Formic	Propylamine	*o*-Chlorophenol	*m*-Chlorobenzoic
4-Methylpyridine	Propanol	3-Chlorobutanoic	*i*-Propylamine	*m*-Cyanophenol	*p*-Chlorobenzoic
2,3-Dimethylpyridine	*i*-Propanol	Benzoic	Butylamine	*m*-Chlorophenol	*p*-Methylbenzoic
2,4-Dimethylpyridine	2-Butanol	4-Chlorobutanoic	2-Butylamine	*m*-Fluorophenol	*m*-Methylbenzoic
3-Bromopyridine	tert-butanol	Hexanoic	*tert*-Butylamine	*p*-Chlorophenol	*p*-Fluorobenzoic
3-Fluoropyridine		Propanoic	Trimethylamine	*p*-Fluorophenol	
3-Cyanopyridine		Pentanoic	Dimethylamine	*m*-Methylphenol	
3-Chloropyridine		Trimethylacetic		*p*-Methylphenol	
				o-Methylphenol	

to the 6-311++G(d,p) basis set. The CBS-4B3* composite method is a simplification of the CBS-QB3 method in which the CCSD(T) calculations and the energetic corrections of spin contamination and empirical corrections are suppressed [22]. This method benefits of lower computational costs and provides as accurate deprotonation free energies [22] and pK_a values [3] as the original CBS-QB3.

The solvent effects have been taken into account in all geometry optimizations and also energy calculations by using the CPCM model [13–15] with the UAKS cavities [23] as implemented in Gaussian 03 [23] and the SMD model [17, 18] as implemented in Gaussian 09 [24].

2.1 Results and discussion

Table 2 shows the mean absolute deviations (MAD) between the experimental and calculated pK_a values of each functional group according to the combination of solvent model (CPCM or SMD) and the quantum chemistry method (B3LYP, M05-2X or CBS-4B3*).

Several important points can be extracted after analyzing the values of Table 2. First of all, it should be noted that no combination of quantum chemistry method and continuum solvent model stand out by providing much lower errors than the other combinations. In fact, the highest difference between MAD values is 0.67 pK_a units for the CPCM/M05-2X and SMD/M05-2X calculations of aliphatic alcohols.

Concerning all the organic functionalities, the MAD values are lower than 1 pKa units in all cases with the exception of the pK_a values predicted with SMD for the aliphatic alcohols, which show MAD values lower than 1.3 pK_a units. The best results correspond to the primary amines and benzoic acids with MAD values lower than 0.35 (Table T3 of Supporting Information) and 0.5 (Table 2) pK_a units, respectively, for the SMD calculations.

Table 2 Mean absolute deviations (MAD), standard deviation (SD) and maximum absolute deviation (AD max) of aqueous pK_a values calculated using different methods (CBS-4B3*/6-311++G(d,p), B3LYP/6-311++G(d,p) and M05-2X/6-311++G(d,p)), and solvent models (CPCM and SMD)

	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Pyridines						
MAD	0.57	0.75	0.83	0.62	0.80	0.78
SD	0.48	0.74	0.55	0.36	0.39	0.44
AD max	1.44	2.17	1.87	1.29	1.34	1.55
Alcohols						
MAD	0.85	0.68	0.62	1.20	1.01	1.29
SD	0.49	0.46	0.51	0.96	0.75	0.97
AD max	1.54	1.19	1.22	2.84	2.27	2.85
Carboxylic acids						
MAD	0.78	0.98	0.79	0.57	0.64	0.67
SD	0.42	0.36	0.47	0.36	0.54	0.59
AD max	1.32	1.44	1.51	1.13	1.54	1.89
Amines						
MAD	0.87	0.72	0.90	0.24	0.35	0.27
SD	0.84	0.74	1.00	0.19	0.25	0.20
AD max	2.51	2.20	2.89	0.63	0.83	0.71
Phenols						
MAD	0.90	1.08	1.02	0.70	0.89	0.87
SD	0.57	0.68	0.68	0.48	0.57	0.56
AD max	2.16	2.71	2.57	1.66	2.18	2.10
Benzoic acids						
MAD	0.45	0.57	0.41	0.41	0.50	0.34
SD	0.29	0.41	0.30	0.35	0.48	0.29
AD max	0.95	1.36	0.82	1.05	1.39	0.69

Considering the MAD values obtained with both CPCM and SMD solvent models, the best results correspond to the SMD model for carboxylic acids, amines, phenols and benzoic acids independently of the quantum chemistry method, whereas CPCM provides lower MAD values for the aliphatic alcohols. Finally, the results of both solvent models for pyridines and primary amines (Table T3 of Supporting Information) show no significant differences.

Regarding the quantum chemistry method, the best results are provided by the CBS-4B3* method for 7 out of 12 combinations, specifically for pyridines, carboxylic acids and phenols with CPCM (MAD values of 0.57, 0.78 and 0.90 pK_a units, respectively) and pyridines, carboxylic acids, amines and phenols with SMD (MAD values of 0.62, 0.57, 0.24 and 0.70 pK_a units, respectively). The M05-2X gives the best results for the alcohols and benzoic acids with CPCM and also benzoic acids with SMD, while the B3LYP functional gives the best results for the amines with CPCM and alcohols with SMD (Table 2).

As noted, the pK_a values obtained with the isodesmic reaction exhibit low errors for all the functionalities, independently of the combination of method and solvent model so, we considered important to compare our results with previously reported theoretical pK_a values calculated with thermodynamic cycles.

Theoretical pK_a values of aliphatic alcohols have been reported in several studies [25–28] among which stand out the one of Pliego and Riveros [25], which reports the pK_a values of methanol and ethanol with MAD values of 0.5 pK_a units with the IPCM model and including 2–3 explicit water molecules. Namazian and Heidary [26] obtained the pK_a of methanol, ethanol, propanol and isopropanol with maximum errors lower than 0.5 pK_a units by using the CPCM model to calculate the solvation energies. In close agreement with such studies, MAD values of 0.62 pK_a units are obtained for the CPCM/M05-2X calculations when using the isodesmic reaction (Table 2).

Carboxylic acids have been also extensively studied in theoretical pK_a calculations [28–38]. Namazian and Halvani [32] calculated the pK_a of 66 acids with the PCM model and B3LYP functional and obtained average errors of 0.5 pK_a units. In another study, Toth et al. [31] calculated the pK_a values of five carboxylic acids at the CPCM/HF/6-31+G(d)//HF/6-31G(d) level with average errors of 0.5 pK_a units. Such average errors are similar to those obtained in this study with the SMD solvent model (Table 2).

The pK_a values of amines were calculated with good accuracy [39–43]. For example, average errors of 0.46 pK_a units calculated with the PCM model at the B3LYP/6-31+G*//MP2/6-31+G* level including an explicit water molecule were reported in the work of Behjatmanesh-Ardakani et al. [42]. In the present work, an accuracy of $\sim$0.3 pK_a units is found for the CPCM model when considering just the primary amines (Table T3 of Supporting Information), although higher errors are obtained when including secondary and tertiary amines (Table 2). Nevertheless, the SMD model provides, in all cases, MAD errors lower to 0.3 pK_a units (Table 2).

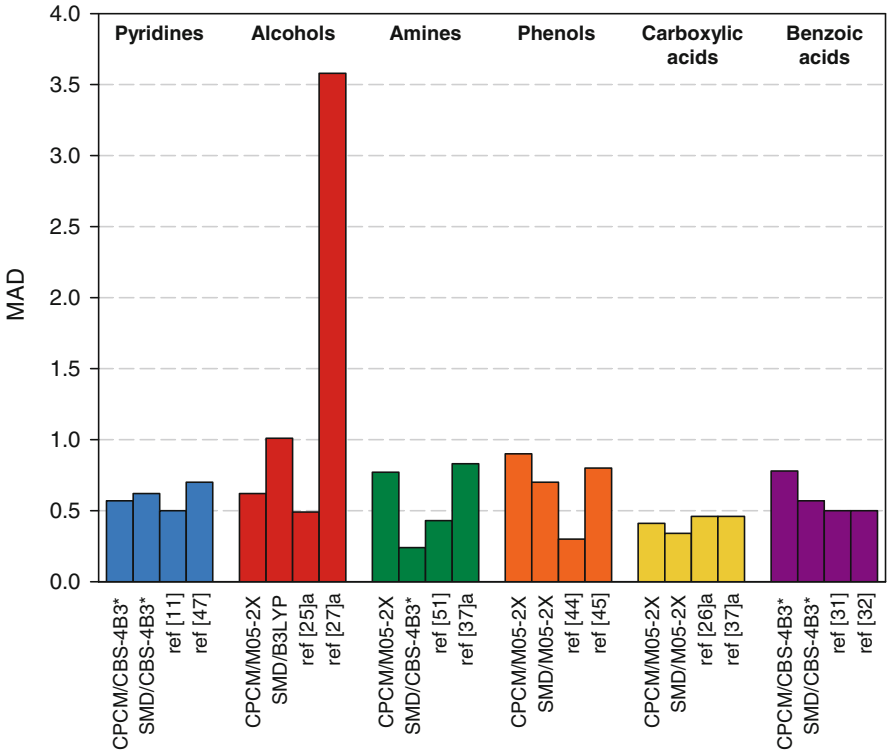
Scmidt am Busch and Knapp [37] calculated the pK_a of benzoic, *p*-methylbenzoic, *m*-methylbenzoic and *p*-chlorobenzoic acids with MAD values of 0.5 pK_a units. As seen in Table 2, the errors obtained from the isodesmic reaction are even lower, particularly those of SMD/M05-2X calculations (i.e., 0.34 pK_a units).

The theoretical prediction of pK_a values of substituted phenols has also been reported in different studies [28, 37, 44, 45]. Liptak et al. [44] reported RMSD errors between 0.4 and 3.9 pK_a units depending on the calculated solvation energies. As can be seen in Table 2, in our case, the MAD values are slightly over 0.5 pK_a units and under 1 pK_a unit for most of the method and solvent model combinations.

Concerning pyridines, mean deviations lower than 1 pK_a unit were reported [11, 46–50]. In a previous work of our group [11], we evaluated the accuracy of different thermodynamic cycles, the importance of explicit water molecules and the influence of using gas-phase or solution-phase optimized geometries when calculating the solvation energies with the CPCM model. In the best cases, mean errors of 0.5 pK_a units were reported when using a single explicit water molecule [11]. The results provided by the isodesmic reaction (Table 2) show the same precision than the best obtained with thermodynamic cycles (i.e., MAD values between 0.57 and 0.83 pK_a units) without requiring explicit solvent molecules.

The lowest MAD values obtained with each solvent model, together with other MAD values of previous works that used thermodynamic cycles are depicted in Fig. 1. As can be seen, the accuracy of the isodesmic reaction predicted pK_a values is of the same order of those predicted with thermodynamic cycles for pyridines, benzoic acids and carboxylic acids for both CPCM and SMD models. Concerning amines, the SMD/CBS-4B3* calculations outperform the errors of thermodynamic cycles. Besides, in the worst-case scenario, the difference between MAD values of isodesmic reaction and thermodynamic cycles is lower than 0.6 pK_a units.

We have chosen the carboxylic acids and amines to evaluate the influence of the reference species on the calculated pK_a values. Tables 3 and 4 show the MAD values obtained when using different references for carboxylic acids and amines, respectively. The resulting MAD values for carboxylic acids vary between 0.5 and 1.0 pK_a units unless chloroacetic and formic acids are used as reference species. Regarding the amines, the MAD values fluctuate between 0.28 and 0.78 pK_a units regardless of the reference is a primary, secondary or tertiary amine.



[a] MADs calculated with the species belonging to the functional group of interest.

Fig. 1 Comparison between the best mean absolute deviations (MAD) obtained with CPCM and SMD solvent models and MAD values reported in previous studies that used thermodynamic cycles

Rebollar-Zepeda and Galano [51] had recently reported a comprehensive assessment of reaction schemes and density functionals for the pK_a calculation of amines. Their evaluation of the isodesmic reaction points out that the precision of the calculated pK_a values is strongly dependent on the reference species [51]. It is well known that the acidity of amines is significantly influenced by the solvent accessibility to the amino group. Since continuum solvent models do not reproduce the solvent structure around the solute, it is not possible to model the steric effects on the solvation of the amino group caused by its substituents. Consequently, not only the electrostatic characteristics have to be taken into account for the selection of the reference species but also the steric behavior of the substituents should be similar to the studied amine species. However, it is not trivial to model such effects by the introduction of explicit water molecules in the calculation because the dynamic contribution of the substituents to the solvation structure would require statistical treatment by monte carlo or molecular dynamics simulations.

Table 3 Effect of the reference species on the accuracy of the calculated pK_a values of carboxylic acids expressed as mean absolute deviations (MAD)

Reference	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Acetic ac.	0.78	0.98	0.79	0.57	0.64	0.67
Chloroacetic ac.	1.45	1.72	1.50	0.84	1.42	1.70
Formic ac.	1.11	1.51	1.13	1.02	1.45	1.17
3-Chlorobutanoic ac.	0.74	1.25	0.80	0.57	0.66	0.79
Benzoic ac.	0.86	0.89	0.86	0.78	0.64	0.88
4-Chlorobutanoic ac.	0.74	0.89	0.79	0.58	0.64	0.67
Hexanoic ac.	0.88	1.09	0.98	0.69	0.82	0.72
Propanoic ac.	0.84	1.00	0.89	0.91	0.96	0.82
Pentanoic ac.	0.88	1.06	0.92	0.66	0.73	0.74
Trimethylacetic ac.	1.07	1.25	1.31	0.87	1.01	0.81

To get a deeper insight into the influence of the reference species, the pK_a of some functional groups has been calculated by using a structurally different species (Table 5). Specifically, the pK_a of benzoic acids, phenols and pyridines were calculated by considering acetic acid, ethanol and ethylamine, respectively, as reference species. In the first case, MAD values of 0.5 pK_a units result when using acetic acid which is similar to the precision obtained if the reference species were the benzoic acid. However, the errors in the pK_a values of phenols are significantly larger (i.e., MAD values between 4.0 and 6.5 pK_a units) when using ethanol as reference species. Such different behavior for benzoic acids and phenols can be attributed to

Table 4 Effect of the reference species on the accuracy of the calculated pK_a values of amines expressed as mean absolute deviations (MAD)

Reference	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Ethylamine	0.87	0.72	0.90	0.24	0.35	0.27
Methylamine	0.61	0.93	0.80	0.61	0.93	0.80
Propylamine	0.35	0.32	0.17	0.35	0.32	0.17
i-Propylamine	0.23	0.34	0.38	0.23	0.34	0.38
Butylamine	0.28	0.35	0.48	0.28	0.35	0.48
2-Butylamine	0.21	0.35	0.22	0.21	0.35	0.22
tert-Butylamine	0.31	0.63	0.30	0.31	0.63	0.30
Trimethylamine	0.27	0.41	0.27	0.27	0.41	0.27
Dimethylamine	0.36	0.58	0.33	0.36	0.58	0.33

the delocalization of the negative charge to the π-system in the phenoxyde species, which causes significant differences in the electrostatic interactions of the deprotonated oxygens of phenoxydes and aliphatic alcoxydes. However, delocalization of the negative charge in benzoate anions is low so the charge distribution of carboxylate groups in benzoates is similar to that of aliphatic carboxylates. When using ethylamine as reference, the MAD values in the pK_a of pyridines increase up to 2.8 pK_a units for the CPCM calculations but remain closer to the MAD values when using pyridine as reference for the SMD calculations (Table 5). Therefore, it is worthy to note that the chemical structure should not be the only criterion to consider when choosing the reference species and that the cancellation of errors due to similarities in the solute–continuum interactions is not always trivial.

Table 5 Effect of the reference species on the accuracy of the calculated pK_a values of aromatic molecules expressed as mean absolute deviations (MAD)

Reference	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Benzoic acids						
Benzoic acid	0.45	0.57	0.41	0.41	0.50	0.34
Acetic acid	0.51	0.47	0.48	1.16	0.65	1.16
Phenols						
Phenol	0.90	1.08	1.02	0.70	0.89	0.87
Ethanol	4.01	4.31	5.01	5.21	5.59	6.46
Pyridines						
Pyridine	0.57	0.75	0.83	0.62	0.80	0.78
Ethylamine	1.86	2.77	1.92	1.22	0.82	1.42

As a suggestion of one of the referees, we considered the inclusion of the rotational and translational contributions, calculated from gas-phase partition functions, to the free energies in solution. Such contributions had a minor effect of 0.01 to 0.1 pK_a units for all the studied functional groups. Only for two species, methanol and methylamine, these energies had higher and significant contribution to the pK_a values (i.e., 0.54 and 0.43 pK_a units, respectively). However, it has to be noted that this result only stresses the fact that rotational and translational free energies of the conjugate acid–base pairs calculated with gas-phase partition functions are very similar and, therefore, largely cancel out in Eq. 1. In addition, this result does not entail the actual rotational and translational free energies of the conjugate acid–base species to be necessarily similar.

In summary, it has been shown that the isodesmic reaction scheme provides reliable results in the pK_a calculation of common organic functionalities, with pK_a values comprised between 1 and 19. The accuracy of the isodesmic reaction predicted pK_a values with combinations of common quantum chemistry methods and continuum solvent models is similar to that provided by thermodynamic cycles. Besides, no explicit water molecules are required to obtain a good accuracy. Although the reference species influences the precision of the calculated pK_a values, the cancellation of errors in the isodesmic reaction allows more flexible criteria in the choice of the reference species. So, taken everything into account, it is worth to consider this procedure for theoretical pK_a predictions, especially for those cases where the thermodynamic cycles show problems related to gas-phase calculations.

Acknowledgments This work was funded by the Spanish Government in the framework of Project CTQ2008-02207/BQU. One of us (R.C.) wishes to acknowledge a fellowship from the Spanish MEC within the FPU program. The authors are grateful to "Centro de Cálculo de Superomputación de Galicia" (CESGA) and to "Centre de Supercomputació de Catalunya" (CESCA) for access to their computational facilities.

References

1. Ho J, Coote ML (2010) A universal approach for continuum solvent pK_a calculations: are we there yet? Theor Chem Acc 125:3–21
2. Ho J, Klamt A, Coote ML (2010) Comment on the correct use of continuum solvent models. J Phys Chem A 114:13442–13444
3. Casasnovas R, Fernández D, Ortega-Castro J, Frau J, Donoso J, Muñoz F (2011) Avoiding gas-phase calculations in theoretical pK_a predictions. Theor Chem Acc 130:1–13
4. Takano Y, Houk KN (2005) Benchmarking the conductor-like polarizable continuum model (CPCM) for aqueous solvation free energies of neutral and ionic organic molecules. J Chem Theory Comput 1:70–77

5. Cramer CJ, Truhlar DG (2008) A universal approach to solvation modeling. Acc Chem Res 41:760–768
6. Wang L, Heard DE, Pilling MJ, Seakins P (2008) A Gaussian-3X prediction on the enthalpies of formation of chlorinated phenols and Dibenzo-*p*-dioxins. J Phys Chem A 112:1832–1840
7. Li GS, Ruiz-López MF, Maigret B (1997) Ab initio study of 4(5)-Methylimidazole in aqueous solution. J Phys Chem A 101:7885–7892
8. Derbel N, Clarot I, Mourer M, Regnouf-de-Vains J, Ruiz-López MF (2012) Intramolecular interactions versus hydration effects on *p*-Guanidinoethyl-phenol structure and pK_a values. J Phys Chem A 116:9404–9411
9. Govender KK, Cukrowski I (2009) Density functional theory in prediction of four stepwise protonation constants for nitrilotripropanoic acid (NTPA). J Phys Chem A 113:3639–3647
10. Govender KK, Cukrowski I (2010) Density functional theory and isodesmic reaction based prediction of four stepwise protonation constants, as log $K_H^{(n)}$, for nitrilotriacetic acid. the importance of a kind and protonated form of a reference molecule used. J Phys Chem A 114:1868–1878
11. Casasnovas R, Frau J, Ortega-Castro J, Salvà A, Donoso J, Muñoz F (2009) Absolute and relative pK_a calculations of mono and diprotic pyridines by quantum methods. J Mol Struct Theochem 912:5–12
12. Gao DQ, Svoronos P, Wong PK, Maddalena D, Hwang J, Walker H (2005) pK_a of acetate in water: a computational study. J Phys Chem A 109:10776–10785
13. Ho J, Coote ML (2009) pK_a calculation of some biologically important carbon acids—an assessment of contemporary theoretical procedures. J Chem Theory Comput 5:295–306
14. Barone V, Cossi M (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. J Phys Chem A 102:1995–2001
15. Barone V, Cossi M, Tomassi J (1998) Geometry optimization of molecular structures in solution by the polarizable continuum model. J Comput Chem 19:404–417
16. Cossi M, Rega M, Scalmani G, Barone V (2003) Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. J Comput Chem 24:669–681
17. Marenich AV, Cramer CJ, Truhlar DG (2009) Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. J Phys Chem B 113:6378–6396
18. Marenich AV, Cramer CJ, Truhlar DG (2009) Performance of SM6, SM8, and SMD on the SAMPL1 test set for the prediction of small-molecule solvation free energies. J Phys Chem B 113:4538–4543
19. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648–5652
20. Zhao Y, Schultz NE, Truhlar DG (2006) Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. J Chem Theory Comput 2:364–382
21. Zhao Y, Truhlar DG (2008) Functionals with broad applicability in chemistry. Acc Chem Res 41:157–167
22. Casasnovas R, Frau J, Ortega-Castro J, Salvà A, Donoso J, Muñoz F (2010) Simplification of the CBS-QB3 method for predicting gas-phase deprotonation free energies. Int J Quantum Chem 110:323–330
23. Gaussian 03, Revision E.01, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick D K, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox D J, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill P MW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA, Gaussian, Inc., Wallingford CT, (2004)
24. Gaussian 09, Revision B.01, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery Jr JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ, Gaussian, Inc., Wallingford CT (2009)
25. Pliego JR, Riveros JM (2002) Theoretical calculation of pK_a using the cluster-continuum model. J Phys Chem A 106:7434–7439
26. Namazian M, Heidary H (2003) Ab initio calculations of pK_a values of some organic acids in aqueous solution. J Mol Struct-Theochem 620:257–263
27. Silva CO, da Silva EC, Nascimento MAC (2000) Ab initio calculations of absolute pK_a values in aqueous solution II. Aliphatic alcohols, thiols, and halogenated carboxylic acids. J Phys Chem A 104:2402–2409
28. Kelly CP, Cramer CJ, Truhlar DG (2006) Adding explicit solvent molecules to continuum solvent calculations for the calculation of aqueous acid dissociation constants. J Phys Chem A 110:2493–2499
29. Liptak MD, Shields GC (2001) Accurate pK_a calculations for carboxylic acids using complete basis set and Gaussian-n models combined with CPCM continuum solvation methods. J Am Chem Soc 123:7314–7319
30. Liptak MD, Shields GC (2001) Experimentation with different thermodynamic cycles used for pK_a calculations on carboxylic acids using complete basis set and gaussian-n models combined with CPCM continuum solvation methods. Int J Quantum Chem 85:727–741
31. Toth AM, Liptak MD, Phillips DL, Shields GC (2001) Accurate relative pK_a calculations for carboxylic acids using complete basis set and Gaussian-n models combined with continuum solvation methods. J Chem Phys 114:4595–4606
32. Namazian M, Halvani S (2006) Calculations of pK_a values of carboxylic acids in aqueous solution using density functional theory. J Chem Thermodyn 38:1495–1502
33. Namazian M, Halvani S, Noorbala MR (2004) Density functional theory response to the calculations of pK_a values of some carboxylic acids in aqueous solution. J Mol Struct Theochem 711:13–18
34. Schuurmann G, Cossi M, Barone V, Tomasi J (1998) Prediction of the pK_a of carboxylic acids using the ab initio continuum-solvation model PCM-UAHF. J Phys Chem A 102:6706–6712
35. da Silva CO, da Silva EC, Nascimento MAC (1999) Ab initio calculations of absolute pK_a values in aqueous solution I. Carboxylic acids. J Phys Chem A 103:11194–11199

36. Saracino GAA, Improta R, Barone V (2003) Absolute pK_a determination for carboxylic acids using density functional theory and the polarizable continuum model. Chem Phys Lett 373: 411–415
37. Schmidt am Busch M, Knapp EW (2004) Accurate pK_a determination for a heterogeneous group of organic molecules. Chem Phys Chem 5:1513–1522
38. Jia Z, Du D, Zhou Z, Zhang A, Hou R (2007) Accurate pK_a determinations for some organic acids using an extended cluster method. Chem Phys Lett 439:374–380
39. da Silva EF, Svendsen HF (2003) Prediction of the pK_a values of amines using ab initio methods and free-energy perturbations. Ind Eng Chem Res 42:4414–4421
40. Eckert F, Klamt A (2006) Accurate prediction of basicity in aqueous solution with COSMO-RS. J Comput Chem 27:11–19
41. Blanco SE, Ferretti FH (2005) Calculation of ionization constants of methylamines in aqueous solution. J Mol Struc Theochem 722: 197–202
42. Behjatmanesh-Ardakani R, Karimi MA, Ebady A (2009) Cavity shape effect on pK_a prediction of small amines. J Mol Struc Theochem 910:99–103
43. Khalili F, Henni A, East ALL (2009) Entropy contributions in pK_a computation: application to alkanolamines and piperazines. J Mol Struc Theochem 916:1–9
44. Liptak MD, Gross KC, Seybold PG, Feldgus S, Shields GC (2002) Absolute pK_a determinations for substituted phenols. J Am Chem Soc 124:6421–6427
45. Rebollar-Zepeda AM, Campos-Hernández T, Ramirez MT, Rojas-Hernández A, Galano A (2011) Searching for computational strategies to accurately predict pK_as of large phenolic derivatives. J Chem Theory Comput 7:2528–2538
46. Caballero NA, Melendez FJ, Muñoz-Caro C, Niño A (2006) Theoretical prediction of relative and absolute pK_a values of aminopyridines. Biophys Chem 124:155–160
47. Chen IJ, MacKerell AD (2000) Computation of the influence of chemical substitution on the pK_a of pyridine using semiempirical and ab initio methods. Theo Chem Acc 103:483–494
48. Guven A (2005) Acidity study on 3-substituted pyridines. Int J Mol Sci 6:257–275
49. Blanco SE, Ferretti FH (2005) Calculation of acidity constants of pyridines in aqueous solution. Chem Phys Lett 403:400–404
50. Ögretir C, Tay NF, Özturk II (2007) A theoretical study on protonation of some halogen substituted pyridine derivatives. J Mol Graph Model 26:740–747
51. Rebollar-Zepeda AM, Galano A (2012) First principles calculations of pk_a values of amines in aqueous solution: application to neurotransmitters. Int J Quantum Chem 112:3449–3460