# Improved pKa Prediction of Substituted Alcohols, Phenols, and Hydroperoxides in Aqueous Medium Using Density Functional Theory and a Cluster-Continuum Solvation Model

Bishnu Thapa and H. Bernhard Schlegel*

Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States

Ⓢ *Supporting Information*

ABSTRACT: Acid dissociation constants (pKa's) are key physicochemical properties that are needed to understand the structure and reactivity of molecules in solution. Theoretical pKa's have been calculated for a set of 72 organic compounds with −OH and −OOH groups (48 with known experimental pKa's). This test set includes 17 aliphatic alcohols, 25 substituted phenols, and 30 hydroperoxides. Calculations in aqueous medium have been carried out with SMD implicit solvation and three hybrid DFT functionals (B3LYP, ωB97XD, and M06-2X) with two basis sets (6-31+G(d,p) and 6-311++G(d,p)). The effect of explicit water molecules on calculated pKa's was assessed by including up to three water molecules. pKa's calculated with only SMD implicit solvation are found to have average errors greater than 6 pKa units. Including one explicit water reduces the error by about 3 pKa units, but the error is

still far from chemical accuracy. With B3LYP/6-311++G(d,p) and three explicit water molecules in SMD solvation, the mean signed error and standard deviation are only −0.02 ± 0.55; a linear fit with zero intercept has a slope of 1.005 and R^2 = 0.97. Thus, this level of theory can be used to calculate pKa's directly without the need for linear correlations or thermodynamic cycles. Estimated pKa values are reported for 24 hydroperoxides that have not yet been determined experimentally.
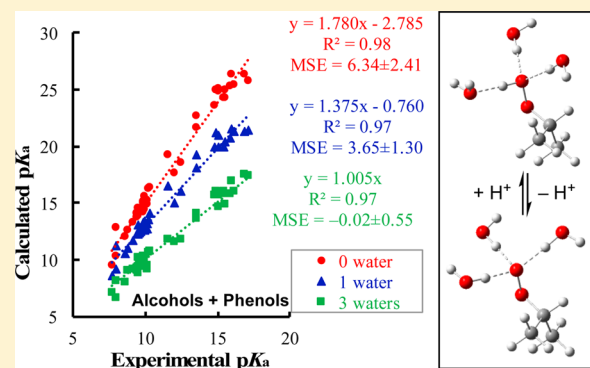
■ INTRODUCTION

Organic compounds with hydroxyl (−OH) and hydroperoxyl (−OOH) functional groups are abundant in nature. Alcohols and phenols constitute a considerable fraction of the building blocks in many biological systems and pharmaceutical compounds, and are important components in organic and inorganic synthesis.[1−5] Similarly, organic hydroperoxides have drawn appreciable attention because they are strong oxidants and radical reservoirs, and they play important roles in combustion, atmospheric chemistry, and biology.[6−13] The protonation state of the oxygen in alcohols, phenols, and hydroperoxides affects their structural and physicochemical properties.[14] The extent of protonation/deprotonation in solution is determined by the acid dissociation constant (pKa). Hence, the knowledge of accurate pKa's of alcohols, phenols, and hydroperoxides is pivotal to understand their chemical reactivity.

The pKa's of a wide range of substituted alcohols and phenols have been studied both experimentally[15−18] and theoretically.[19−35] Nevertheless, there are many more −OH containing compounds that do not have experimentally determined pKa's. Contrary to the case for alcohols, there are far fewer experimental studies of the pKa's of hydroperoxides.[17,18] To our best knowledge, there are no systematic theoretical studies of pKa's of hydroperoxides in aqueous solution. Furthermore, pKa's of short-lived intermediates and compounds in complex

chemical environments are difficult to measure, despite technological advances in experimental techniques.[36] Consequently, there is always a need for computational methods for estimating pKa's.

On the theoretical side, considerable effort has been dedicated to predicting the pKa's of alcohols and phenols from first principles.[19−34] One of the biggest challenges in computing free energies in solution, particularly in water, is accurately incorporating solvation effects for the neutral and ionic species involved in proton-transfer processes.[37−40] Although the use of explicit solvent can provide an accurate description of solute−solvent interactions and dynamics, such calculations are still expensive for chemically interesting systems.[37−43] A more cost-effective alternative is to use a continuum solvation model such as COSMO, PCM, SMx (x = 1−8) or SMD, where the solvation effects are modeled with a dielectric medium. For example, Shields and co-workers have been quite successful in calculating pKa's of a number of carboxylic acids and substituted phenols using various thermodynamic cycles to combine gas-phase CBS and Gn energy differences with CPCM solvation.[34,44,45] Ho and co-workers[39,46] have shown that thermodynamic cycles are no

**Scheme 1. Organic Alcohols and Phenols Considered in This Study**
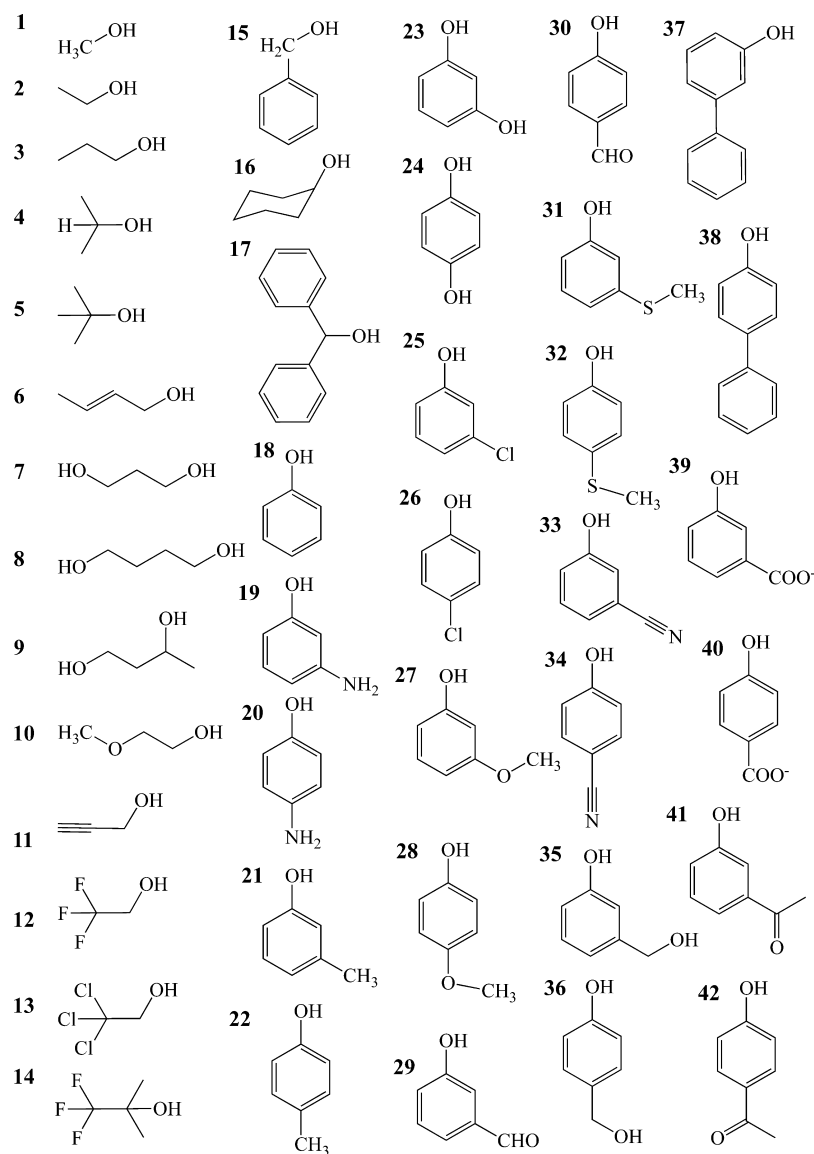
longer necessary with current implicit solvation methods, and that pK_a's can be calculated directly by optimizing the structures in solution. Some of the newer solvation models are well calibrated against large test sets and yield balanced treatments of the relative solvation energies of neutrals and ions. However, implicit solvation models do not incorporate specific short-range solute−solvent interactions such as hydrogen bonding, which can be particularly important in pK_a calculations.[37−43,47,48] The necessity of explicitly addressing such solute−solvent interactions is particularly important in pK_a calculations to achieve an accuracy of $\Delta pK_a \leq 1$ with respect to experiment.[37−40,47,48]

Usually, the error introduced by using implicit solvation in calculated free energies is quite systematic within a class of chemical compounds (e.g., alcohols, thiols).[38−41,47,48] Often, the error arises because of a difference in the accuracy of solvation energies for the neutrals and the protonated or deprotonated forms. Several research groups have successfully tackled the shortcomings of implicit solvation by using linear fits of computed deprotonation energies to the pK_a's for a large group of molecules. For example, Muckerman et al. have

recently explored the correlation between the calculated pK_a's for a number of organic alcohols and demonstrated that implicit-only solvation has a systematic error, which can be corrected by using the linear fit equation.[33] Many other researchers have also calculated the pK_a's of alcohols and phenols using the linear-fit methods.[23−30] The linear regression equations can then be used to predict pK_a's of unknown compounds within chemical accuracy.[38−41,47,48] However, different linear-fit equations are required for compounds with different functional groups, sometimes even for compounds within the same class.[29,38] For example, the pK_a's of protonated aromatic amines vs aliphatic amines require different linear regression equations to obtain reliable pK_a's despite having same the $-NH_3^+$ protonated group.[29,32] Furthermore, the linear fitting equations depend on the level of theory used for the electronic structure calculations, and even more on the choice of continuum solvation model. In addition, a fairly large set of compounds with accurate experimental pK_a's distributed over a sufficiently wide range is required to obtain a reliable linear-fit equation. Because of a lack of experimental data to

Scheme 2. Organic Hydroperoxides Considered in This Study
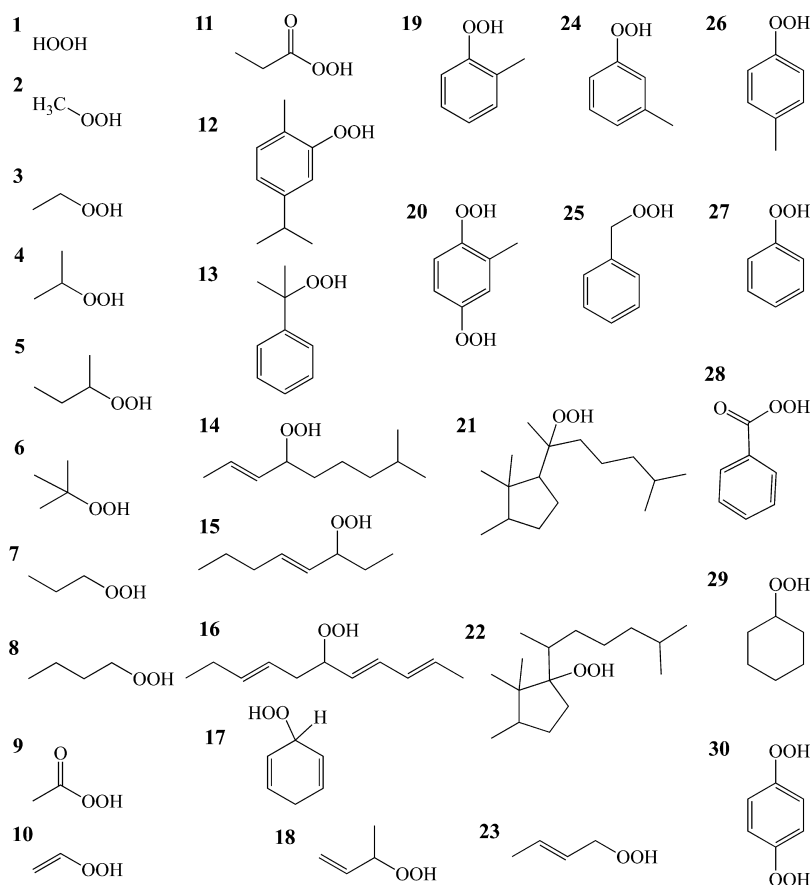
calibrate the correlations, it may not be possible to develop suitable linear fits for all groups of interest.

Though linear regression may be sufficient to obtain computational estimates of pK_a's, it is difficult to use this approach to calculate the free energy profiles of reactions in solution that involve protonation, deprotonation, or proton coupled electron transfer. All of the compounds in a reaction mechanism need to be calculated on an equal footing. This requires direct calculations of free energies in solution without scaling or linear fits.

The importance of including explicit solvent molecules while calculating solvation free energies of charged species such as oxy-anions has been discussed in the literature[38,39] and pursued by many researchers.[19−23] Some pK_a studies included explicit solvent molecules only for the charged species[19,20,31] whereas the others[21−23] included explicit water molecules for both the hydroxyl (−OH) and oxy-anion (−O⁻). Eckert et al.[22] used the BP/TZPV level of theory with the COSMO-RS implicit solvation model and added up to two explicit water molecules to calculate the pK_a's of a variety of organic compounds, including five alcohols. Their study showed that the slope of the linear free energy regression (LFER) correlation between the experimental pK_a's and the calculated free energy change could be systematically improved by including explicit waters. However, the authors concluded that the performance of the method was not general enough to be used for everyday production work. Zhang[23] computed the pK_a's of some aliphatic alcohols, substituted phenols and carboxylic acids by including zero, one, and two water molecules along with the COSMOS solvation model. Despite the use of explicit waters

forming hydrogen bonds, the study still needed to use a linear fit for calculating pK_a's and argued that including explicit waters deteriorates the accuracy of the computed pK_a's. Cunningham et al.[21] employed an implicit−explicit solvation approach, similar to the present study, to calculate absolute pK_a values for substituted phenols with B3LYP/6-31+G(d) method and PCM solvation. They used up to four explicit water molecules forming hydrogen bonds with the OH group of phenol and O⁻ of phenolate along with the PCM solvation model to calculate pK_a's in aqueous medium. Compared to the case when the pK_a's are calculated with PCM only, including one and two explicit waters improved the calculated pK_a's significantly. The authors proposed a model to calculate pK_a's of substituted phenols that include an empirical correction factor of −1.4 in addition to two explicit water molecules and PCM solvation. They also reported that with more than three explicit water molecules, the calculated pK_a's had errors larger than the values calculated with PCM solvation with no explicit water. However, the study did not offer any further explanation about such unexpectedly large deviations (mean deviation >5) in calculated pK_a's with more than three explicit water molecules.

In the present study, we have calculated the pK_a's of alcohols, phenols, and hydroperoxides in aqueous solution from first principles. In previous publications,[49−51] we have shown that the systematic error in calculated free energies attributable to the implicit solvation model can be significantly reduced by incorporating a small number of explicit solvent molecules. For nucleobases and a wide range of substituted thiols and selenols, we have used this protocol to directly calculate the pK_a's within 1 pK_a unit without the need for linear fits or thermodynamic

cycles. These studies suggest that the pK_a's of compounds without experimentally measured pK_a's can be predicted accurately via a careful selection of DFT methods and appropriate number of explicit water molecules without fitting any linear equations or thermodynamic cycles. In this study, we aim to extend this protocol to compute the pK_a's of chemically interesting, and biologically as well as pharmaceutically important alcohols, phenols, and hydroperoxides. As in our previous studies, the free energy of deprotonation is calculated by a direct approach using the SMD solvation model with and without explicit water molecules. We believe that this study provides a practical and systematic way to directly calculate the pK_a's of a wide range of alcohols, phenols, and hydroperoxides, and to explore reaction mechanisms involving protonation, deprotonation, and proton coupled electron transfer.

■ COMPUTATIONAL METHODS

Calculations were performed with the development version of the Gaussian series of programs.[52] A total of 72 organic compounds were used in this study: 17 organic alcohols and 25 phenols (Scheme 1), all with experimentally measured pK_a's, and 30 hydroperoxides (Scheme 2), 6 with experimentally measured pK_a values. The compounds were selected on the basis of their practical relevance, structural variance, availability of accurate experimental pK_a's, and the range of the pK_a's. On the basis of a more extensive set of methods tested for the calculation of pK_a's for thiols,[51] we selected three hybrid density functionals (B3LYP,[53−56] ωB97XD,[57] and M06-2X[58]), two basis sets (6-31+G(d,p) and 6-311++G(d,p))[59−63] and the SMD[43] implicit solvation model. Geometries were optimized in aqueous solution. Harmonic frequencies were calculated to confirm the structures were minima on the potential energy surface and to obtain thermal and entropic contributions to the free energies. The effect of explicit waters on the calculated pK_a's was assessed by including one and three explicit water molecules directly hydrogen-bonded to the site being protonated/deprotonated. Typical arrangements of water near the protonation/deprotonation site are shown in Figure 1a−c. For the alcohols, phenols, and hydroperoxides with one explicit water, the OH group provides a H atom to form a hydrogen bond with the O of the explicit water molecule (R−OH---OH$_2$ or R−OOH---OH$_2$) whereas in the deprotonated form the O$^-$ forms a hydrogen bond with the H atom of the explicitly placed water molecule (e.g., R−O$^-$---H−OH). When there are three explicit water molecules near the OH group, the O of one of the three water molecules accepts H from the OH group of R−OH and the other two donate H to form three hydrogen bonds (i.e., one R−OH---OH$_2$ and two R−O$^-$---H−OH hydrogen bonds). In the deprotonated form, R−O$^-$ forms a hydrogen bond with one hydrogen atom from each of the three explicit water molecules (R−O$^-$---H−OH). The arrangements of explicit water molecules near R−OH and R−O$^-$ were kept as consistent as possible (i.e., similar conformations and the same number of hydrogen bonds) to avoid any bias that may result in unsystematic contributions to the calculated free energy (e.g., avoiding extraneous hydrogen bonds with neighboring groups on the molecule). Conformers with the energy difference within 1 kcal/mol are found to have pK_a's within about 0.5 units. For the molecule possessing more than one stable conformer, the one with the lowest energy was used in the final calculations.

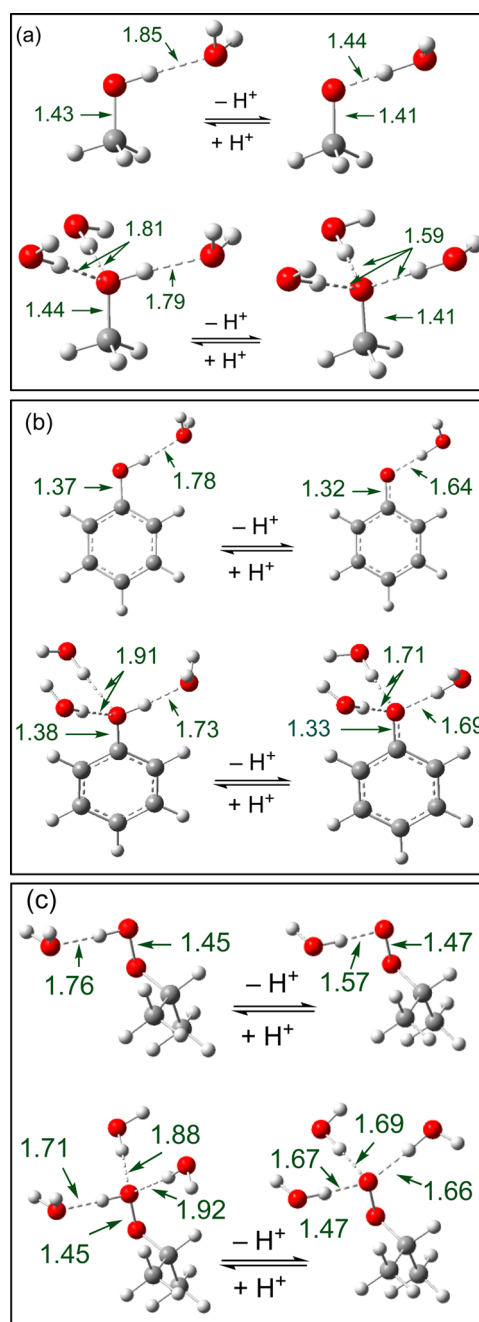The pK_a for a reaction ROH ⇌ RO$^-$ + H$^+$ is given by

Figure 1. Arrangement of explicit water molecules near the OH and O$^-$ groups in (a) alcohols, (b) phenols, and (c) hydroperoxides. Some of the key bond lengths are shown in green (in Å).

$$pK_a = \frac{\Delta G^*_{aq}}{2.303RT} = \frac{G^*_{aq,RO^-} + G^*_{aq,H^+} - G^*_{aq,ROH}}{2.303RT}$$

where $G^*_{aq,ROH}$ and G^*_{aq,RO^-} are the standard free energy of protonated and deprotonated species in the aqueous medium, respectively. The free energy of a proton in aqueous solution is calculated as

$$G^*_{aq,H^+} = G^\circ_{g,H^+} + \Delta G^*_{H^+,aq} + \Delta G^{1atm \rightarrow 1M}$$

where, $G^\circ_{g,H^+} = H^\circ_{g,H^+} - TS^\circ_{g,H^+}$ is the gas-phase free energy of the proton at 298.15 K obtained by using $H^\circ_{g,H^+} = 5RT/2 = 1.48$ kcal/mol and $S^\circ_{g,H^+} = 26.05$ cal/(mol·K). $\Delta G^*_{aq,H^+} = -265.9$ kcal/mol is the aqueous-phase solvation free energy of the proton,

Table 1. Averages and Standard Deviation of Errors in Calculated pK_a's for Alcohols, Phenols, and Hydroperoxides

		$\Delta pK_a = pK_{a,calculated} - pK_{a,experiment}$		
functional	basis set	SMD	SMD + 1 water	SMD + 3 waters
B3LYP	6-31+G(d,p)	6.62 ± 2.41	3.68 ± 1.05	−0.21 ± 0.49
	6-311++G(d,p)	6.43 ± 2.41	3.65 ± 1.27	−0.02 ± 0.55
ωB97XD	6-31+G(d,p)	8.12 ± 2.33	5.45 ± 1.16	0.96 ± 0.55
	6-311++G(d,p)	7.85 ± 2.38	5.39 ± 1.23	0.95 ± 0.72
M06-2X	6-31+G(d,p)	6.46 ± 2.32	3.23 ± 0.72	−1.19 ± 0.53
	6-311++G(d,p)	6.23 ± 2.55	3.12 ± 0.98	−1.40 ± 0.63

taken from the literature.[20,42,64−66] $\Delta G^{1atm \rightarrow 1M} = RT \log(24.4)$ = 1.89 kcal/mol is a correction term for the change in standard state of 1 atm to 1 mol/L. The symbols * and ° denote the standard state of 1 mol/L and 1 atm, respectively.

RESULT AND DISCUSSION

The pK_a's of 17 alcohols, 25 phenols, and 30 hydroperoxides have been investigated using three hybrid DFT functionals (B3LYP, ωB97XD, and M06-2X) and two basis sets (6-31+G(d,p) and 6-311++G(d,p)) with the SMD solvation model. The experimentally known pK_a's of organic alcohols and phenols ranges from 5 to 16. The experimental data for the hydroperoxides are rather meager. Standard tables[15−18] list pK_a's for only 6 of the 30 hydroperoxides considered in this study. The effect of hydrogen-bonded water molecules on the calculated pK_a's has been assessed by including one and three explicit water molecules coordinated to the R−OH/R−O⁻ group.

Table 1 summarizes the calculated pK_a errors for the alcohols, phenols, and hydroperoxides using various density functionals with SMD solvation with zero, one, and three explicit water molecules. One water hydrogen bonded to ROH/RO⁻ accounts for the largest fraction of the effect of explicit solvent; three waters complete the first solvation shell around the site. The error in the calculated pK_a's using SMD solvation with no explicit water molecules is rather large. For B3LYP and M06-2X, the mean unsigned error (MUE) is 6.4−6.6 pK_a units whereas the MUE is even larger for ωB97XD ($\Delta pK_a \sim 8$). The magnitudes of these errors are similar to errors calculated for thiols reported in our previous study.[51] Using a larger basis set has only a small effect on the calculated pK_a's (improves the pK_a's by only about 0.1−0.2 pK_a units). When one explicit water is included along with SMD solvation, the MUE becomes smaller by about 3 pK_a units. For B3LYP and M06-2X, the inclusion of one water brings the calculated error close to 3 pK_a units whereas the error for ωB97XD is still higher than 5 pK_a units. Even though there is substantial improvement in the calculated pK_a's, the error is still significantly higher than 1 pK_a unit. The improvement in the computed pK_a's is greater when three explicit water molecules are included. For B3LYP/6-31+G(d,p), the mean signed error (MSE) is calculated to be −0.21 pK_a units with a standard deviation (SD) of 0.49 (MUE = 0.44 ± 0.49). B3LYP with the larger 6-311++G(d,p) basis set brings the MSE to −0.02 with SD = 0.55 (MUE = 0.45 ± 0.55). It is interesting to see that although ωB97XD was quite good for predicting pK_a's of thiols (MSE = −0.11 ± 0.50) and selenols (MSE = 0.36 ± 0.24),[50,51] it is found to have a MSE of 0.96 ± 0.55 pK_a units with three explicit water molecules and 6-31+G(d,p) basis set. With three explicit waters along with SMD solvation, M06-2X is found to slightly overstabilize the anionic form of alcohols, phenols and hydroperoxides (MSE = −1.19 ± 0.53). However, the

calculated MSE using the M06-2X/6-31+G(d,p) is significantly better than for thiols (MSE = −3.59 ± 0.45) and selenols (MSE = −3.21 ± 1.04). It is very gratifying to see that, with three explicit waters and SMD solvation, the mean unsigned error (MUE) for all the three functionals and two basis sets used in this study ranges from 0.4 to 1.4 pK_a units (MSE from −1.4 to +1.0 pK_a units), indicating that an accuracy of ±1.5 pK_a units could be achieved by using any of the functionals and basis sets used in this study. The use of larger basis sets with density functional methods or the use of compound methods such as CBS and Gn for calculating energy differences may improve the of pK_a calculations, but the present protocol already achieves a satisfactory level of accuracy and is inexpensive and broadly applicable.

Figure 2 shows the linear correlation between the experimental pK_a's and the values calculated by B3LYP/6-
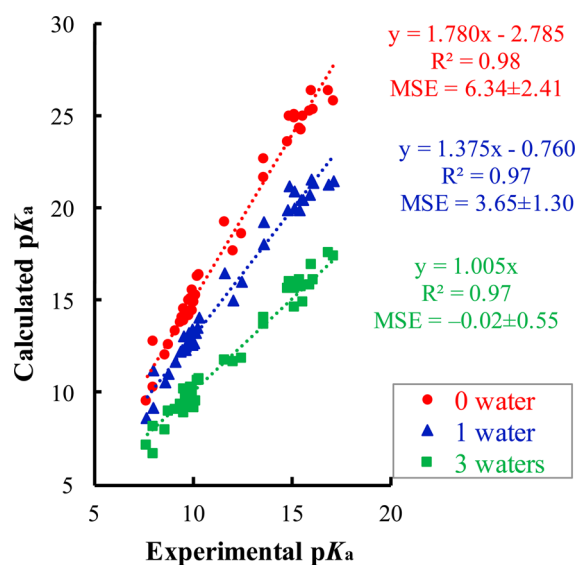

Figure 2. Linear correlations of experimental pK_a's and calculated pK_a's using the B3LYP/6-311++G(d,p) level of theory with SMD solvation model with no explicit water (red circles), one explicit water (blue triangle), and three explicit waters (green square).

311++G(d,p) for the alcohols, phenols, and 6 of the peroxides. Similar correlation diagrams for other functionals are shown in Figure S1 in the Supporting Information. As shown in Figure 2, there is a good correlation between the experimental and computed pK_a's (with and without explicit water molecules) with the linear regression coefficients (R^2) of 0.97. The linear fits for zero, one, and three explicit water molecules have slopes of 1.78, 1.37, and 1.00, respectively. For SMD without explicit waters, the slope of the linear fit line is significantly larger than the theoretical value of 1. Including one explicit water improves the slope by 0.4. Including three explicit waters improves the

Table 2. Experimental pK_a's[a] and Calculated pK_a's Using the B3LYP/6-311++G(d,p) Level of Theory for Substituted Alcohols and Phenols

SN	compound	pK_a (exp)[a]	0 water		1 water		3 waters	
			pK_a	ΔpK_a	pK_a	ΔpK_a	pK_a	ΔpK_a
1	CH$_3$OH	15.54	24.98	9.44	20.42	4.88	14.88	−0.66
2	CH$_3$CH$_2$OH	15.90	25.25	9.35	20.75	4.85	15.84	−0.06
3	CH$_3$CH$_2$CH$_2$OH	16.10	25.38	9.28	21.35	5.25	16.06	−0.04
4	(CH$_3$)$_2$−CH−OH	17.10	25.77	8.67	21.46	4.36	17.40	0.30
5	(CH$_3$)$_3$−C−OH	16.00	26.36	10.36	21.57	5.57	16.91	0.91
6	CH$_2$CHCH$_2$OH	15.52	24.27	8.75	20.40	4.88	15.77	0.25
7	HOCH$_2$CH$_2$ CH$_2$ OH	15.10	24.86	9.76	19.97	4.87	15.66	0.56
8	CH$_2$(OH)CH$_2$CH$_2$CH$_2$OH	15.10	25.09	9.99	20.91	5.81	14.62	−0.48
9	CH2(OH)CH$_2$CH(OH)CH$_3$	14.90	24.94	10.04	21.21	6.31	15.98	1.08
10	CH$_3$OCH$_2$CH$_2$OH	14.80	23.58	8.78	19.93	5.13	15.63	0.83
11	HCCCH$_2$OH	13.55	21.61	8.06	18.07	4.52	13.67	0.12
12	CF$_3$CH$_2$OH	12.43	18.58	6.15	15.99	3.56	11.85	−0.58
13	CCl$_3$CH$_2$OH	12.02	17.63	5.61	14.96	2.94	11.61	−0.41
14	CF$_3$C(CH$_3$)$_2$OH	11.60	19.27	7.67	16.49	4.89	11.76	0.16
15	C$_6$H$_5$−CH$_2$OH	15.40	24.31	8.91	19.87	4.47	16.07	0.67
16	C$_6$H$_{11}$OH	16.84	26.37	9.53	21.25	4.41	17.57	0.73
17	C$_6$H$_5$−CH(OH)−C$_6$H$_5$	13.54	22.68	9.14	19.25	5.71	14.03	0.49
	Phenols							
18	H−	9.98	14.43	4.45	12.57	2.59	9.77	−0.21
19	m-NH$_2$−	9.87	15.11	5.24	13.02	3.15	10.31	0.44
20	p-NH$_2$−	10.30	16.37	6.07	14.05	3.75	10.73	0.43
21	m-CH$_3$−	10.08	14.87	4.79	12.71	2.63	9.20	−0.88
22	p-CH$_3$−	10.14	15.23	5.09	13.21	3.07	9.51	−0.63
23	m-HO−	9.44	14.09	4.65	12.30	2.86	9.05	−0.39
24	p-HO−	9.96	15.57	5.61	13.44	3.48	10.17	0.21
25	m-Cl−	8.78	12.55	3.77	11.03	2.25	9.02	0.24
26	p-Cl−	9.14	13.32	4.18	11.63	2.49	9.10	−0.04
27	m-CH$_3$O−	9.65	14.38	4.73	12.52	2.87	9.45	−0.20
28	p-CH$_3$O−	10.20	16.25	6.05	13.54	3.34	10.69	0.49
29	m-CHO−	8.00	12.80	4.80	11.19	3.19	8.16	0.16
30	p-CHO−	7.66	9.50	1.84	8.57	0.91	7.15	−0.51
31	m-CH$_3$S−	9.53	13.96	4.43	12.29	2.76	9.36	−0.17
32	p-CH$_3$S−	9.53	14.50	4.97	13.03	3.50	10.21	0.68
33	m-CN−	8.57	11.99	3.42	10.53	1.96	7.97	−0.60
34	p-CN−	7.95	10.25	2.30	9.17	1.22	6.64	−1.31
35	m-HOCH$_2$−	9.82	14.18	4.36	12.80	2.98	9.96	0.14
36	p-HOCH$_2$−	9.83	14.98	5.15	13.26	3.43	9.42	−0.41
37	m-C$_6$H$_5$−	9.59	14.29	4.70	12.32	2.73	9.73	0.14
38	p-C$_6$H$_5$−	9.51	13.83	4.32	12.34	2.83	8.85	−0.66
39	m-COO−	9.94	14.98	5.04	12.78	2.84	9.75	−0.19
40	p-COO−	9.39	13.75	4.36	12.20	2.81	9.31	−0.08
41	m-CH$_3$CO−	9.19	13.63	4.44	11.47	2.28	8.51	−0.68
42	p-CH$_3$CO−	8.05	10.46	2.41	9.63	1.58	7.28	−0.77
	mean signed error (MSE)			6.34		3.65		−0.02
	mean unsigned error (MUE)			6.34		3.65		0.45
	standard deviation (STDEV)			2.41		1.27		0.55

[a]References 15−18.

slope by an additional 0.4, for a total of 0.8 compared to the slope for implicit-only. Zhang also observed similar improvements in the slope of the linear fit when one and two explicit waters were included.[32] The improvement in the MUE is proportional to the improvement in the slope.

A list of experimental pK_a's along with the calculated values at the B3LYP/6-311++G(d,p) level of theory using SMD implicit solvation model with and without explicit water molecules are presented in Table 2. The pK_a's calculated with only the implicit solvation model have large errors (MSE = 6.34

± 2.41). The errors become smaller when one explicit water is included in the calculation (MSE = 3.65 ± 1.27). When three explicit water molecules are included, the calculated pK_a's are very close to the experimental values with MSE = −0.02 ± 0.55 (MSE = 0.23 ± 0.54 for the alcohols and −0.19 ± 0.49 for phenols). The scatter in the calculated pK_a's also improves significantly from SD = 2.41 for SMD solvation only to SD = 0.55 for SMD solvation with three explicit water molecules. The largest positive and negative errors are +1.08 and −1.31, respectively.

Table 3 shows a list of estimated pK_a values for various organic hydroperoxides with B3LYP/6-311++G(d,p) and

Table 3. Experimental pK_a's[a] and Calculated pK_a's of Some Substituted Organic Hydroperoxides with the B3LYP/6-311++G(d,p) and ωB97XD/6-311++G(d,p) Levels of Theory and SMD Solvation with Three Explicit Waters

SN	pK_a (exp)[a]	B3LYP/6-311++G(d,p)		ωB97XD/6311+G(d,p)	
		pK_a	ΔpK_a	pK_a	ΔpK_a
1	11.6	10.29	−1.31	12.16	0.56
2	11.5	11.22	−0.28	11.88	0.38
3	11.8	11.34	−0.46	12.22	0.42
4	11.8	10.60	−1.20	12.85	1.05
5	12.8	12.56	−0.24	12.93	0.13
6	12.8	12.22	−0.58	13.05	0.25
7		11.11		12.58	
8		10.54		11.99	
9		6.53		8.53	
10		8.42		10.37	
11		6.38		8.74	
12		7.20		9.74	
13		11.41		12.96	
14		11.46		12.09	
15		11.51		12.45	
16		11.57		13.24	
17		10.41		12.50	
18		11.37		11.26	
19		8.89		9.92	
20		8.34		9.68	
21		11.64		11.73	
22		12.59		13.92	
23		11.92		12.17	
24		9.09		10.05	
25		11.36		12.25	
26		9.44		10.43	
27		8.09		9.84	
28		6.47		7.32	
29		11.66		13.31	
30		7.91		9.44	
mean signed error (MSE)			−0.68		0.46
mean unsigned error (MUE)			0.68		0.46
standard deviation (STDEV)			0.46		0.32

[a]Reference 18.

ωB97XD/6-311++G(d,p) with SMD solvation and three explicit water molecules. The results for other methods with zero, one, and three explicit waters in SMD solvation are given in the Supporting Information. For the 6 hydroperoxides with experimentally measured pK_a's, the errors in the values calculated with SMD solvation and no water molecule are larger than 5 pK_a units. Including one explicit water reduces the error by 2−2.5 pK_a units. As in the case of alcohols and phenols, including three explicit waters further reduces the error by about 3 pK_a units. It would be desirable to have more experimental data to see if this holds for a boarder, more representative test set. In keeping with the pK_a's calculated for alcohols and phenols, the values calculated with B3LYP/6-311++G(d,p) using SMD and three explicit waters have slightly smaller errors (MUE = 0.68) than those calculated with

B3LYP/6-31+G(d,p) (MUE = 1.04). With ωB97XD and SMD plus three explicit waters, the calculated MSE are 0.41 and 0.46 pK_a units for the 6-31+G(d,p) and 6-311+G(d,p) basis sets, respectively. The pK_a's estimated with SMD and three explicit waters are found to be a little smaller than the experimental values with B3LYP/6-311++G(d,p) (MSE = −0.68) and a bit larger with ωB97XD/6-311++G(d,p) (MSE = 0.46).

■ SUMMARY

This study presents a reliable protocol for calculating the pK_a's of organic alcohols, phenols, and hydroperoxides in aqueous solution. A total of 72 compounds were calculated, including 17 alcohols and 25 phenols with experimentally known pK_a's and 30 hydroperoxides (6 with experimentally known pK_a's). Three hybrid DFT functionals (B3LYP, ωB97XD, and M06-2X) with two different basis sets (6-31+G(d,p) and 6-311++G(d,p)) and SMD implicit solvation with zero, one, and three explicit waters were used to calculate the pK_a's. The pK_a's of the alcohols, phenols, and peroxides calculated with only implicit solvation have large errors (greater than 6 pK_a units) and have a large slope and intercept in the experimental vs calculated linear correlations. Including one explicit water forming a hydrogen bond reduces the error by about 3 pK_a units and improves the slope and intercept; however, the calculated values are still far from chemical accuracy (MSE > 3 pK_a units). Three explicit water molecules with SMD solvation improves the calculated pK_a's by an average of 7 pK_a units compared to the SMD only results. With three explicit waters, the B3LYP functional performs the best. The MSE for the B3LYP/6-31+G(d,p) method is only −0.21 ± 0.49, which is improved slightly by using a larger basis set (MSE = −0.02 ± 0.55 for 6-311+ +G(d,p)). This study suggests that B3LYP with a reasonable basis set and the SMD solvation model with three explicit waters can be used to calculate reliable pK_a's for alcohols, phenols, and peroxides. The mean signed error in the pK_a's calculated with the ωB97XD functional is 1.0 pK_a units whereas for M06-2X the MSE is −1.2 to −1.4 units. It is very rewarding to see that with three explicit waters and SMD solvation, the error for all the three functionals and two basis sets used in this study ranges from 0.4 to 1.4 pK_a units for MUE and −1.4 to +1.0 pK_a units for MSE. This indicates that an accuracy of ±1.5 pK_a units can be achieved with any of the functionals and basis sets used in this study. However, care must be taken in placing the explicit water molecules to avoid any unsystematic energy contributions. We believe that this study provides a well-defined and reliable method for predicting the pK_a's of alcohols, phenols, and hydroperoxides with good accuracy and low computational cost.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.7b03907.

Linear correlation plots of calculated and experimental pK_a's and coordinates for the optimized geometries of protonated and deprotonated alcohols, phenols, and hydroperoxides with three explicit water molecules (PDF)

Absolute free energies of the various compounds and details of the pK_a calculations (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*H. B. Schlegel. E-mail: hbs@chem.wayne.edu. Phone: 313-577-2562.

ORCID ⦿

H. Bernhard Schlegel: 0000-0001-7114-2821

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Rappoport, Z. *The chemistry of phenols*; John Wiley & Sons: Chichester, England, 2003.

(2) Zheng, W.; Wang, S. Y. Antioxidant activity and phenolic compounds in selected herbs. *J. Agric. Food Chem.* **2001**, *49*, 5165−5170.

(3) Pantelidis, G.; Vasilakakis, M.; Manganaris, G.; Diamantidis, G. Antioxidant capacity, phenol, anthocyanin and ascorbic acid contents in raspberries, blackberries, red currants, gooseberries and cornelian cherries. *Food Chem.* **2007**, *102*, 777−783.

(4) Szmant, H. H. *Organic building blocks of the chemical industry*; John Wiley & Sons: New York, 1989.

(5) Monick, J. A. *Alcohols; their chemistry, properties and manufacture*; Reinhold Book Corporation: New York, 1968.

(6) Calvert, J. G.; Lazrus, A.; Kok, G. L.; Heikes, B. G.; Walega, J. G.; Lind, J.; Cantrell, C. A. Chemical mechanisms of acid generation in the troposphere. *Nature* **1985**, *317*, 27−35.

(7) Ravetta, F.; Jacob, D. J.; Brune, W.; Heikes, B.; Anderson, B.; Blake, D.; Gregory, G.; Sachse, G.; Sandholm, S.; Shetter, R. Experimental evidence for the importance of convected methylhydroperoxide as a source of hydrogen oxide (HOx) radicals in the tropical upper troposphere. *J. Geophys. Res. Atmos.* **2001**, *106*, 32709−32716.

(8) Hallquist, M.; Wenger, J.; Baltensperger, U.; Rudich, Y.; Simpson, D.; Claeys, M.; Dommen, J.; Donahue, N.; George, C.; Goldstein, A. The formation, properties and impact of secondary organic aerosol: Current and emerging issues. *Atmos. Chem. Phys.* **2009**, *9*, 5155−5236.

(9) Chen, Q.; Liu, Y.; Donahue, N. M.; Shilling, J. E.; Martin, S. T. Particle-phase chemistry of secondary organic material: Modeled compared to measured O:C and H:C elemental ratios provide constraints. *Environ. Sci. Technol.* **2011**, *45*, 4763−4770.

(10) Lee, M.; Heikes, B. G.; Jacob, D. J.; Sachse, G.; Anderson, B. Hydrogen peroxide, organic hydroperoxide, and formaldehyde as primary pollutants from biomass burning. *J. Geophys. Res. Atmos.* **1997**, *102*, 1301−1309.

(11) Miyamoto, S.; Ronsein, G. E.; Prado, F. M.; Uemi, M.; Corrêa, T. C.; Toma, I. N.; Bertolucci, A.; Oliveira, M. C. B.; Motta, F. D.; Medeiros, M. H. G.; Mascio, P. D. Biological hydroperoxides and singlet molecular oxygen generation. *IUBMB Life* **2007**, *59*, 322−331.

(12) Ullrich, V.; Schatz, M.; Meisch, F. Physiological and toxicological roles of hydroperoxides. In *Biological reactive intermediates IV: Molecular and cellular effects and their impact on human health*; Witmer, C. M., Snyder, R. R., Jollow, D. J., Kalf, G. F., Kocsis, J. J., Sipes, I. G., Eds.; Springer: New York: Boston, 1991; pp 79−86.

(13) Sies, H. Hydroperoxides and thiol oxidants in the study of oxidative stress in intact cells and organs. In *Oxidative stress*; Sies, H., Ed.; Academic Press: Boca Raton, Florida, USA, 1985; pp 73−90.

(14) Bell, R. P. *The Proton in Chemistry*; Chapman and Hall: London, U.K., 1973.

(15) Dobos, D. *Electrochemical Data*; Elsevier Co.: Amsterdam, Oxford, New York, 1975.

(16) Lide, D. R. *CRC Handbook of Chemistry and Physics*, 82 ed.; CRC Press: New York, 2001.

(17) Serjeant, E. P.; Dempsey, B. *Ionisation Constants of Organic Acids in Aqueous Solution*; Pergamon: New York, 1979; Vol. 23.

(18) Williams, R. http://www.chem.wisc.edu/areas/reich/pkatable/pKa_compilation-1-williams.Pdf (accessed May 25, 2017).

(19) Pliego, J. R.; Riveros, J. M. Theoretical calculation of pKa using the cluster−continuum model. *J. Phys. Chem. A* **2002**, *106*, 7434−7439.

(20) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. 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* **2006**, *110*, 16066−16081.

(21) Cunningham, I. D.; Bhaila, K.; Povey, D. C. Computational calculation of absolute aqueous pKa values for phenols. *Comput. Theor. Chem.* **2013**, *1019*, 55−60.

(22) Eckert, F.; Diedenhofen, M.; Klamt, A. Towards a first principles prediction of pKa: COSMO-RS and the cluster-continuum approach. *Mol. Phys.* **2010**, *108*, 229−241.

(23) Zhang, S. A reliable and efficient first principles - based method for predicting pKa values. III. Adding explicit water molecules: Can the theoretical slope be reproduced and pKa values predicted more accurately? *J. Comput. Chem.* **2012**, *33*, 517−526.

(24) Adam, K. R. New density functional and atoms in molecules method of computing relative pKa values in solution. *J. Phys. Chem. A* **2002**, *106*, 11963−11972.

(25) Klamt, A.; Eckert, F.; Diedenhofen, M.; Beck, M. E. First principles calculations of aqueous pKa values for organic and inorganic acids using COSMO−RS reveal an inconsistency in the slope of the pKa scale. *J. Phys. Chem. A* **2003**, *107*, 9380−9386.

(26) Schmidt am Busch, M.; Knapp, E. W. Accurate pKa determination for a heterogeneous group of organic molecules. *ChemPhysChem* **2004**, *5*, 1513−1522.

(27) Chipman, D. M. Computation of pKa from dielectric continuum theory. *J. Phys. Chem. A* **2002**, *106*, 7413−7422.

(28) Zhang, S.; Baker, J.; Pulay, P. A reliable and efficient first principles-based method for predicting pKa values. 1. Methodology. *J. Phys. Chem. A* **2010**, *114*, 425−431.

(29) Klicić, J. J.; Friesner, R. A.; Liu, S.-Y.; Guida, W. C. Accurate prediction of acidity constants in aqueous solution via density functional theory and self-consistent reaction field methods. *J. Phys. Chem. A* **2002**, *106*, 1327−1335.

(30) Zeng, Y.; Chen, X.; Zhao, D.; Li, H.; Zhang, Y.; Xiao, X. Estimation of pKa values for carboxylic acids, alcohols, phenols and amines using changes in the relative gibbs free energy. *Fluid Phase Equilib.* **2012**, *313*, 148−155.

(31) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Adding explicit solvent molecules to continuum solvent calculations for the calculation of aqueous acid dissociation constants. *J. Phys. Chem. A* **2006**, *110*, 2493−2499.

(32) Zhang, S. A reliable and efficient first principles-based method for predicting pKa values. 4. Organic bases. *J. Comput. Chem.* **2012**, *33*, 2469−2482.

(33) Muckerman, J. T.; Skone, J. H.; Ning, M.; Wasada-Tsutsui, Y. Toward the accurate calculation of pKa values in water and acetonitrile. *Biochim. Biophys. Acta, Bioenerg.* **2013**, *1827*, 882−91.

(34) Liptak, M. D.; Gross, K. C.; Seybold, P. G.; Feldgus, S.; Shields, G. C. Absolute pKa determinations for substituted phenols. *J. Am. Chem. Soc.* **2002**, *124*, 6421−6427.

(35) Seybold, P. G.; Kreye, W. C. Theoretical estimation of the acidities of alcohols and azoles in gas phase, dmso, and water. *Int. J. Quantum Chem.* **2012**, *112*, 3769−3776.

(36) Reijenga, J.; Van Hoof, A.; Van Loon, A.; Teunissen, B. Development of methods for the determination of pKa values. *Anal. Chem. Insights* **2013**, *8*, 53.

(37) Seybold, P. G.; Shields, G. C. Computational estimation of pKa values. *WIREs Comput. Mol. Sci.* **2015**, *5*, 290−297.

(38) Alongi, K. S.; Shields, G. C. Theoretical calculations of acid dissociation constants: A review article. *Annu. Rep. Comput. Chem.* **2010**, *6*, 113−138.

(39) Ho, J. Predicting pKa in implicit solvents: Current status and future directions. *Aust. J. Chem.* **2014**, *67*, 1441−1460.

(40) Shields, G. C.; Seybold, P. G. *Computational approaches for the prediction of pKa values*; CRC Press: Boca Raton, London, New York, 2013.

(41) Ratkova, E. L.; Palmer, D. S.; Fedorov, M. V. Solvation thermodynamics of organic molecules by the molecular integral equation theory: Approaching chemical accuracy. *Chem. Rev.* **2015**, *115*, 6312−6356.

(42) Marenich, A. V.; Ho, J.; Coote, M. L.; Cramer, C. J.; Truhlar, D. G. Computational electrochemistry: Prediction of liquid-phase reduction potentials. *Phys. Chem. Chem. Phys.* **2014**, *16*, 15068−15106.

(43) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. 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* **2009**, *113*, 6378−6396.

(44) Liptak, M. D.; Shields, G. C. Accurate pKa calculations for carboxylic acids using complete basis set and Gaussian-n models combined with CPCM continuum solvation methods. *J. Am. Chem. Soc.* **2001**, *123*, 7314−7319.

(45) Liptak, M. D.; Shields, G. C. Experimentation with different thermodynamic cycles used for pKa calculations on carboxylic acids using complete basis set and Gaussian - n models combined with CPCM continuum solvation methods. *Int. J. Quantum Chem.* **2001**, *85*, 727−741.

(46) Ho, J. Are thermodynamic cycles necessary for continuum solvent calculation of pKas and reduction potentials? *Phys. Chem. Chem. Phys.* **2015**, *17*, 2859−2868.

(47) Ho, J.; Coote, M. L. A universal approach for continuum solvent pKa calculations: Are we there yet? *Theor. Chem. Acc.* **2010**, *125*, 3−21.

(48) Ho, J.; Ertem, M. Z. Calculating free energy changes in continuum solvation models. *J. Phys. Chem. B* **2016**, *120*, 1319−1329.

(49) Thapa, B.; Schlegel, H. B. Calculations of pKa's and redox potentials of nucleobases with explicit waters and polarizable continuum solvation. *J. Phys. Chem. A* **2015**, *119*, 5134−44.

(50) Thapa, B.; Schlegel, H. B. Theoretical calculation of pKa's of selenols in aqueous solution using an implicit solvation model and explicit water molecules. *J. Phys. Chem. A* **2016**, *120*, 8916−8922.

(51) Thapa, B.; Schlegel, H. B. Density functional theory calculation of pKa's of thiols in aqueous solution using explicit water molecules and the polarizable continuum model. *J. Phys. Chem. A* **2016**, *120*, 5726−35.

(52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian development version*: Revision I.06; Gaussian, Inc.: Wallingford, CT, 2016.

(53) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098−3100.

(54) Becke, A. D. Density - functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(55) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.

(56) Stephens, P.; Devlin, F.; Chabalowski, C.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

(57) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom−atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.

(58) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.

(59) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function - augmented basis sets for anion calculations. III. The 3 - 21+G basis set for first - row elements, Li-F. *J. Comput. Chem.* **1983**, *4*, 294−301.

(60) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self - consistent molecular - orbital methods. IX. An extended Gaussian-type basis for molecular - orbital studies of organic molecules. *J. Chem. Phys.* **1971**, *54*, 724−728.

(61) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self - consistent molecular orbital methods. XXIII. A polarization - type basis set for second - row elements. *J. Chem. Phys.* **1982**, *77*, 3654−3665.

(62) Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theoret. Chim. Acta* **1973**, *28*, 213−222.

(63) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **1972**, *56*, 2257−2261.

(64) Camaioni, D. M.; Schwerdtfeger, C. A. Comment on "Accurate experimental values for the free energies of hydration of H⁺, OH⁻, and H₃O⁺". *J. Phys. Chem. A* **2005**, *109*, 10795−10797.

(65) Isse, A. A.; Gennaro, A. Absolute potential of the standard hydrogen electrode and the problem of interconversion of potentials in different solvents. *J. Phys. Chem. B* **2010**, *114*, 7894−7899.

(66) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. The proton's absolute aqueous enthalpy and gibbs free energy of solvation from cluster-ion solvation data. *J. Phys. Chem. A* **1998**, *102*, 7787−7794.