THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

Cite This: J. Phys. Chem. A 2018, 122, 4366−4374

pubs.acs.org/JPCA

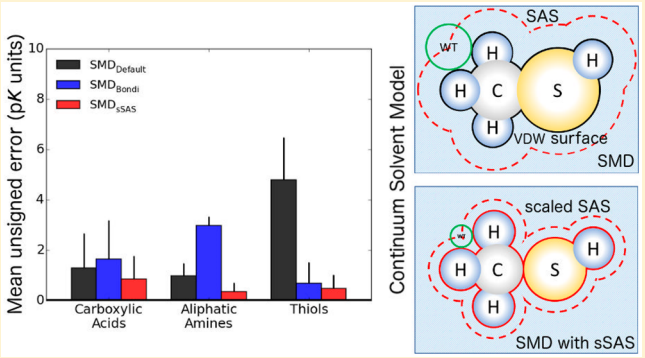
Quantum Chemical Calculation of pK_as of Environmentally Relevant Functional Groups: Carboxylic Acids, Amines, and Thiols in Aqueous Solution

Peng Lian, Ryne C. Johnston, Jerry M. Parks, and Jeremy C. Smith*

UT/ORNL Center for Molecular Biophysics, Biosciences Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, Tennessee 37831-6309, United States

Supporting Information

ABSTRACT: Developing accurate quantum chemical approaches for calculating pK_as is of broad interest. Useful accuracy can be obtained by using density functional theory (DFT) in combination with a polarizable continuum solvent model. However, some classes of molecules present problems for this approach, yielding errors greater than 5 pK units. Various methods have been developed to improve the accuracy of the combined strategy. These methods perform well but either do not generalize or introduce additional degrees of freedom, increasing the computational cost. The Solvation Model based on Density (SMD) has emerged as one of the most commonly used continuum solvent models. Nevertheless, for some classes of organic compounds, e.g., thiols, the pK_as calculated with the original SMD model show errors of 6−10 pK units, and we traced these errors to inaccuracies in the solvation free energies of the anions. To improve the accuracy of pK_as calculated with DFT and the SMD model, we developed a scaled solvent-accessible surface approach for constructing the solute−solvent boundary. By using a "direct" approach, in which all quantities are computed in the presence of the continuum solvent, the use of thermodynamic cycles is avoided. Furthermore, no explicit water molecules are required. Three benchmark data sets of experimentally measured pK_a values, including 28 carboxylic acids, 10 aliphatic amines, and 45 thiols, were used to assess the optimized SMD model, which we call SMD with a scaled solvent-accessible surface (SMD$_{sSAS}$). Of the methods tested, the M06-2X density functional approximation, 6-31+G(d,p) basis set, and SMD$_{sSAS}$ solvent model provided the most accurate pK_as for each set, yielding mean unsigned errors of 0.9, 0.4, and 0.5 pK units, respectively, for carboxylic acids, aliphatic amines, and thiols. This approach is therefore useful for efficiently calculating the pK_as of environmentally relevant functional groups.



INTRODUCTION

Acid dissociation constants (pK_as) are essential for evaluating the acid−base properties of organic compounds in chemistry,[1−3] drug discovery,[4−6] and environmental[7−9] and material sciences.[10−12] Furthermore, pK_as and stability constants (log β) of metal complexes are important for understanding and predicting metal ion speciation[13,14] in the environment, because metal ions and protons "compete" for metal-binding functional groups, such as carboxylic acids, amines, and thiols. Direct measurement of pK_a values with conventional experimental approaches for each functional group of all organic compounds in complex aqueous systems can be challenging.[15] Therefore, the accurate calculation of pK_as with theoretical methods is of great interest.

A widely used method for calculating pK_as involves performing quantum chemical calculations, most often density functional theory (DFT), with a polarizable continuum solvation model. Introducing the continuum approximation for the solvent, a solute is embedded in a solute-shaped cavity surrounded by a polarizable dielectric medium. Continuum models are less computationally demanding than models including explicit solvent molecules. In many continuum solvation models, the total electrostatic potential satisfies the nonhomogeneous Poisson equation (NPE) for electrostatics. For example, the polarizable continuum model (PCM)[16−18] and its variants, e.g., IEF-PCM,[19] SCI-PCM,[20] I-PCM,[20] and C-PCM,[21,22] the conductor-like screening model (COSMO),[23] and COSMO-RS,[24] all solve the NPE based on the continuous charge density of the solute.

As an electron-density-based solvation model, the Solvation Model based on Density (SMD)[25] separates the solvation free energy into two components, the bulk electrostatic contribution and the cavity-dispersion-solvent-structure (CDS) term. The first component is calculated from a molecular orbital self-consistent reaction field calculation, which requires an iterative solution of the NPE to obtain the reaction field. The second

Received: February 20, 2018
Revised: April 6, 2018
Published: April 10, 2018

DOI: 10.1021/acs.jpca.8b01751
J. Phys. Chem. A 2018, 122, 4366−4374

component describes shorter-range polarization and non-electrostatic effects such as cavitation, dispersion, and local solvent-structure changes. For the SMD model, a set of parameters including intrinsic atomic Coulomb radii and atomic surface tension coefficients was introduced and optimized against 2821 experimentally measured solvation data using 6 electronic structure methods. SMD is a universal solvation model, as it was designed to be applied to any charged or uncharged solute in any solvent or liquid medium.[25] Therefore, SMD is widely used in the calculation of condensed-phase properties including, but not limited to, solvation free energies and pK_as.[26−29]

Many strategies have been developed to calculate pK_as with quantum chemical approaches. One of the most common methods involves constructing an appropriate thermodynamic cycle to calculate the aqueous Gibbs free energy of deprotonation. In this approach, the geometries of the acid and conjugate base are optimized in the gas phase, and then, solvation free energies are computed at those geometries with a continuum model. Various formulations of thermodynamic cycles and their advantages and disadvantages have been summarized in recent reviews.[2,3,30] Another increasingly common method is the "direct" approach, in which the geometries are optimized, and deprotonation free energies are calculated directly in the presence of the continuum solvent without the need for gas-phase calculations. A series of studies comparing thermodynamic-cycle-based methods and the direct approach showed that the two approaches can achieve similar accuracy.[28,29] However, for systems that display significant differences between gas-phase and solution-phase geometries, the direct approach may perform better than the corresponding thermodynamic-cycle method.

The combination of DFT and a continuum solvation model is widely used in pK_a calculations. This method performs well for many species, as DFT is capable of calculating accurate electronic structures of many classes of molecules. However, for some molecules, e.g., thiols, the use of continuum solvation results in errors of 6−10 pK units.[30−33] One way to improve the accuracy of pK_a calculations involves applying corrections based on linear regression. This method assumes that for molecules in the same class, there is a linear relationship between the pK_a and some property of the molecule, such as the atomic charge of the anionic form[34] or the deprotonation energy.[35,36] By applying linear regression fitting on the training data set, which contains the experimental pK_a values and the calculated numerical property values, the parameters for the linear model can be obtained. Then, the pK_a for a molecule in the same class can be predicted with the linear mode and the calculated numerical property. The accuracy of this method is good, with errors below 0.5 pK units for some finely tuned models. However, separate linear fits are required for each functional group. Thus, it is difficult to apply this method to sets that include multiple classes of molecules.

An alternative strategy for improving the accuracy of calculated pK_as involves the use of a reference molecule with an experimentally known pK_a.[30,31,37,38] Instead of calculating the absolute free energy of the deprotonation reaction, a proton exchange reaction is constructed with the reference molecule. Then, the difference in pK_a relative to the computed value for the reference acid is calculated. The final pK_a is obtained by adding the pK_a difference to the experimentally measured pK_a value of the reference acid. The accuracy of this approach derives from systematic error cancellation originating primarily from the interaction between the solutes and the continuum solvent.[3]

Including first-shell water molecules explicitly is a further way to improve the accuracy of pK_a calculations. Using this approach, it was shown that including two to four explicit water molecules as hydrogen bond donors or acceptors to dicarboxylic acids reduced the mean unsigned error (MUE) by 2.4−9.0 pK units compared to continuum solvation alone, yielding final MUE values of 0.6−0.8 pK units.[39] Later, it was shown that, for a set of thiols, alcohols, and hydroperoxides, compared to a fully implicit solvation model, for which the MUE > 6 pK units, including one explicit water molecule reduced the MUE of pK_a calculations by ∼3 pK units.[33] By including three explicit water molecules, further improvements were obtained, yielding a MUE less than 1 pK unit.[32,33] However, this approach introduces additional degrees of freedom, requiring substantial effort and surveillance to place the water molecules optimally and identify low-energy minima.[40]

Another effective way to optimize a solvation model is to scale the solvent cavity or the individual radii of the solute atoms.[41−48] Because continuum solvation models have limitations in describing the short-range interactions between the solute and first-shell solvent molecules, the solute−solvent surface and the cavity size are critical for overcoming this deficiency in pK_a calculations. As shown in a recent assessment of continuum methods for calculating pK_as of several chemical groups, the solute cavity is the dominant factor for reproducing experimental values.[49] Substantial efforts have been made in the last decades to reduce errors in computed solvation free energies and pK_as by scaling the solvent cavity or the atomic radii.[41−48,50]

In the present work, we used DFT together with the SMD continuum solvent model to develop an accurate and efficient approach for calculating the pK_as of molecules that include environmentally relevant functional groups, carboxylic acids, aliphatic amines, and thiols. The choice of these functional groups was based on the above-mentioned existing inaccuracy for thiols and the central role of these groups in aqueous environmental science. Carboxylic acids, aliphatic amines, and thiols are three important metal-binding functional groups of organic molecules present in the environment,[7−9] forming strong complexes with many different metal cations, including Zn^{2+}, Cu^{2+}, Cd^{2+}, and Hg^{2+}.[50,51] Acid dissociation constants are intimately related to the metal-binding affinities of these functional groups and are therefore important for investigating the transport, speciation, and toxicity of metals in aquatic systems.

Based on the current implementation of SMD in Gaussian 09,[52] we developed an alternative approach to construct the solute cavity. All free energies were computed directly in the continuum solvent, thus eliminating the need for gas-phase calculations and thermodynamic cycles. Moreover, no explicit water molecules were used, which simplifies the calculation substantially compared to mixed discrete-continuum approaches. Three sets of molecules with experimentally measured pK_a values (28 carboxylic acids, 10 aliphatic amines, and 45 thiols) were used to assess the performance of the optimized SMD model. MUEs of 0.86, 0.35, and 0.48 pK units were obtained for carboxylic acids, aliphatic amines, and thiols, respectively. Moreover, we tested the optimized SMD model by comparing calculated solvation free energies to experimentally measured reference values for 83 monoanions, 60 monocations,

DOI: 10.1021/acs.jpca.8b01751
J. Phys. Chem. A 2018, 122, 4366−4374

and 256 neutral molecules available in the Minnesota Solvation Database (MNSol-v2012).[53]

METHODS

The data sets used in this study are listed in Table 1. Details of the molecules in each data set are provided in Scheme S1. Initial coordinates of the molecules were obtained from the references listed in Table 1.

Table 1. Benchmark Data Sets Used in This Study

molecule	number of molecules	refs.
carboxylic acids	28	54−56
aliphatic amines	10	57,58
thiols	45	59−70

Using the direct approach, given the proton dissociation reaction shown in eq 1, the pK_a of molecule AH was calculated according to eq 2. The free energy difference in the 1 M standard state, ΔG^*_{aq}, was calculated directly from the aqueous Gibbs free energies of the acid and conjugate base with eq 3. The gas-phase standard state correction terms, $\Delta G^{\circ\rightarrow *}$ for both the acid and conjugate base, cancel in this equation. The standard state aqueous free energy of a proton, $G^*_{aq}(H^+)$, was calculated according to eq 4, in which the gas-phase free energy ($G^\circ_g(H^+) = -6.29$ kcal/mol) and the experimentally measured hydration free energy ($G_{aq,solv}(H^+) = -265.9$ kcal/mol) were taken from the literature.[32,71−74]

$$AH(aq) \xrightarrow{\Delta G^*_{aq}} A^-(aq) + H^+(aq) \quad (1)$$

$$pK_a = \frac{\Delta G^*_{aq}}{2.303RT} \quad (2)$$

$$\Delta G^*_{aq} = G^*_{aq}(A^-) + G^*_{aq}(H^+) - G^*_{aq}(AH) \quad (3)$$

$$G^*_{aq}(H^+) = G^\circ_g(H^+) + \Delta G_{aq,solv}(H^+) + \Delta G^{\circ\rightarrow *} \quad (4)$$

$$\Delta G^*_{Solv} = \Delta G_{EP} + G_{CDS} + \Delta G^{\circ\rightarrow *} \quad (5)$$

The SMD model calculates the free energy of solvation in three parts (eq 5): the bulk electrostatic energy, ΔG_{ENP}, the cavitation-dispersion-solvent-structure (CDS) contribution, G_{CDS}, and the gas-phase standard state correction, $\Delta G^{\circ\rightarrow *}$. ΔG_{ENP} includes electronic (E), nuclear (N), and polarization (P) components of the solvation free energy and is evaluated by self-consistently performing numerical integration of the NPE coupled with the QM electron density of the solute molecule. G_{CDS} includes solvent cavitation (C), the dispersion (D) energy, and changes in the local solvent structure (S). $\Delta G^{\circ\rightarrow *}$ is used to convert from the 1 atm ideal gas standard state to the 1 M standard state, where the superscripts $^\circ$ and * indicate the 1 atm and 1 M standard states, respectively. $\Delta G^{\circ\rightarrow *} = RT\ln 24.46 = 1.89$ kcal/mol at 298 K.[25]

An optimized description of the solute−solvent boundary is a key aspect of all continuum solvation models. For the default SMD model (denoted as $SMD_{Default}$), intrinsic Coulomb radii were used to create the cavity and calculate the bulk electrostatic contribution.[25] Also, the SMD model with Bondi radii,[75] denoted here as SMD_{Bondi}, has shown good performance in the solvation free energy calculations of small organic compounds.[47,76] Therefore, the SMD_{Bondi} method was investigated for comparison. To optimize the solute−solvent boundary and the cavity used in the SMD continuum model, we used the scaled solvent-accessible surface (sSAS) to construct the cavity (denoted as SMD_{sSAS}). For SMD_{sSAS}, the surface type and the scaling factor options in the SCRF section were tuned simultaneously. By choosing SAS as the solute−solvent boundary, the solvent radius (1.385 Å for water[19]) was added to the intrinsic Coulomb radii to construct the cavity. The scaling factor α was introduced to tune the size of the SAS cavity. For the CDS contributions, as in the default SMD model, Bondi's values for the van der Waals radii plus 0.4 Å were used to calculate the solvent-accessible surface area (SASA) of the individual atoms of the solute.[25]

Gaussian 09 revision E.01[52] was used to perform all electronic structure calculations. All geometries were optimized at the M06-2X/6-31+G(d,p) level of theory with differently tuned SMD models. Input file samples for SMD variants are provided in the Supporting Information. Vibrational frequencies were computed in each case to confirm that each structure was a local minimum on the potential energy surface and to compute thermodynamic quantities. The standard rigid-rotor harmonic approximation, as implemented in Gaussian 09, calculated on the solution-phase potential energy surface (PES), was applied to the calculation of the thermal correction terms.

RESULTS AND DISCUSSION

Performance of the Default SMD Approach. Three electronic structure methods were used in the original development of the SMD model: M05-2X, B3LYP,[77−79] and Hartree−Fock.[25] Here, we used B3LYP and M06-2X, the latter of which is an updated version of M05-2X. We initially evaluated the accuracy of the default SMD model ($SMD_{Default}$). The mean unsigned error (MUE) of the calculated pK_as for carboxylic acids, aliphatic amines, and thiols are 1.4, 0.7, and 7.1 pK units for B3LYP compared to 1.3, 1.0, and 4.8 pK units for M06-2X (Table 2). The calculated pK_a values for all molecules are listed in Table S1.

The MUEs for both B3LYP and M06-2X on the carboxylic acid set are <2 pK units, and the maximum error (MaxE), the mean signed error (MSE), and standard deviation (SD) have similar magnitudes. For the aliphatic amines, both functionals exhibit similar performance, with MSEs, MUEs, and SDs less

Table 2. Performance of the Standard SMD Model with Two Hybrid Density Functionals for Calculating pK_a of Carboxylic Acids, Aliphatic Amines, and Thiols

	B3LYP				M06-2X			
molecule	MaxE	MUE	MSE	SD	MaxE	MUE	MSE	SD
carboxylic acids	−2.99	1.38	0.82	1.35	−2.98	1.29	0.61	1.37
aliphatic amines	1.95	0.68	0.48	0.76	−1.52	0.99	−0.98	0.50
thiols	9.27	7.12	7.12	1.80	7.14	4.80	4.80	1.69

DOI: 10.1021/acs.jpca.8b01751
J. Phys. Chem. A 2018, 122, 4366−4374

than 1 pK unit and MaxE > 1.5 pK unit. However, for the thiols, B3LYP with the default SMD model yields a MaxE of 9.3 pK units, MSE (and MUE) of 7.1 pK units, and an SD of 1.8 pK units, while with M06-2X, the MaxE is 7.1 pK units, the MSE and MUE are both 4.8 pK units, and the SD is 1.7 pK units. Although M06-2X performs better than B3LYP with the default SMD model, the error of ∼5 pK units remains unacceptable for useful pK_a estimation. Thus, we chose to use M06-2X in all subsequent work aimed at improving the accuracy of the SMD model.

Optimization of the SMD$_{sSAS}$ Model. In a continuum solvation model, the cavity formed by the solute−solvent boundary is one of the most important factors for the accuracy of the model. The essence of the SMD$_{sSAS}$ approach is to improve the accuracy of the SMD continuum model by optimizing this solute−solvent boundary. The SAS is a smooth surface that depends on both the geometry of the solute and the effective radius of the solvent molecule. However, using the SAS directly as the solute−solvent boundary is improper, because the SAS is constructed by combining the atomic radii of the solute and the solvent radius. With no scaling, the cavity defined by the SAS is too large to describe short-range solvent−solute interactions. For example, for the SMD model in the aqueous phase, the SAS is constructed by using a set of radii that combines the intrinsic Coulomb radii and the effective radius of a water molecule (1.385 Å). These combined radii are significantly larger than the intrinsic Coulomb radii and the widely used Bondi radii. Therefore, we introduced a scaling factor, α, into the SMD$_{sSAS}$ model to tune the size of the cavity. To determine the optimal value of α, we carried out a systematic optimization by minimizing the MUE with respect to α for each of the three data sets (Figure 1). In each case, the

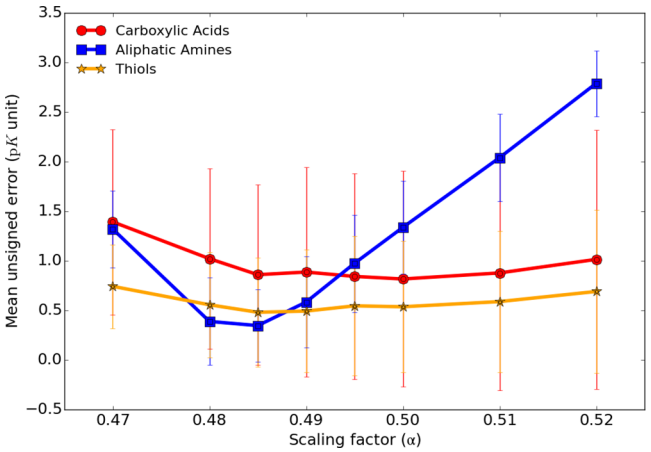


Figure 1. MUE as a function of the scaling factor α calculated at the M06-2X/6-31+G(d,p) level of theory with the SMD$_{sSAS}$ solvent model.

MUE displayed a roughly quadratic relationship with α. Carboxylic acids and thiols are both relatively insensitive to the scaling factor, as the curves are broad and shallow. This behavior agrees with a previous study indicating a relatively weak dependence on the solute cavity for these functional groups.[49] However, the amines display a deeper well, with a minimal MUE value of 0.35 pK units when α = 0.485. An additional test performed on a larger data set consisting of 44 amines with 54 pK_as[80] showed that SMD$_{sSAS}$ performs similarly well (Figure S1), with the MUE reaching a minimal value of 0.4 pK units when α = 0.480 and 0.5 pK units when α = 0.485. However, as the confidence intervals overlap at both points, we concluded that α = 0.485 gives the best combined performance across the three data sets, and this single value was thus applied to all 83 compounds.

In Gaussian 09, the GEPOL[81−84] surface-building algorithm is used to discretize the solute−solvent boundary surface for the calculation of the reaction field of the solute, which is required to solve the NPE. The values of the radii used by GEPOL are critical for the accuracy of the SMD model. Gaussian 09 provides a scaling factor, α, to tune the radii used by GEPOL. In the default implementation, scaling factors of 1.0 and 1.1 are used for SMD$_{Default}$ and SMD$_{Bondi}$, respectively, with GEPOL (Table 3). In the SMD$_{sSAS}$ approach, an atomic radius is defined as the sum of the intrinsic Coulomb radius and the solvent radius, multiplied by the factor α = 0.485. A radius of 1.385 Å is used for a water molecule in Gaussian 09, which makes the unscaled SAS radii about twice as large as the corresponding Bondi radius. With α = 0.485, the scaled SAS radii used by GEPOL for SMD$_{sSAS}$ are smaller than those used by either the SMD default model or SMD$_{Bondi}$, which suggests that stronger short-range interactions between the solute and solvent may be required for an accurate description of these species in aqueous solution. The differences in these radii are illustrated in Figure S2. Although the scaled SAS radii are smaller than the intrinsic Coulomb radii, they are similar to Bondi radii. For atoms such as H, N, F, S, and Cl, the scaled SAS radii are intermediate between Bondi and intrinsic Coulomb radii. Thus, although the scaled SAS radii are relatively small, they are still in a physically reasonable range.

Performance of the SMD$_{sSAS}$ Model. The performance of different SMD variants on the benchmark data sets in Table 1 was assessed by comparing the MaxE, MSE, MUE, and SD of the calculated pK_as with the experimental values (Figure 2).

The carboxylic acid data set contains experimentally measured pK_a values for 28 carboxylic acids. The SMD variants perform similarly for this data set, with all MUEs being less than 2 pK units (Figure 2). SMD$_{Bondi}$ yields a MUE of 1.7 pK units, which is slightly worse than the MUE of 1.3 pK units for SMD$_{Default}$. SMD$_{sSAS}$ performs the best, reducing the MUE to 0.9 pK units.

For the amines, differing performance was observed for the methods tested. The default SMD model yields a MUE of 1.0 pK units. SMD$_{Bondi}$ performs the worst, as its MUE is ∼3 pK units versus less than 1 for the other methods. Compared to the default SMD model, SMD$_{Bondi}$ changes the atomic radii from intrinsic Coulomb radii to smaller values. The poor performance of SMD$_{Bondi}$ indicates that simply decreasing the radii is not enough to improve upon the default SMD model for the amines. SMD$_{sSAS}$, which not only decreases the radii but also replaces the cavity surface with a scaled SAS, shows excellent agreement with the experimentally measured pK_as, yielding a MUE of 0.4 pK units.

For the data set of 45 thiols, SMD$_{Default}$ shows the worst performance, producing a MUE of ∼5 pK units. For SMD$_{Bondi}$, the performance improves significantly. The MUE for SMD$_{Bondi}$ drops to less than 1 pK unit, primarily because of the smaller sulfur radius used by the GEPOL surface-building algorithm (Table 3). These findings agree with previous work, which showed that the inadequate consideration of short-range interactions between the solute and the solvent in the SMD model results in inaccurate thiol pK_as.[32] In general, including first-shell water molecules explicitly in a mixed discrete-continuum approach results in superior calculated pK_as. For

DOI: 10.1021/acs.jpca.8b01751
J. Phys. Chem. A 2018, 122, 4366−4374

Table 3. Parameters Used for Different SMD Variants

		$SMD_{Default}$			SMD_{Bondi}			SMD_{sSAS}			
element	Z	intrinsic Coulomb radius	α	GEPOL	Bondi radius	α	GEPOL	solvent radius[a]	sSAS radius[b]	α	GEPOL
H	1	**1.20**	1.0	1.20	**1.20**	1.1	1.32	1.385	2.585	0.485	**1.254**
C	6	**1.85**	1.0	1.85	**1.70**	1.1	1.87	1.385	3.235	0.485	**1.569**
N	7	**1.89**	1.0	1.89	**1.55**	1.1	1.71	1.385	3.275	0.485	**1.588**
O	8	**1.52**	1.0	1.52	**1.52**	1.1	1.67	1.385	2.905	0.485	**1.409**
F	9	**1.73**	1.0	1.73	**1.47**	1.1	1.62	1.385	3.115	0.485	**1.511**
S	16	**2.49**	1.0	2.49	**1.80**	1.1	1.98	1.385	3.875	0.485	**1.879**
Cl	17	**2.38**	1.0	2.38	**1.75**	1.1	1.93	1.385	3.765	0.485	**1.826**

[a]The default radius for water was used to construct the SAS surface. [b]Sum of the intrinsic Coulomb radius and the solvent radius for each atom.

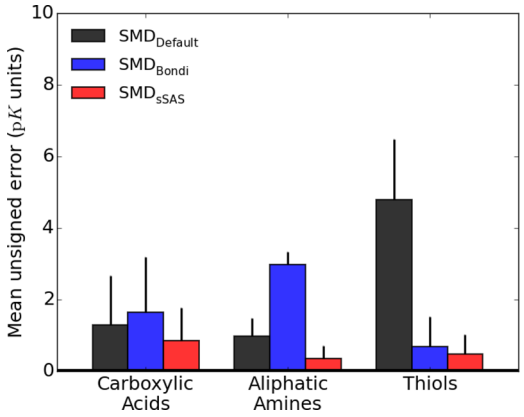


Figure 2. Mean unsigned errors and standard deviations for pK_as of carboxylic acids, aliphatic amines, and thiols calculated with different SMD variants. The M06-2X functional and 6-31+G(d,p) basis set were used in all calculations.

the thiols, as the electron density of the sulfur atom is polarizable and diffuse, smaller radii such as Bondi and scaled SAS radii reduce the size of the solvent cavity around sulfur and enhance the interaction between the thiols and the continuum solvent, thus improving the accuracy of the calculations. SMD_{sSAS} yields a MUE < 0.5 pK units for the thiols, which is the best of all the SMD variants. The good performance of SMD_{sSAS} is mainly the combined result of not only reducing the size of the cavity but at the same time constructing a solvent-accessible surface as the solute−solvent boundary instead of using the van der Waals surface directly.

Tables S2−S4 list the errors in pK_as calculated with different SMD variants for all individual molecules in the data set of carboxylic acids, aliphatic amines, and thiols, respectively. The chemical structures in XYZ format are given in the Supporting Information. SMD_{Bondi} performs better than $SMD_{Default}$ for thiols, with a MUE of 0.7 pK units compared to 4.8 pK units for $SMD_{Default}$ and a MaxE of −2.1 pK units versus 7.1 pK units. For the carboxylic acids and aliphatic amines, SMD_{Bondi} is worse than $SDM_{Default}$. SMD_{sSAS} exhibits the best performance on all three data sets with a MUE of 0.9 pK units for carboxylic acids, 0.4 pK units for aliphatic amines, and 0.5 pK units for the thiols. In terms of the maximum error, SMD_{sSAS} is also the best with MaxE of −2.3, −0.5, and −1.6 pK units for carboxylic acids, aliphatic amines, and thiols, respectively. SMD_{sSAS} yields the smallest MSE for the carboxylic acids (0.5 pK units) and aliphatic amines (−0.1) but is slightly larger (0.3) than SMD_{Bondi} (−0.1) for the thiols. Using Bondi radii reduces the error for thiols but not for the carboxylic acids or aliphatic amines. Even worse, SMD_{Bondi} destabilizes the energies of ammonium cations, as it systematically underestimates the pK_as of amines by ∼3 pK units (Table S3). The balanced performance of SMD_{sSAS} on these three data sets demonstrates the generalizability of this approach to different organic functional groups.

The pK_as calculated with different SMD variants show good linear correlations with the corresponding experimental values, with R^2 values of 0.95 for all variants (Figure 3). However, for

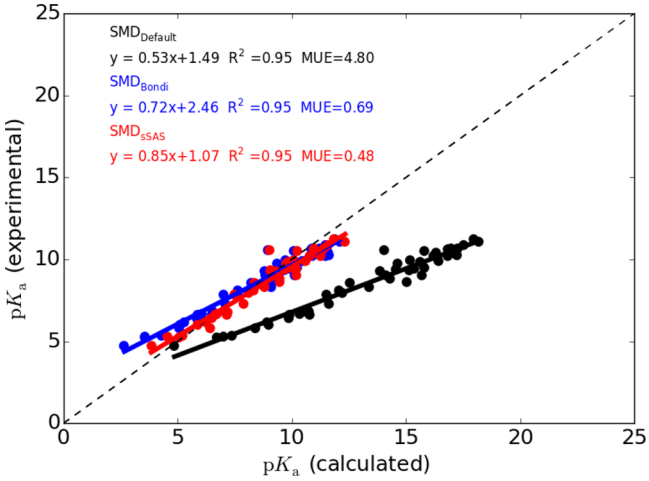


Figure 3. Comparison of thiol pK_as calculated with different SMD variants and corresponding experimental values. Experimental data were obtained from refs 59−70.

$SMD_{Default}$, the slopes are around 0.5, which is far from ideal (i.e., m = 1.0). This behavior implies systematic error in the method, and the error increases for molecules with high pK_as. Using Bondi radii to generate the solute−solvent boundary improved the performance of the SMD model substantially, as the slope for SMD_{Bondi} increased to 0.72. The better performance of SMD_{Bondi} than $SMD_{Default}$ is underscored by the MUE values (Figure 2). However, for molecules with low pK_as, SMD_{Bondi} gives larger errors, and the fitted line begins to deviate from the diagonal. SMD_{sSAS} further improves the performance of the SMD model for the thiols compared to SMD_{Bondi}, increasing the slope to 0.85 and reducing the MUE to 0.48. The accuracy obtained with SMD_{sSAS} is comparable to the explicit-water approach obtained for thiols by Thapa and Schlegel,[32] which achieved a MUE of 0.5 pK units on the same data set. However, SMD_{sSAS} is computationally less demanding, because no explicit water molecules are required. The success of the SMD_{sSAS} approach in correcting systematic errors in pK_a calculations of thiols indicates the feasibility of accounting for

DOI: 10.1021/acs.jpca.8b01751
J. Phys. Chem. A 2018, 122, 4366−4374

short-range solute−solvent interactions by optimizing the solute cavity.

For carboxylic acids, all SMD variants show similar performance (Figure S3), with SMD_{sSAS} performing slightly better than $SMD_{Default}$ and SMD_{Bondi}. Compared to SMD_{Bondi}, the slope for SMD_{sSAS} is increased from 0.43 to 0.57, and the MUE is reduced from 1.65 to 0.86 pK units. The accuracy of SMD_{sSAS} is better than that obtained with a thermodynamic-cycle-based method using acetic acid as reference molecule and calculated at the PCM/B3LYP/6-31+G(d,p) level of theory, which yielded a MUE of 1.3 pK units.[56] For aliphatic amines, SMD_{sSAS} performs the best with a MUE of 0.35 pK units (Figure S4). In contrast, SMD_{Bondi} introduces large errors for the amines, increasing the MUE from 0.35 to 2.98 pK units compared to SMD_{sSAS}. The accuracy of SMD_{sSAS} is comparable to that calculated using a thermodynamic cycle together with the B3LYP functional, a Poisson−Boltzmann continuum solvent model, which yielded MUEs of 0.24 pK units with aug-cc-pVTZ(-f) and 0.28 pK units with 6-31++G(d,p).[58] Although the performance improvements obtained with SMD_{sSAS} for the carboxylic acids and aliphatic amines are not as significant as for the thiols, this approach is still the most accurate of the methods considered here.

Improvement of the Solvation Free Energies of the Anions by SMD_{sSAS}. To identify potential sources of error in the pK_a calculations, we compared the calculated solvation free energies to 408 experimentally well-characterized neutral and monoionic molecules in the MNSol-v2012 database[53] (Tables 4 and S5). $SMD_{Default}$ exhibits the best performance for calculating the solvation free energies of 256 neutral molecules with MUE < 1 kcal/mol and a MaxE of −4.1 kcal/mol. SMD_{Bondi} is similar to $SMD_{Default}$ on the neutral molecules. SMD_{sSAS} yields a MUE of 1.7 kcal/mol, MaxE of −7.5 kcal/mol, and MSE of −1.5 kcal/mol, indicating that this approach underestimates the solvation free energies of neutral molecules by ∼1.5 kcal/mol on average. For the cation data set, SMD_{sSAS} is slightly better than $SMD_{Default}$, with a MUE of 3.3 kcal/mol and MSE of −0.4 kcal/mol compared 3.4 and −2.7 kcal/mol, respectively, for $SMD_{Default}$. SMD_{Bondi} yields the largest errors for the cations with MUE, MaxE, and MSE values of 6.5, 22.3, and −6.4 kcal/mol, respectively. Anions are the most challenging data set for the SMD models. All three models underestimate the solvation free energies of the anions. $SMD_{Default}$ produces a MUE of 6.8 kcal/mol, MaxE of 16.1 kcal/mol, and MSE of −6.7 kcal/mol on the 83 anions, which is much worse than its performance on cations and neutral molecules. SMD_{Bondi} is inaccurate for the anions as the MaxE, MUE, and MSE are 17.6, 8.4, and −8.3 kcal/mol, respectively. However, compared to the other two models, SMD_{sSAS} has the highest accuracy for the solvation free energies of the anions. The MaxE, MUE, and MSE with SMD_{sSAS} on the anions are 12.3, 4.6, and −4.0 kcal/mol, respectively. Compared to $SMD_{Default}$, SMD_{sSAS} improves the accuracy of the solvation free energies of the anions by ∼33% in terms of MUE, while maintaining essentially the same accuracy for cations and yielding slightly lower accuracy for neutral molecules.

Table 4. Errors in the Solvation Free Energies (kcal/mol) of Neutral and Monoionic Molecules in the MNSol-v2012 Database Calculated with Different SMD Variants

SMD variant	data set	no. of compounds	MaxE	MUE	MSE	SD
$SMD_{Default}$	anions	83	16.1	6.8	−6.7	3.8
SMD_{Bondi}			17.6	8.4	−8.3	4.4
SMD_{sSAS}			12.3	4.6	−4.0	3.6
$SMD_{Default}$	cations	60	13.9	3.4	−2.7	3.9
SMD_{Bondi}			22.3	6.5	−6.4	4.8
SMD_{sSAS}			16.2	3.3	−0.4	4.8
$SMD_{Default}$	neutral molecules	256	−4.1	0.8	−0.5	0.9
SMD_{Bondi}			4.2	1.0	0.4	1.1
SMD_{sSAS}			−7.5	1.7	−1.5	1.6

To investigate further the sources of error in pK_a calculations, we compared the performance of the three SMD variants on a selected set of representative neutral and ionic molecules (Figure 4). All the solvation free energy calculation data are provided in Table S5. The default SMD model performs accurately for many neutral molecules, e.g., CH_3COOH, NH_3, CH_3NH_2, CH_3SH, Ph-SH, and H_2S, but is inaccurate for anions, e.g., CH_3S^-, Ph-S^-, SH^-, OH^-, F^-, and Cl^-. Notably, $SMD_{Default}$ yields an absolute error >9 kcal/mol for the sulfur-containing anions. Therefore, the poor performance of $SMD_{Default}$ on the solvation free energy calculation of sulfur-containing anions is the main reason for its inaccuracy on thiol pK_as. Using Bondi radii improves the accuracy for thiolates but results in larger errors for other ions such as CH_3COO^-, Ph−COO^-, OH^-, F^-, NH_4^+, and $CH_3NH_3^+$. Moreover, for carboxylic acids, such as CH_3COOH, SMD_{Bondi} yields a positive error, whereas the other methods tend to underestimate the solvation free energies. Overall, the poor behavior of SMD_{Bondi} for solvation free energies explains its poor performance on the carboxylic acid and amine data sets. On the contrary, SMD_{sSAS} improves the accuracy for both anions and cations. Although it introduces some negative errors for the neutral molecules, e.g., CH_3COOH, NH_3, H_2S, and H_2O, the absolute errors are <3 kcal/mol. In terms of the accuracy of pK_a calculations, the solvation free energies of both the acid and conjugate base are equally important. Therefore, analyzing the performance of SMD_{sSAS} on all three solvation free energy data sets reveals that it achieves good performance for pK_a calculations by improving the accuracy of the solvation free energies of the anions but also benefits from error cancellation due to a systematic underestimation of solvation free energies.

CONCLUSIONS

The default SMD model achieves accurate pK_a values in many cases[26−29] but produces large errors for thiols (Figure 2).[32,34] Here, we have developed a simple modification of the default SMD model, SMD_{sSAS}, which greatly improves the accuracy of pK_a calculations with DFT and the SMD continuum solvation model, without requiring the inclusion of explicit water molecules. The approach was optimized for functional groups of particular interest in environmental chemistry: carboxylic acids, amines, and thiols. Using the M06-2X density functional, 6-31+G(d,p) basis set, and the SMD_{sSAS} model, which adds only a single scaling parameter, agreement with experiment to within well under 1.0 pK unit was obtained for all three classes of molecule. By comparing the calculated solvation free energies to experimentally well-characterized neutral and monoionic molecules in the MNSol-v2012 database, we have shown that SMD_{sSAS} increases the accuracy of computed pK_a values primarily by improving the accuracy of the solvation free energies of the anions. For the three functional groups investigated here, which are among the most dominant species in the environment and are also important biologically, the

DOI: 10.1021/acs.jpca.8b01751
J. Phys. Chem. A 2018, 122, 4366−4374

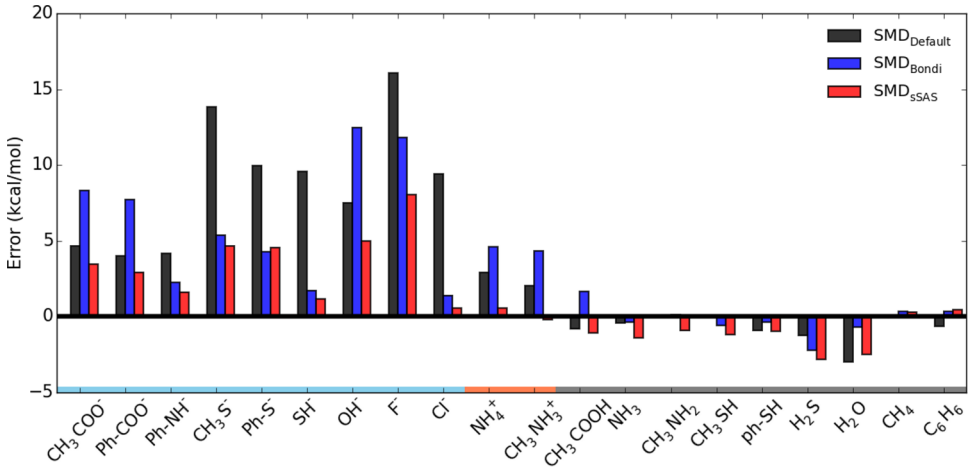


Figure 4. Absolute error in the solvation free energies of experimentally well-characterized small molecules and ions calculated with different SMD variants.

approach performs quite well and should be of wide applicability.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.8b01751.

Structures, coordinates, experimentally measured and calculated pK_a values of the molecules. Performance of the scaling factor, schematic diagram of different radii, and sample input files for SMD variants. Experimentally measured and calculated solvation free energies for the molecules in the MNSol-v2012 database (PDF)

XYZ files of structures used in this study (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: smithjc@ornl.gov (J.C.S.)

ORCID

Peng Lian: 0000-0003-1701-5025

Ryne C. Johnston: 0000-0002-6606-9401

Jerry M. Parks: 0000-0002-3103-9333

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research, Subsurface Biogeochemical Research Program through award DE-SC0016478 and through the Mercury Science Focus Area Program (SFA) at Oak Ridge National Laboratory (ORNL). ORNL is managed by UT-Battelle, LLC for the U.S. DOE under contract DE-AC05- 00OR22725. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy.

■ REFERENCES

(1) Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99* (8), 2161−2200.

(2) Alongi, K. S.; Shields, G. C. Theoretical Calculations of Acid Dissociation Constants: A Review Article. *Annu. Rep. Comput. Chem.* **2010**, *6*, 113−138.

(3) Casasnovas, R.; Ortega-Castro, J.; Frau, J.; Donoso, J.; Muñoz, F. Theoretical pK_a Calculations with Continuum Model Solvents, Alternative Protocols to Thermodynamic Cycles. *Int. J. Quantum Chem.* **2014**, *114* (20), 1350−1363.

(4) Lanevskij, K.; Japertas, P.; Didziapetris, R.; Petrauskas, A. Ionization-Specific Prediction of Blood−Brain Permeability. *J. Pharm. Sci.* **2009**, *98* (1), 122−134.

(5) Rajapakse, H. A.; Nantermet, P. G.; Selnick, H. G.; Barrow, J. C.; McGaughey, G. B.; Munshi, S.; Lindsley, S. R.; Young, M. B.; Ngo, P. L.; Katherine Holloway, M.; et al. SAR of Tertiary Carbinamine Derived BACE1 Inhibitors: Role of Aspartate Ligand Amine pK_a in Enzyme Inhibition. *Bioorg. Med. Chem. Lett.* **2010**, *20* (6), 1885−1889.

(6) Zhang, J.; Liu, Z.; Lian, P.; Qian, J.; Li, X.; Wang, L.; Fu, W.; Chen, L.; Wei, X.; Li, C. Selective Imaging and Cancer Cell Death via pH-Switchable Near-Infrared Fluorescence and Photothermal Effects. *Chem. Sci.* **2016**, *7* (9), 5995−6005.

(7) Aiken, G. R.; Hsu-Kim, H.; Ryan, J. N. Influence of Dissolved Organic Matter on the Environmental Fate of Metals, Nanoparticles, and Colloids. *Environ. Sci. Technol.* **2011**, *45* (8), 3196−3201.

(8) Weng, L.; Temminghoff, E. J. M.; Lofts, S.; Tipping, E.; Van Riemsdijk, W. H. Complexation with Dissolved Organic Matter and Solubility Control of Heavy Metals in a Sandy Soil. *Environ. Sci. Technol.* **2002**, *36* (22), 4804−4810.

(9) Rao, B.; Simpson, C.; Lin, H.; Liang, L.; Gu, B. Determination of Thiol Functional Groups on Bacteria and Natural Organic Matter in Environmental Systems. *Talanta* **2014**, *119*, 240−247.

(10) Cheng, J.; Sprik, M. Acidity of the Aqueous Rutile TiO_2(110) Surface from Density Functional Theory Based Molecular Dynamics. *J. Chem. Theory Comput.* **2010**, *6* (3), 880−889.

(11) Bryantsev, V. S. Predicting the Stability of Aprotic Solvents in Li-Air Batteries: pK_a Calculations of Aliphatic C−H Acids in Dimethyl Sulfoxide. *Chem. Phys. Lett.* **2013**, *558*, 42−47.

(12) Gallus, D. R.; Wagner, R.; Wiemers-Meyer, S.; Winter, M.; Cekic-Laskovic, I. New Insights into the Structure-Property Relationship of High-Voltage Electrolyte Components for Lithium-Ion Batteries Using the pK_a Value. *Electrochim. Acta* **2015**, *184*, 410−416.

(13) Gutten, O.; Rulíšek, L. Predicting the Stability Constants of Metal-Ion Complexes from First Principles. *Inorg. Chem.* **2013**, *52* (18), 10347−10355.

(14) Casasnovas, R.; Ortega-Castro, J.; Donoso, J.; Frau, J.; Muñoz, F. Theoretical Calculations of Stability Constants and pK_a Values of Metal Complexes in Solution: Application to Pyridoxamine−Copper-(II) Complexes and Their Biological Implications in AGE Inhibition. *Phys. Chem. Chem. Phys.* **2013**, *15* (38), 16303−16313.

DOI: 10.1021/acs.jpca.8b01751
J. Phys. Chem. A 2018, 122, 4366−4374

(15) Nambu, K.; Yonebayashi, K. Acidic Properties of Dissolved Organic Matter Leached from Organic Layers in Temperate Forests. *Soil Sci. Plant Nutr.* **1999**, *45* (1), 65−77.
(16) Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of Ab Initio Molecular Potentials for the Prevision of Solvent Effects. *Chem. Phys.* **1981**, *55* (1), 117−129.
(17) Miertuš, S.; Tomasi, J. Approximate Evaluations of the Electrostatic Free Energy and Internal Energy Changes in Solution Processes. *Chem. Phys.* **1982**, *65* (2), 239−245.
(18) Pascual-Ahuir, J. L.; Silla, E.; Tuñon, I. GEPOL: An Improved Description of Molecular Surfaces. III. A New Algorithm for the Computation of a Solvent-Excluding Surface. *J. Comput. Chem.* **1994**, *15* (10), 1127−1138.
(19) Scalmani, G.; Frisch, M. J. Continuous Surface Charge Polarizable Continuum Models of Solvation. I. General Formalism. *J. Chem. Phys.* **2010**, *132* (11), 114110.
(20) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. Solvent Effects. 5. Influence of Cavity Shape, Truncation of Electrostatics, and Electron Correlation on Ab Initio Reaction Field Calculations. *J. Phys. Chem.* **1996**, *100* (40), 16098−16104.
(21) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102* (11), 1995−2001.
(22) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model. *J. Comput. Chem.* **2003**, *24* (6), 669−681.
(23) Klamt, A.; Schüürmann, G. 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**, *0* (5), 799−805.
(24) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99* (7), 2224−2235.
(25) 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* (18), 6378−6396.
(26) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Performance of SM6, SM8, and SMD on the SAMPL1 Test Set for the Prediction of Small-Molecule Solvation Free Energies. *J. Phys. Chem. B* **2009**, *113* (14), 4538−4543.
(27) Zhao, Y.; Truhlar, D. G. Applications and Validations of the Minnesota Density Functionals. *Chem. Phys. Lett.* **2011**, *502* (1), 1−13.
(28) Ho, J. Are Thermodynamic Cycles Necessary for Continuum Solvent Calculation of PKas and Reduction Potentials? *Phys. Chem. Chem. Phys.* **2015**, *17* (4), 2859−2868.
(29) Ho, J.; Ertem, M. Z. Calculating Free Energy Changes in Continuum Solvation Models. *J. Phys. Chem. B* **2016**, *120* (7), 1319−1329.
(30) Ho, J.; Coote, M. L. A Universal Approach for Continuum Solvent pK_a Calculations: Are We There Yet? *Theor. Chem. Acc.* **2010**, *125* (1−2), 3.
(31) Ho, J.; Coote, M. L. pK_a Calculation of Some Biologically Important Carbon Acids - An Assessment of Contemporary Theoretical Procedures. *J. Chem. Theory Comput.* **2009**, *5* (2), 295−306.
(32) Thapa, B.; Schlegel, H. B. Density Functional Theory Calculation of pK_as of Thiols in Aqueous Solution Using Explicit Water Molecules and the Polarizable Continuum Model. *J. Phys. Chem. A* **2016**, *120* (28), 5726−5735.
(33) Thapa, B.; Schlegel, H. B. Improved pK_a Prediction of Substituted Alcohols, Phenols, and Hydroperoxides in Aqueous Medium Using Density Functional Theory and a Cluster-Continuum Solvation Model. *J. Phys. Chem. A* **2017**, *121* (24), 4698−4706.
(34) Ugur, I.; Marion, A.; Parant, S.; Jensen, J. H.; Monard, G. Rationalization of the pK_a Values of Alcohols and Thiols Using Atomic Charge Descriptors and Its Application to the Prediction of Amino Acid pK_as. *J. Chem. Inf. Model.* **2014**, *54* (8), 2200−2213.
(35) Zhang, S.; Baker, J.; Pulay, P. A Reliable and Efficient First Principles-Based Method for Predicting pK_a Values. 1. Methodology. *J. Phys. Chem. A* **2010**, *114* (1), 425−431.
(36) Zhang, S.; Baker, J.; Pulay, P. A Reliable and Efficient First Principles-Based Method for Predicting pK_a Values. 2. Organic Acids. *J. Phys. Chem. A* **2010**, *114* (1), 432−442.
(37) Namazian, M.; Kalantary-Fotooh, F.; Noorbala, M. R.; Searles, D. J.; Coote, M. L. Møller−Plesset Perturbation Theory Calculations of the pK_a Values for a Range of Carboxylic Acids. *J. Mol. Struct.: THEOCHEM* **2006**, *758* (2), 275−278.
(38) Burger, S. K.; Liu, S.; Ayers, P. W. Practical Calculation of Molecular Acidity with the Aid of a Reference Molecule. *J. Phys. Chem. A* **2011**, *115* (7), 1293−1304.
(39) Marenich, A. V.; Ding, W.; Cramer, C. J.; Truhlar, D. G. Resolution of a Challenge for Solvation Modeling: Calculation of Dicarboxylic Acid Dissociation Constants Using Mixed Discrete−Continuum Solvation Models. *J. Phys. Chem. Lett.* **2012**, *3* (11), 1437−1442.
(40) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. Are Mixed Explicit/Implicit Solvation Models Reliable for Studying Phosphate Hydrolysis? A Comparative Study of Continuum, Explicit and Mixed Solvation Models. *ChemPhysChem* **2009**, *10* (7), 1125−1134.
(41) Bachs, M.; Luque, F. J.; Orozco, M. Optimization of Solute Cavities and van Der Waals Parameters in Ab Initio MST-SCRF Calculations of Neutral Molecules. *J. Comput. Chem.* **1994**, *15* (4), 446−454.
(42) Orozco, M.; Luque, F. J. Optimization of the Cavity Size for Ab Initio MST-SCRF Calculations of Monovalent Ions. *Chem. Phys.* **1994**, *182* (2), 237−248.
(43) Fu, Y.; Liu, L.; Yu, H.-Z.; Wang, Y.-M.; Guo, Q.-X. Quantum-Chemical Predictions of Absolute Standard Redox Potentials of Diverse Organic Molecules and Free Radicals in Acetonitrile. *J. Am. Chem. Soc.* **2005**, *127* (19), 7227−7234.
(44) Caricato, M.; Mennucci, B.; Tomasi, J. Solvent Polarity Scales Revisited: A ZINDO-PCM Study of the Solvatochromism of Betaine-30. *Mol. Phys.* **2006**, *104* (5−7), 875−887.
(45) Ginovska, B.; Camaioni, D. M.; Dupuis, M. The H_2O_2 + OH → HO_2 + H_2O Reaction in Aqueous Solution from a Charge-Dependent Continuum Model of Solvation. *J. Chem. Phys.* **2008**, *129* (1), 014506.
(46) Verdolino, V.; Cammi, R.; Munk, B. H.; Schlegel, H. B. Calculation of pK_a Values of Nucleobases and the Guanine Oxidation Products Guanidinohydantoin and Spiroiminodihydantoin Using Density Functional Theory and a Polarizable Continuum Model. *J. Phys. Chem. B* **2008**, *112* (51), 16860−16873.
(47) Sviatenko, L.; Isayev, O.; Gorb, L.; Hill, F.; Leszczynski, J. Toward Robust Computational Electrochemical Predicting the Environmental Fate of Organic Pollutants. *J. Comput. Chem.* **2011**, *32* (10), 2195−2203.
(48) Psciuk, B. T.; Lord, R. L.; Munk, B. H.; Schlegel, H. B. Theoretical Determination of One-Electron Oxidation Potentials for Nucleic Acid Bases. *J. Chem. Theory Comput.* **2012**, *8* (12), 5107−5123.
(49) Matsui, T.; Shigeta, Y.; Morihashi, K. Assessment of Methodology and Chemical Group Dependences in the Calculation of the pK_a for Several Chemical Groups. *J. Chem. Theory Comput.* **2017**, *13*, 4791.
(50) Riccardi, D.; Guo, H.-B.; Parks, J. M.; Gu, B.; Liang, L.; Smith, J. C. Cluster-Continuum Calculations of Hydration Free Energies of Anions and Group 12 Divalent Cations. *J. Chem. Theory Comput.* **2013**, *9* (1), 555−569.
(51) Lian, P.; Guo, H.-B.; Riccardi, D.; Dong, A.; Parks, J. M.; Xu, Q.; Pai, E. F.; Miller, S. M.; Wei, D.-Q.; Smith, J. C.; Guo, H. X-ray Structure of a Hg^{2+} Complex of Mercuric Reductase (MerA) and Quantum Mechanical/Molecular Mechanical Study of Hg^{2+} Transfer between the C-Terminal and Buried Catalytic Site Cysteine Pairs. *Biochemistry* **2014**, *53* (46), 7211−7222.

DOI: 10.1021/acs.jpca.8b01751
J. Phys. Chem. A 2018, 122, 4366−4374

(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.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision E.01; Gaussian, Inc.: Wallingford, CT, 2009.
(53) Marenich, A. V.; Kelly, C. P.; Thompson, J. D.; Hawkins, G. D.; Chambers, C. C.; Giesen, D. J.; Winget, P.; Cramer, C. J.; Truhlar, D. G. *Minnesota Solvation Database*—version 2012; University of Minnesota: Minneapolis, 2012.
(54) Schüürmann, G.; Cossi, M.; Barone, V.; Tomasi, J. Prediction of the pK_a of Carboxylic Acids Using the Ab Initio Continuum-Solvation Model PCM-UAHF. *J. Phys. Chem. A* **1998**, *102* (33), 6706−6712.
(55) da Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. Ab Initio Calculations of Absolute pK_a Values in Aqueous Solution I. Carboxylic Acids. *J. Phys. Chem. A* **1999**, *103* (50), 11194−11199.
(56) Namazian, M.; Halvani, S.; Noorbala, M. R. Density Functional Theory Response to the Calculations of pK_a Values of Some Carboxylic Acids in Aqueous Solution. *J. Mol. Struct.: THEOCHEM* **2004**, *711* (1−3), 13−18.
(57) Perrin, D. D.; Dempsey, B.; Serjeant, E. P. Prediction of pK_a Values of Substituted Aliphatic Acids and Bases. In *pKa Prediction for Organic Acids and Bases*; Springer: Dordrecht, 1981; pp 27−43.
(58) Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A. pK_a Calculations of Aliphatic Amines, Diamines, and Aminoamides via Density Functional Theory with a Poisson−Boltzmann Continuum Solvent Model. *J. Phys. Chem. A* **2007**, *111* (20), 4422−4430.
(59) Kreevoy, M. M.; Harper, E. T.; Duvall, R. E.; Wilgus, H. S.; Ditsch, L. T. Inductive Effects on the Acid Dissociation Constants of Mercaptans. *J. Am. Chem. Soc.* **1960**, *82* (18), 4899−4902.
(60) Kortüm, G.; Andrussow, K.; Vogel, W. *Dissociation constants of organic acids in aqueous solution*; Butterworth: London, 1961.
(61) Kreevoy, M. M.; Eichinger, B. E.; Stary, F. E.; Katz, E. A.; Sellstedt, J. H. The Effect of Structure on Mercaptan Dissociation Constants. *J. Org. Chem.* **1964**, *29* (6), 1641−1642.
(62) Irving, R. J.; Nelander, L.; Wadsö, I.; Halvarson, H.; Nilsson, L. Thermodynamics of the Ionization of Some Thiols in Aqueous Solution. *Acta Chem. Scand.* **1964**, *18*, 769−787.
(63) Perrin, D. D. *Dissociation Constants of Organic Bases in Aqueous Solution*; Butterworths: London, 1965.
(64) Perrin, D. D. *Dissociation Constants of Organic Bases in Aqueous Solution: Supplement*; Butterworths: London, 1972.
(65) De Maria, P.; Fini, A.; Hall, F. M. Thermodynamic Acid Dissociation Constants of Aromatic Thiols. *J. Chem. Soc., Perkin Trans. 2* **1973**, *0* (14), 1969−1971.
(66) Tsonopoulos, C.; Coulson, D. M.; Inman, L. B. Ionization Constants of Water Pollutants. *J. Chem. Eng. Data* **1976**, *21* (2), 190−193.
(67) Hupe, D. J.; Jencks, W. P. Nonlinear Structure-Reactivity Correlations. Acyl Transfer between Sulfur and Oxygen Nucleophiles. *J. Am. Chem. Soc.* **1977**, *99* (2), 451−464.
(68) Serjeant, E. P.; Dempsey, B. *Ionisation Constants of Organic Acids in Aqueous Solution*; Pergamon Press, 1979.
(69) Arnold, A. P.; Canty, A. J. Methylmercury(II) Sulfhydryl Interactions. Potentiometric Determination of the Formation Constants for Complexation of Methylmercury(II) by Sulfhydryl Containing Amino Acids and Related Molecules, Including Glutathione. *Can. J. Chem.* **1983**, *61* (7), 1428−1434.
(70) Pettit, L.D.; Powell, K. J.; Solov'ev, V. *The IUPAC Stability Constants Database*; Academic Software: Otley, West Yorkshire, U.K., 2006.
(71) Camaioni, D. M.; Schwerdtfeger, C. A. Comment on "Accurate Experimental Values for the Free Energies of Hydration of H^+, OH^-, and H_3O^+. *J. Phys. Chem. A* **2005**, *109* (47), 10795−10797.
(72) 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* (23), 7894−7899.
(73) 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* (32), 16066−16081.
(74) 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* (29), 15068−15106.
(75) Bondi, A. Van Der Waals Volumes and Radii. *J. Phys. Chem.* **1964**, *68* (3), 441−451.
(76) Sviatenko, L. K.; Gorb, L.; Hill, F. C.; Leszczynska, D.; Leszczynski, J. Theoretical Study of One-Electron Reduction and Oxidation Potentials of N-Heterocyclic Compounds. *Chem. Heterocycl. Compd.* **2014**, *50* (3), 311−318.
(77) 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* (2), 785−789.
(78) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648−5652.
(79) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98* (45), 11623−11627.
(80) Haworth, N. L.; Wang, Q.; Coote, M. L. Modeling Flexible Molecules in Solution: A pK_a Case Study. *J. Phys. Chem. A* **2017**, *121* (27), 5217−5225.
(81) Pascual-Ahuir, J. L.; Silla, E. GEPOL: An Improved Description of Molecular Surfaces. I. Building the Spherical Surface Set. *J. Comput. Chem.* **1990**, *11* (9), 1047−1060.
(82) Pomelli, C. S.; Tomasi, J.; Cammi, R. A Symmetry Adapted Tessellation of the GEPOL Surface: Applications to Molecular Properties in Solution. *J. Comput. Chem.* **2001**, *22* (12), 1262−1272.
(83) Frediani, L.; Cammi, R.; Pomelli, C. S.; Tomasi, J.; Ruud, K. New Developments in the Symmetry-Adapted Algorithm of the Polarizable Continuum Model. *J. Comput. Chem.* **2004**, *25* (3), 375−385.
(84) Scalmani, G.; Rega, N.; Cossi, M.; Barone, V. Finite Elements Molecular Surfaces in Continuum Solvent Models for Large Chemical Systems. *J. Comput. Methods Sci. Eng.* **2002**, *2* (3−4), 469−474, DOI: 10.3233/JCM-2002-23-423.

DOI: 10.1021/acs.jpca.8b01751
J. Phys. Chem. A 2018, 122, 4366−4374