

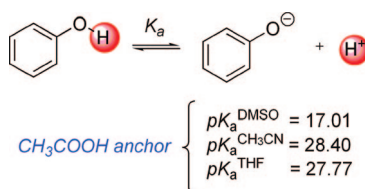
# First-Principles Calculation of $pK_a$ Values for Organic Acids in Nonaqueous Solution

Feizhi Ding, Jeremy M. Smith,\* and Haobin Wang\*

Department of Chemistry and Biochemistry, New Mexico State University, Las Cruces, New Mexico 88003

jesmith@nmsu.edu; haobin@nmsu.edu

Received December 1, 2008



Electronic structure theory, mainly the density functional theory (DFT), is applied to calculate the  $pK_a$  values for a variety of organic acids in several nonaqueous solvents: namely DMSO, MeCN, and THF. Following the supermolecule approach the solute molecule, together with a few solvent molecules in close proximity, is treated explicitly by the electronic structure theory, and the remaining solvent environment is approximated by using a standard dielectric continuum model. It is found that in most cases including only one explicit solvent molecule gives satisfactory results for  $pK_a$  estimations. Next, the equilibrium position free energy difference is calculated between a reference acid–base pair whose  $pK_a$  is known experimentally and the acid–base pair whose  $pK_a$  is to be determined theoretically. This bypasses the step of treating the solvated proton that most of the current theories have difficulty with and, to a large extent, induces favorable error cancelations in the final theoretical results. Accurate theoretical predictions of  $pK_a$  values are thus obtained at a moderate level of theory (MP2 single point on B3LYP/6-31+G(d) optimized geometry) for a series of organic acids spanning a wide range of acidities in DMSO, MeCN, and THF. Furthermore, the correlation between the  $pK_a$  values of these acids in different solutions is investigated theoretically, and excellent agreement is found with the experimental results.

## 1. Introduction

Knowledge of the  $pK_a$  value of a compound is important for understanding many chemical and biochemical processes that depend on proton transfer,<sup>1</sup> e.g., interpreting the functional mechanism of ionizable groups in a protein at a molecular level.<sup>2</sup> Accurate  $pK_a$  data are also important for understanding reaction mechanisms that do not involve proton transfer. For example, thermodynamic cycles that use  $pK_a$  data are used to determine the hydricities<sup>3</sup> and bond dissociation energies<sup>4</sup> of metal hydrides and other compounds.<sup>5</sup> These data are critical for

elucidating reaction mechanisms that involve hydrogen atom transfer/proton-coupled electron transfer.<sup>6,7</sup>

However, the accurate experimental determination of  $pK_a$  values is not a trivial task in many cases and may pose a

\* To whom all correspondence should be addressed.

(1) Bell, R. P. *The Proton in Chemistry*; Chapman and Hall: London, UK, 1973.

(2) Ohno, K.; Kamiya, N.; Asakawa, N.; Inoune, Y.; Sakurai, M. *Chem. Phys. Lett.* **2001**, *341*, 387.

(3) (a) Qi, X.; Fu, Y.; Liu, L.; Guo, Q. *Organometallics* **2007**, *26*, 4197. (b) Fraze, K.; Wilson, A. D.; Appel, A. M.; Rakowski DuBois, M.; DuBois, D. L. *Organometallics* **2007**, *26*, 3918. (c) Raebiger, J. W.; Miedaner, A.; Curtis, C. J.; Miller, S. M.; Anderson, O. P.; DuBois, D. L. *J. Am. Chem. Soc.* **2004**, *126*, 5502. (d) Ellis, W. W.; Ciancanelli, R.; Miller, S. M.; Raebiger, J. W.; Rakowski DuBois, M.; DuBois, D. L. *J. Am. Chem. Soc.* **2003**, *125*, 12230. (e) Sarker, N.; Bruno, J. W. *Organometallics* **2001**, *20*, 55.

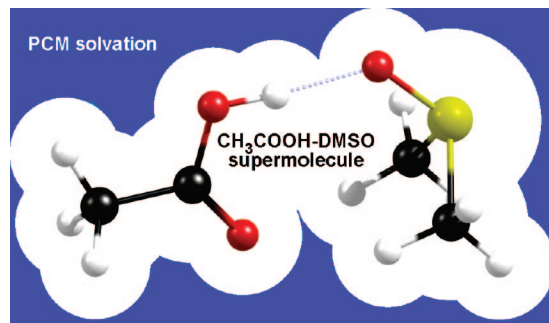
(4) (a) Choi, J.; Pulling, M. E.; Smith, D. M.; Norton, J. R. *J. Am. Chem. Soc.* **2008**, *130*, 4250. (b) Fu, Y.; Liu, L.; Wang, Y.; Li, J.; Yu, T.; Guo, Q. *J. Phys. Chem. A* **2006**, *110*, 5874. (c) Ciancanelli, R.; Noll, B. C.; DuBois, D. L.; DuBois, M. R. *J. Am. Chem. Soc.* **2002**, *124*, 2984. (d) Protasiewicz, J. D.; Theopold, K. H. *J. Am. Chem. Soc.* **1993**, *115*, 5559. (e) Tilsted, M.; Parker, V. D. *J. Am. Chem. Soc.* **1989**, *111*, 6711.

(5) (a) Nieto, I.; Ding, F.; Bontchev, R. P.; Wang, H.; Smith, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 2716. (b) Appel, A. M.; DuBois, D. L.; Rakowski DuBois, M. *J. Am. Chem. Soc.* **2005**, *127*, 12717. (c) Ellis, W. W.; Raebiger, J. W.; Curtis, C. J.; Bruno, J. W.; DuBois, D. L. *J. Am. Chem. Soc.* **2004**, *126*, 2738. (d) Soper, J. D.; Mayer, J. M. *J. Am. Chem. Soc.* **2003**, *125*, 12217. (e) Roth, J. P.; Lovell, S.; Mayer, J. M. *J. Am. Chem. Soc.* **2000**, *122*, 5486. (f) Ellis, W. W.; Miedaner, A.; Curtis, C. J.; Gibson, D. H.; DuBois, D. L. *J. Am. Chem. Soc.* **2002**, *124*, 1926. (g) Bordwell, F. G.; Cheng, J. P.; Ji, G.-Z.; Satish, A. V.; Zhang, X. M. *J. Am. Chem. Soc.* **1991**, *113*, 9790.

(6) (a) Costentin, C. *Chem. Rev.* **2008**, *108*, 2145. (b) Liu, F.; Concepcion, J. J.; Jurss, J. W.; Cardolaccia, T.; Templeton, J. L.; Meyer, T. J. *Inorg. Chem.* **2008**, *47*, 1227. (c) Huynh, M. H. V.; Meyer, T. J. *Chem. Rev.* **2007**, *107*, 5004. (d) Mayer, J. M. *Annu. Rev. Phys. Chem.* **2004**, *55*, 363. (e) Mayer, J. M. *Acc. Chem. Res.* **1998**, *31*, 441.

significant experimental challenge.<sup>8</sup> This is the case for species such as reaction intermediates, very strong acids, and very weak acids.<sup>9</sup> As a result, determining  $pK_a$  values computationally has become an issue of great interest. Over the years, considerable theoretical effort<sup>9–21,62</sup> has been devoted to accurately modeling acid–base equilibria in solution and calculating the associated  $pK_a$  values. The general goal is to improve the reliability of the computational approach so as to have a realistic comparison with the available experimental data. Along this line Jorgensen et al.<sup>62</sup> performed Monte Carlo simulations of the relative  $pK_a$  values for two weak organic acids, acetonitrile and ethane, in water, and good agreement with experiments was found. Schuurmann et al.<sup>10</sup> calculated the  $pK_a$  values of 16 aliphatic carboxylic acids in aqueous solution, and found a reasonable correlation between the theoretical and experimental results. Nascimento et al.<sup>11</sup> proposed a thermodynamic cycle to calculate the absolute  $pK_a$  values of a number of aliphatic alcohols, thiols, and carboxylic acids in aqueous solution, and good agreement with the experimental  $pK_a$  values was found. Moreover, Shields et al.<sup>12</sup> carried out systematic study in predicting  $pK_a$  values for carboxylic acids and phenols in water, achieving an accuracy of less than one-half of a  $pK_a$  unit. Namazian et al.<sup>13</sup> and Barone et al.<sup>14</sup> also demonstrated the successful  $pK_a$  calculations for carboxylic acids in aqueous solution.

Advances in electronic structure theory and computer speed have made it possible to obtain free energies for small gas-phase molecules at the level that is approaching the chemical accuracy.<sup>22</sup> The situation is less satisfactory with respect to calculating the solvation free energy for a molecular system, in which the solvent molecules are either treated explicitly within the electronic structure theory or molecular dynamics (MD)



**FIGURE 1.** Schematic representation of the supermolecule solvation approach.

simulation framework<sup>23</sup> or implicitly by using a continuum-based solvation model. The explicit solvent model requires a relatively high computational cost and is only applicable to a limited number of small molecule systems. On the other hand, various implicit solvation models have been developed based on dielectric continuum theories, which are applicable to larger molecule systems with reasonable computational effort.<sup>24,25</sup> In the latter methods the solute molecule is put in a cavity that is surrounded by the solvent continuum. The size of the cavity and other solvent-related quantities are treated as adjustable parameters and obtained from physically motivated modeling procedures. Differences between these models are often distinguished by the specific definition of the cavity and how the solute–solvent interaction is treated.

Among the implicit solvent models the polarizable continuum model (PCM)<sup>25</sup> is one of the most popular approaches. It has been extensively used to calculate the  $pK_a$  values for acid/base equilibria in aqueous solution.<sup>10–14</sup> Although reasonable results were obtained with the PCM method, all these works focused on only a few closely related acids. A more extensive study by Pliego and Riveros<sup>9</sup> found that the error of using the PCM method in  $pK_a$  predictions can become as large as 7  $pK_a$  units when a larger number of acids with different functional groups are considered. It was proposed that the solute–solvent interactions (e.g., hydrogen bonding) in the first solvation shell may cause such a sizable error.<sup>9</sup>

From a more fundamental point of view, the contribution to the solvation free energy is not only restricted to the bulk polarization effects, which may be described by a dielectric continuum model, but should also include the first-layer solvation effect, which is caused by the short-range solute–solvent interactions (i.e., hydrogen-bonding and charge-dipole interactions). One approach along this direction is to add a few solvent molecules in close proximity to the solute molecule, which forms a composite “supermolecule”. A dielectric continuum model such as the PCM approach is then applied to this composite entity. A schematic representation of this supermolecule solvation approach is shown in Figure 1. This hybrid explicit/implicit solvation model has been called the discrete-continuum,<sup>26</sup> the supermolecule-reaction-field,<sup>27</sup> and the cluster-

(7) (a) Markle, T. F.; Rhile, I. J.; DiPasquale, A. G.; Mayer, J. M. *Proc. Natl. Acad. Sci.* **2008**, *105*, 8185. (b) Fecenko, C. J.; Thorp, H. H.; Meyer, T. J. *J. Am. Chem. Soc.* **2007**, *129*, 15098. (c) Zdilla, M. J.; Dexheimer, J. L.; Abu-Omar, M. M. *J. Am. Chem. Soc.* **2007**, *129*, 11505. (d) Cowley, R. E.; Bontchev, R. P.; Sorrell, J.; Sarracino, O.; Feng, Y.; Wang, H.; Smith, J. M. *J. Am. Chem. Soc.* **2007**, *129*, 2424.

(8) Albert, A.; Serjeant, E. P. *The Determination of Ionization Constants*; Chapman and Hall: New York, 1984.

(9) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2002**, *106*, 7434.

(10) Schuurmann, G. S.; Cossi, M.; Barone, V.; Tomasi, J. *J. Phys. Chem. A* **1998**, *102*, 6706.

(11) Silva, C. O.; Silva, M. A.; Nascimento, M. A. C. *J. Phys. Chem. A* **1999**, *103*, 11194.

(12) (a) Toth, A. M.; Liptak, M. D.; Phillips, D. L.; Shields, G. C. *J. Chem. Phys.* **2001**, *114*, 4595. (b) Liptak, M. D.; Shields, G. C. *J. Am. Chem. Soc.* **2001**, *123*, 7314. (c) Liptak, M. D.; Gross, K. C.; Seybold, P. G.; Feldgus, S.; Shields, G. C. *J. Am. Chem. Soc.* **2002**, *124*, 6421.

(13) (a) Namazian, M.; Heidary, H. *THEOCHEM* **2003**, *620*, 257. (b) Namazian, M.; Halvani, S. *J. Chem. Thermodyn.* **2006**, *38*, 1495.

(14) Saracino, G. A. A.; Improta, R.; Barone, V. *Chem. Phys. Lett.* **2003**, *373*, 411.

(15) Chen, J.; McAllister, M. A.; Lee, J. K.; Houk, K. N. *J. Org. Chem.* **1998**, *63*, 4611.

(16) Adam, K. R. *J. Phys. Chem. A* **2002**, *106*, 11963.

(17) Klamt, A.; Eckert, F.; Diedenhofen, M.; Beck, M. E. *J. Phys. Chem. A* **2003**, *107*, 9380.

(18) (a) Vianello, R.; Maksic, Z. B. *Eur. J. Org. Chem.* **2005**, *16*, 3571. (b) Vianello, R.; Maksic, Z. B. *Tetrahedron Lett.* **2005**, *46*, 3711. (c) Vianello, R.; Maksic, Z. B. *Tetrahedron* **2005**, *61*, 9381.

(19) Chipman, D. M. *J. Phys. Chem. A* **2002**, *106*, 7413.

(20) Almerindo, G. I.; Tondo, D. W.; Pliego, J. R., Jr. *J. Phys. Chem. A* **2004**, *108*, 166.

(21) Magill, A. M.; Cavell, K. J.; Yates, B. F. *J. Am. Chem. Soc.* **2004**, *126*, 8717.

(22) (a) Burk, P.; Koppel, I. A.; Koppel, I.; Leito, I.; Travnikova, O. *Chem. Phys. Lett.* **2000**, *323*, 482. (b) Seo, Y.; Kim, Y.; Kim, Y. *Chem. Phys. Lett.* **2001**, *340*, 186. (c) Pokon, E. K.; Liptak, M. D.; Feldgus, S.; Shields, G. C. *J. Phys. Chem. A* **2001**, *105*, 10483. (d) Ervin, K. M.; DeTuri, V. F. *J. Phys. Chem. A* **2002**, *106*, 9947. (e) Wiberg, K. B. *J. Org. Chem.* **2002**, *67*, 4787. (f) Remko, M. *J. Phys. Chem. A* **2002**, *106*, 5005.

(23) Sprik, M. *Chem. Phys.* **2000**, *258*, 139.

(24) (a) Cramer, C. J.; Truhlar, D. G. *Rev. Comput. Chem.* **1995**, *6*, 1. (b) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.

(25) (a) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210. (b) Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **1998**, *102*, 870. (c) Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **2000**, *104*, 4690.

(26) Claverie, P.; Daudey, J. P.; Langlet, J.; Pullman, B.; Piazzola, D.; Huron, M. *J. Phys. Chem.* **1978**, *82*, 405.

(27) Freitas, L. C. G.; Longo, R. L.; Simas, A. M. *J. Chem. Soc., Faraday Trans.* **1992**, *88*, 189.

continuum model.<sup>28</sup> Using this type of approach one may calculate the solvation free energies with a much greater accuracy than those obtained by using a traditional dielectric continuum approach. For instance, Pliego et al. have successfully predicted the solvation free energies for a series of ions in aqueous solutions using the cluster-continuum model.<sup>28</sup> Furthermore, Pliego and Riveros<sup>9</sup> apply this model in pK<sub>a</sub> predications in aqueous solution. A total number of 17 acids with different functional groups were considered in their study and the calculated pK<sub>a</sub> values are in good agreement with the experimental values. While these ideas have been explored mainly on the solvation of ions in aqueous solutions, less work has been done on the solvation of neutral molecules or in nonaqueous solutions.<sup>29</sup>

The objective of this paper is to investigate the reliability of using the above approach to calculate the pK<sub>a</sub> values of various organic acids in nonaqueous solutions. Instead of calculating the solvation free energy of each individual species, we compute the net solvation free energy change in an acid–base reaction. The hybrid explicit/implicit solvation method is employed to both neutral acidic molecules and their conjugate anionic bases. A series of common acids with different functional groups or substituents are treated in this work. Experimental pK<sub>a</sub> values of these chosen acids in DMSO span a wide range (0 to 35 pK<sub>a</sub> units) and have gradually changing acidities. These acids are representatives from all of the conventional families of OH acids (e.g., alcohols, phenols, carboxylic acids, and sulfonic acids), NH acids (e.g., amines and ammonium ions), and CH acids (e.g., hydrocarbons and ketones). This series therefore provides an interesting and challenging set of solutes, allowing for systematic and nonbiased evaluation of solvation effects in nonaqueous solution.

Another goal of this work is to predict the pK<sub>a</sub> values which are not experimentally available. Acidity data in nonaqueous solutions are of continuing interest to chemists. Among these nonaqueous solvents, acetonitrile (MeCN) and tetrahydrofuran (THF) are widely used. The low polarity and solvating abilities of these solvents introduces complications associated with ion-pairing that can complicate experimental pK<sub>a</sub> determinations.<sup>30</sup> Although acidity data for different compounds in the pK<sub>a</sub> range of 4 to 28 exist in MeCN, to the best of our knowledge, there are no data for acids with pK<sub>a</sub><sup>MeCN</sup> > 28. Furthermore, large discrepancies exist in the pK<sub>a</sub> measurements for some acids. For example, the experimental pK<sub>a</sub><sup>MeCN</sup> value for 2,4-(NO<sub>2</sub>)<sub>2</sub>-phenol ranges from 15.3<sup>31</sup> to 18.4.<sup>32</sup> The situation is even worse in THF where pK<sub>a</sub> data are quite limited.<sup>33</sup> In this paper, we predict the pK<sub>a</sub> values in MeCN and in THF for the same set of acids calculated in DMSO. Good agreement is found between our calculated pK<sub>a</sub> values and experimental data. Moreover,

reliable acidity scale correlations of pK<sub>a</sub><sup>MeCN</sup>/pK<sub>a</sub><sup>DMSO</sup> and pK<sub>a</sub><sup>THF</sup>/pK<sub>a</sub><sup>DMSO</sup> are established, aiming to serve as a useful tool for pK<sub>a</sub>-related studies in MeCN and THF.

## 2. Methodology and Computational Details

**2.1. The Relative Determination of pK<sub>a</sub> Based on a Reference Acid/Base Pair.** A seemingly straightforward way to evaluate pK<sub>a</sub> is to compute the total change in free energy for the solution-phase deprotonation reaction



where “(solv)” denotes the appropriate solvation environment. However, problems arise when treating the solvated proton whose chemical structure is often unknown and difficult to model theoretically. This is largely due to the fact that the proton is much smaller than any other cations in acid–base reactions and it has strong interactions with the surrounding solvent molecules. The PCM and other similar continuum models underestimate the proton solvation energy.<sup>34</sup> Use of the experimental solvation free energy for the proton<sup>35</sup> is also problematic because there are considerable experimental uncertainties with respect to this information.<sup>36</sup> Such challenges to experiment and theory are reflected by the fact that in water the experimental value for the proton solvation energy ranges from –254 to –261 kcal/mol,<sup>37</sup> whereas theoretically it ranges from –245.6 to –262.4 kcal/mol.<sup>38,39</sup> Even less information is available for the proton solvation energy in nonaqueous solvents. These inadequacies in treating the solvated proton may induce large uncertainties in the calculated pK<sub>a</sub> values, making eq 1a somewhat less useful for acidity predictions. To illustrate, for pK<sub>a</sub> calculations in water, DMSO, and MeCN using the experimental proton solvation energies, Chipman<sup>19</sup> observed a significant shift (~6 pK<sub>a</sub> units) between the predicted pK<sub>a</sub> values and the experimental pK<sub>a</sub> values.

The approach we used in this paper is to consider a reference deprotonation reaction whose pK<sub>a</sub> is known experimentally



Since free energy is a state function, we can consider the following generic acid–base reaction by combining eqs 1a and 1b

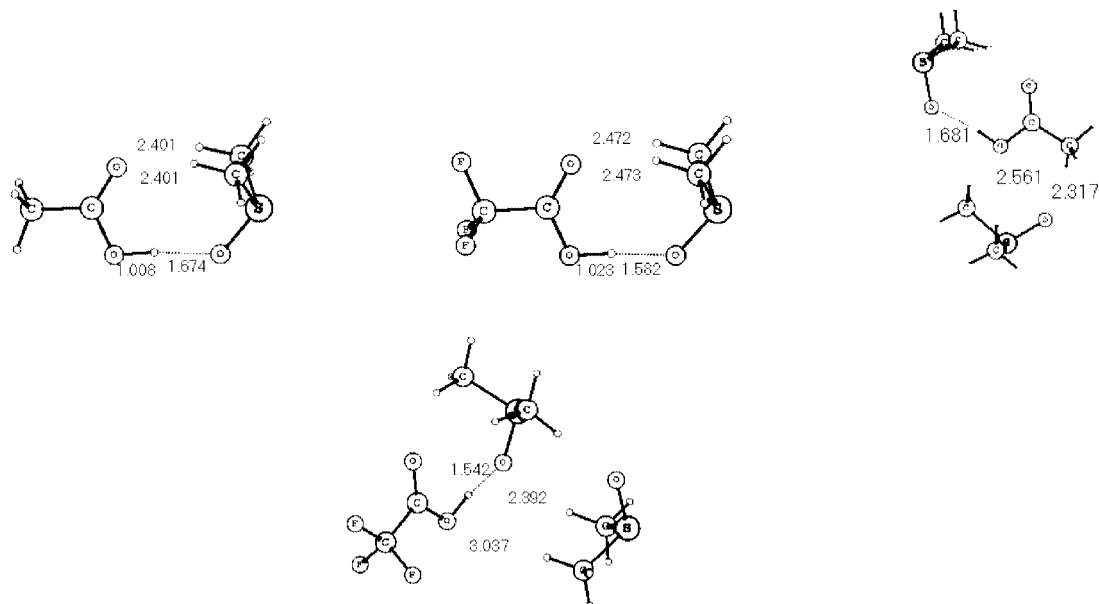


The free energy of eq 1b (denoted by ΔG<sub>1b</sub>) for the anchor acid HB can be derived from its experimental pK<sub>a</sub> value, ΔG<sub>1b</sub> = (2.303RT)pK<sub>a</sub>(HB); whereas the free energy of eq 2 (denoted by ΔG) can be obtained from a theoretical calculation. The pK<sub>a</sub> of the unknown acid HA in eq 1a can thus be extracted from the free energy of eqs 1b and 2 as

$$\text{pK}_a(\text{HA}) = \frac{\Delta G_{1b} + \Delta G}{2.303RT} = \text{pK}_a(\text{HB}) + \frac{\Delta G}{2.303RT} \quad (3)$$

(28) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2001**, *105*, 7241.  
 (29) (a) Fu, Y.; Liu, L.; Li, R.; Liu, R.; Guo, Q. *J. Am. Chem. Soc.* **2004**, *126*, 814. (b) Fu, Y.; Shen, K.; Liu, L.; Guo, Q. *J. Am. Chem. Soc.* **2007**, *129*, 13510. (c) Qi, X.; Liu, L.; Fu, Y.; Guo, Q. *Organometallics* **2006**, *25*, 5879. (d) Shen, K.; Fu, Y.; Li, J.; Liu, L.; Guo, Q. *Tetrahedron* **2007**, *63*, 1568. (e) Li, J.; Fu, Y.; Liu, L.; Guo, Q. *Tetrahedron* **2006**, *62*, 11801. (f) Li, J.; Liu, L.; Fu, Y.; Guo, Q. *Tetrahedron* **2006**, *62*, 4453.  
 (30) Sooväli, L.; Kaljurand, I.; Kütt, A.; Leito, I. *Anal. Chim. Acta* **2006**, *566*, 290.  
 (31) Juillard, J.; Loubinoux, B. *C. R. Acad. Sci., Ser. C* **1967**, *264*, 1680.  
 (32) Jasinski, T.; El-Harakany, A. A.; Halaka, F. G.; Sadek, H. *Croat. Chem. Acta* **1987**, *51*, 1.  
 (33) (a) Garrido, G.; Rosés, M.; Ràfols, C.; Bosch, E. *J. Solution Chem.* **2008**, *37*, 689. (b) Rõõm, E.-I.; Kütt, A.; Kaljurand, I.; Koppel, I.; Leito, I.; Koppel, I.; Mishima, M.; Goto, K.; Miyahara, Y. *Chem.–Eur. J.* **2007**, *13*, 7631. (c) Garrido, G.; Koort, E.; Ràfols, C.; Bosch, E.; Rodima, T.; Leito, I.; Rosés, M. *J. Org. Chem.* **2006**, *71*, 9062.

(34) Pliego, J. R., Jr. *Chem. Phys. Lett.* **2003**, *367*, 145.  
 (35) Namazian, M.; Halvani, S. *J. Chem. Thermodyn.* **2006**, *38*, 1495.  
 (36) Pliego, J. R.; Riveros, J. M. *Chem. Phys. Lett.* **2000**, *332*, 597.  
 (37) Reiss, H.; Heller, A. *J. Phys. Chem.* **1985**, *89*, 4207.  
 (38) Mejia, J. A.; Lago, S. *J. Chem. Phys.* **2000**, *113*, 7306.  
 (39) Zhan, C.-G.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 11534.



**FIGURE 2.** Optimized geometries for  $\text{CH}_3\text{COOH}(\text{DMSO})$ ,  $\text{CF}_3\text{COOH}(\text{DMSO})$ ,  $\text{CH}_3\text{COOH}(\text{DMSO})_2$ , and  $\text{CF}_3\text{COOH}(\text{DMSO})_2$ .

This approach bypasses the treatment of the solvated proton by using a reference anchor acid, which we term the relative determination method for the  $\text{p}K_a$  evaluation. Since the charges are the same on both sides of eq 2 and each conjugate acid/base pair ( $\text{HA}/\text{A}^-$  or  $\text{HB}/\text{B}^-$ ) is of similar size and structure, we expect favorable error cancellations in the theoretical calculation of  $\Delta G$ .

The free energy of the acid–base reaction in eq 2 can be evaluated as

$$\Delta G = [\Delta G_g(\text{HA}) - \Delta G_g(\text{HB})] + \Delta G_{\text{solv}}(\text{A}^-) + \Delta G_{\text{solv}}(\text{HB}) - \Delta G_{\text{solv}}(\text{HA}) - \Delta G_{\text{solv}}(\text{B}^-) \quad (4a)$$

It includes contributions from the gas-phase free energy and the solvation free energy, which can be written as

$$\Delta G = [\Delta G_g(\text{HA}) - \Delta G_g(\text{HB})] + \Delta\Delta G_{\text{solv}} \quad (4b)$$

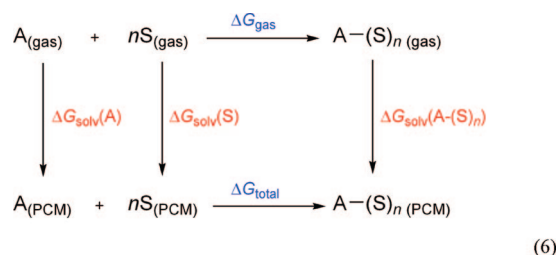
Here,  $\Delta G_g(\text{HA})$  and  $\Delta G_g(\text{HB})$  represent the gas-phase acidities of HA and HB.<sup>40</sup> The method of calculating these quantities has been given previously.<sup>12,19,21</sup>  $\Delta\Delta G_{\text{solv}}$  is the net solvation free energy change in the overall reaction in eq 2, i.e.

$$\Delta\Delta G_{\text{solv}} = \Delta G_{\text{solv}}(\text{A}^-) + \Delta G_{\text{solv}}(\text{HB}) - \Delta G_{\text{solv}}(\text{HA}) - \Delta G_{\text{solv}}(\text{B}^-) \quad (5)$$

Thus, a successful prediction of  $\text{p}K_a$  values requires the accurate computation of gas-phase acidities, which can be obtained by high-level ab initio methods, and also the solvation free energies. In this work, we use a supermolecule approach to calculate the solvation contribution.

**2.2. Solvation Energy Change: The Supermolecule/PCM Approach.** The idea of the supermolecule approach is to represent the solvated species as a cluster formed by attaching several solvent molecules to that species, and then to solvate

the cluster by using a dielectric continuum model (e.g., PCM). This process can be modeled as:



where A is the solute particle, which can be either neutral or ionic. S stands for the solvent molecule.  $\text{A}-(\text{S})_n$  represents the resulting cluster.  $\Delta G_{\text{gas}}$  refers to the gas-phase free energy change when forming the cluster.  $\Delta G_{\text{solv}}(\text{A})$ ,  $\Delta G_{\text{solv}}(\text{S})$ , and  $\Delta G_{\text{solv}}(\text{A}-(\text{S})_n)$  represent the solvation free energy of the solute A, the solvent molecule S, and the cluster  $\text{A}-(\text{S})_n$ , respectively.  $\Delta G_{\text{total}}$  is the total free energy change of the solution-phase reaction.

In the present work,  $\Delta G_{\text{gas}}$  is obtained at the MP2/6-311++G(d,p)//B3LYP/6-31+G(d) level of theory: geometry optimizations and harmonic vibrational frequency calculations are performed with the B3LYP/6-31+G(d) method, and the single-point electronic energy for an optimized geometry is computed at the MP2 level of theory with the larger 6-311++G(d,p) basis sets. For clusters (i.e.,  $\text{A}-(\text{S})_n$ ) with more than one possible conformation, geometry optimizations are performed with several possible starting structures. In the final result the one with the lowest total free energy (i.e., gas-phase free energy plus the solvation free energy calculated with the PCM approach) is taken. It is found that for neutral and cationic solutes, the most stable structure of the cluster  $\text{A}-(\text{S})_n$  usually involves hydrogen bonding between the solute and the solvent molecules, as illustrated in Figure 2. For anionic solutes, the most stable conformation of  $\text{A}-(\text{S})_n$  is formed through charge-dipole interactions as shown in Figure 3.

The number of attached solvent molecules ( $n$ ) is chosen to minimize the total free energy change (i.e.,  $\Delta G_{\text{total}}$ ) of the solution-phase reaction depicted in eq 6. The rationale for using

(40) The gas-phase acidity is the Gibbs free energy change for the following gas-phase dissociation reaction  $\text{HA}(\text{g}) \rightarrow \text{H}^+(\text{g}) + \text{A}^-(\text{g})$  at a standard state of 1 atm and 298.15 K.

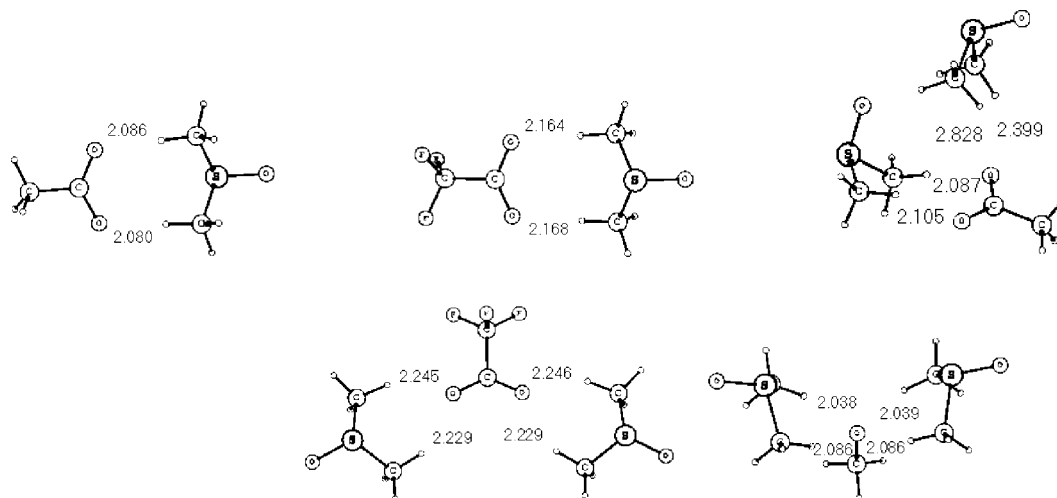


FIGURE 3. Optimized geometries for CH<sub>3</sub>COO<sup>-</sup>(DMSO), CF<sub>3</sub>COO<sup>-</sup>(DMSO), CH<sub>3</sub>COO<sup>-</sup>(DMSO)<sub>2</sub>, CF<sub>3</sub>COOH(DMSO)<sub>2</sub>, and CH<sub>3</sub>O<sup>-</sup>(DMSO)<sub>2</sub>.

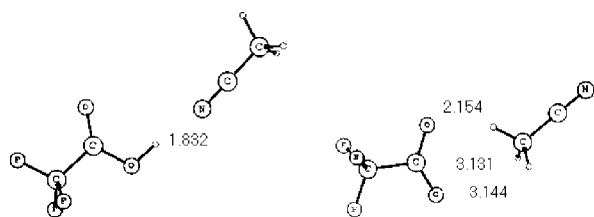


FIGURE 4. Optimized geometries for CF<sub>3</sub>COOH(MeCN), and CF<sub>3</sub>COO<sup>-</sup>(MeCN).

this criterion is to achieve the maximum calculated thermodynamic stability for the solvated species within the PCM framework. This practical approximation captures the most important interactions between the solute and the solvent molecules, which depends on the specific nature of the studied chemical species. It leaves the more averaged bulk solvent effect to the continuum solvent model and also neglects secondary effects such as clustering among solvent molecules. Other justifications of using the supermolecule approach have been given previously.<sup>28</sup>

Thus, starting with  $n = 0$ , we have performed a series of computations for all studied species in DMSO, MeCN, and THF, searching for the lowest  $\Delta G_{\text{total}}$ . To illustrate this, Table 1 lists some examples of the calculated total free energy changes when different numbers of solvent molecules are included in the supermolecule/PCM approach (also see eq 6). Optimized structures for some of the supermolecules are shown in Figures 25.

In DMSO, it is found that for a relatively strong acid, such as CF<sub>3</sub>COOH, the calculated  $\Delta G_{\text{total}}$  reaches the minimum when  $n = 1$ . In this case, the acid usually forms hydrogen bonds with the DMSO molecule, as displayed in Figure 2. This may imply that hydrogen bonding should be included in the first solvation shell when dealing with these acids. For a conjugate base, the A<sup>-</sup>(DMSO) cluster is usually formed through charge-dipole interactions, as shown in Figure 3. On the other hand,  $n = 0$  was found for a weaker acid such as CH<sub>3</sub>COOH.

In MeCN and THF, the number  $n$  is found to be zero for most species. For instance, the lowest total free energy change in forming the CF<sub>3</sub>COO<sup>-</sup>-MeCN cluster is more positive than that obtained when no MeCN molecules were attached. This indicates that within the PCM approach the effect of the explicit solvent molecules is not as significant as that in DMSO, which

TABLE 1. Calculation of  $\Delta G_{\text{total}}$  (kcal/mol), Using the Supermolecule/PCM Approach

A-(S) <sub>n</sub>	$\Delta G_{\text{total}}^a$
DMSO	
CF <sub>3</sub> COOH	0
<b>CF<sub>3</sub>COOH-(DMSO)</b>	<b>-0.89</b>
CF <sub>3</sub> COOH-(DMSO) <sub>2</sub>	6.73
CF <sub>3</sub> COO <sup>-</sup>	0
<b>CF<sub>3</sub>COO<sup>-</sup>-(DMSO)</b>	<b>-2.52</b>
CF <sub>3</sub> COO <sup>-</sup> -(DMSO) <sub>2</sub>	5.19
<b>CH<sub>3</sub>COOH</b>	<b>0</b>
CH <sub>3</sub> COOH-(DMSO)	2.88
CH <sub>3</sub> COOH-(DMSO) <sub>2</sub>	6.24
<b>CH<sub>3</sub>COO<sup>-</sup></b>	<b>0</b>
CH <sub>3</sub> COO <sup>-</sup> -(DMSO)	1.29
CH <sub>3</sub> COO <sup>-</sup> -(DMSO) <sub>2</sub>	5.34
MeCN	
picric acid	0
<b>picric acid-(MeCN)</b>	<b>-2.33</b>
picrate	0
<b>picrate-(MeCN)</b>	<b>-2.56</b>
<b>Cl<sub>2</sub>CHCOOH</b>	<b>0</b>
Cl <sub>2</sub> CHCOOH-(MeCN)	1.95
Cl <sub>2</sub> CHCOOH-(MeCN) <sub>2</sub>	9.58
<b>Cl<sub>2</sub>CHCOO<sup>-</sup></b>	<b>0</b>
Cl <sub>2</sub> CHCOO <sup>-</sup> -(MeCN)	3.85
Cl <sub>2</sub> CHCOO <sup>-</sup> -(MeCN) <sub>2</sub>	10.41
<b>CF<sub>3</sub>COO<sup>-</sup></b>	<b>0</b>
CF <sub>3</sub> COO <sup>-</sup> -(MeCN)	1.44
CH <sub>3</sub> COO <sup>-</sup> -(MeCN) <sub>2</sub>	3.46
THF	
<b>CF<sub>3</sub>COOH</b>	<b>0</b>
CF <sub>3</sub> COOH-(THF)	3.84
<b>CF<sub>3</sub>COO<sup>-</sup></b>	<b>0</b>
CF <sub>3</sub> COO <sup>-</sup> -(THF)	6.52

<sup>a</sup>  $\Delta G_{\text{total}}$  is calculated from eq 6. Units of kcal/mol and  $T = 298.15$  K. The minimum  $\Delta G_{\text{total}}$  is labeled in bold.

may be due to the smaller dielectric constants of MeCN and THF. In this case, the standard PCM model works better. It should be noted that for some very strong acids such as picric acid and MeSO<sub>2</sub>OH, the calculated solvation free energies reach a minimum when  $n = 1$ .

In summary, the use of the supermolecule/PCM approach has an important effect on the solvation of strong acids in DMSO,



FIGURE 5. Optimized geometries for  $\text{CF}_3\text{COOH}(\text{THF})$  and  $\text{CF}_3\text{COO}^-(\text{THF})$ .

while the standard PCM method works better for most acids in MeCN and THF.

**2.3. Computational Details.** All electronic structure calculations in this study were performed with the quantum chemical program package Gaussian 03.<sup>42</sup> The single-point electronic energies for a gas-phase system was calculated at the MP2/6-311++G(d,p)//B3LYP/6-31+G(d) level of theory: the standard 6-31+G(d) basis set (split-valence plus heavy-atom polarization and diffuse functions) was used for the geometry optimization of a neutral molecule, an ion, or a cluster, and the electronic energy for the optimized geometry was obtained at the MP2 level of theory with the larger 6-311++G(d,p) basis sets. The thus obtained electronic energy was then combined with the entropic contribution, calculated at the B3LYP/6-31+G(d) level of theory, to give the corresponding gas-phase free energy of the system. The latter contains vibration (with a harmonic approximation, using the calculated vibrational frequencies), rotation (with a rigid rotor approximation), and translation, which is part of the standard output of Gaussian 03 after a frequency analysis. For comparison purposes, the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) method and the B3LYP/6-311++G(d,p) method were also employed. In the former method, the geometry optimization and harmonic vibrational frequency calculation were performed at the B3LYP/6-31+G(d) level of theory and the single-point electronic energy for the optimized geometry was obtained at the same level of theory but with a large basis, i.e., 6-311++G(d,p). In the latter method, geometry optimizations and harmonic vibrational frequency calculations were performed at the B3LYP/6-311++G(d,p) level of theory (single-point electronic energy was also obtained at this level).

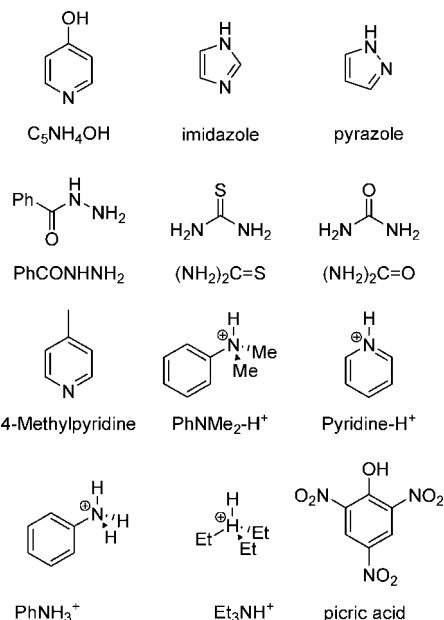
Solvation free energies for the solvent continuum were calculated by the standard PCM approach implemented in Gaussian 03 with use of the B3LYP/6-31+G(d) method. Bondi radii<sup>43</sup> were employed with the default scale factors for the atomic radii in DMSO, MeCN, and THF solvents. The gas-phase optimized geometries are used for all of the solution-phase calculations. It has been demonstrated that the change of geometry by the solvation effect is usually not significant,<sup>11–14</sup> which is also confirmed by our test calculations.

(41) Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 792.

(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. *Gaussian 03*, Revision C.02; Gaussian, Inc., Wallingford, CT, 2004.

(43) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

CHART 1. Structures of Selected Acids



### 3. Results and Discussion

**3.1. Gas-Phase Acidities.** To assess the accuracy of gas-phase free energy calculations, we have calculated the gas-phase acidities, eq 4b, of all the compounds considered in this study using the MP2/6-311++G(d,p)//B3LYP/6-31+G(d) method. The calculated gas-phase acidities<sup>40</sup> for 33 acids are listed in Table 2 (the structures of some of these acids are shown in Chart 1). Available experimental values are also included, taken from the NIST WebBook.<sup>44</sup> As evidenced in Table 2, the MP2/6-311++G(d,p)//B3LYP/6-31+G(d) method gives excellent gas-phase acidity prediction with a root-mean-square error of 2.0 kcal/mol and a mean-absolute error of 1.5 kcal/mol. To further evaluate the accuracy of using the present method, a linear regression is made by plotting all the calculated values versus experimental data. As can be seen in Figure 6, the slope of the regression line is close to unity, and the correlation coefficient ( $r = 0.999$ ) is very high. Careful inspection of Table 2 shows that the deviation of the calculated result usually ranges from about 0 to 2 kcal/mol from experiment. Considering that the typical error bar of experimental values is around 2 kcal/mol, our calculated results are acceptable. However, it is noteworthy that there are three instances of large deviation: picric acid (4.9 kcal/mol), imidazole (4.1 kcal/mol), and MeCN (3.2 kcal/mol). These large deviations may lead to an error of 2.3–4.6 pK<sub>a</sub> units in later pK<sub>a</sub> calculations.

As a comparison, we also applied the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) method in the gas-phase acidity calculations. Geometry optimizations and harmonic vibrational frequency calculations were carried out with the B3LYP/6-31+G(d) method, followed by the single-point energy calculation at the B3LYP/6-311++G(d,p) level of theory. The calculated results are also listed in Table 2. Compared to the previous approach, the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) method leads to a greater deviation from experiment, with a root-mean-square error of 3.4 kcal/mol and a mean-absolute error of 2.4 kcal/

(44) Bartmess, J. E. Negative Ion Energetics Data In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 20899; July 2001 (<http://webbook.nist.gov>).

TABLE 2. Experimental and Theoretical Gas-Phase Acidities

acid	exptl <sup>a</sup>	gas-phase acidity (kcal/mol)		
		MP2/6-311++G**// B3LYP/6-31+G* <sup>b</sup>	B3LYP/6-311++G**// B3LYP/6-31+G* <sup>c</sup>	B3LYP/6-311++G** <sup>d</sup>
picric acid	302.8 ± 2.0	297.90	293.28	293.72
MeSO <sub>2</sub> OH	315 ± 2.0	313.99	312.43	312.67
CF <sub>3</sub> COOH	317.4 ± 2.0	316.69	311.05	311.47
2,4-(NO <sub>2</sub> ) <sub>2</sub> -phenol		311.58	306.84	304.78
PhCOSH		326.02	323.95	322.72
Cl <sub>3</sub> CHCOOH	321.9 ± 2.0	320.62	316.57	316.35
<i>p</i> -NO <sub>2</sub> -C <sub>6</sub> H <sub>4</sub> OH	320.9 ± 2.0	320.55	314.16	314.71
PhCOOH	333 ± 2.0	331.98	331.10	330.96
CH <sub>3</sub> COOH	341.1 ± 2.0	341.15	337.62	338.69
HCN	343.6 ± 0.30	343.71	341.88	341.86
(CH <sub>3</sub> CO) <sub>2</sub> CH <sub>2</sub>		334.88	330.90	330.62
C <sub>5</sub> NH <sub>4</sub> OH		326.31	328.28	328.29
PhCH <sub>2</sub> SH		344.06	340.58	340.85
NC-NH <sub>2</sub>	344.1 ± 2.0	342.40	339.64	339.81
<i>n</i> -BuSH	347.4 ± 2.0	349.28	345.80	345.89
PhOH	342.3 ± 2.0	341.20	339.99	339.00
imidazole	342.8 ± 2.0	338.69	342.11	342.10
PhCONHNH <sub>2</sub>		335.61	335.51	335.38
pyrazole	346.4 ± 2.0	344.28	347.24	347.20
(NH <sub>2</sub> ) <sub>2</sub> C=S		339.54	337.40	339.63
CF <sub>3</sub> CH <sub>2</sub> OH	354.1 ± 2.0	353.94	348.96	349.86
H <sub>2</sub> NCOOEt	354.9 ± 2.2	354.46	353.93	353.97
PhCOCH <sub>3</sub>	354.5 ± 2.0	356.06	353.99	354.11
CH <sub>3</sub> C(O)NH <sub>2</sub>	355.0 ± 2.0	354.08	355.58	355.96
CH <sub>3</sub> C(O)CH <sub>3</sub>	361.9 ± 2.0	364.04	361.18	361.37
(NH <sub>2</sub> ) <sub>2</sub> C=O	355.7 ± 3.2	355.40	355.19	355.27
CH <sub>3</sub> OH	375.1 ± 1.1	375.94	371.80	371.60
MeCN	365.2 ± 2.0	368.45	363.38	363.47
DMSO	366.4 ± 2.0	369.10	368.37	368.51
4-methylpyridine	362.9 ± 5.0	366.81	363.27	362.95
PhNMe <sub>2</sub> -H <sup>+</sup>	217.3	219.02	216.25	216.25
pyridine-H <sup>+</sup>	214.7	213.06	215.72	215.70
PhNH <sub>3</sub> <sup>+</sup>	203.3	204.39	203.23	203.89
Et <sub>3</sub> N-H <sup>+</sup>	227.0	226.27	226.43	226.44
root-mean-square error		2.0	3.4	3.2
mean absolute error		1.5	2.4	2.4

<sup>a</sup> Experimental gas-phase acidities are taken from NIST Standard Reference Database 69, March 1998 Release: *NIST Chemistry WebBook* (data compiled by J. E. Bartmess). <sup>b</sup> MP2/6-311++G(d,p) single-point calculation based on optimized geometries obtained by the B3LYP/6-31+G(d) method. <sup>c</sup> B3LYP/6-311++G(d,p) single-point calculation based on optimized geometries obtained by the B3LYP/6-31+G(d) method. <sup>d</sup> Geometry optimization and harmonic vibrational frequency calculation at the B3LYP/6-311++G(d,p) level of theory.

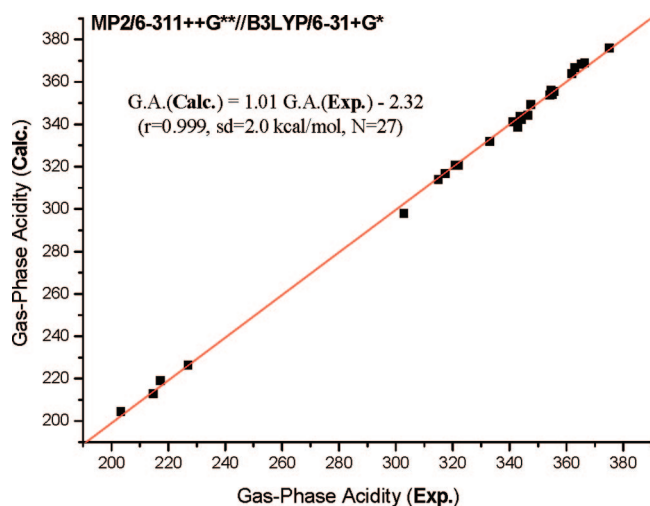


FIGURE 6. Correlation of the calculated gas-phase acidities (GA) and experimental values.

mol. Moreover, the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) method gives erroneous gas-phase acidity predictions for

carboxylic acids and alcohols: the calculated values are systematically lowered by 2 to 6 kcal/mol, and the largest deviation is observed to be 9.5 kcal/mol for picric acid. To assess the effect of using different level theories on the geometry optimizations and the vibrational frequency calculations, we have also performed full B3LYP/6-311++G(d,p) calculations. As demonstrated in Table 2, this method gives similar results as compared to the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) method. All these indicate that the B3LYP/6-311++G(d,p) level of theory is not as reliable for gas-phase acidity predictions as the MP2/6-311++G(d,p)//B3LYP/6-31+G(d) method.

The accuracy of calculated gas-phase acidities can be further improved by employing higher level theories and larger basis sets, such as G3 and CBS-Q approaches. However, these methods are quite demanding computationally and will not significantly increase the accuracy. For example, the error of the G3 and CBS-Q calculations is about 1–2 kcal/mol,<sup>45</sup> which is not much smaller than the 2.0 kcal/mol obtained from our approach.

(45) Curtiss, L. A.; Redfern, K. P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374.

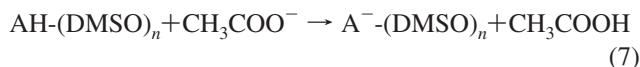
In summary, the economical MP2/6-311++G(d,p)//B3LYP/6-31+G(d) method allows us to accurately estimate the gas-phase acidities, which can be finally used in the  $pK_a$  calculations.

**3.2. Theoretical  $pK_a$  Values in DMSO.** The largest set of  $pK_a$  data in nonaqueous solutions is for DMSO solvent, mainly due to the efforts of Bordwell.<sup>46</sup> The importance of these data is illustrated by the many thermodynamic investigations that use these  $pK_a$  values, e.g., bond energy determinations for organic molecules,<sup>5g</sup> organometallic complexes<sup>47</sup> and metalloenzyme model complexes.<sup>48</sup>

With reliable theoretical gas-phase acidities in hand, the next crucial step for the  $pK_a$  estimation is the solution-phase calculation. In this work, we apply the supermolecule/PCM approach (Figure 1) to calculate the  $pK_a$  values in DMSO. As discussed in section 2.2, a crucial part of this approach is to determine the number of attached DMSO molecules.

For carboxylic acids (i.e., picric acid,  $\text{MeSO}_2\text{OH}$ ,  $\text{CF}_3\text{COOH}$ ,  $\text{Cl}_2\text{CHCOOH}$ ,  $p\text{-NO}_2\text{C}_6\text{H}_4\text{OH}$ , and  $\text{PhCOOH}$ ), some alcohols (i.e.,  $\text{PhOH}$  and  $\text{CF}_3\text{CH}_2\text{OH}$ ), and some other acids (i.e.,  $\text{PhCOSH}$ ,  $\text{HCN}$ ,  $\text{NH}_2\text{CSNH}_2$ ,  $\text{NH}_2\text{CONH}_2$ , and pyrazole), the calculated total free energy change when forming the cluster reaches a minimum for  $n = 1$ . In this case, the supermolecule/PCM method is employed by attaching one DMSO molecule onto each conjugate acid/base pair. For other acids, the number of attached solvent molecules was found to be  $n = 0$ . In this case, the standard PCM approach is used.

For the specific case of DMSO as solvent, we choose acetic acid ( $\text{CH}_3\text{COOH}$ ) as the reference acid for  $pK_a$  estimation ( $\text{HB} = \text{CH}_3\text{COOH}$  in eq 2).



where  $n = 1$  or 0 depending on the number of attached DMSO molecules determined in the supermolecule/PCM approach.

Thus, we focus on the net solvation free energy change ( $\Delta\Delta G_{\text{solv}}$ ) of eq 7, which can be obtained by calculating the gas-phase clustering energies and the solvation free energies of the clusters. The  $pK_a$  of any unknown acid HA is then determined by combining the gas-phase acidities and the net solvation free energy change ( $\Delta\Delta G_{\text{solv}}$ ) together (also see eqs 4a, 4b, and 5).

With the supermolecule/PCM approach, we have calculated the  $pK_a$  values of 33 organic acids in DMSO. The detailed results are summarized in Table 3 (column 3). Experimental values taken from Bordwell's  $pK_a$  data<sup>46</sup> are also listed for comparison. As shown in Table 3, the predicted  $pK_a$  values are in good agreement with the experimental data. The absolute error ranges from  $\sim 0$  to 1.6  $pK_a$  units. The root-mean-square error is observed to be 0.91  $pK_a$  units and the mean-absolute error is only 0.79  $pK_a$  units. To further analyze the correlation of the predicted  $pK_a$  values and experimental values, a linear regression is established in Figure 7. As evidenced in Figure 7, excellent correlation is found between the theoretical and experimental data. The straight line indicates that the best linear regression fits to an expression (see eq 8) with a slope close to unity and a very high correlation coefficient ( $r = 0.996$ ).

$$pK_a^{\text{DMSO}}(\text{calcd}) = 1.02 pK_a^{\text{DMSO}}(\text{exptl}) - 0.11 \quad (8)$$

The accuracy of the calculated results, also referred to as systematic error, is another important factor that affects the ability of the theory to predict  $pK_a$  values. Previous calculations

**TABLE 3.** Comparison between Calculated and Experimental  $pK_a$  Values in DMSO (Theoretical  $pK_a$  Values Are Obtained by Using the Supermolecule/PCM Approach)

acid	$pK_a^{\text{DMSO}}(\text{exptl})^a$	$pK_a^{\text{DMSO}}(\text{calcd})$	
		$\text{CH}_3\text{COOH}$ anchor <sup>b</sup>	$\text{H}^+$ anchor <sup>c</sup>
picric acid	$\sim 0$	-1.02	-2.86
$\text{MeSO}_2\text{OH}$	1.6	0.68	-1.16
$\text{CF}_3\text{COOH}$	3.45	4.00	2.16
2,4-( $\text{NO}_2$ ) <sub>2</sub> -phenol	5.1	5.20	3.37
$\text{PhCOSH}$	5.2	6.46	4.62
$\text{Cl}_2\text{CHCOOH}$	6.4	7.29	5.45
$p\text{-NO}_2\text{-C}_6\text{H}_4\text{OH}$	10.8	10.89	9.06
$\text{PhCOOH}$	11.0	11.98	10.14
$\text{CH}_3\text{COOH}$	12.3		10.46
$\text{HCN}$	12.9	14.31	12.47
$(\text{CH}_3\text{CO})_2\text{CH}_2$	13.3	14.07	12.23
$\text{C}_5\text{NH}_4\text{OH}$	14.8	13.71	11.87
$\text{PhCH}_2\text{SH}$	15.4	15.32	13.48
$\text{NC-NH}_2$	16.9	16.82	14.98
$n\text{-BuSH}$	17.0	16.92	15.08
$\text{PhOH}$	18.0	17.01	15.18
imidazole	18.6	17.97	16.13
$\text{PhCONHNH}_2$	18.9	16.98	15.14
pyrazole	19.8	19.29	17.45
$(\text{NH}_2)_2\text{C=S}$	21.1	22.48	20.64
$\text{CF}_3\text{CH}_2\text{OH}$	23.45	23.18	21.34
$\text{H}_2\text{NCOOEt}$	24.6	25.30	23.46
$\text{PhCOCH}_3$	24.7	25.72	23.89
$\text{CH}_3\text{C(O)NH}_2$	25.5	24.56	22.72
$\text{CH}_3\text{C(O)CH}_3$	26.5	27.56	25.73
$(\text{NH}_2)_2\text{C=O}$	26.95	26.58	24.75
$\text{CH}_3\text{OH}$	29.0	29.87	28.03
$\text{MeCN}$	31.3	30.99	29.15
DMSO	35	36.57	34.73
4-methylpyridine	35	36.01	34.17
$\text{PhNMe}_2\text{-H}^+$	2.4	2.42	0.59
pyridine- $\text{H}^+$	3.4	4.13	2.30
$\text{PhNH}_3^+$	3.6	2.19	0.35
$\text{Et}_3\text{N-H}^+$	9	10.03	8.20
root-mean-square error		0.91	1.9
mean absolute error		0.79	1.7

<sup>a</sup> Experimental  $pK_a^{\text{DMSO}}$  values taken from ref 46. <sup>b</sup> Calculated  $pK_a^{\text{DMSO}}$  values with use of the relative determination method (eq 7). Acetic acid is chosen as the reference acid. <sup>c</sup> Absolute determination method,  $pK_a$  calculated with eq 9.

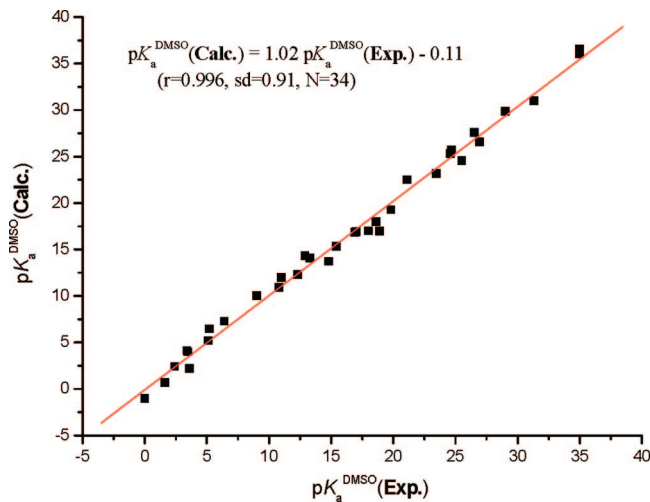
in DMSO solution usually observe a systematic deviation of several  $pK_a$  units from experimental data.<sup>19,29</sup> This systematic deviation usually comes from the computational error of the chosen reference. For instance, in Chipman's study<sup>19</sup> of several acids in DMSO and MeCN using the absolute determination method (i.e., eqs 1a and 1b), a significant shift of about 5  $pK_a$  units was observed between the calculated and the experimental values. He assigned this systematic error to the uncertainties in the proton solvation free energy. It should also be mentioned that the polarizable continuum method was used in his study. To improve the accuracy, Guo et al.<sup>29</sup> applied the cluster-continuum model proposed by Pliego and Riveros<sup>28</sup> in the  $pK_a$  estimation in DMSO. Although good correlation was found between their calculated  $pK_a$  values and experimental data, there still exists a systematic deviation of about 2  $pK_a$  units which, in our point of view, is mainly related to the use of DMSO/

(46) Bordwell, F. G. *Acc. Chem. Res.* **1988**, *21*, 456.

(47) Trujillo, H. A.; Casado, C. M.; Ruiz, J.; Astruc, D. *J. Am. Chem. Soc.* **1999**, *121*, 5674.

(48) (a) Goldsmith, C. R.; Stack, T. D. P. *Inorg. Chem.* **2006**, *45*, 6048. (b) Gupta, R.; Borovik, A. S. *J. Am. Chem. Soc.* **2003**, *125*, 13234.





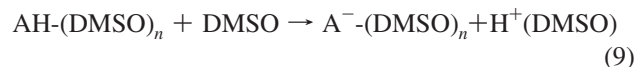
**FIGURE 7.** Correlation of the calculated pK<sub>a</sub> and experimental pK<sub>a</sub> values in DMSO. Theoretical values are obtained by using the *relative* determination method. The reference acid is acetic acid (pK<sub>a</sub><sup>DMSO</sup>(CH<sub>3</sub>COOH) = 12.3).

MeSOCH<sub>2</sub><sup>−</sup> as the reference. Since in their work the gas-phase acidity of DMSO was overestimated by 2.7 kcal/mol, the calculated pK<sub>a</sub> values were thus underestimated by 2 pK<sub>a</sub> units.

In an ideal situation, if all the quantities (i.e., gas-phase acidities and solvation-free energies) can be evaluated exactly, there will be no systematic shift in calculated results. Due to the limitations of the current theories and models, the systematic error cannot be eliminated completely. However, the computational error can be minimized to some extent, using either higher level theories or appropriate modeling strategies. Thus, to induce appropriate, favorable error cancelations, we use the relative determination method in which the experimental pK<sub>a</sub> value of the anchor compound is known. For the pK<sub>a</sub> prediction in DMSO solvent, acetic acid is chosen as the reference since it has a moderate pK<sub>a</sub>, which allows us to evaluate its gas-phase acidity (using the MP2/6-311++G(d,p)//B3LYP/6-31+G(d) method) and solvation free energies (using the supermolecule/PCM approach) with reasonable accuracy. As can be noticed in Figure 7, a small systematic shift of −0.11 pK<sub>a</sub> units is obtained by using this strategy. On the other hand, the computational results do not depend sensitively on the choice of the anchor acid in DMSO. As demonstrated in Table 3, the largest absolute error between simulation and experiment is 1.6 pK<sub>a</sub> units, which indicates that changing the anchor acid may induce at most a systematic shift of 1.6 pK<sub>a</sub> units.

To summarize, in this study experimental pK<sub>a</sub> values in DMSO are reproduced with high precision (0.91 pK<sub>a</sub> units) and high accuracy (−0.11 pK<sub>a</sub> units) by using a consistent supermolecule/PCM approach. It is important to note that a total of 33 organic acids with various functional groups are considered in this work and there is no empirical fitting or utilization of extra approximations in our calculations. The experimental pK<sub>a</sub> of acetic acid is the only external value used in the above investigations.

A more challenging task is to predict pK<sub>a</sub> purely by computations, i.e., without any external or experimental data. In the present work, we consider the simplest solvated form of the proton, i.e., H<sup>+</sup>(DMSO). The pK<sub>a</sub> of any acid HA is thus calculated via eq 9.



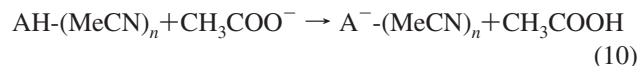
where  $n = 1$  or 0 depending on the optimal number of attached DMSO molecules determined in the supermolecule approach.

The calculated results are collected in Table 3 (last column). Compared to the previous relative determination method, this approach leads to a larger deviation from the experiment. The root-mean-square error and the mean-absolute error are observed to be 1.9 and 1.7 pK<sub>a</sub> units correspondingly. All calculated values are systematically lowered by about 1.7 pK<sub>a</sub> units from experimental data. This systematic shift may mainly come from the inaccurate calculation of the solvation free energies: the solvation free energy of DMSO is underestimated by 0.9 kcal/mol and the solvation free energy of H<sup>+</sup>(DMSO) is underestimated by 3.2 kcal/mol.<sup>49</sup>

**3.3. Theoretical pK<sub>a</sub> Values in Acetonitrile.** Acetonitrile (MeCN) has many properties that make it a suitable medium for acid–base studies. It has a low basicity and a very low ability to solvate anions.<sup>50</sup> The low basicity gives acetonitrile an advantage in studies of strong acids over DMSO, which is considerably more basic (stronger hydrogen-bond acceptor). It also has a high dielectric constant ( $\epsilon = 36.0$ ), which favors the dissociation of ion pairs into free ions. Thus, acetonitrile is a widely used solvent in chemistry, and is commonly used as a solvent for thermodynamic studies that require pK<sub>a</sub> data.<sup>3–5</sup>

There is less experimental pK<sub>a</sub> data in MeCN than in DMSO and for some acids there are discrepancies in the experimental pK<sub>a</sub>.<sup>31,32,51</sup> Little computational work has been done on the prediction of pK<sub>a</sub> values in MeCN,<sup>19,21</sup> and those calculated pK<sub>a</sub> values are not as accurate as experimental data. Therefore, a reliable, self-consistent and universally applicable prediction method is highly desirable, yet challenging.

With the success of using the relative determination method in DMSO, we further apply this protocol to calculate pK<sub>a</sub> values in MeCN. The reference acid is again chosen as acetic acid (CH<sub>3</sub>COOH), whose experimental pK<sub>a</sub><sup>MeCN</sup> is found to be 22.3.<sup>52</sup> The supermolecule/PCM method is employed to determine the net solvation free energy change of the exchange reaction (eq 10).



where  $n = 1$  for picric acid, MeSO<sub>2</sub>OH, and 2,4-(NO<sub>2</sub>)<sub>2</sub>-phenol; and  $n = 0$  for the other acids.

The pK<sub>a</sub> of each acid is obtained by combining the calculated net solvation free energy change and the gas-phase acidities. Values calculated from this approach are summarized in Table 4 (column 3). Experimental pK<sub>a</sub> values are found for picric acid (11.0),<sup>53</sup> *p*-NO<sub>2</sub>-C<sub>6</sub>H<sub>4</sub>OH (22.85),<sup>51</sup> benzoic acid (20.7),<sup>52</sup> phenol (29.1),<sup>54</sup> PhNMe<sub>2</sub>-H<sup>+</sup> (11.4), pyridine-H<sup>+</sup> (12.5), PhNH<sub>3</sub><sup>+</sup> (10.6), and Et<sub>3</sub>N-H<sup>+</sup> (18.8).<sup>55</sup> From Table 4, it is clearly seen that the

(49) The solvation free energy of H<sup>+</sup>(DMSO) is calculated as −66.4 kcal/mol, while the experimental value is −69.6 kcal/mol (taken from ref 22). The solvation free energy of DMSO is calculated as −6.8 kcal/mol, with a deviation of 0.9 kcal/mol from the experimental value of −7.7 kcal/mol (ref 22). Totally, the solvation free energy is underestimated by (3.2 kcal/mol − 0.9 kcal/mol) = 2.3 kcal/mol, which leads to a shift of 1.7 pK<sub>a</sub> units.

(50) Coetzee, J. F. *Prog. Phys. Org. Chem.* **1967**, *4*, 45.

(51) Kütt, A.; Leito, I.; Kaljurand, I.; Sooväli, L.; Vlasov, V. M.; Yagupolskii, L. M.; Koppel, I. A. *J. Org. Chem.* **2006**, *71*, 2829.

(52) Kolthoff, I. M.; Chantooni, M. K., Jr. *J. Am. Chem. Soc.* **1975**, *97*, 1376.

(53) Kolthoff, I. M.; Chantooni, M. K., Jr. *J. Am. Chem. Soc.* **1965**, *87*, 4428.

**TABLE 4.** Calculated  $pK_a$  Values in MeCN and Corresponding Experimental Data (Theoretical  $pK_a$  Values Are Obtained by Using the Supermolecule/PCM Approach)

acid	$pK_a^{\text{MeCN}}(\text{exptl})$	$pK_a^{\text{MeCN}}(\text{calcd})$	
		$\text{CH}_3\text{COOH}$ anchor <sup>a</sup>	$\text{H}^+$ anchor <sup>b</sup>
picric acid	11.0 <sup>c</sup>	11.44	14.92
MeSO <sub>2</sub> OH		12.06	13.12
CF <sub>3</sub> COOH		15.82	16.88
2,4-(NO <sub>2</sub> ) <sub>2</sub> -phenol	16.66 <sup>d</sup>	16.33	19.81
PhCOSH		18.37	19.43
Cl <sub>2</sub> CHCOOH		19.23	20.29
<i>p</i> -NO <sub>2</sub> -C <sub>6</sub> H <sub>4</sub> OH	22.85 <sup>d</sup>	22.60	23.66
PhCOOH	20.7 <sup>e</sup>	20.32	21.38
CH <sub>3</sub> COOH	22.3 <sup>e</sup>		23.36
HCN		23.40	24.46
(CH <sub>3</sub> CO) <sub>2</sub> CH <sub>2</sub>		25.88	26.94
C <sub>5</sub> NH <sub>4</sub> OH		25.65	26.71
PhCH <sub>2</sub> SH		26.32	27.38
NC-NH <sub>2</sub>		26.07	27.13
<i>n</i> -BuSH		27.61	28.67
PhOH	29.1 <sup>f</sup>	28.40	28.37
imidazole		29.26	30.32
PhCONHNH <sub>2</sub>		29.02	30.08
pyrazole		29.61	30.67
(NH <sub>2</sub> ) <sub>2</sub> C=S		29.56	30.62
CF <sub>3</sub> CH <sub>2</sub> OH		32.94	34.00
H <sub>2</sub> NCOOEt		34.01	35.07
PhCOCH <sub>3</sub>		36.62	37.68
CH <sub>3</sub> C(O)NH <sub>2</sub>		35.31	36.37
CH <sub>3</sub> C(O)CH <sub>3</sub>		38.29	39.35
(NH <sub>2</sub> ) <sub>2</sub> C=O		34.79	35.85
CH <sub>3</sub> OH		39.37	40.43
MeCN		42.32	43.38
DMSO		48.20	49.26
4-methylpyridine		47.19	48.25
PhNMe <sub>2</sub> -H <sup>+</sup>	11.4 <sup>g</sup>	13.39	14.45
pyridine-H <sup>+</sup>	12.5 <sup>g</sup>	15.37	16.43
PhNH <sub>3</sub> <sup>+</sup>	10.6 <sup>g</sup>	12.99	14.05
Et <sub>3</sub> N-H <sup>+</sup>	18.8 <sup>g</sup>	21.07	22.13
root-mean-square error		1.5	2.7
mean absolute error		1.2	2.4

<sup>a</sup> Calculated  $pK_a^{\text{MeCN}}$  values from the relative determination method. Acetic acid ( $\text{CH}_3\text{COOH}$ ) is chosen as the reference acid. <sup>b</sup> Absolute determination method. <sup>c</sup> Reference 53. <sup>d</sup> Reference 51. <sup>e</sup> Reference 52. <sup>f</sup> Reference 54. <sup>g</sup> Reference 55.

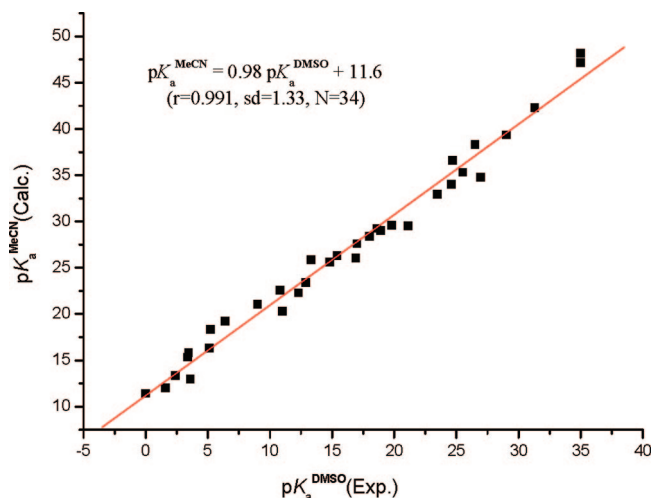
theoretical results agree well with experiment for neutral acids. In particular, the calculated  $pK_a$  values for picric acid (11.4), *p*-NO<sub>2</sub>-C<sub>6</sub>H<sub>4</sub>OH (22.6), benzoic acid (20.3), and phenol (28.4) are quite consistent with the experimental values, with the average error for these neutral acids is less than 0.4  $pK_a$  units. For 2,4-(NO<sub>2</sub>)<sub>2</sub>-phenol, experimental  $pK_a$  value ranges from 15.3<sup>31</sup> to 18.4,<sup>32</sup> and our calculated value ( $pK_a^{\text{calcd}} = 16.3$ ) falls into this range. Most recently, Leito et al.<sup>51</sup> have reported a  $pK_a$  value of 16.66 for 2,4-(NO<sub>2</sub>)<sub>2</sub>-phenol, based on a self-consistent UV-vis spectrophotometric measurement in MeCN. Again, our predicted result agrees well with the experiment.

For cationic acids, i.e., PhNMe<sub>2</sub>-H<sup>+</sup>, pyridine-H<sup>+</sup>, PhNH<sub>3</sub><sup>+</sup>, and Et<sub>3</sub>N-H<sup>+</sup>, the predicted  $pK_a$  values have somewhat large deviations from the experiments by ~2.2  $pK_a$  units. Although this is not bad agreement when the error bars in both theory

**TABLE 5.** Calculated  $pK_a$  Values in MeCN for Cationic Acids with Use of the reference of PhNH<sub>3</sub><sup>+</sup>/PhNH<sub>2</sub>

cationic acid	$pK_a^{\text{MeCN}}(\text{calcd})$	
	$pK_a^{\text{MeCN}}(\text{exptl})^a$	PhNH <sub>3</sub> <sup>+</sup> anchor <sup>b</sup>
PhNMe <sub>2</sub> -H <sup>+</sup>	11.4	11.02
pyridine-H <sup>+</sup>	12.5	13.00
PhNH <sub>3</sub> <sup>+</sup>	10.6	
Et <sub>3</sub> N-H <sup>+</sup>	18.8	18.70

<sup>a</sup> Reference 55. <sup>b</sup> Calculated  $pK_a^{\text{MeCN}}$  with the relative determination method. PhNH<sub>3</sub><sup>+</sup> is chosen as the reference acid.

**FIGURE 8.** Correlation between  $pK_a^{\text{MeCN}}(\text{calcd})$  and  $pK_a^{\text{DMSO}}(\text{exptl})$ .

and experiment are considered, the situation can be further improved by noting that the systematic shift may be assigned to the inaccurate calculation for solvation free energies of  $\text{CH}_3\text{COOH}/\text{CH}_3\text{COO}^-$ . As depicted in eq 10, if HA is a neutral acid, we can expect a favorable error cancelation since the charges are the same on both sides of eq 10 and each conjugate acid/base pair is of similar size and structure. For cationic acids, the error in the calculated solvation free energies of  $\text{CH}_3\text{COOH}/\text{CH}_3\text{COO}^-$  is not eliminated due to the asymmetry in charge distribution in eq 10. This problem can be avoided by using a cationic acid as the reference instead of  $\text{CH}_3\text{COOH}$ . Thus we have performed additional calculations for cationic acids using PhNH<sub>3</sub><sup>+</sup> as the reference, and the results are listed in Table 5. As evidenced in Table 5, the calculated values for cationic acids are in excellent agreement with experiment.

The ultimate goal of our study is to provide a valuable predictive tool for acid/base chemistry in nonaqueous solutions. In this work, we try to establish a reliable correlation of the acidity scales in MeCN and DMSO, since experimental  $pK_a$  values in DMSO are usually available and there are many unknowns for experimental  $pK_a$  values in MeCN. With confidence in our theoretical approach, we plot the calculated  $pK_a^{\text{MeCN}}$  values for all acids ( $\text{CH}_3\text{COOH}$  anchor) against the corresponding experimental  $pK_a^{\text{DMSO}}$  values (see Figure 8). A least-squares fit line is found, giving the following correlation expression (eq 11):

$$pK_a^{\text{MeCN}} = 0.98 pK_a^{\text{DMSO}} + 11.6 \quad (11)$$

As shown in Figure 8, the correlation coefficient is high ( $r = 0.991$ ) and the standard deviation is small (1.33  $pK_a$  units). This correlation can be verified by comparing the values

(54) Kütt, A.; Movchun, V.; Rodima, T.; Dansauer, T.; Rusanov, E. B.; Leito, I.; Kaljurand, I.; Koppel, J.; Pihl, V.; Koppel, I.; Ovsjannikov, G.; Toom, L.; Mishima, M.; Medebielle, M.; Lork, E.; Rösenthaller, G.; Koppel, I. A.; Kolomeitsev, A. A. *J. Org. Chem.* **2008**, *73*, 2607.

(55) Kaljurand, I.; Kütt, A.; Sooväli, L.; Rodima, T.; Mäemets, V.; Leito, I.; Koppel, I. A. *J. Org. Chem.* **2005**, *70*, 1019.

obtained from eq 11 with the experimental pK<sub>a</sub><sup>MeCN</sup> values. For example, the pK<sub>a</sub> values of picric acid, acetic acid, and phenol in MeCN are estimated as 11.6, 23.6, and 29.2 correspondingly, agreeing well with the experimental values (11.0, 22.3, and 29.1). Our results can be compared with experimental measurements of pK<sub>a</sub> values in MeCN using UV-vis spectrophotometry by Leito et al.,<sup>51</sup> who established a correlation between pK<sub>a</sub><sup>MeCN</sup> and pK<sub>a</sub><sup>DMSO</sup> with an expression of pK<sub>a</sub><sup>MeCN</sup> = 0.98 pK<sub>a</sub><sup>DMSO</sup> + 12.3. Gratifyingly, our correlation (eq 11) is quite consistent with that proposed from experimental results.

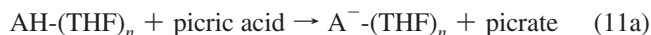
For comparison purposes, the absolute determination method (see eq 1a) is also employed to calculate the pK<sub>a</sub><sup>MeCN</sup> values. In this approach, a simple solvated form of the proton, H<sup>+</sup>-(MeCN), is considered, and the results are summarized in Table 4 (last column). As seen from Table 4, the error ranges from 1 to about 4 pK<sub>a</sub> units, and the largest deviation is observed to be 3.9 pK<sub>a</sub> units for picric acid. The root-mean-square error is as large as 2.7 pK<sub>a</sub> units. This method leads to large deviation from the experimental data.

**3.4. Theoretical pK<sub>a</sub> Values in THF.** There is considerable interest in experimental acidity studies in THF (tetrahydrofuran)<sup>33,56–58</sup> due to its relatively low reactivity, especially toward basic compounds. However, THF has a very low dielectric constant ( $\epsilon = 7.6$ ) and thus the ion-pairs are not dissociated as completely as those in DMSO and MeCN. Experimentally the absolute pK<sub>a</sub><sup>THF</sup> values are very difficult to obtain due to the complication of ion-pairing in THF. A number of measures for acidity in THF have been proposed. Streitwieser and Antipin have introduced pK<sub>fi</sub>, which measures ion-pair acidities in THF.<sup>57</sup> Morris introduced pK <sub>$\alpha$</sub> <sup>THF</sup> as a measure of the acidity of metal hydrides and phosphines in THF; pK <sub>$\alpha$</sub> <sup>THF</sup> is an approximation to pK<sub>a</sub><sup>THF</sup> that uses the Fuoss equation<sup>59</sup> to account for the ion-pairing that occurs in THF solution.<sup>58</sup> Although in principle pK<sub>a</sub><sup>THF</sup> and pK <sub>$\alpha$</sub> <sup>THF</sup> should be equal, there may be disagreement because of errors in the corrections for ion-pairing.<sup>60</sup> Absolute pK<sub>a</sub><sup>THF</sup> measurements for a number of acids have been used to anchor the pK <sub>$\alpha$</sub> <sup>THF</sup> scale to the absolute pK<sub>a</sub><sup>THF</sup> scale.<sup>61</sup>

In contrast, pK<sub>a</sub><sup>THF</sup> values can be determined theoretically since the ion-pairs can be treated in complete dissociation form. To the best of our knowledge, there are no theoretical predictions of the pK<sub>a</sub> values in THF. At this stage of development, calculation of pK<sub>a</sub><sup>THF</sup> is still challenging due to the lack of experimental data for comparison.

In this present study, we apply the previous approach to calculate the pK<sub>a</sub> values in THF. Furthermore, we want to establish in this paper a reliable correlation between pK<sub>a</sub><sup>THF</sup> and pK<sub>a</sub><sup>DMSO</sup>, aiming to serve as a valuable tool for future researchers in acid–base chemistry in THF.

The same set of acids as studied in DMSO and MeCN is also investigated in THF. The relative determination method is employed: picric acid is chosen as the reference for calculating pK<sub>a</sub><sup>THF</sup> values of *neutral acids*; anilinium (PhNH<sub>3</sub><sup>+</sup>) is chosen as the reference for *cationic acids*. The experimental pK<sub>a</sub><sup>THF</sup> of picric acid is found to be 11.6,<sup>56</sup> and the experimental pK<sub>a</sub><sup>THF</sup> of PhNH<sub>3</sub><sup>+</sup> is found to be 7.97.<sup>33a</sup> The supermolecule/PCM approach is used to calculate the net solvation free energy change for the exchange reaction (eqs 11a and 11b).



where AH stands for neutral acids,  $n = 1$  for MeSO<sub>2</sub>OH, and  $n = 0$  for the other acids.



where BH<sup>+</sup> stands for cationic acids.

The calculated pK<sub>a</sub><sup>THF</sup> values, as obtained by combining the gas-phase acidities and the net solvation-free energy changes, are listed in Table 6 (column 3). Unfortunately, very limited experimental pK<sub>a</sub><sup>THF</sup> data are available for comparison. Among these, the picric acid is the only *neutral acid* for which the pK<sub>a</sub><sup>THF</sup> value is determined (pK<sub>a</sub><sup>THF</sup> = 11.6).<sup>56</sup> On the other hand, the pK<sub>a</sub><sup>THF</sup> values of a few cationic acids have been measured by Garrido et al.,<sup>61</sup> and thus we adopt their results for PhNMe<sub>2</sub>-H<sup>+</sup>, pyridine-H<sup>+</sup>, and Et<sub>3</sub>N-H<sup>+</sup>. These experimental pK<sub>a</sub><sup>THF</sup> values are summarized in Table 6.

To our satisfaction, the calculated results are in good agreement with available experimental data, with a root-mean-square error of 1.0 pK<sub>a</sub> units and a mean-absolute error of 0.88 pK<sub>a</sub> units. For example, the pK<sub>a</sub><sup>THF</sup> values for Et<sub>3</sub>N-H<sup>+</sup> and PhNMe<sub>2</sub>-H<sup>+</sup> are calculated to be 12.92 and 7.71, respectively, as compared with the experimental pK <sub>$\alpha$</sub>  values (13.66 and 7.39) in THF. The largest deviation is observed to be 1.6 pK<sub>a</sub> units for pyridine-H<sup>+</sup>. Thus, all evidence suggests that our calculated pK<sub>a</sub><sup>THF</sup> values are reliable, which gives us confidence in establishing a reasonable correlation of pK<sub>a</sub><sup>THF</sup> with pK<sub>a</sub><sup>DMSO</sup>.

An interesting observation found by Abdur-Rashid et al.<sup>58e</sup> is that there are two separate pK<sub>a</sub><sup>THF</sup>/pK<sub>a</sub><sup>DMSO</sup> correlations for *cationic acids* and *neutral acids*. The reason is probably that in THF cationic acids and neutral acids behave quite differently. Since most of our studied species are *neutral acids* and there are only *four* cationic acids considered in the present work, here we only build the pK<sub>a</sub><sup>THF</sup>/pK<sub>a</sub><sup>DMSO</sup> correlation for *neutral acids*. Figure 9 shows the plot of our calculated pK<sub>a</sub><sup>THF</sup> values for the studied neutral acids against the experimental pK<sub>a</sub><sup>DMSO</sup> values. A least-squares fit line through all the points gives a correlation expression (eq 12) with a very high correlation coefficient ( $r = 0.992$ ) and a standard deviation of 1.3 pK<sub>a</sub> units.

$$\text{pK}_a^{\text{THF}} = 1.03 \text{pK}_a^{\text{DMSO}} + 11.3 \quad (12)$$

To evaluate the reliability of our pK<sub>a</sub><sup>THF</sup>/pK<sub>a</sub><sup>DMSO</sup> correlation, we compare it with an experimental correlation for neutral phosphorus-containing acids: pK <sub>$\alpha$</sub> <sup>THF</sup> = 1.18 pK<sub>a</sub><sup>DMSO</sup> + 11.3, which is proposed by Morris et al.,<sup>58e</sup> where pK <sub>$\alpha$</sub> <sup>THF</sup> is an approximation to pK<sub>a</sub><sup>THF</sup> (as mentioned above). Given that the

(56) Coetzee, J. F.; Deshmukh, B. K.; Liao, C.-C. *Chem. Rev.* **1990**, *90*, 827.

(57) See for example: (a) Kaufman, M. J.; Gronert, S.; Streitwieser, A. *J. Am. Chem. Soc.* **1988**, *110*, 2829. (b) Krom, J. A.; Petty, J. T.; Streitwieser, A. *J. Am. Chem. Soc.* **1993**, *115*, 8024. (c) Antipin, I. S.; Garayev, R. F.; Verdernikov, A. N.; Konovalov, A. I. *J. Phys. Org. Chem.* **1994**, *7*, 181. (d) Facchetti, A.; Streitwieser, A. *J. Org. Chem.* **2004**, *69*, 8345.

(58) (a) Hinman, J. G.; Lough, A. J.; Morris, R. H. *Inorg. Chem.* **2007**, *46*, 4392. (b) Kolomeitsev, A. A.; Koppel, I. A.; Rodima, T.; Barten, J.; Lork, E.; Rösenthaller, G.-V.; Kaljurand, I.; Kütt, A.; Koppel, I.; Mäemets, V.; Leito, I. *J. Am. Chem. Soc.* **2005**, *127*, 17656. (c) Kaljurand, I.; Rodima, T.; Pihl, A.; Mäemets, V.; Leito, I.; Koppel, I. A.; Mishima, M. *J. Org. Chem.* **2003**, *68*, 9988. (d) Rodima, T.; Kaljurand, I.; Pihl, A.; Mäemets, V.; Leito, I.; Koppel, I. A. *J. Org. Chem.* **2002**, *67*, 1873. (e) Abdur-Rashid, K.; Fong, T. P.; Greaves, B.; Gusev, D. G.; Hinman, J. G.; Landau, S. E.; Lough, A. J.; Morris, R. H. *J. Am. Chem. Soc.* **2000**, *122*, 9155.

(59) Fuoss, R. M. *J. Am. Chem. Soc.* **1958**, *80*, 5059.

(60) (a) Streitwieser, A.; Kim, Y.-J. *J. Am. Chem. Soc.* **2000**, *122*, 11783. (b) Streitwieser, A.; McKeown, A. E.; Hasanayn, F.; Davis, N. R. *Org. Lett.* **2005**, *7*, 1259.

(61) Garrido, G.; Koort, E.; Råfols, C.; Bosch, E.; Rodima, T.; Leito, I.; Rosés, M. *J. Org. Chem.* **2006**, *71*, 9062.

(62) Jorgensen, W. L.; Briggs, J. M.; Gao, J. *J. Am. Chem. Soc.* **1987**, *109*, 6857.

**TABLE 6.** Calculated  $pK_a$  Values and Corresponding Experimental  $pK_a$  Values in THF

acid	$pK_a^{\text{THF}}(\text{exptl})$	$pK_a^{\text{THF}}(\text{calcd})$	
		Picric acid anchor <sup>a</sup>	H <sup>+</sup> anchor <sup>b</sup>
picric acid	11.6 <sup>c</sup>		17.61
MeSO <sub>2</sub> OH		13.48	13.53
CF <sub>3</sub> COOH		14.67	17.42
2,4-(NO <sub>2</sub> ) <sub>2</sub> -phenol		15.44	20.88
PhCOSH		16.29	22.30
Cl <sub>2</sub> CHCOOH		15.04	21.05
<i>p</i> -NO <sub>2</sub> -C <sub>6</sub> H <sub>4</sub> OH		21.36	27.37
PhCOOH		19.46	25.47
CH <sub>3</sub> COOH		22.08	28.09
HCN		23.80	29.81
(CH <sub>3</sub> CO) <sub>2</sub> CH <sub>2</sub>		25.48	31.49
C <sub>5</sub> NH <sub>4</sub> OH		24.37	30.38
PhCH <sub>2</sub> SH		27.64	33.65
NC-NH <sub>2</sub>		26.12	32.13
<i>n</i> -BuSH		28.99	35.00
PhOH		27.77	33.78
imidazole		29.06	35.07
PhCONHNH <sub>2</sub>		28.31	34.32
Pyrazole		28.67	34.68
(NH <sub>2</sub> ) <sub>2</sub> C=S		29.16	35.17
CF <sub>3</sub> CH <sub>2</sub> OH		33.63	39.64
H <sub>2</sub> NCOOEt		34.69	40.70
PhCOCH <sub>3</sub>		37.36	43.37
CH <sub>3</sub> C(O)NH <sub>2</sub>		35.71	41.72
CH <sub>3</sub> C(O)CH <sub>3</sub>		38.99	45.00
(NH <sub>2</sub> ) <sub>2</sub> C=O		38.14	44.15
CH <sub>3</sub> OH		41.20	47.21
MeCN		42.51	48.52
DMSO		46.62	52.63
4-methylpyridine		47.08	53.09
		PhNH <sub>3</sub> <sup>+</sup> anchor <sup>d</sup>	
PhNMe <sub>2</sub> -H <sup>+</sup>	7.39 <sup>e</sup>	7.71	10.38
pyridine-H <sup>+</sup>	8.25 <sup>e</sup>	9.84	12.51
PhNH <sub>3</sub> <sup>+</sup>	7.97 <sup>f</sup>		10.64
Et <sub>3</sub> N-H <sup>+</sup>	13.66 <sup>e</sup>	12.92	18.36
root-mean-square error		1.0	4.3
mean absolute error		0.88	4.1

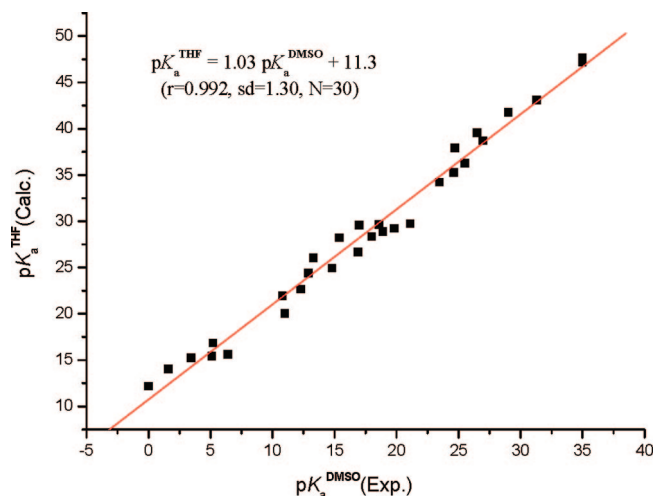
<sup>a</sup> Calculated  $pK_a$  values with the relative determination method. Picric acid is chosen as the reference for *neutral acids*, anilinium (PhNH<sub>3</sub><sup>+</sup>) is chosen as the reference for *cationic acids*. <sup>b</sup> Absolute determination method. <sup>c</sup> Reference 56. <sup>d</sup> Calculated  $pK_a$  values using the relative determination method. Picric acid is chosen as the reference for *neutral acids*, anilinium (PhNH<sub>3</sub><sup>+</sup>) is chosen as the reference for *cationic acids*. <sup>e</sup> Reference 61. <sup>f</sup> Reference 33a.

errors in  $pK_a^{\text{THF}}$  measurements are between 2 and 4  $pK_a$  units,<sup>58e</sup> there is good agreement between the computationally and experimentally derived extrapolations. Thus, this gives us confidence that our calculations may have useful predictive power for  $pK_a$  estimations in THF.

To compare, the absolute determination method (see eq 1a) is also employed to calculate the  $pK_a^{\text{THF}}$  values. In this approach, a simple solvated form of the proton, H<sup>+</sup>-(THF), is considered, and the results are summarized in Table 6 (last column). As seen from Table 6, this approach leads to very large deviation from the experimental data. The root-mean-square error is as large as 4.3  $pK_a$  units. Thus, this method is not successful in the prediction of the  $pK_a^{\text{THF}}$  values.

#### 4. Conclusions

A supermolecule/PCM solvation approach together with a relative determination method are applied to improve the accuracy of  $pK_a$  predictions in nonaqueous solvents. This work considers a total number of 33 organic acids with



**FIGURE 9.** Correlation between  $pK_a^{\text{THF}}(\text{calcd})$  and  $pK_a^{\text{DMSO}}(\text{exptl})$  for neutral acids.

various functional groups, including the representatives from all of the conventional families of OH acids (e.g., alcohols, phenols, carboxylic acids, and sulfonic acids), NH acids (e.g., amines and ammonium ions), and CH acids (e.g., hydrocarbons and ketons). The economical MP2/6-311++G(d,p)//B3LYP/6-31+G(d) method is employed and found to be sufficiently accurate in predicting the gas-phase acidities. Combining the supermolecule/PCM approach with the MP2/6-311++G(d,p)//B3LYP/6-31+G(d) method, we have successfully predicted the  $pK_a$  values of the studied acids in DMSO, MeCN, and THF. The calculated results show very good agreement with experimental data in all studied solvents. The precision of theoretical  $pK_a$  values in DMSO, MeCN, and THF is observed to be 0.93, 1.5, and 0.63  $pK_a$  units, correspondingly. Furthermore, we have established reliable  $pK_a^{\text{MeCN}}/pK_a^{\text{DMSO}}$  and  $pK_a^{\text{THF}}/pK_a^{\text{DMSO}}$  correlations, allowing a reasonable estimate of the  $pK_a^{\text{MeCN}}$  and  $pK_a^{\text{THF}}$  values. We envisage that these correlations can serve as a valuable tool for future researchers in acid–base chemistry in nonaqueous solvents.

The motivation for using the relative determination method is to introduce appropriate, favorable error cancellation in the final results. In this approach the  $pK_a$  value of an unknown acid is calculated by anchoring to a reference acid whose  $pK_a$  value is already known experimentally. We have shown that in the present study of small acids in DMSO and THF, the computational results do not depend sensitively on the specific choice of the anchor acid, i.e., changing the anchor acid may induce at most a shift of only 1.6  $pK_a$  units. The situation is slightly different in MeCN, where using a neutral anchor acid may induce a systematic error of  $\sim 2.3$   $pK_a$  units for the  $pK_a$  prediction of cationic acids. Although this is comparable to the error range of the experimental measurement and the electronic structure calculation, we can further improve the accuracy by adopting the following physically motivated strategy: using a neutral acid as an anchor for the  $pK_a$  prediction of neutral acids and using a cationic acid as an anchor for cationic acids.

A computational approach using the absolute determination method is also tested. This approach does not use any external parameters or empirical fittings. In the present study, we focus on a simplified treatment of proton solvation in nonaqueous solvents, with the combination of the supermolecule method.

Calculations show that this approach does not work well in the studied solvents. Future work should be directed toward developing a more applicable computational method for non-aqueous solvents.

Finally, the acids studied in this paper are relatively small. It is an interesting future direction to extend our approach to treat more complex, multifunctional compounds and compounds with multiple dissociation sites. The specific methodology, however, may need significant improvement beyond a simple application of the electronic structure theory within the PCM framework, i.e., to take into account multiple free-energy minima as well as dynamical effects.

**Acknowledgment.** This study is supported by the National Science Foundation CAREER award CHE-0348956 (H.W.) and the U.S. Department of Energy, Office of Basic Energy Sciences DE-FG02-08ER15996 (J.M.S.).

**Supporting Information Available:** Free energies for all acids in the gas phase and solvents (DMSO, MeCN, and THF) and Cartesian coordinates for the optimized structures of all acids. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JO802641R