Biochimica et Biophysica Acta 1827 (2013) 882–891

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbabio

Toward the accurate calculation of pK_a values in water and acetonitrile☆

James T. Muckerman [a,*], Jonathan H. Skone [a], Ming Ning [a], Yuko Wasada-Tsutsui [b]

[a] Chemistry Department, Brookhaven National Laboratory, Upton, NY, 11973-5000, USA
[b] Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

ARTICLE INFO

Article history:
Received 25 October 2012
Received in revised form 25 February 2013
Accepted 26 March 2013
Available online 6 April 2013

Keywords:
pK_a value
Solvation model
Acidity

ABSTRACT

We present a simple approach for the calculation of accurate pK_a values in water and acetonitrile based on the straightforward calculation of the gas-phase absolute free energies of the acid and conjugate base with use of only a continuum solvation model to obtain the corresponding solution-phase free energies. Most of the error in such an approach arises from inaccurate differential solvation free energies of the acid and conjugate base which is removed in our approach using a correction based on the realization that the gas-phase acidities have only a small systematic error relative to the dominant systematic error in the differential solvation. The methodology is outlined in the context of the calculation of a set of neutral acids with water as the solvent for a reasonably accurate electronic structure level of theory (DFT), basis set, and implicit solvation model. It is then applied to the comparison of results for three different hybrid density functionals to illustrate the insensitivity to the functional. Finally, the approach is applied to the comparison of results for sets of neutral acids and protonated amine cationic acids in both aqueous (water) and nonaqueous (acetonitrile) solvents. The methodology is shown to generally predict the pK_a values for all the cases investigated to within 1 pH unit so long as the differential solvation error is larger than the systematic error in the gas-phase acidity calculations. Such an approach is rather general and does not have additional complications that would arise in a cluster-continuum method, thus giving it strength as a simple high-throughput means to calculate absolute pK_a values. This article is part of a Special Issue entitled: Metals in Bioenergetics and Biomimetics Systems.

1. Introduction

The theoretical prediction of accurate pK_a values is of considerable importance in a variety of subfields of chemistry and biology, yet it is in general an elusive goal [1–17]. The acid dissociation reaction,

$$HA_{(s)} \rightarrow H^+_{(s)} + A^-_{(s)} \quad (1a)$$

where the subscript "(s)" indicates a solvated species, is deceptively simple, especially considering that the product $H^+_{(s)}$ is common to all acids in a given solvent, and has a known (constant) free energy in water and several other solvents. The pK_a value is given by

$$pK_a = \Delta G^*_a / RT\ln(10) \quad (1b)$$

where

$$\Delta G^*_a = G^*\left(A^-_{(s)}\right) + G^*\left(H^+_{(s)}\right) - G^*\left(HA_{(s)}\right) \quad (1c)$$

and G represents an absolute Gibbs free energy in the quantum chemical sense, and the superscript "*" represents a standard state of one mole per liter and 298.15 K in solution or gas phase. Closely associated with the solution phase acid dissociation reaction is the gas-phase acidity

$$HA_{(g)} \rightarrow H^+_{(g)} + A^-_{(g)} \quad (2a)$$

and corresponding free energy change

$$\Delta G^o_{a,(g)} = G^o\left(A^-_{(g)}\right) + G^o\left(H^+_{(g)}\right) - G^o\left(HA_{(g)}\right) \quad (2b)$$

where the superscript "o" represents a standard state of a gas at one atmosphere of pressure and 298.15 K.

The direct calculation of pK_a values via Eqs. (1b) and (1c) is generally avoided because the absolute free energy of the solvated proton is difficult to calculate accurately. On the other hand, the absolute free

☆ This article is part of a Special Issue entitled: Metals in Bioenergetics and Biomimetics Systems.

* Corresponding author. Tel.: +1 631 344 4368; fax: +1 631 344 4358.
E-mail address: muckerma@bnl.gov (J.T. Muckerman).

0005-2728/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbabio.2013.03.011

energy of the gas-phase proton at standard temperature and pressure is easily calculated using the Sackur–Tetrode equation [18], and has the value −6.28 kcal/mol, rendering it straightforward to calculate gas-phase acidities. For purposes of the present discussion, it is convenient to define a solution-phase and gas-phase "acid half reaction", analogous to electrochemical half reactions, as

$$HA_{(s)} \rightarrow A^-_{(s)} \quad (3a)$$

and

$$HA_{(g)} \rightarrow A^-_{(g)} \quad (3b)$$

which results in expressions for either phase in terms of the computationally calculated quantities only, independent of the absolute value of the proton. These acid half reactions have several useful features: (1) they incorporate all the quantum chemical input into the calculated pK_a values that we actually compute; (2) the reactant and product are isoelectronic, and usually have quite similar geometric and electronic structures; (3) they contain all the species that are unique to a given acid; and (4) they can be used in isodesmic reaction schemes [16] in exactly the same way as the corresponding acid dissociation reactions, i.e.,

$$HA_{(s)} + B^-_{(s)} \rightarrow HB_{(s)} + A^-_{(s)} \quad (4a)$$

and

$$\Delta G^* = \Delta G^*_a(HA) - \Delta G^*_a(HB) = \Delta G^*_{AHR,(s)}(HA) - \Delta G^*_{AHR,(s)}(HB) \quad (4b)$$

where the subscript "AHR" refers to an acid half reaction. Equations analogous to Eqs. (4a) and (4b) can also be written for the gas-phase acidities and acid half reactions, $\Delta G^o_{AHR,(g)}$. We will make use of the acid half reaction (AHR) later when evaluating the systematic error contribution from the differential solvation. The acid half reaction in solution is defined as

$$\Delta G^*_{AHR,(s)}(HA) = \Delta G^*_{AHR,(g)}(HA) + \Delta\Delta G^*_{solv}(HA) \quad (5a)$$

where

$$\Delta\Delta G^*_{solv}(HA) = \Delta G^*_{solv}(A^-) - \Delta G^*_{solv}(HA) \quad (5b)$$

and

$$\Delta G^*_{solv} = \Delta G^o_{solv} + \Delta G^{o\rightarrow *}. \quad (5c)$$

The term $\Delta G^{o\rightarrow *}$ is the free energy associated with the change of standard state from one atmosphere of pressure to one mole per liter in the gas phase.

2. Preliminary considerations

We begin our analysis of computing accurate pK_a values by the straightforward calculation of the absolute free energies of the species in Eq. (3a) for the homologous series of ten neutral ("training set") acids in water listed in the top section of Table 1 [19–26]. As detailed below, the geometry of each species was optimized (with energy $E_{(g)}$) and a vibrational frequency calculation was then carried out (yielding $G^o_{(g)}$) using the B3LYP hybrid density functional [27–30] in the gas phase. A single-point energy calculation (with energy $E_{(g)} + \Delta G^o_{solv}$) was then performed using the Conductor-like Polarizable Continuum solvation Model (CPCM) [31–33] with default parameters (e.g., universal force field, UFF, radii [34]) as implemented in Gaussian 09 [35] to obtain the solvation free energy, ΔG^o_{solv}. We then use these free energies ($G^o_{(g)} + \Delta G^o_{solv}$), along with the literature value (−272.2 kcal/mol) of the absolute free energy of the proton in aqueous solution [36–39], to compute the pK_a using Eqs. (1c) and (1b). The standard state correction of $\Delta G^{o\rightarrow *} = 1.894$ kcal/mol [39] does not need to be applied for the AHR because the correction would cancel in Eq. (1c) for all the cases considered in this work. Fig. 1A shows the computed pK_a values (red points) as a function of the corresponding experimental ones.

Table 1
Comparison of final $pK_{a,lfit}$ values for the neutral acid training and test sets in water using UFF, UAHF and Bondi radii with the B3LYP functional.[a]

Acid	$pK_{a,expt}$	UFF $pK_{a,lfit}$	UAHF $pK_{a,lfit}$	Bondi $pK_{a,lfit}$
Sulfuric	−3.00	−4.81	−5.05	−4.54
Benzenesulfonic	−2.80	−2.16	−1.85	−2.11
Methanesulfonic	−1.90	−1.32	−1.25	−1.09
Trifluoroacetic	0.23	−0.71	−0.69	−0.58
Formic	3.77	4.12	4.06	3.95
Benzoic	4.21	5.18	4.16	5.02
Acetic	4.79	5.48	5.05	5.17
Phenol	9.95	9.24	10.79	9.56
Ethanol	15.90	16.53	16.08	16.36
Isopropanol	17.10	16.64	15.95	16.50
	RMS error	0.85	0.97	0.76
Nitric	−1.40	−2.34	−2.21	−2.53
Oxalic	1.38	1.53	2.79	1.96
Phosphoric	2.15	1.50	2.11	2.23
Citric	3.09	2.30	4.02	2.67
Peracetic	8.20	8.75	9.62	8.05
Propargyl alcohol	13.55	13.73	11.73	13.57
	RMS error	0.56	1.21	0.55

[a] The linear fit to training set $pK_{a,lfit}$ vs. $pK_{a,expt}$ for all three solvation models has slope 1.00 and intercept 0.00.

Several features of this straightforward calculation of the pK_a values are obvious in Fig. 1A. The first is that the agreement between calculated and experimental values is rather poor (Table S1, Supporting Information); the root-mean-square (RMS) error is 6.85 pH units, four calculated values have error greater than 5 pH units, and only one calculated value has an error of less than 2 pH units. The second is that the slope of the least squares linear fit to the calculated vs. experimental values is 1.92, indicating that the range of the calculated values is much larger than that of the experimental values, and that (because of the non-unit slope) an isodesmic scheme employing only an implicit solvation model would not work well. Third, and most interestingly, the linear least squares fit shows a very strong linear correlation (correlation coefficient 0.99) with very little scatter of the calculated points from the fitted line. This indicates that the error in the computed values is systematic, and strongly dependent on the experimental pK_a values. There are only two possible sources for this systematic error in the *slope* of the best fit line: a systematic error in the gas-phase acid half reaction arising from the electronic structure method and/or basis, or a systematic error in the $\Delta\Delta G^*_{solv}$ values arising from the solvation model (Eq. (5a)). Because of the aforementioned isoelectronic property and usually quite similar geometric and electronic structures of the two species involved in each gas-phase acid half reaction, one would expect any systematic error in their calculation to cancel to a large extent, resulting in more or less random error as evidenced by the displacement of the red points from the red line in Fig. 1A [2,40–53]. On the other hand, the neutral acids in Table 1 have no net charge while their conjugate bases have a net charge of −1. This suggests that the $\Delta\Delta G^*_{solv}$ values of these acids will be dominated by the solvation free energy of the anion, and that any systematic error in the solvation model will not tend to cancel. We can test this hypothesis by comparing how the solvation model *should* behave as a function of experimental pK_a, and how it actually *does* behave. We first determine the extent of the error in the calculated gas-phase acid half reaction as shown in Fig. S1 (Supporting Information) and establish that it contributes little systematic error relative to the systematic error in the pK_a calculation. This allows us then to estimate how the solvation model should behave by using Eqs. (1b), (1c) and (5a) to solve for $\Delta\Delta G^*_{solv}$ for each of the ten training set acids

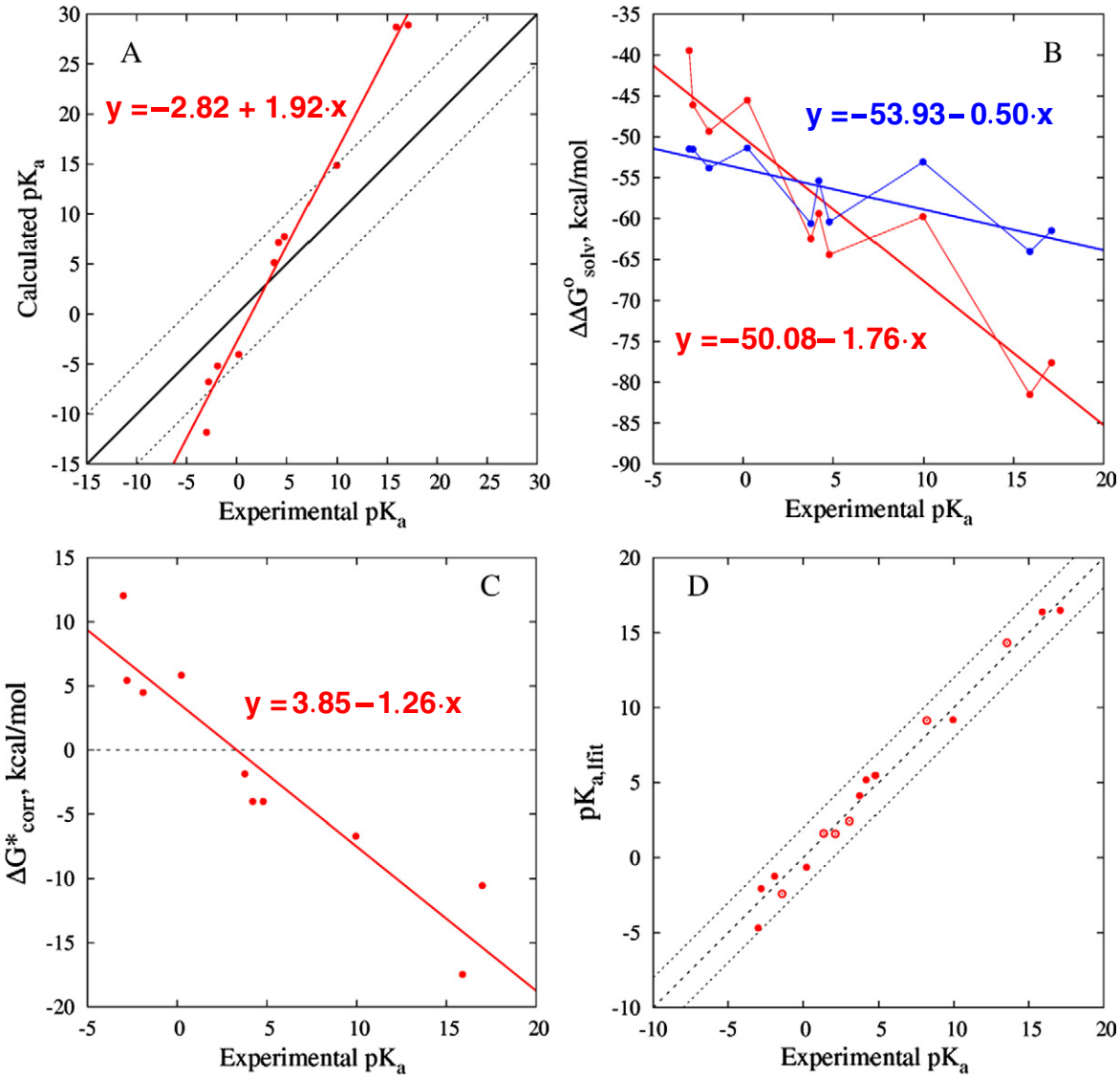

Fig. 1. Calculated results for neutral training and test set acids in water: (A) calculated pK_a values as a function of experimental pK_a values (red points) for the ten training set neutral acids in aqueous solution listed in Table 1, the red line is the calculated linear least squares fit $y = -2.82 + 1.92 \cdot x$, and the black line indicates the sought after perfect agreement, with the dashed lines indicating errors of ±5 pH units; (B) comparison of the desired behavior of the solvation model (red points connected by light red lines) and the actual behavior of the CPCM method with UFF radii (blue points connected by light blue lines) as a function of experimental pK_a values for acids in panel (A), the heavy red line is the calculated linear least squares fit $y = -50.08 - 1.76 \cdot x$ to the red points, and the heavy blue line is the calculated least squares fit $y = -53.93 - 0.50 \cdot x$ to the blue points; (C) correction free energy term required to remove the error introduced by the CPCM solvation model with UFF radii (red points), and the heavy red line is the calculated linear least squares fit $y = 3.85 - 1.26 \cdot x$ to the red points. (D) Comparison of $pK_{a,lfit}$ values with $pK_{a,expt}$ for the training set (red points) and the test set (red open circles with central points) acids in Table 1, the black line indicates the sought after perfect agreement, and the dashed lines indicate errors of ±2 pH units, the RMS error of the training set is 0.87 pH units, and that of the test set is 0.61 pH units.

listed in the top part of Table 1 given the experimental pK_a value and the literature value of $G^*(H^+_{(s)})$. The result is shown by the red points connected by thin red lines in Fig. 1B. The corresponding $\Delta\Delta G^*_{solv}$ values resulting from single-point solvation calculations for each training set acid and conjugate base using the CPCM method with UFF radii are shown by the blue points connected by thin blue lines in Fig. 1B. The heavy red and blue lines are the linear least-squares fits to the red and blue points, respectively.

It is clear from Fig. 1B that the solvation model under-solvates the acid half reactions (gives a more positive $\Delta\Delta G^*_{solv}$ than derived from the experimental pK_a values) for weak acids (strong conjugate bases) and over-solvates the acid half reactions for strong acids (weak conjugate bases), and this is responsible for the systematic error in the calculated pK_a values. We can therefore define a correction term, $\Delta G^*_{corr}(HA)$, for each of the acids that, on the average (since the systematic error in the calculated gas-phase acidities is small relative to the error in the pK_a in solution), corrects the calculated pK_a ($pK_{a,calc}$) for the deficiency of the solvent model as a function of experimental pK_a ($pK_{a,expt}$),

$$\Delta G^*_{corr}(HA) \equiv RT\ln(10) \cdot \left(pK_{a,expt} - pK_{a,calc}\right) \tag{6}$$

as shown by the red points in Fig. 1C. The solid red line indicates the linear least-squares fit to this correction free energy, which is equivalent to the heavy blue line minus the heavy red line in Fig. 1B. The red line in Fig. 1C indicates the average solvation model error as a linear function of experimental pK_a,

$$\Delta G^*_{corr,lfit} = a_0 + a_1 \cdot pK_{a,expt} \tag{7a}$$

with correlation coefficient $r = -0.97$, where $a_0 = 3.85$, $a_1 = -1.26$, and the subscript "lfit" stands for "linear fit". An expression for $\Delta G^*_{corr,lfit}$ as a function of something that can be calculated with reasonable accuracy would provide a straightforward way to apply the correction to acids of unknown pK_a. From Fig. 1C we can see that on average we can approximate $pK_{a,expt}$ on the right-hand side of Eq. (7a) by

$$pK_{a,expt} \approx pK_{a,calc} + \Delta G^*_{corr,lfit}/RT\ln(10) \quad (7b)$$

so that

$$\Delta G^*_{corr,lfit} = a_0 + a_1 \cdot \left[pK_{a,calc} + \Delta G^*_{corr,lfit}/RT\ln(10)\right] \quad (7c)$$

which can be solved for $\Delta G^*_{corr,lfit}$ to yield

$$\Delta G^*_{corr,lfit} = c_0 + c_1 \cdot pK_{a,calc} \quad (8)$$

where $c_0 = a_0 \cdot RT\ln(10) / [RT\ln(10) - a_1]$ and $c_1 = a_1 \cdot RT\ln(10) / [RT\ln(10) - a_1]$. Defining $pK_{a,lfit} \equiv pK_{a,calc} + \Delta G^*_{corr,lfit}/RT\ln(10)$, we obtain

$$pK_{a,lfit} = c_0/RT\ln(10) + [1 + c_1/RT\ln(10)] \cdot pK_{a,calc}. \quad (9)$$

Substituting the numerical values of a_0 and a_1 obtained from Fig. 1C, we have for the neutral acid training set in water

$$pK_{a,lfit} = 1.47 + 0.52 \cdot pK_{a,calc} \quad (10)$$

which is exactly the same relation we would obtain by solving for x in terms of y in the least-squares linear fit in Fig. 1A, but that would not have provided the insight into the source of the error arising from the deficiency of the solvation model.

Fig. 1D shows the comparison between the $pK_{a,lfit}$ values of the acids listed in Table 1 (red filled points), which have served as a training set for parameterization of the fitted values, and the corresponding experimental values. Also shown in Fig. 1D are the $pK_{a,lfit}$ values of the six neutral acids listed in the bottom part of Table 1 (red open points), which serve as a test set. The RMS error of the training set is 0.85 pH units, with a maximum error of 2 pH units. The RMS error of the test set is 0.56 pH units with a maximum error of ~1 pH unit.

Before proceeding with a more systematic study of the effect of solvation models on the calculated pK_a values of different homologous classes of acids and solvent, we should further examine whether the systematic error arising from the electronic structure method employed in the calculation of gas-phase acidities for a given solvation method effectively cancels. Fig. 2A presents the linear least-squares fits to $pK_{a,calc}$ vs. $pK_{a,expt}$ for the neutral acid training set in water using three different hybrid functionals, B3LYP (as in Fig. 1A), M06 [54] and B3P86 [27,55], for the CPCM solvation model with UFF radii (Table 2). While the calculated pK_a values for a given acid vary with the DFT functional used (Table S2), the deviations appear to be random. This results in the three least-squares lines being quite similar; the B3LYP and M06 lines are almost identical, with the B3P86 line having nearly the same slope and a slightly higher intercept. Fig. 2B shows that the three subsequent sets of calculated $pK_{a,lfit}$ values vs. $pK_{a,expt}$ distributions analogous to Fig. 1D are of comparable accuracy, having RMS errors of 0.85, 0.91 and 0.85, respectively (Table 2). This behavior is consistent with accurate gas-phase acid half-reaction calculations.

Unless otherwise noted, the B3LYP hybrid density functional and 6-311+G(d,p) 5d basis set [56,57] were used throughout this work for both classes of acids, (neutral and cationic). The geometry of each acid and its conjugate base was optimized in the gas phase, from which the vibrational frequencies and unrelaxed single-point solvation energies were determined. The solvation energy was evaluated using a

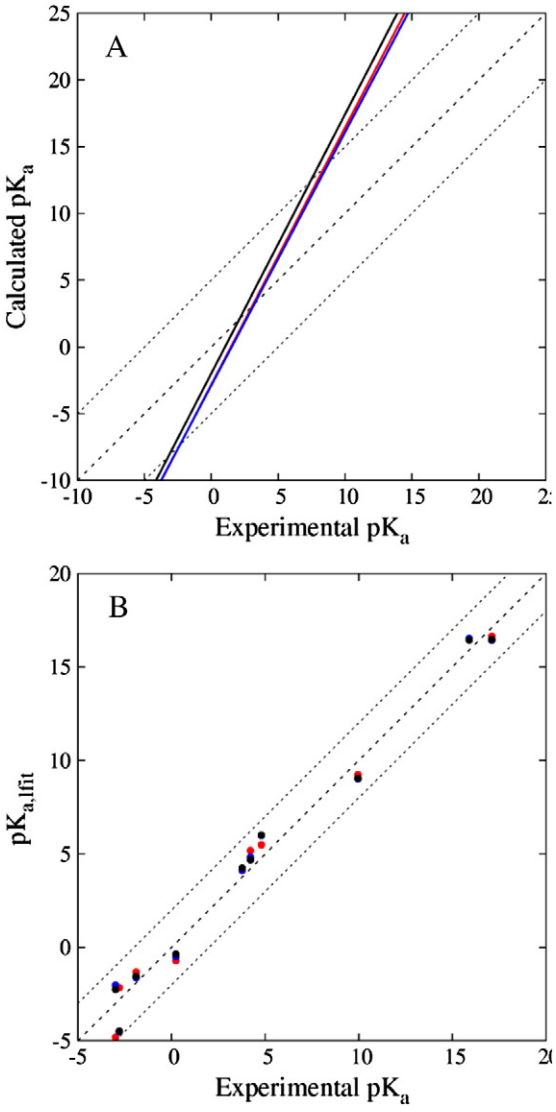

Fig. 2. Dependence of calculated and lfit p*K*a values on choice of density functional: (A) linear least-squares fits to three sets of $pK_{a,calc}$ vs. $pK_{a,expt}$ for the training set neutral acids in water with the CPCM solvation method using UFF radii, with the results for the B3LYP, M06 and B3P86 functionals shown by the red, blue and black lines, respectively; (B) the final $pK_{a,lfit}$ vs. $pK_{a,expt}$ distributions for the three functionals B3LYP (red points), M06 (blue points) and B3P86 (black points), where the heavy dashed line indicates the ideal behavior, and the lighter dashed lines indicate ±2 pH unit error.

Table 2
Comparison of final $pK_{a,lfit}$ values for the neutral acid training set in water using UFF radii with different hybrid density functionals.[a]

Acid	$pK_{a,expt}$	B3LYP $pK_{a,lfit}$	M06 $pK_{a,lfit}$	B3P86 $pK_{a,lfit}$
Sulfuric	−3.00	−4.81	−2.02	−2.25
Benzenesulfonic	−2.80	−2.16	−4.53	−4.48
Methanesulfonic	−1.90	−1.32	−1.62	−1.55
Trifluoroacetic	0.23	−0.71	−0.50	−0.36
Formic	3.77	4.12	4.13	4.25
Benzoic	4.21	5.18	4.83	4.69
Acetic	4.79	5.48	5.97	6.00
Phenol	9.95	9.24	9.01	9.04
Ethanol	15.90	16.53	16.53	16.44
Isopropanol	17.10	16.64	16.43	16.47
	RMS error	0.85	0.91	0.85

[a] The linear fit to $pK_{a,lfit}$ vs. $pK_{a,expt}$ for all three functionals has slope 1.00 and intercept 0.00.

conductor-like polarizable continuum model (CPCM) with three different choices of atomic radii: UFF, UAHF [58] and Bondi [59] (note: Bondi employed here used the non-default values of OFAC = 0.8 and RMIN = 0.5). Each of the three solvation methods was used with two homologous classes of acids, comprised of both a training set and a test set, in both water and acetonitrile solvents. The two homologous classes of acids were neutral acids (including alcohols, carboxylic acids, and sulfonic acids) and protonated amine cationic acids (including pyridines and anilines). The neutral acids and the protonated amine cationic acids are listed in Tables 1 and 4, and Tables 3 and 5, respectively. See Figs. S2 and S3 (Supporting Information) for the chemical structures of each acid class. For neutral acids, the solvation of the anionic conjugate base dominates the $\Delta\Delta G^*_{solv}$, whereas for protonated amine cationic acids, the solvation of the cationic acid dominates the differential solvation. We would therefore expect the neutral acids to have increasingly negative $\Delta\Delta G^*_{solv}$ values with increasing anion solvation (more negative ΔG^*_{solv}, Fig. 1B), and the cationic acids to have increasingly positive $\Delta\Delta G^*_{solv}$ with increasing cation solvation.

3. Results and discussion

3.1. Neutral acids in water

We begin by extending the calculations of neutral acids in water discussed above for the case of UFF radii to the other two solvation models being considered. The scatter plots of $pK_{a,calc}$ vs. $pK_{a,expt}$ for all three solvation models are shown in Fig. 3A along with the three linear least-squares correlations. While the two new linear fits for UAHF and Bondi radii resemble that for UFF radii shown in Fig. 1A, there is a larger quantitative difference among them than in the case of the differences arising from employing different density functionals (Fig. 2A, Table 2 and Table S2). Nevertheless, the scatter in each set of points from the corresponding linear correlation is small (± less than 1 pH unit, Table 1 and Table S1). The behavior of $\Delta\Delta G^*_{solv}$ (Fig. 3B) is, however, very different for the three solvation models. The "experimental" $\Delta\Delta G^*_{solv}$ (i.e., derived from the experimental pK_a values and the experimental value of $G^*(H^+_{(s)})$, and depending on accurate calculated ΔG^o_{AHR} values) exhibits a much steeper negative slope (−1.76 kcal/mol per pH) than any of the CPCM calculations with different radii. The UAHF radii yield the line (slope −1.13) that most closely approximates the "experimental" one, with the Bondi line (slope −0.71) next best, and the UFF line (slope −0.50) worst. All three solvent models overestimate the solvation of the strong acids (weak conjugate bases) and underestimate the solvation of weak acids (strong conjugate bases). Interestingly, this difference in the performance of the solvation models doesn't appear to make much difference in the calculated $pK_{a,lfit}$ values (shown in Fig. 3C and listed in Table 1) either for the training set or the test set acids. Only one training set $pK_{a,lfit}$ has an error of slightly over 2 pH units, and that corresponds to the best (UAHF) uncorrected solvation model, which also has the largest RMS error in both training set and test set.

Table 4
Comparison of final $pK_{a,lfit}$ values for the neutral acid training and test sets in acetonitrile using UFF, UAHF and Bondi radii with the B3LYP functional.[a]

Acid	$pK_{a,expt}$	UFF $pK_{a,lfit}$	UAHF $pK_{a,lfit}$	Bondi $pK_{a,lfit}$
Perchloric	1.57	−0.50	−0.41	−0.26
Nitrobenzenesulfonic	6.40	7.42	8.11	7.55
Chlorobenzenesulfonic	7.20	8.67	9.23	8.55
Nitric	8.80	9.57	9.90	8.86
Methanesulfonic	9.97	11.43	11.53	11.46
Trichloroacetic	10.75	10.91	9.47	11.35
Trifluoroacetic	12.65	10.64	10.02	10.61
Chloroacetic	18.80	17.23	17.75	17.56
Benzoic	21.51	22.14	21.09	21.55
Phenol	29.14	29.30	30.11	29.56
	RMS error	1.31	1.59	1.22
Trifluoromethanesulfonic	2.60	1.97	3.01	2.62
Sulfuric	7.20	5.01	5.43	5.05
Fumaric	19.20	19.40	19.46	19.15
Acetic	23.51	21.28	19.60	20.81
	RMS error	1.61	2.16	1.73

[a] The linear fit to training set $pK_{a,lfit}$ vs. $pK_{a,expt}$ for all three solvation models has slope 1.00 and intercept 0.00.

3.2. Protonated amine cationic acids in water

We turn now to the case of protonated amine cationic acids in water for which it is the acid, not the conjugate base that is charged and should dominate $\Delta\Delta G^*_{solv}$. Fig. 4A shows a scatter plot of calculated vs. experimental pK_a values (Table S3) [60–63]. In this case, the points corresponding to the use of the three different radii fall into distinct groups, with all of the UAHF and Bondi calculated values within 5 pH units of the corresponding experimental ones. Only a few of the UFF calculated values have error less than 5 pH units, and all are lower than their corresponding experimental values. As in Fig. 3A, for all three data sets, the deviation of all the calculated points

Table 3
Comparison of final $pK_{a,lfit}$ values for the cationic acid training and test sets in water using UFF, UAHF and Bondi radii with the B3LYP functional.[a]

Acid	$pK_{a,expt}$	UFF $pK_{a,lfit}$	UAHF $pK_{a,lfit}$	Bondi $pK_{a,lfit}$
2-chloro-pyridinium	0.49	1.74	1.42	1.16
4-cyano-pyridinium	1.86	2.95	2.19	2.82
4-bromo-anilinium	3.89	1.47	3.21	2.43
Anilinium	4.62	2.35	3.88	3.03
Pyridinium	5.24	6.47	5.36	5.81
2,4,6-collidinium	7.33	8.61	7.51	7.68
Benzylammonium	9.30	8.25	8.00	9.20
Triethylammonium	10.72	11.06	11.17	11.17
Pyrrolidinium	11.27	11.10	10.99	11.45
Guanidinium	13.60	14.41	14.79	13.76
	RMS error	1.37	0.74	0.82
2,5-dichloro-anilinium	1.53	−1.61	−0.17	−0.61
p-anisidinium	5.36	3.60	4.98	4.22
2,6-dimethyl-pyridinium	6.70	8.41	7.40	7.63
Hydrazinium	8.12	5.53	7.68	6.86
DMAP	9.60	11.29	9.55	10.10
	RMS error	2.26	0.86	1.31

[a] The linear fit to training set $pK_{a,lfit}$ vs. $pK_{a,expt}$ for all three solvation models has slope 1.00 and intercept 0.00.

Table 5
Comparison of final $pK_{a,lfit}$ values for the cationic acid training and test sets in acetonitrile using UFF, UAHF and Bondi radii with the B3LYP functional.[a]

Acid	$pK_{a,expt}$	UFF $pK_{a,lfit}$	UAHF $pK_{a,lfit}$	Bondi $pK_{a,lfit}$
2,6-Dichloro-anilinium	5.06	3.67	4.72	4.48
2-Chloro-pyridinium	6.80	8.32	7.34	7.43
4-Cyano-pyridinium	8.10	9.65	8.23	9.26
4-Bromo-anilinium	9.40	7.90	9.46	8.75
Anilinium	10.62	8.94	10.37	9.49
Pyridinium	12.53	13.83	12.40	12.89
2,4,6-Collidinium	14.98	16.38	15.24	15.15
Benzylammonium	16.91	15.83	15.70	16.75
Triethylammonium	18.82	19.19	19.85	19.18
Pyrrolidinium	19.70	19.17	19.56	19.44
	RMS error	1.30	0.56	0.65
2,5-Dichloro-anilinium	6.20	4.32	5.15	5.19
p-Cyano-anilinium	7.00	4.30	5.69	5.21
p-Anisidinium	11.86	10.43	11.81	10.92
2,6-Dimethyl-pyridinium	14.13	16.13	15.07	15.07
DMAP	17.95	19.55	17.92	18.05
	RMS error	1.62	0.86	1.10

[a] The linear fit to training set $pK_{a,lfit}$ vs. $pK_{a,expt}$ for all three solvation models has slope 1.00 and intercept 0.00.

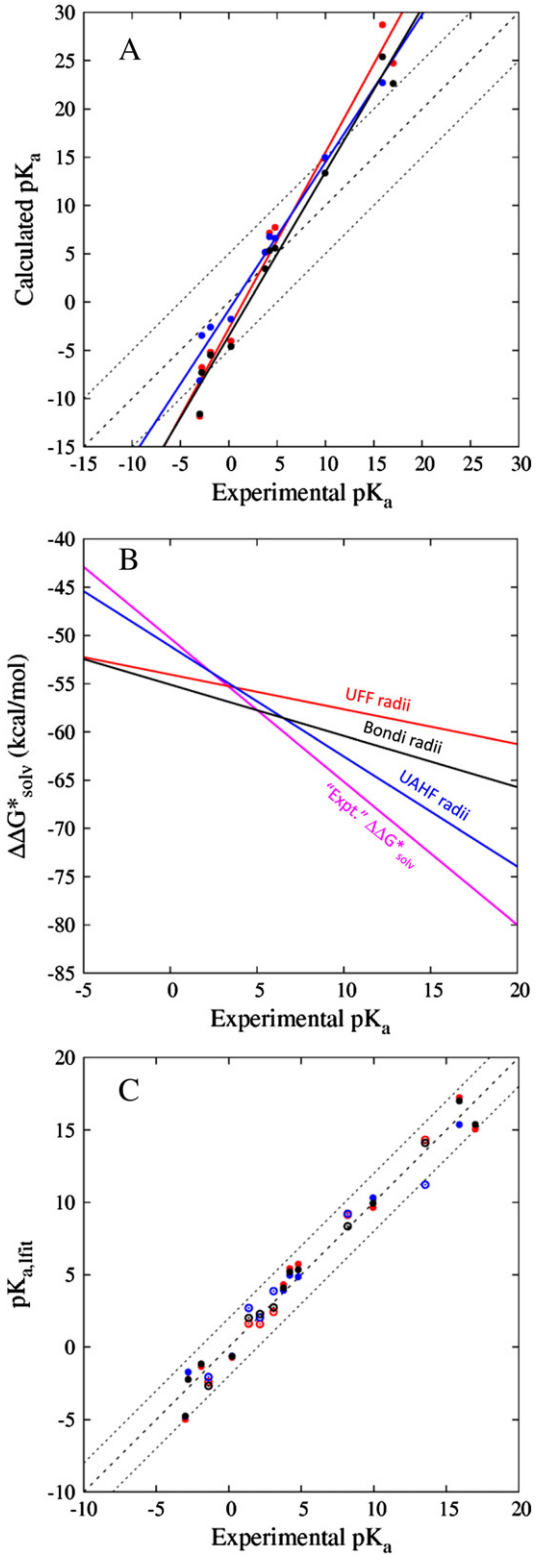

Fig. 3. Comparison of calculated results for neutral acids in water for CPCM solvation model with UFF, (red), UAHF (blue) and Bondi (black) radii: (A) scatter plot of p$K_{a,calc}$ vs. p$K_{a,expt}$ for training set with linear least-squares fits to points; (B) $\Delta\Delta G^*_{solv}$ vs. p$K_{a,expt}$ with "Expt." $\Delta\Delta G^*_{solv}$ (magenta); and (C) p$K_{a,lfit}$ vs. p$K_{a,expt}$ for training set (filled points) and test set (open circles with central points).

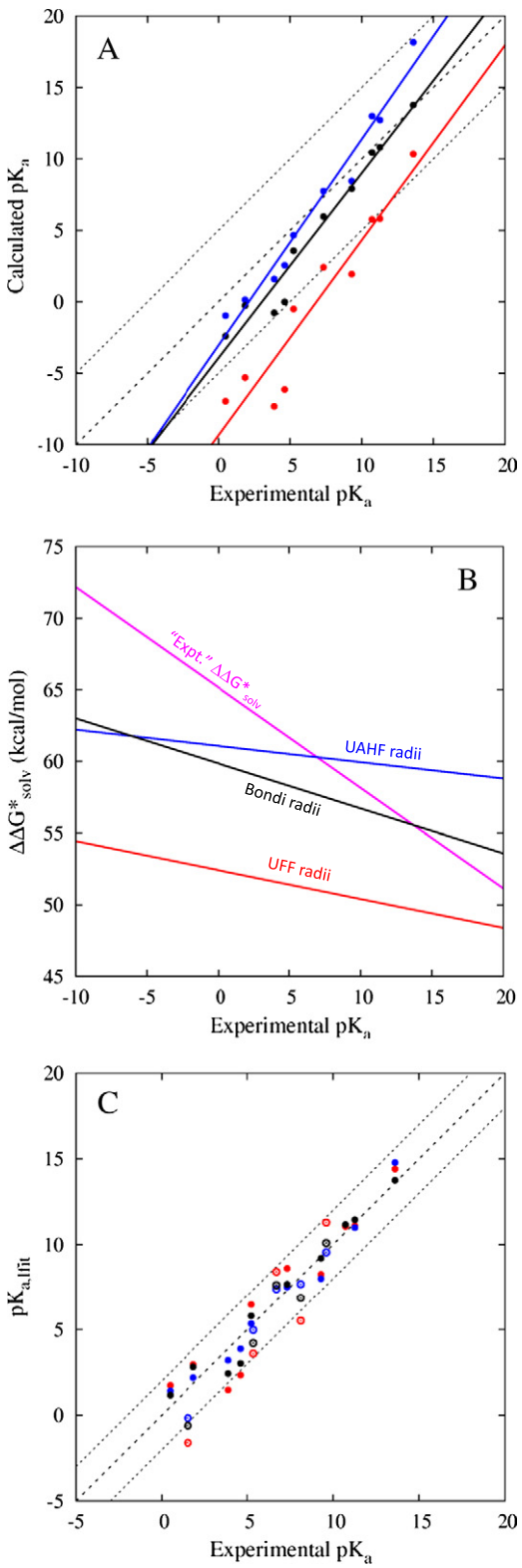

Fig. 4. Comparison of calculated results for protonated amine cationic acids in water for CPCM solvation model with UFF, (red), UAHF (blue) and Bondi (black) radii: (A) scatter plot of p$K_{a,calc}$ vs. p$K_{a,expt}$ for training set with linear least-squares fits to points; (B) $\Delta\Delta G^*_{solv}$ vs. p$K_{a,expt}$ with "Expt." $\Delta\Delta G^*_{solv}$ (magenta); and (C) p$K_{a,lfit}$ vs. p$K_{a,expt}$ for training set (filled points) and test set (open circles with central points).

from the least-squares linear fit is small, especially for the Bondi and UAHF radii, indicating that the majority of any error is not random.

The slope (−0.70 kcal/mol per pH) of the "experimental" $\Delta\Delta G^*_{solv}$ line in Fig. 4B is again more negative than that of any of the CPCM lines corresponding to the different radii, but ca. 2.5 times smaller in magnitude than for the case of neutral acids in water. All three solvation models under-solvate the stronger acids, and the calculations with UAHF and Bondi radii over-solvate the weaker acids while those with

UFF radii under-solvate the entire cationic acid training set. The slopes of the linear least-squares lines for the Bondi, UAHF and UFF data are −0.31, −0.11 and −0.20 kcal/mol per pH, respectively, and the Bondi and UAHF lines cross the "experimental" line in the pK_a region (0 to 15) of interest, so neither the Bondi nor UAHF CPCM data for the training set diverge from the "experimental" line very much.

Fig. 4C shows that the p$K_{a,lfit}$ values derived from the use of the UAHF and Bondi data (including the test set) fairly accurately fit the individual experimental values, having RMS errors of 0.74 and 0.82, respectively, for the training set and 0.86 and 1.31, respectively, for the test set, with one error (2.14 for 2,5-dichloroanilinium) larger than 2 pH units. The p$K_{a,lfit}$ values (Table 3) derived from the UFF data have RMS errors of 1.37 and 2.26 for the training and test sets, respectively, with two training set and two test set points having error greater than 2 pH units.

3.3. Neutral acids in acetonitrile

The aqueous environment is a ubiquitous medium for biological and chemical processes, but often it is desirable or even necessary to study processes in a nonaqueous environment. In particular, inorganic and bioinorganic reactions often require the chemistry to take place in a nonaqueous medium such as acetonitrile or dimethylformamide, principally because the species are unstable in an aqueous environment. It is therefore useful to study the applicability of the current approach to other, nonaqueous, solvents and compare and contrast the observed errors in terms of the underlying fundamental principles. For this we apply our methodology to the prediction of pK_a values in acetonitrile to a different training set and test set of neutral acids using the same three solvation models, with the only difference now that the dielectric constant in the continuum reflects that of acetonitrile and not water. Table S4 lists the experimental and uncorrected calculated pK_a values, while Table 4 compares the experimental [64–67] and derived p$K_{a,lfit}$ values. Fig. 5A shows scatter plots of the three calculated vs. experimental pK_a data sets. All three data sets underestimate the pK_a, and no set has a calculated value within 5 pH units of the experimental values. As in the case of neutral acids in water (Fig. 1A), the three sets of calculated values agree with each other fairly well, and the deviation in the linear least-squares line within each data set is again relatively small.

The "experimental" $\Delta\Delta G^*_{solv}$ (Fig. 5B) has a much weaker dependence on the experimental pK_a values (slope −0.34 kcal/mol per pH) of these neutral acids in acetonitrile than was the case for the neutral acid test set in water solvent, and all three solvation models oversolvate the anionic conjugate bases of these neutral acids. Acetonitrile appears to be a more benign solvent than water considering that the three solvation models give a very similar dependence of the CPCM $\Delta\Delta G^*_{solv}$ on the experimental pK_a (slopes of −0.38, −0.42 and −0.29 kcal/mol per pH for UFF, UAHF and Bondi radii, respectively), so that an isodesmic approach in which acetonitrile is simply modeled as a dielectric continuum would yield fairly accurate results over a wide range of pK_a, which is not the case for water in such a simplistic continuum-only approach. The scatter plot of p$K_{a,lfit}$ vs. experimental pK_a values in Fig. 5C indicates that the current methodology is generally successful for this case, but the UAHF results are surprisingly less accurate than those obtained with the other two solvation models. The RMS errors in p$K_{a,lfit}$ values for the training and test sets are 1.31 and 1.61 (UFF), 1.59 and 2.16 (UAHF), and 1.22 and 1.73 (Bondi), respectively. Only one of the UAHF test set points has unusually large error, while most of the other points have error less than 2 pH units.

3.4. Protonated amine cationic acids in acetonitrile

Finally, we turn to protonated amine cationic acids in acetonitrile. The calculated pK_a values for the various solvation models are listed in Table S5, and are shown in the scatter plot in Fig. 6A (Table 5). The three data sets are completely separated, and the linear least-squares

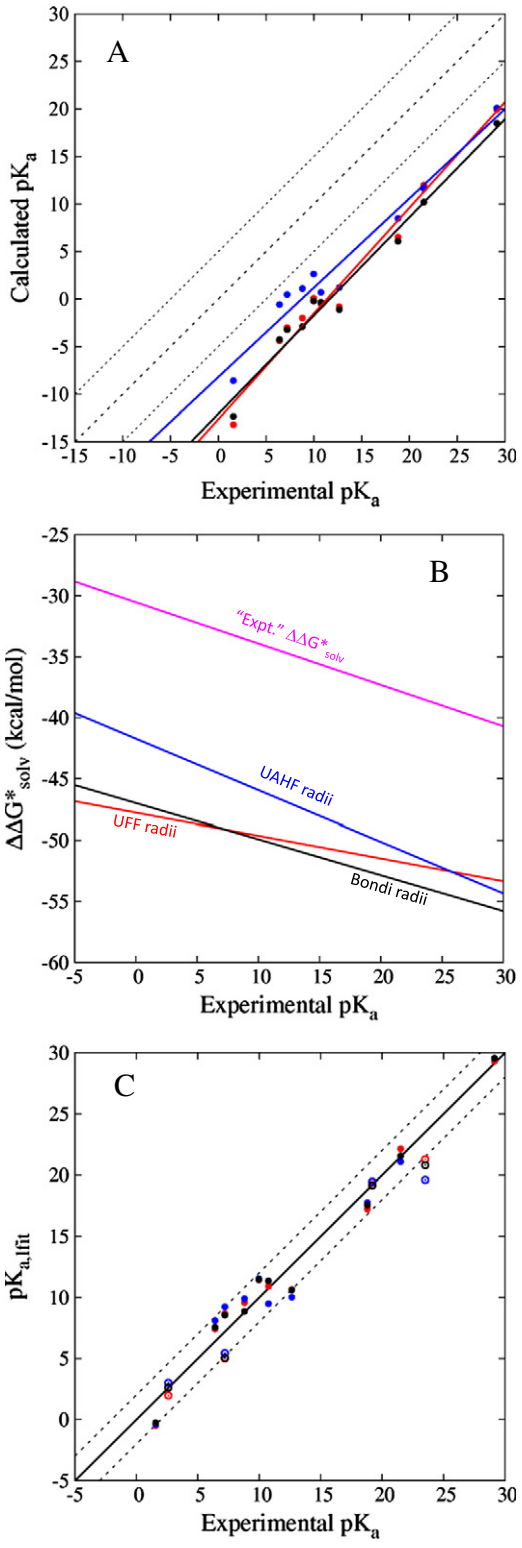

Fig. 5. Comparison of calculated results for neutral acids in acetonitrile for CPCM solvation model with UFF, (red), UAHF (blue) and Bondi (black) radii: (A) scatter plot of p$K_{a,calc}$ vs. p$K_{a,expt}$ for training set with linear least-squares fits to points; (B) $\Delta\Delta G^*_{solv}$ vs. p$K_{a,expt}$ with "Expt." $\Delta\Delta G^*_{solv}$ (magenta); and (C) p$K_{a,lfit}$ vs. p$K_{a,expt}$ for training set (filled points) and test set (open circles with central points).

fit lines have slopes near unity (1.17, 1.11 and 1.09 for UFF, UAHF and Bondi radii, respectively) so that an isodesmic approach for calculating pK_a would likely, as with the neutral acids, be quite accurate. The deviations of the UAHF and Bondi points from the corresponding least-squares line are particularly small. The comparison of "experimental" and CPCM

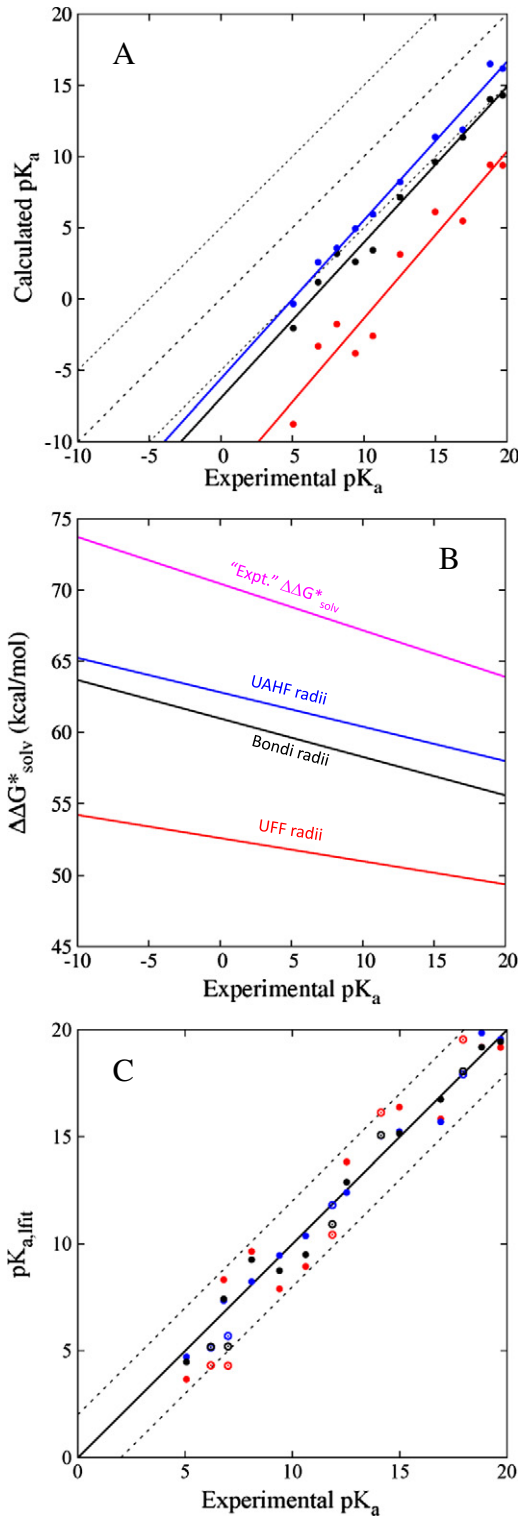

Fig. 6. Comparison of calculated results for protonated amine cationic acids in acetonitrile for CPCM solvation model with UFF, (red), UAHF (blue) and Bondi (black) radii: (A) scatter plot of p$K_{a,calc}$ vs. p$K_{a,expt}$ for training set with linear least-squares fits to points; (B) $\Delta\Delta G^*_{solv}$ vs. p$K_{a,expt}$ with "Expt." $\Delta\Delta G^*_{solv}$ (magenta); and (C) p$K_{a,lfit}$ vs. p$K_{a,expt}$ for training set (filled points) and test set (open circles with central points).

derived values of $\Delta\Delta G^*_{solv}$ in Fig. 6B shows that all three solvation models under-solvate the cationic acids, as they tended to do for the most part in water. The dependence of the "experimental" $\Delta\Delta G^*_{solv}$ on pK_a (−0.33 kcal/mol per pH) is quite similar to that (−0.34) for the neutral acids in acetonitrile, and in the case of these protonated amine cationic acids is reasonably well approximated by all three CPCM solvation models (−0.16, −0.24 and −0.27 kcal/mol per pH for UFF, UAHF and Bondi radii, respectively), underscoring the appropriateness of an isodesmic scheme for this case [16].

The scatter plot of p$K_{a,lfit}$ vs. p$K_{a,expt}$ in Fig. 6C reflects the general success of the method, with RMS errors of 1.30, 0.56 and 0.65 in the training set and 1.62, 0.86 and 1.10 in the test set for UFF, UAHF and Bondi radii, respectively. The UFF results are seen to be the least accurate as reflected by the largest RMS errors and the largest scatter from the ideal line, including one point with error larger than 2 pH units.

3.5. General observations

Figure S4 shows the linearized error in the CPCM solvation models under consideration for the four acid/solvent cases [acids: neutral, cationic; solvent: water, MeCN] treated in this work. These linearized errors are simply the differences between the CPCM line for each type of radii and the "experimental" line in panel (B) of Figs. 3–6. The largest pK_a-dependent solvation errors found in this work, as evidenced by larger slopes as opposed to intercepts, were for water as a solvent, and were larger for neutral acids than for protonated amine cationic acids. This suggests that there is more error associated with solvation models of anions than for cations. In changing the solvent from water to acetonitrile, which solvates the ionic species less strongly than water (as evidenced by smaller slopes and larger contributions from the intercepts), absolute errors in the $\Delta\Delta G^*_{solv}$ values become smaller and approach those in the gas-phase acidity calculations, e.g., the scatter in the "experimental" and CPCM points in Fig. 1B become comparable to the difference between the two lines. Consequently, the RMS and maximum errors in the p$K_{a,lfit}$ values become larger.

Also shown in Fig. S4 are horizontal lines indicating the average signed error and mean unsigned error computed for the range of experimental pK_a values employed in each case. It is noteworthy that in the case of neutral acids, the average signed errors and mean unsigned errors are both positive and relatively close in magnitude. This indicates that on the average, the anionic conjugate bases of neutral acids are under-solvated so that their average error is positive and the signed and unsigned errors are similar. For the cationic acids, unsigned errors are essentially the negative of the (negative valued) signed errors reflecting significant under-solvation of the positively charged cationic acid species.

It is worth mentioning that for the amine cationic acids in acetonitrile there is very little linear dependence in the error of the differential solvation as can be seen from the small slope in Fig. S4d and similarly from the parallelism of the computed $\Delta\Delta G_{solv}$ and $\Delta\Delta G_{solv}$ experimental in Fig. 6B. The error is simply a constant over the full range of pH making an isodesmic scheme, where this constant error can be easily canceled, a favorable approach.

The presence of a large positive slope of $\Delta\Delta G^*_{solv}$ error for the neutral acids both in water and acetonitrile, and even for the case of cationic acids in water, indicates a strong dependence in the solvation error on the pK_a, which renders an isodesmic approach inappropriate for those cases. A better solution would be to represent the interaction of the solvent with the active site of the acid or base by a more realistic model of the solvent than is offered by a polarizable continuum by itself. A cluster-continuum approach in which *N* explicit molecules of the solvent are included as part of the solute, thereby offering a hydrogen bond acceptor for the acid and a hydrogen bond donor to the conjugate base, with this newly defined solute species embedded in a dielectric continuum, in addition to the polarizable continuum would likely reduce the strong pK_a dependence of the $\Delta\Delta G^*_{solv}$ error observed in this work [8]. Nevertheless, the current methodology appears to be capable of calculating pK_a values in both water and acetonitrile solvents that are generally within 1 pH unit and almost always within 2 pH units of the experimental values. The training sets employed here, especially the two neutral acid ones (containing

inorganic and organic acids as well as alcohols), were intentionally rather diverse, and selected to span a wide range of pK_a values for neutral or cationic acids while otherwise being as widely representative as possible. The usefulness of the current approach depends on there not being too many classes of acids to calibrate. The most efficient course of action in this regard would be to choose an electronic structure method, a reasonably large basis set (e.g., B3LYP and 6-311+G(d,p) as used here), and a solvation model (e.g., CPCM with UAHF or Bondi radii) to be used in all situations. After successfully calibrating several classes of acids to obtain $pK_{a,lfit}$ from calculated pK_a values, predicting the value of an unknown pK_a for an acid belonging to one of those classes would then be as simple as an isodesmic scheme, and far more general.

4. Conclusions

We have presented an approach to the calculation of pK_a values in water and acetonitrile based on the straightforward calculation of the gas-phase absolute free energies of the acid and conjugate base, and then using a continuum solvation model to obtain the corresponding solution phase free energies. Most of the error in such an approach arises from inaccurate differential solvation free energies of the acid and conjugate base, which can be corrected, provided the gas-phase acidities have only small systematic error, which was shown here to be the case. The methodology was applied to the comparison of predicted pK_a values for a set of neutral acids in water using three different density functionals, showing that the results were insensitive to the choice of functional despite what appeared to be different random error in the calculated gas-phase acidities. Finally, the approach was applied to the comparison of predicted pK_a values for sets of neutral acids and protonated amine cationic acids in both water and acetonitrile solvents. The methodology was shown to generally predict the pK_a values for all the cases investigated to within 1 pH unit so long as the differential solvation error is larger than the random error in the gas-phase acidity calculations.

For the case of neutral acids in water, the different solvation models solvated the anionic conjugate bases fairly accurately on the average as evidenced by Fig. 3B, but the dependence of the basicity of the anion was underestimated to varying degrees by the different models. The three continuum solvent models in water tended to undersolvate the cationic acids, and the dependence on the acidity of the cations was underestimated. The anionic conjugate bases of neutral acids in acetonitrile were significantly oversolvated by all the continuum solvation models, but the dependence on the basicity of the anion was modeled better. In the case of protonated amine cationic acids in acetonitrile, all the continuum solvation models significantly undersolvated the cations, but the dependence on the acidity of the cations was modeled well. However, one of the take-home messages of the present work is that even rather poor initial estimates can be systematically corrected.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbabio.2013.03.011.

References

[1] K. Adam, New density functional and atoms in molecules method of computing relative pK_a values in solution, J. Phys. Chem. A 106 (2002) 11963–11972.
[2] M.D. Liptak, K.C. Gross, P.G. Seybold, S. Feldgus, G.C. Shields, Absolute pK_a determinations for substituted phenols, J. Am. Chem. Soc. 124 (2002) 6421–6427.
[3] J.R. Pliego Jr., J.M. Riveros, Theoretical calculation of pK_a using the cluster-continuum model, J. Phys. Chem. A 106 (2002) 7434–7439.
[4] M. Namazian, J. Heidary, Ab initio calculations of pK_a values of some organic acids in aqueous solution, Theochem 620 (2003) 257–263.
[5] Y. Fu, L. Liu, R.-Q. Li, R. Liu, Q.-X. Guo, First-principle predictions of absolute pK_a's of organic acids in dimethyl sulfoxide solution, J. Am. Chem. Soc. 126 (2004) 814–822.
[6] A.M. Magill, K.J. Cavell, B.F. Yates, Basicity of nucleophilic carbenes in aqueous and nonaqueous solvents—theoretical predictions, J. Am. Chem. Soc. 126 (2004) 8717–8724.
[7] M. Namazian, S. Halvani, M.R. Noorbala, Density functional theory response to the calculations of pK_a values of some carboxylic acids in aqueous solution, Theochem 711 (2004) 13–18.
[8] C.P. Kelly, C.J. Cramer, D.G. Truhlar, Adding explicit solvent molecules to continuum solvent calculations for the calculation of aqueous acid dissociation constants, J. Phys. Chem. A 110 (2006) 2493–2499.
[9] M. Namazian, S. Halvani, Calculations of pK_a values of carboxylic acids in aqueous solution using density functional theory, J. Chem. Thermodyn. 38 (2006) 14951502.
[10] V.S. Bryantsev, M.S. Diallo, W.A. Goddard III, pK_a calculations of aliphatic amines, diamines, and aminoamides via density functional theory with a Poisson–Boltzmann continuum solvent model, J. Phys. Chem. A 111 (2007) 4422–4430.
[11] M. Namazian, M. Zkery, M.R. Noorbala, M.L. Coote, Accurate calculation of the pK_a of trifluoroacetic acid using high-level ab initio calculations, Chem. Phys. Lett. 451 (2008) 163–168.
[12] J. Ho, M.L. Coote, pK_a calculation of some biologically important carbon acids — an assessment of contemporary theoretical procedures, J. Chem. Theor. Comput. 5 (2009) 295–305.
[13] J. Ho, M.L. Coote, A universal approach to continuum solvent pK_a calculations: are we there yet? Theor. Chem. Acc. 125 (2010) 3–21.
[14] S. Zhang, J. Baker, P. Pulay, A reliable and efficient first principles-based method for predicting pK_a values. 1. Methodology, J. Phys. Chem. A 114 (2010) 425–431.
[15] S. Zhang, J. Baker, P. Pulay, A reliable and efficient first principles-based method for predicting pK_a values. 2. Organic acids, J. Phys. Chem. A 114 (2010) 432–442.
[16] J. Ho, M.L. Coote, First-principles prediction of acidities in the gas and solution phase, WIREs Comput. Mol. Sci. 1 (2011) 649–660.
[17] J.A. Keith, E.A. Carter, Theoretical insights into pyridinium-based photoelectrocatalytic reduction of CO_2, J. Am. Chem. Soc. 134 (2012) 7580–7583.
[18] D.A. McQuarrie, Statistical Mechanics, Harper and Row, New York, 1973.
[19] J.P. Guthrie, Hydrolysis of esters of oxy acids: pK_a values for strong acids; Brønsted relationship for attack of water at methyl; free energies of hydrolysis of esters of oxy acids; and a linear relationship between free energy of hydrolysis and pK_a holding over a range of 20 pK units, Can. J. Chem. 56 (1978) 2342–2354.
[20] J.F.J. Dippy, S.R.C. Hughes, A. Rozanski, The dissociation constants of some symmetrically disubstituted succinic acids, J. Chem. Soc. 498 (1959) 2492–2498.
[21] H.C. Brown, D.H. McDaniel, O. Häfliger, Dissociation constants, chapter 14, in: E.A. Braude, F.C. Nachod (Eds.), Determination of Organic Structures by Physical Methods, 561, Academic Press, New York, 1955, pp. 567–662.
[22] E.P. Serjeant, B. Dempsey, Ionisation Constants of Organic Acids in Aqueous Solution, IUPAC Chemical Data Series No, 23rd ed. Pergamon Press, Oxford, UK, 1979.
[23] P. Ballinger, F.A. Long, Acid ionization constants of alcohols. II. Acidities of some substituted methanols and related compounds, J. Am. Chem. Soc. 82 (1960) 795–798.
[24] J. Bjerrum, G. Schwarzenback, L.G. Sillen, Stability Constants of Metal Complexes, with Solubility Products of Inorganic Substances: Part I, Organic Ligands and Part II, Inorganic Ligands, Chemical Society, London, 1957, 1958.
[25] R.M.C. Dawson, et al., Data for Biochemical Research, Clarendon Press, Oxford, 1959.
[26] B. Elvers, S. Hawkins, G. Schulz, Ullmann's Encyclopedia of Industrial Chemistry, 5th ed. Weinheim, VCH, 1991.
[27] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (1993) 5648–5652.
[28] C.T. Lee, W.T. Yang, R.G. Parr, Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density, Phys. Rev. B 37 (1988) 785–789.
[29] S.H. Vosko, L. Wilk, M. Nusair, Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis, Can. J. Phys. 58 (1980) 1200–1211.
[30] A.D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, Phys. Rev. A 38 (1988) 3098–3100.
[31] A. Klamt, G. Schűűrmann, COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, J. Chem. Soc., Perkin Trans. 2 (1993) 799–805.
[32] V. Barone, M. Cossi, Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model, J. Phys. Chem. A 102 (1998) 1995–2001.
[33] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[34] A.K. Rappe, C.J. Casewit, K.S. Colwell, W.A. Goddard III, W.M. Skiff, UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, J. Am. Chem. Soc. 114 (1992) 10024–10035.
[35] M.J. Frisch, et al., Gaussian09, Revision B.01, Gaussian Inc., Wallingford, CT, 2010.
[36] M.D. Tissandier, K.A. Cowen, W.Y. Feng, E. Gundlach, M.H. Cohen, A.D. Earhart, J.V. Coe, T.R. Tuttle Jr., The proton's absolute aqueous enthalpy and Gibbs free energy of solvation from cluster-ion solvation data, J. Phys. Chem. A 102 (1998) 7787–7794.
[37] D.M. Camaioni, C.A. Schwerdtfeger, Comment on "accurate experimental values for the free energies of hydration of H^+, OH^-, and H_3O^+", J. Phys. Chem. A 109 (2005) 10795–10797.

[38] C.P. Kelly, C.J. Cramer, D.G. Truhlar, Aqueous solvation free energies of ions and ion–water clusters based on an accurate value for the absolute aqueous solvation free energy of the proton, J. Phys. Chem. B 110 (2006) 16066–16081.
[39] C.P. Kelly, C.J. Cramer, D.G. Truhlar, Single-ion solvation free energies and the normal hydrogen electrode potential in methanol, acetonitrile, and dimethyl sulfoxide, J. Phys. Chem. B 111 (2007) 408–422.
[40] M.D. Liptak, G.C. Shields, 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 (2001) 7314–7319.
[41] P. Burk, I.A. Koppel, I. Koppel, I. Leito, O. Travnikova, Critical test of performance of B3LYP functional for prediction of gas-phase acidities and basicities, Chem. Phys. Lett. 323 (2000) 482–489.
[42] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople, Gaussian-3 (G3) theory for molecules containing first and second-row atoms, J. Chem. Phys. 109 (1998) 7764–7776.
[43] S. Hammerum, Heats of formation and proton affinities by the G3 method, Chem. Phys. Lett. 300 (1999) 529–532.
[44] J.M. Martin, T.J. Lee, The atomization energy and proton affinity of NH3. An ab initio calibration study, Chem. Phys. Lett. 258 (1996) 136–143.
[45] J.W. Ochterski, G.A. Petersson, K.B. Wiberg, A comparison of model chemistries, J. Am. Chem. Soc. 117 (1995) 11299–11308.
[46] K.A. Peterson, S.S. Xantheas, D.A. Dixon, T.H.J. Dunning Jr., Predicting the proton affinities of H_2O and NH_3, J. Phys. Chem. A 102 (1998) 2449–2454.
[47] E.K. Pokon, M.D. Liptak, S. Feldgus, G.C. Shields, Comparison of CBS-QB3, CBS-APNO, and G3 predictions of gas phase deprotonation data, J. Phys. Chem. A 105 (2001) 10483–10487.
[48] M. Remko, The gas-phase acidities of substituted hydroxamic and silahydroxamic acids: a comparative ab initio study, J. Phys. Chem. A 106 (2002) 5005–5010.
[49] Y. Seo, Y. Kim, Y. Kim, MC-QCISD calculations for proton affinities of molecules and geometries, Chem. Phys. Lett. 340 (2001) 186–193.
[50] B.J. Smith, L. Radom, Evaluation of accurate gas-phase acidities, J. Phys. Chem. 95 (1991) 10549–10551.
[51] B.J. Smith, L. Radom, Assigning absolute values to proton affinities: a differentiation between competing scales, J. Am. Chem. Soc. 115 (1993) 4885–4888.
[52] B.J. Smith, L. Radom, An evaluation of the performance of density functional theory, MP2, MP4, F4, G2(MP2) and G2 procedures in predicting gas-phase proton affinities, Chem. Phys. Lett. 231 (1994) 345–351.
[53] K.B. Wiberg, Substituent effects on the acidity of weak acids. 2. Calculated gas-phase acidities of substituted benzoic acids, J. Org. Chem. 67 (2002) 4787–4794.
[54] Y. Zhao, D.G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, Theor. Chem. Acc. 120 (2008) 215–241.
[55] J.P. Perdew, Density-functional approximation for the correlation energy of the inhomogeneous electron gas, Phys. Rev. B 33 (1986) 8822–8824.
[56] A.D. McLean, G.S. Chandler, Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z = 11–18, J. Chem. Phys. 72 (1980) 5639–5648.
[57] R. Krishnan, J.S. Binkley, R. Seeger, J.A. Pople, Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions, J. Chem. Phys. 72 (1980) 650–654.
[58] V. Barone, M. Cossi, J. Tomasi, A new definition of cavities for the computation of solvation free energies by the polarizable continuum model, J. Chem. Phys. 107 (1997) 3210–3221.
[59] A. Bondi, van der Waals volumes and radii, J. Phys. Chem. 68 (1964) 441–451.
[60] I. Kaljurand, A. Kütt, L. Sooväli, T. Rodima, V. Mäemets, I. Leito, I.A. Koppel, Extension of the self-consistent spectrophotometric basicity scale in acetonitrile to a full span of 28 pK_a units: unification of different basicity scales, J. Org. Chem. 70 (2005) 1019–1028.
[61] D.D. Perrin, Dissociation Constants of Organic Bases in Aqueous Solution, Butterworths, London, 1965. (Supplement, 1972).
[62] A.M. Appel, S.-J. Lee, J.A. Franz, D.L. DuBois, M. Rakowski DuBois, B. Twamley, Determination of S – H bond strengths in dimolybdenum tetrasulfide complexes, Organometallics 28 (2009) 749–754.
[63] H.K. Hall Jr., Correlation of the base strengths of amines, J. Am. Chem. Soc. 79 (1957) 5441–5444.
[64] A. Kütt, I. Leito, I. Kaljurand, L. Sooväli, V.M. Vlasov, L.M. Yagupolskii, I.A. Koppel, A comprehensive self-consistent spectrophotometric acidity scale of neutral Brønsted acids in acetonitrile, J. Org. Chem. 71 (2006) 2829–2838.
[65] A. Kütt, V. Movchun, T. Rodima, T. Dansauer, E.B. Rusanov, I. Leito, I. Kaljurand, J. Koppel, V. Pihl, I. Koppel, G. Ovsjannikov, L. Toom, M. Mishima, M. Medebielle, E. Lork, G.-V. Röschenthaler, I.A. Koppel, A.A. Kolomeitsev, Pentakis(trifluoromethyl) phenyl, a sterically crowded and electron-withdrawing group: synthesis and acidity of pentakis(trifluoromethyl)benzene, -toluene, -phenol, and -aniline, J. Org. Chem. 73 (2008) 2607–2620.
[66] O.W. Webster, Polycyanation. The reaction of cyanogen chloride, cyclopentadiene, and sodium hydride, J. Am. Chem. Soc. 88 (1966) 3046–3050.
[67] A. Kütt, T. Rodima, J. Saame, E. Raamat, V. Mäemets, I. Kaljurand, I.A. Koppel, R.Y. Garlyauskayte, Y.L. Yagupolskii, L.M. Yagupolskii, E. Bernhardt, H. Willner, I. Leito, Equilibrium acidities of superacids, J. Org. Chem. 76 (2011) 391–395.