Calculation of the Solvation Free Energy of Neutral and Ionic Molecules in Diverse Solvents

Sehan Lee,[†] Kwang-Hwi Cho,[‡] Chang Joon Lee,[†] Go Eun Kim,[†] Chul Hee Na,[†] Youngyong In,[§] and Kyoung Tai No*[,†,§]

Department of Biotechnology, Yonsei University, Seoul, 120-749, Korea, Department of Bioinformatics, Soongsil University, Seoul, 156-743, Korea, and Bioinformatics and Molecular Design Research Center, Seoul, 120-749, Korea

Received August 6, 2010

The solvation free energy density (SFED) model was modified to extend its applicability and predictability. The parametrization process was performed with a large, diverse set of solvation free energies that included highly polar and ionic molecules. The mean absolute error for 1200 solvation free energies of the 379 neutral molecules in 9 organic solvents and water was 0.40 kcal/mol, and for 90 hydration free energies of ions was 1.7 kcal/mol. Overall, the calculated solvation free energies of a wide range of solute functional groups in diverse solvents were consistent with experimental data.

INTRODUCTION

Over the past five decades, the accuracy of computational descriptions of molecules in the gas phase has improved such that many approaches can be used to understand the rates, mechanisms, and other features of isolated and clustered molecules. However, since most chemical and biological phenomena take place in solvents, a comprehensive understanding of solutions is one of the most important goals in theoretical chemistry.[1,2]

The standard state free energy of solvation, ΔG^{o}_{solv}, is a fundamental quantity that represents the effect of a solution on a chemical or biological process. The description of solvent effect in molecular scale is very hard task, because of enormous degree of freedom in solution, which leads to various approximated approaches. The empirical additive approaches are based on the assumption that the solvation free energy of a given molecule is the sum of the solvation free energies of all fragments of the molecule.[3−6] The key advantages of additive methods are their simplicity and speed. However, these phenomenological parametrizations critically rely on the quality of the training database and cannot give real insight of solvent effect on solvation. These inherent problems make a more fundamental approach to solvation phenomena desirable.

The explicit solvent approach[7,8] employs thousands of discrete solvent molecules and performs thermodynamic averaging based on the simulations of liquids, such as Monte Carlo[9] and molecular dynamics.[10,11] It provides microscopic descriptions of solutions, however, the computational effort of performing reliable simulations compelled the simplification of potential energy functions, which in turn emphasized the physical problems associated with this approach. Therefore, application of explicit solvent approaches is limited to simplifications that do not affect the overall simulation accuracy.

The original concept of a continuum solvent approach given by Born[12] and Bell[13] was formally extended by Kirkwood[14] to quantum description of solute, and provided an interpretative tool by Onsanger.[15] These focuses primary on the description of solute and tries to represent the influence of the solvent by an effective continuum surrounding the solute.[16,17] In general, continuum solvent approaches decompose solvation free energy into polar and nonpolar contributions.[18−21] Due to the outstanding importance of the electrostatic interactions in solvation phenomena, the main task of these models is the proper description of the electrostatic interaction[22−24] within this continuum approximation. Poisson−Boltzmann[25] (PB) equation gives numerical solutions for molecules of arbitrary shape and charge distribution. The Langevin dipoles[26,27] model treats the solvent molecules as polarizable point dipoles fixed on cubic grid and their strength and orientation are optimized simultaneously with the solute wave function.

No et al.[28] proposed an empirical continuum solvation free energy calculation approach called the Solvation Free Energy Density (SFED) model to estimate the ΔG^{o}_{solv} of a molecule in a solvent. The SFED model has been tuned such that the accuracy of predicting hydration and 1-octanol solvation free energies of typical neutral molecules is greater than 0.5 kcal/mol.[29] However, the SFED model has some shortcomings: (1) The shell model adapted to generate a grid space of solute is computationally expensive and demonstrates some difficulty because of overlapping grids between molecules in close contact with each other. (2) The effective charge used to express stabilization resulting from a hydrogen bond[28] shows limited accuracy and does not reflect the characteristics of these bonds properly. (3) Since the parametrization procedure was performed with a small number of neutral molecules, the applicability domain of the SFED model is limited.

* To whom correspondence should be addressed. Phone: 82-2-393-9551. Fax: 82-2-393-9554. E-mail: ktno@yonsei.ac.kr.
[†] Yonsei University.
[‡] Soongsil University.
[§] Bioinformatics and Molecular Design Research Center.

10.1021/ci100299m
Published on Web 12/06/2010

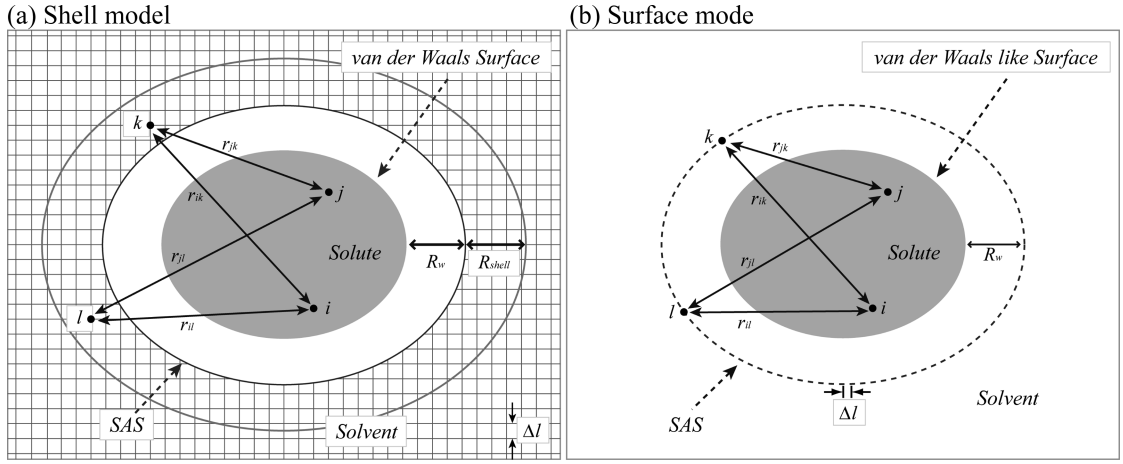


Figure 1. Solute and solvent of a solution are described as an assemblage of interacting compartments. Furthermore, the solvent accessible surface (SAS), R_W, R_{shell}, Δl, and van der Waals-like surface of the model are described.

Here we present advanced solvation free energy models for 10 solvents developed based on the SFED model. The purpose of this work was to overcome the shortcomings of the previous model and extend its applicability and predictability. To accomplish this, we modified the basic functions and grid space of the model. Then, the parametrizations were performed with a larger, more diverse data set.

METHOD

A. Description of the SFED Model. *Basic Theory of the SFED Model.* Details of the SFED model have been published previously.[28,29] In the shell model (Figure 1), the grid points are located between the solvent accessible surface (SAS) and a defined distance from the SAS area, namely R_{shell}. The SAS of the solute is defined by summing the area of the atomic SASs except where they overlap. The atomic SAS is the spherical surface, and the radius of the sphere is defined by the sum of its van der Waals radius, R_{vdw}, and the effective solvent shell thickness, R_W. The standard state solvation free energy of solvation, ΔG°_{solv}, was expressed as

$$\Delta G^{\circ}_{solv} = \Delta G_{inter} + \Delta G_{cav} = \sum_{k}^{Q}\sum_{j}^{m} C_j h_j(r_{ik}) + \Delta G_{cav} \quad (1)$$

where $\sum_k^Q$ is the sum of the grid points Q in the shell and $\sum_j^m$ is the linear combination of m basis functions, $h_j(r_{ik})$. $\sum_j^m C_j h_j(r_{ik})$ is the free energy of the interaction between ith interacting compartment of the solute and the kth grid point in the cavity.

The free energy of cavity formation by the solute was expressed by a term proportional to the SAS area, S_S, of the solute as follows

$$\Delta G_{cav} = C_S S_S + C \quad (2)$$

The coefficients C_j, C_S, and C of each solvent model were determined by minimizing the difference between the calculated and experimental solvation free energies. Selected molecules that are typically neutral appear frequently as fragments of biomolecules, and are not too flexible, were used as constraints for determining the coefficients and the form of basis functions in the SFED.

In our previous work,[28] we found that four basis functions contribute significantly to ΔG_{sol}. The basis function vector $\vec{X} = \{h_1^{\circ}, h_2^{\circ}, h_3^{\circ}, h_4^{\circ}\}$ was

$$\vec{X} = \left\{ \left| \sum_{i=1}^{N_A} \frac{q_i}{r_{ik}} \right|, \sum_{i=1}^{N_A} \frac{q_i^2}{r_{ik}}, \sum_{i=1}^{N_A} \frac{\alpha_i}{r_{ik}^3}, \sum_{i=1}^{N_A} \frac{D_i}{r_{ik}^6} \right\} \quad (3)$$

where N_A is the number of atoms in the solute. The atom-centered net atomic charges q_i were calculated using MPEOE, an empirical net atomic charge calculation method.[30,31] The effective atomic polarizability α_i was calculated from the charge dependent effective atomic polarizabiltiy (CDEAP) method,[32] an empirical formula in which the effective atomic polarizability in a molecule is described as a function of its net atomic charge. The dispersion coefficients D_i are proportional to the square of α_i, α_i^2.

These basis functions work well on neutral solutes containing C, H, O, N, S, and F. However, for molecules that contain highly polarizable atoms, such as Cl or Br, the fourth basis function contributes too much, and the solvation free energies of these molecules are overestimated. To overcome this problem, the second and forth basis functions were modified as follows:[29]

$$\vec{X} = \left\{ \left| \sum_{i=1}^{N_A} \frac{q_i}{r_{ik}} \right|, \sum_{i=1}^{N_A} \frac{q_i^2}{r_{ik}^3}, \sum_{i=1}^{N_A} \frac{\alpha_i}{r_{ik}^3}, \sum_{i=1}^{N_A} \frac{\alpha_i}{r_{ik}^6} \right\} \quad (4)$$

In the modified basis function set, r_{ik} of h_2° was replaced by r_{ik}^3 and the atomic dispersion coefficient D_i of h_4° was substituted with the atomic polarizability α_i.

Refinement of the SFED Model. Bumping of the cavity in the Shell-SFED model can cause some problems, such as intrusion of grid points of one molecule into the atomic radius of another molecule. Thus the extra computation is required to regenerate the cavity including both molecules. In this study, the surface model (Figure 1) was adapted instead of the shell model to solve the problem of cavity overlapping and save computational resources. The grid points were located on the SAS, and the solvation free energy was obtained by summing the SFED at these surface grid points $\sum_k^S$, where S is the number of grid points on the cavity surface. The SAS is defined by the same way explained above. The

atomic radii used to construct the cavity surface were optimized based on functionality. We refer to each model as either the Shell-SFED model of the Surface-SFED model from now on.

With small changes to the previous model, the Surface-SFED model has been expanded to assess ionic molecules with no additional basis functions. The basis functions work well for ionic species overall. However, the first term increased rapidly with increasing ionic molecule surface area, resulting in some inverse correlations in a single ionic species. For this reason, r_{ik} of h_1° was replaced by r_{ik}^2. The basis function set for the Surface-SFED model was transformed to become

$$\vec{X} = \left\{ \left| \sum_{i=1}^{N_A} \frac{q_i}{r_{ik}^2} \right|, \sum_{i=1}^{N_A} \frac{q_i^2}{r_{ik}^3}, \sum_{i=1}^{N_A} \frac{\alpha_i}{r_{ik}^3}, \sum_{i=1}^{N_A} \frac{\alpha_i}{r_{ik}^6} \right\} \quad (5)$$

Stabilization because of hydrogen bonding, ΔG_{HB}, is partially a result of electrostatic interactions between the positively polar hydrogen of the donor and the lone pair of the acceptor, ΔG_{HB}^{ele}. The tunneling effect which redistributes electron densities of the acceptor and donor, $\Delta G_{HB}^{tunneling}$, also contributes to this stabilization. Thus

$$\Delta G_{HB} = \Delta G_{HB}^{ele} + \Delta G_{HB}^{tunneling} \quad (6)$$

ΔG_{HB}^{ele} can be described by two basis functions, h_1° and h_2° in the SFED model. However, $\Delta G_{HB}^{tunneling}$ is not properly represented in the basis functions. Therefore, the calculated solvation free energies of polar molecules were smaller than the experimental solvation free energies.[28] In our previous publication,[28] the magnitude of the dipole moment of a functional group, X−H, that forms hydrogen bonds was increased until the error between the calculated and experimental values of ΔG_{solv}° was minimal.

In doing so, predicting hydration free energy improved. Nevertheless, there were problems with this approach. First, there was limited accuracy in functional groups that form strong hydrogen bonds, such as amines and acetic acids. Moreover, the directional dependency of hydrogen bonding was not reflected. Finally, this approach raises questions about the transferability of such parametrization to other solvents.

To understand and interpret the effects of hydrogen bonding quantitatively, the contribution of hydrogen bonds to solvation is calculated independent of other possible solvent−solute interactions.

$$\Delta G_{inter} = \Delta G_{inter}^* + \Delta G_{HB} \quad (7)$$

By applying the multiplicative principle to describe hydrogen bonding,[33,34] ΔG_{HB} could be expressed as the product of parameters characteristic of the acid, C_A, and base, C_B, components.

$$\Delta G_{HB} = 5.46 C_A C_B \quad (8)$$

In this study, ΔG_{HB} was divided into two terms based on the role of the solute molecule in hydrogen bonding, with the donor and acceptor represented as subscript "d" and "a", respectively.

$$\Delta G_{HB} = \Delta G_d + \Delta G_a = c_d \alpha_{solute} \beta_{solvent} + c_a \beta_{solute} \alpha_{solvent} \quad (9)$$

Hydrogen bond acidity α and basicity β are numerical parameters characteristic of the proton donor and acceptor components. Thus, the contribution of solvent properties to solvation is implied as linear expansion coefficients.

$$\Delta G_{HB} = \Delta G_d + \Delta G_a = C_d \alpha_{solute} + C_a \beta_{solute} \quad (10)$$

The experimental α and β values of neutral solutes were obtained from Abraham et al.[35−37] Application of these parameters to the SFED model produced results in good agreement with the experimental solvation free energy. The experimental α and β values are available for 283 out of 379 neutral solutes. In addition, 1025 free energies of solvation of neutral molecules in 10 solvents yielded a mean absolute error (MAE) of 0.36 kcal/mol (Table 9).

For the general purpose of the SFED model, primitive empirical α and β of neutral molecules are expressed by

$$\alpha = \sum_i \alpha_i \text{ and } \beta = \sum_i \beta_i \quad (11)$$

where α_i and β_i represent the acidity and basicity of the acidic and basic functional groups of the solute and are meant to reproduce the values published by Abraham et al.

The ionic hydrogen bond can be formed between ions and molecules with a bond strength of 5−35 kcal/mol, up to one-third the strength of covalent bonds.

$$BH^+ + B' \leftrightarrow BH^+ \cdot B' \text{ and}$$
$$A^- + HA' \leftrightarrow A^- \cdot HA' \quad (12)$$

Because of limited experimental data on ionic hydrogen bonds[38−45] and their variation within a functional group, other properties were considered. The linear correlations between proton affinity difference, ΔPA, and ionic hydrogen bond dissociation enthalpies (eq 12), ΔH_D°, were found over ΔPA ranges up to 60 kcal/mol in $NH^+\cdots O$ and $NH^+\cdots S$ bonds.[38] The gas phase basicity, GB, and ΔPA of a molecule, M, are defined in terms of a hypothetical gas phase reaction:

$$M(g) + H^+(g) \rightarrow MH^+(g) \quad (13)$$

The gas basicity of M at temperature T, GB(M,T), and the proton affinity, PA(M,T), are the negative of the Gibbs free energy and enthalpy change for this reaction, respectively

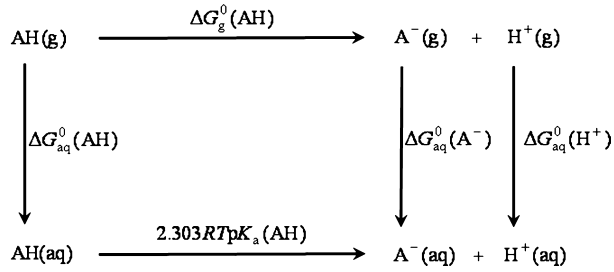
$$GB(M,T) \equiv -\Delta G_{Rnl}^\circ(T) \text{ and}$$
$$PA(M,T) \equiv -\Delta H_{Rnl}^\circ(T) \quad (14)$$

GB and free energy of the ionic hydrogen bond displayed a linear correlation (unpublished data). The acidity and basicity of ions were computed as

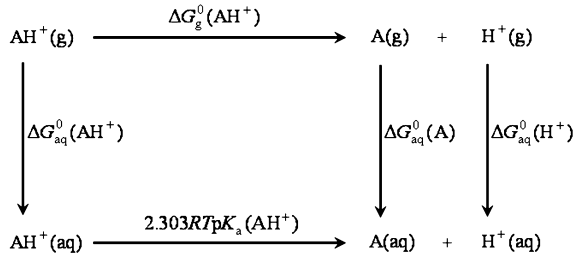
$$\alpha = \sum_i (a_i GB + b_i) + \sum_i \alpha_i \text{ and}$$
$$\beta = \sum_i (a_i GB + b_i) + \sum_i \beta_i \quad (15)$$

where a_i and b_i are optimized parameters depending on the type of ionic hydrogen bond component, namely, amine,

Scheme 1. Thermodynamic Cycle Relating the Solvation Free Energy of an Anion



Scheme 2. Thermodynamic Cycle Relating the Solvation Free Energy of a Cation



aniline and pyridine, and alcohol and carboxylic acid (Table 6). The final formula for the SFED computation of a solute is

$$\Delta G^{o}_{solv} = \Delta G^{*}_{inter} + \Delta G_{HB} + \Delta G_{cav} = \sum_{k}^{S}\sum_{j}^{m} C_j h_j(r_{ik}) + C_d\alpha + C_a\beta + C_S S_S + C \quad (16)$$

The linear expansion coefficients, C_j, C_d, C_a, C_S, and C, and atomic radii, R_i, were determined by minimizing the difference between experimental and calculated solvation free energies.

B. Surface-SFED Model Training Set for Parametrization. *Experimental Data for Neutral Solutes.* The experimental solvation free energies of neutral molecules were obtained from two sources, the MedChem database[46] and publications from Wang et al.[47] The collected experimental data set encompassed a broad range of common functional groups present in biological and drugable small molecules that possess H, C, N, O, F, P, S, Cl, Br, and I atoms. The data set contained 2193 free energies of solvation for 387 solutes in 91 solvents. Since many of the solvents did not have sufficient free energy of solvation data to determine the linear expansion coefficients, only 10 solvents, water, 1-octanol, hexadecane, heptanes, hexane, cyclohexane, chloroform, carbon tetrachloride, benzene, and diethyl ether were investigated. Solutes containing a phosphorus atom were excluded because proper experimental data for solvation free energy were unavailable. When these solvation data are available, we will optimize the atomic solvation parameters for variations of the phosphorus atom.

Determination of the Hydration Free Energy of Ionic Solutes. The hydration free energies of ionic solutes were obtained using the thermodynamic cycles,[48] presented in Schemes 1 and 2. To investigate biological molecules, especially protein and small organic molecules, the ionic molecule data set was focused on four chemical species, namely alcohol, carboxylic acid, amine, and pyridine, as well as 37 anions and 53 cations (Tables 1 and 2).

Table 1. Hydration Free Energies[a] of Bare Anions

A^-	$\Delta G^{\circ}_{water}(AH)$[a]	$pK_a(AH)$[b]	$\Delta G^{\circ}_{g}(AH)$[c]	$\Delta G^{\circ}_{water}(A^-)$
HCO_2^-	−7.00[d]	3.80	338.30	−76.11
$CH_3CO_2^-$	−6.70	4.76	341.49	−77.69
$CH_3CH_2CO_2^-$	−6.46	4.86	340.39	−76.21
$CH_3(CH)_2CO_2^-$	−6.35	4.83	339.00	−74.75
$CH_3(CH)_3CO_2^-$	−6.16	4.84	339.70	−75.25
$CH_3(CH)_4CO_2^-$	−6.46	4.85	339.10	−74.93
$H_2C{=}CHCO_2^-$	−6.60[d]	4.30	337.20	−73.92
$CH_3COCO_2^-$	−9.40[d]	2.39	326.51	−68.64
$CF_3CO_2^-$	−7.30[d]	0.50	316.70	−59.32
$CH_2ClCO_2^-$	−8.70[d]	2.90	328.90	−69.64
$CHCl_2CO_2^-$	−6.60[d]	1.40	321.50	−62.19
$C_6H_5CO_2^-$	−7.90[d]	4.20	333.00	−71.16
HO^-	−6.31	15.70	383.70	−104.56
CH_3O^-	−5.07	15.50	375.00	−94.89
$CH_3CH_2O^-$	−4.90	15.90	371.30	−90.47
$CH_3CH_2CH_2O^-$	−4.85	16.10	369.40	−88.25
$CH(CH_3)_2O^-$	−4.75	17.10	368.80	−86.18
$CF_3CHOCH_3^-$	−4.16	11.80	353.70	−77.74
$CH_3CH_2CHOCH_3$	−4.61	17.60	367.50	−84.06
$C(CH_3)_3O^-$	−4.50	19.20	367.90	−82.17
$H_2C{=}CHCH_2O^-$	−5.10	15.50	366.60	−86.52
$C_6H_5CH_2O^-$	−6.60[d]	15.40	363.40	−84.96
$CH_3OCH_2CH_2O^-$	−6.80	14.80	366.80	−89.38
$CH_2OHCH_2O^-$	−9.30	15.40	360.90	−85.16
$CF_3CH_2O^-$	−4.10	12.37	354.10	−77.30
$CH(CF_3)_2O^-$	−3.80	9.30	338.40	−65.49
$C_6H_5O^-$	−6.53	10.00	342.90	−71.77
o-$CH_3C_6H_4O^-$	−5.86	10.30	342.40	−70.19
m-$CH_3C_6H_4O^-$	−5.49	10.10	343.30	−70.99
p-$CH_3C_6H_4O^-$	−6.12	10.30	343.80	−71.85
m-$HOC_6H_4O^-$	−11.40[d]	9.30[e]	339.10	−73.79
p-$HOC_6H_4O^-$	−11.90[d]	9.90[e]	343.10	−77.47
o-$NO_2C_6H_4O^-$	−4.50[d]	7.20	329.50	−60.16
m-$NO_2C_6H_4O^-$	−9.60[d]	8.40	327.60	−61.72
p-$NO_2C_6H_4O^-$	−10.60[d]	7.10	320.90	−57.80
o-$ClC_6H_4O^-$	−4.50[d]	8.50	337.10	−65.99
p-$ClC_6H_4O^-$	−6.20[d]	9.40	336.50	−65.86

[a] ΔG° in kcal mol^{-1}. [b] From refs 52 and 53. [c] Gas-phase acidities taken form ref 50. [d] Ref 54. [e] Ref 55.

Using the thermodynamic cycle illustrated in Scheme 1, the hydration free energy of anion A^-, $\Delta G^{\circ}_{aq}(A^-)$, can be calculated as

$$\Delta G^{\circ}_{aq}(A^-) = \Delta G^{\circ}_{aq}(AH) + 2.303RTpK_a(AH) - \Delta G^{\circ}_{g}(AH) - \Delta G^{\circ}_{aq}(H^+) \quad (17)$$

where $\Delta G^{\circ}_{aq}(AH)$ is the hydration free energy of the neutral specious AH, pK_a is the negative common logarithm of the solution phase acid dissociation constant of AH, and $\Delta G^{\circ}_{aq}(H^+)$ is the hydration free energy of the proton. Zhan and Dixon's value of −264 kcal/mol[49] was used. The gas phase acidity of AH, $\Delta G^{\circ}_{g}(AH)$, is the Gibbs energy change of the reaction $AH(g) \rightarrow A^-(g) + H^+(g)$ and equal to

$$\Delta G^{\circ}_{g}(AH) = G^{\circ}_{g}(A^-) + G^{\circ}_{g}(H^+) - G^{\circ}_{g}(AH) \quad (18)$$

In the case of cations, the thermodynamic cycle illustrated in Scheme 2 was used to calculate the hydration free energy of cation AH^+, $\Delta G^{\circ}_{aq}(AH^+)$ according to

$$\Delta G^{\circ}_{aq}(AH^+) = \Delta G^{\circ}_{g}(AH^+) + \Delta G^{\circ}_{aq}(A) + \Delta G^{\circ}_{aq}(H^+) - 2.303RTpK_a(AH^+) \quad (19)$$

where $\Delta G^{\circ}_{aq}(A)$ is the hydration free energy of the neutral specious A and pK_a is the negative common logarithm of

Table 2. Hydration Free Energies[a] of Bare Cation

AH^+	$\Delta G^\circ_{water}(A)$	$pK_a(AH^+)$[b]	$\Delta G^\circ_g(AH^+)$[c]	$\Delta G^\circ_{water}(AH^+)$
NH_4^+	−4.29	9.30	195.40	−85.58
$CH_3NH_3^+$	−4.60	10.63	206.62	−76.48
$CH_3(CH_2)_2NH_3^+$	−4.50	10.71	211.26	−71.86
$(CH_3)_2CHNH_3^+$	−3.70[d]	10.60	212.50	−69.66
$(CH_3)_3CNH_3^+$	−3.90	10.70	215.10	−67.40
c-$C_6H_{11}NH_3^+$	−5.10[d]	10.70	215.00	−68.70
$H_2C{=}CHCH_2NH_3^+$	−4.30[d]	9.50	209.20	−72.06
$(CH_3)_2NH_2^+$	−4.28	10.68	214.27	−68.58
$(CH_3CH_2)_2NH_2^+$	−4.06	11.09	219.74	−63.45
n-$(CH_3CH_2CH_2)_2NH_2^+$	−3.65	11.00	222.11	−60.55
$(H_2C{=}CHCH_2)NH_2^+$	−4.00[d]	9.30	219.00	−61.69
$(CH_3)_3NH^+$	−3.20	9.80	219.40	−61.17
$(CH_3CH_2)_3NH^+$	−3.03	10.78	227.29	−54.44
$(n$-$CH_3CH_2CH_2)_3NH^+$	−2.50[d]	10.30	229.50	−51.05
$C_6H_5NH_3^+$	−5.49	4.63	203.30	−72.51
o-$CH_3C_6H_4NH_3^+$	−5.60[d]	4.50	205.30	−70.44
m-$CH_3C_6H_4NH_3^+$	−5.70[d]	4.70	206.50	−69.61
p-$CH_3C_6H_4NH_3^+$	−5.60[d]	5.10	206.70	−69.86
m-$NH_2C_6H_4NH_3^+$	−9.90[d]	5.00	214.90	−65.82
$C_6H_5NH_2CH_3^+$	−4.70[d]	4.90[e]	212.70	−62.69
$C_6H_5NH_2CH_2CH_3^+$	−4.60[d]	5.10[e]	213.40	−62.16
$C_6H_5NH(CH_3)_2^+$	−2.90[d]	5.10	217.30	−56.56
p-$CH_3C_6H_4NH(CH_3)_2^+$	−3.70[d]	5.60	219.40	−55.94
$C_6H_5NH(CH_3CH_2)_2^+$	−2.90[d]	6.60	221.80	−54.11
$C_{10}H_7NH_3^+$	−7.30[d]	3.90	209.20	−67.42
p-$NO_2C_6H_4NH_3^+$	−9.90[d]	1.00	199.40	−75.86
p-$CH_3OC_6H_4NH_3^+$	−7.60[d]	5.30	207.60	−71.23
m-$ClC_6H_4NH_3^+$	−5.80[d]	3.50	199.90	−74.68
p-$ClC_6H_4NH_3^+$	−5.90[d]	4.00	201.20	−74.16
$CH_3(CH_2)_3NH_3^+$	−4.38	10.78	211.90	−71.19
$(CH_3(CH_2)_3)_2NH_2^+$	−3.31	11.39	223.54	−59.31
$CH_3(CH_2)_5NH_3^+$	−4.04	10.56	213.55	−68.90
$CH_3(CH_2)_4NH_3^+$	−4.09	10.63	212.60	−70.00
$CH_3CH_2NH_3^+$	−4.61	10.70	209.85	−73.36
$C_2H_4NH_2^+$	−4.50[d]	8.00	208.50	−70.92
$C_3H_6NH_2^+$	−5.60	11.30	217.20	−67.82
$C_4H_8NH_2^+$	−5.50	11.30	218.80	−66.12
$C_5H_{10}NH_2^+$	−5.10	11.10	220.20	−64.04
$C_6H_{12}NH_2^+$	−4.90[d]	11.10	220.70	−63.34
$C_9H_{10}NH^+$	−5.70	4.80	220.20	−56.05
$C_4H_8NHNH_2^+$	−7.40	9.70	218.60	−66.03
$C_4H_8NHNH_2^+$	−7.40	9.70	218.60	−66.03
$C_5H_5NH^+$	−4.70	5.20	214.70	−61.09
2,3-$(CH_3)_2C_5H_5NH^+$	−4.81	6.57	221.56	−56.21
2,4-$(CH_3)_2C_5H_5NH^+$	−4.85	6.99	222.47	−55.92
2,5-$(CH_3)_2C_5H_5NH^+$	−4.70	6.40	221.53	−55.90
2,6-$(CH_3)_2C_5H_5NH^+$	−4.60	6.65	222.54	−55.13
3,4-$(CH_3)_2C_5H_5NH^+$	−5.21	6.46	221.20	−56.82
3,5-$(CH_3)_2C_5H_5NH^+$	−4.84	6.15	220.72	−56.51
m-$CH_3C_5H_5NH^+$	−4.77	5.63	217.88	−58.57
p-$CH_3C_5H_5NH^+$	−4.92	5.98	218.76	−58.32
o-$CH_3CH_2C_5H_5NH^+$	−4.32	5.89	220.03	−56.33
m-$CH_3CH_2C_5H_5NH^+$	−4.60	5.56	218.81	−57.38
p-$CH_3CH_2C_5H_5NH^+$	−4.72	5.87	219.69	−57.04

[a] ΔG° in kcal mol^{-1}. [b] Refs 52 and 53. [c] Gas-phase acidities taken form ref 51. [d] Ref 54. [e] Ref 55.

the solution phase acid dissociation constant of AH^+. The gas phase acidity of AH^+, $\Delta G^\circ_g(AH^+)$, which is equal to the gas phase basicity of A, is the Gibbs energy change of the reaction $AH^+(g) \rightarrow A(g) + H^+(g)$ and equal to

$$\Delta G^\circ_g(AH^+) = G^\circ_g(A) + G^\circ_g(H^+) - G^\circ_g(AH^+) \quad (20)$$

Experimental gas-phase acidity and basicity values of neutral species were obtained from the National Institute of Standards and Technology (NIST) database[50] and compilation of Hunter and Lias.[51] Experimental aqueous pK_a data were obtained from publications by Tehan et al.[52,53] Geometries of the molecules were obtained by energy minimization in the gas-phase, using CVFF and B3LYP/3-21G(d,p) for neutrals and ionics, respectively.

Parameter Optimization. Three different types of parameters were optimized, namely, the atomic radii R_i, the linear expansion coefficients, C_j, C_d, C_a, C_S, and C, and the ionic hydrogen bond parameters a_i and b_i. The parametrizations of atomic radii and linear expansion coefficients were performed with an iterative procedure that attains convergence. Sets of parameters were optimized sequentially with a neutral data set in order of their linear expansion coefficients and atomic radii. Then the parameters were reoptimized using values optimized previously as constraints. This process composes one unit cycle and was repeated until the difference between the calculated and experimental solvation free energies was minimized. Ionic hydrogen bond parameters were optimized with preoptimized linear expansion coefficients and atomic radii parameters. α_i and β_i were determined from regenerated Abraham's values.

Table 3. Atomic Species Introduced in This Work and the Atomic Radius Optimized for the SFED Model

atomic species	definition	atomic radius (Å)
C_{sp3}	sp3 carbon	1.948
C_{ar}	sp2 carbon in conjugate system	1.881
C═O	sp2 carbon in carbonyl	1.833
C_{sp2}	sp2 carbon in alkene	1.858
C_{sp}	sp carbon	1.817
CO_2^-	sp2 carbon in carboxylate	1.743
$C_{sp3}{-}O^-$	sp3 carbon bonded to alcohol ion	1.742
$C_{ar}{-}O^-$	C_{ar} bonded to alcohol ion	1.801
H	aliphatic H	1.198
$H{-}X_{ar}$	H bonded to aromatic system	1.181
$H{-}O_{sp3}$	hydroxyl H in alcohol	1.052
H−COO	hydroxyl H in carboxylic acid	1.038
H−C(═O)N	H H bonded to amide N	1.080
$H{-}N_{sp3}$	H bonded to amine N	1.096
$H{-}S_{sp3}$	H bonded to thiol S	1.240
$H{-}N_{sp3}^+$	H bonded to ammonium	0.900
$H{-}N_{ar}^+$	H bonded to pyridine cation	0.950
O_{sp3}	sp3 hydroxyl oxygen	1.545
O_{sp2}	sp2 oxygen	1.518
$O_2{-}N^+$	sp2 oxygen in nitro group	2.201
$O_2{-}C$	oxygen in carboxylate	1.517
O_{sp3}^-	oxygen in alcohol anion	1.429
$N_{ar,2}$	2-valent conjugated nitrogen	1.670
$N_{ar,3}$	3-valent conjugated nitrogen	1.696
N_{sp3}	sp3 nitrogen in amine	1.704
$N_{sp2}C{=}O$	sp2 nitrogen in amide	1.660
N_{sp2}	sp2 nitrogen	1.680
N_{sp}	sp nitrogen	1.645
$N_{sp2}{}^+O_2^-$	sp2 nitrogen in nitro group	1.920
N_{sp3}^+	ammonium nitrogen	1.495
N_{ar}^+	cationic nitrogen in pyridine	1.690
S_{sp3}	sp3 sulfur	2.022
F	fluoride	1.669
Cl	chloride	2.195
Br	bromide	2.41
I	iodide	2.632

RESULTS AND DISCUSSION

The shape and size of the cavity that defines the boundary between the charge distribution of a solute and the solvent are critical factors for predicting the electrostatic effects involved in ionic solvation. In this study, a set of atomic radii that depend only on the local chemical environment (i.e., the function groups) of the atom were optimized to construct the cavity surface of the solute. Values proportional to the atomic radii were defined previously[28] and used as initial values for the optimization

Table 4. Hydrogen Bond Basicity β_i and the Parameter Values Optimized for the SFED Model

atomic species	definition	values
$R_2C{=}CR_2$	alkene	0.035
c-$R_2C{=}CR_2$	cycloalkene	0.050
$RC{\equiv}CR'$	alkyne	0.050
XR	monohalogenated alkane	0.100
X_2R	dihalogenated alkane	0.050
X_3R	trihalogenated alkane	0.020
RCHO	aldehyde	0.450
RCOR′	ketone	0.510
O=(c-R)	cycloketone	0.560
HCOOR	methanoate	0.190
RCOOR′	ester	0.225
RCN	nitrile	0.310
NH_3	ammonia	0.620
RNH_2	primary amine	0.610
R_2NH	secondary amine	0.690
R_3N	tertiary amine	0.790
RNO_2	nitro	0.150
$RCONH_2$	primary amide	0.340
RCONRH	secondary amide	0.360
$RCONR_2$	tertiary amide	0.390
H_2O	water	0.350
ROH	alcohol	0.480
RSH	thiol	0.240
RSR′	sulfide	0.320
Ph	benzene	0.020
$HCONH_2$	primary formamide	0.600
HCONRH	secondary formamide	0.550
$HCONR_2$	tertiary formamide	0.740
RCOOH	carboxylic acid	0.225
RSSR′	disulfide	0.140
ROR′	ether	0.450
c-COH	cycloalcohol	0.570
PhOH	phenol	0.250
C_5H_4N	pyridine	0.420
$PhNR_2$	aniline	0.290
$PhNO_2$	nitrobenzene	0.080

Table 5. Hydrogen Bond Acidity α_i and the Parameter Values Optimized for the SFED Model

atomic species	definition	values
$HC{\equiv}CR$	hydrogen bonded to alkyne	0.120
HRX_2	hydrogen bonded to dihalogenated carbon	0.100
HRX_3	hydrogen bonded to trihalogenated carbon	0.150
H_3N	hydrogen in ammonia	0.047
RNH_2	hydrogen bonded to primary amine	0.080
R_2NH	hydrogen bonded to secondary amine	0.080
$RCONH_2$	hydrogen bonded to primary amide	0.280
RCONRH	hydrogen bonded to secondary amide	0.400
RCOOH	hydrogen in carboxylic acid	0.600
$HCONH_2$	hydrogen bonded to primary formamide	0.310
HCONRH	hydrogen bonded to secondary formamide	0.400
H_2O	hydrogen in water	0.190
ROH	hydrogen in alcohol	0.370
c-COH	hydrogen in cycloalcohol	0.320
PhOH	hydrogen in phenol	0.600
$PhNR_2$	hydrogen in aniline	0.130

procedure. To cover a wide range of molecules, including highly polarizable solutes, we defined more elaborate radii. The optimized atomic radii used in this model are summarized in Table 3. The optimum value for grid interval, Δl, and effective solvent shell thickness, R_W, were determined to be 0.5 and 1.4 Å, respectively, as a compromise between computing time and the accuracy of the calculation.[28]

Table 6. Ionic Hydrogen Bond Parameters a_i and b_i and the Parameter Values Optimized for the SFED Model

atomic species	definition	a_i	b_i
R_3NH^+	amine cation	−0.276	76.997
$PhNR_2H^+$ and $C_5H_4NH^+$	aniline and pyridine cation	−0.354	90.988
$RCOO^-$ and RO^-	carboxylic acid and alcohol anion	0.292	81.420

Tables 4 and 5 list the hydrogen bond parameters and their determined values. The optimized linear expansion coefficients for each solvent and the contribution of each basis function in water, 1-octanol, and hexadecane model are summarized in Tables 7 and 8.

Two functions were introduced to describe the contribution from electrostatic interaction, $|\sum_{i=1}^{N_A}(q_i)/(r_{ik}^2)|$ and $\sum_{i=1}^{N_A}(q_i^2)/(r_{ik}^3)$. The former is the magnitude of the electrostatic potential at the surface point k. This function is responsible for the electrostatic stabilization of the dipole of the solvent by the electric potential produced by the solute. $\sum_{i=1}^{N_A}(q_i^2)/(r_{ik}^3)$ is used to mimic the interaction between the point charges on the solute and its image charge induced in the solvent, especially in the solvation shell. Combining these two functions is pertinent to polarization, dispersion, and repulsion interaction of a solute with a solvent.

Linear expansion coefficients correlate with solvent properties. Since all basis functions have positive values, the sign and magnitude of linear expansion coefficients indicate the contribution of a characteristic feature of the solvent to solvation. C_1 and especially C_2 of polar solvents were larger than those of nonpolar solvents because the major solute−solvent interaction is due to electrostatic interaction. In contrast, C_3 and C_4 had similar values in different solvents. However, when the polarity of solvents increased, C_3 and C_4 became smaller, even with a positive C_4 value in water. C_d and C_a of polar solvents, which form a hydrogen bond, became larger as the polarity of the solvent grew, implying formation of stronger hydrogen bonds. The acidity and basicity of molecules were computed by summing the acidity and basicity components that participated in hydrogen bonding. However, these parameters did not reflect the influence of substituent near hydrogen-bond acceptors or donors. Thus, the error of hydrogen bonding increased depending on the substituent.

The calculated and experimental solvation free energies of neutral and ionic (e.g., cations and anions) solutes are listed in Tables S1−S3 of the Supporting Information and compared in Figure 2. MAE for solvation free energies of 1200 neutral molecules was 0.40 kcal/mol while MAE for solvation free energies of 90 ionic molecules was 1.78 kcal/mol. The calculated solvation free energies were mostly consistent with experimental results. A deviation greater than 1 kcal/mol from the corresponding experimental values was observed in 105 out of 1200 neutral solutes in 10 solvents. Nearly half were calculated for water and one-third was for perhalogenated molecules. The greatest deviation in solvation free energy from experimentally derived data occurred with a complex compound, 4-amino-3−5−6-trichloropyridine-2-carboxylic acid, in water. The prediction was overestimated in small molecules, especially methane and its derivatives, since the hydrophobic effect was underestimated due to the first hydration shell effect.[56,57]

Table 7. Optimized Linear Expansion Coefficients for Each Solvent

solvent	$C_1 \times 10$	$C_2 \times 10$	$C_3 \times 10^2$	$C_4 \times 10$	$C_S \times 10^3$	$C \times 10$	C_d	C_a
water	−2.887	−5.290	−0.110	1.285	2.972	0.069	−2.730	−2.848
n-hexadecane	−1.542	−0.755	−3.015	−1.800	1.329	8.504	0.000	0.000
1-octanol	−2.087	−1.070	−2.479	−1.385	0.678	9.192	−4.023	−1.269
chloroform	−1.933	−2.235	−2.630	−1.952	0.416	6.507	−0.222	−0.539
cyclohexane	−1.413	−0.642	−2.823	−2.211	0.803	9.043	0.000	0.000
carbon tetrachloride	−1.935	−1.020	−2.469	−2.723	0.387	9.042	0.000	0.000
diethyl ether	−1.928	−0.831	−2.666	−2.022	0.589	8.757	−3.710	0.000
benzene	−1.833	−0.920	−2.785	−1.207	0.439	6.279	−1.086	0.000
n-heptane	−1.619	−0.754	−2.803	−2.471	1.129	8.993	0.000	0.000
n-hexane	−1.492	−0.764	−2.515	−2.658	0.663	9.028	0.000	0.000

Table 8. Contributions of Each Basis Function

water solvation free energy, ΔG_{hyd} (kcal/mol)									
mol name	$\sum\|C_1h_1\|$	$\sum C_2h_2$	$\sum C_3h_3$	$\sum C_4h_4$	C_SS_S	C	$C_d\alpha$	$C_a\beta$	ΔG_{hyd}^{cal}
propane	−0.043	−0.014	−1.008	0.310	2.502	0.007	0.000	0.000	1.754
1-propanol	−1.540	−2.713	−1.122	0.343	2.696	0.007	−1.010	−1.366	−4.706
propylamine	−1.199	−2.211	−1.270	0.382	2.803	0.007	−0.437	−1.736	−3.663
propionic acid	−1.890	−3.381	−1.133	0.338	2.761	0.007	−1.638	−1.281	−6.217
1-octanol solvation free energy, ΔG_{oct} (kcal/mol)									
mol name	$\sum\|C_1h_1\|$	$\sum C_2h_2$	$\sum C_3h_3$	$\sum C_4h_4$	C_SS_S	C	$C_d\alpha$	$C_a\beta$	ΔG_{oct}^{cal}
propane	−0.031	−0.003	−2.274	−0.334	0.571	0.919	0.000	0.000	−1.152
1-propanol	−1.114	−0.549	−2.532	−0.369	0.615	0.919	−1.488	−0.609	−5.126
propylamine	−0.867	−0.447	−2.866	−0.411	0.639	0.919	−0.644	−0.774	−4.451
propionic acid	−1.366	−0.684	−2.556	−0.364	0.630	0.919	−2.414	−0.571	−6.406
hexadecane solvation free energy, ΔG_{hex} (kcal/mol)									
mol name	$\sum\|C_1h_1\|$	$\sum C_2h_2$	$\sum C_3h_3$	$\sum C_4h_4$	C_SS_S	C	$C_d\alpha$	$C_a\beta$	ΔG_{hex}^{cal}
propane	−0.023	−0.002	−2.767	−0.434	1.119	0.850	0.000	0.000	−1.256
1-propanol	−0.823	−0.387	−3.080	−0.480	1.205	0.850	0.000	0.000	−2.714
propylamine	−0.641	−0.316	−3.486	−0.535	1.253	0.850	0.000	0.000	−2.874
propionic acid	−1.009	−0.483	−3.110	−0.473	1.235	0.850	0.000	0.000	−2.990

Table 9. Performance of the SFED Model by Solvent for Neutral Solute

	exptl HB		empirical HB	
solvent	number of data	MAE[a]	number of data	MAE[a]
water	274	0.40	365	0.49
n-hexadecane	171	0.32	187	0.32
1-octanol	147	0.36	177	0.43
chloroform	79	0.48	91	0.49
cyclohexane	74	0.36	81	0.36
carbon tetrachloride	64	0.31	68	0.31
diethyl ether	56	0.37	61	0.45
benzene	57	0.32	61	0.36
n-heptane	54	0.30	57	0.32
n-hexane	49	0.28	52	0.32
total	1025	0.36	1200	0.40

[a] Mean absolute error.

Performance of the surface-SFED model in water is compared with the Shell-SFED model[29] in Table 10. The surface-SFED is adapted almost 4 times larger and much more diverse training set for its parametrization and MAEs of the models are comparable. However, there are considerable improvements for some functional classes. MAE of the amide in surface-SFED model is decreased almost half of the Shell-SFED model and of halo

Table 10. Performance of the Surface-SFED Model in Water Compared with Shell-SFED Model by Solute Functional Class

	shell-SFED model[a]		Surface-SFED model	
solute class	number of solute	MAE[b]	number of solute	MAE[b]
alkanes	9	0.23	15	0.55
cycloalkanes			6	0.38
alkenes	4	0.19	20	0.32
alkynes	3	0.19	8	0.35
alcohols	13	0.17	32	0.27
arenes	20	0.14	16	0.24
ethers	9	0.54	19	0.75
aldehydes			9	0.31
ketones	7	0.28	17	0.27
carboxylic acids	3	0.11	5	0.29
esters	4	0.43	29	0.17
aliphatic amines	8	1.52	23	1.08
aromatic amines	4	1	16	0.11
nitriles			4	0.14
nitrohydrocarbons			7	0.87
amides and ureas	3	1.89	6	0.99
thiols			4	0.54
sulfides			5	0.54
disulfides			2	0.54
fluorinatd hydrocarbons	2	2.13	7	0.82
chlorinated hydrocarbons	4	0.18	38	0.48
brominated hydrocarbons	2	0.26	19	0.33
iodinated hydrocarbons			8	0.25
other halo compounds			16	0.46
bifunctional compounds			31	1.12
inorganic compound			3	0.64
total	95	0.47	365	0.49

[a] Ref 29. [b] Mean absolute error.

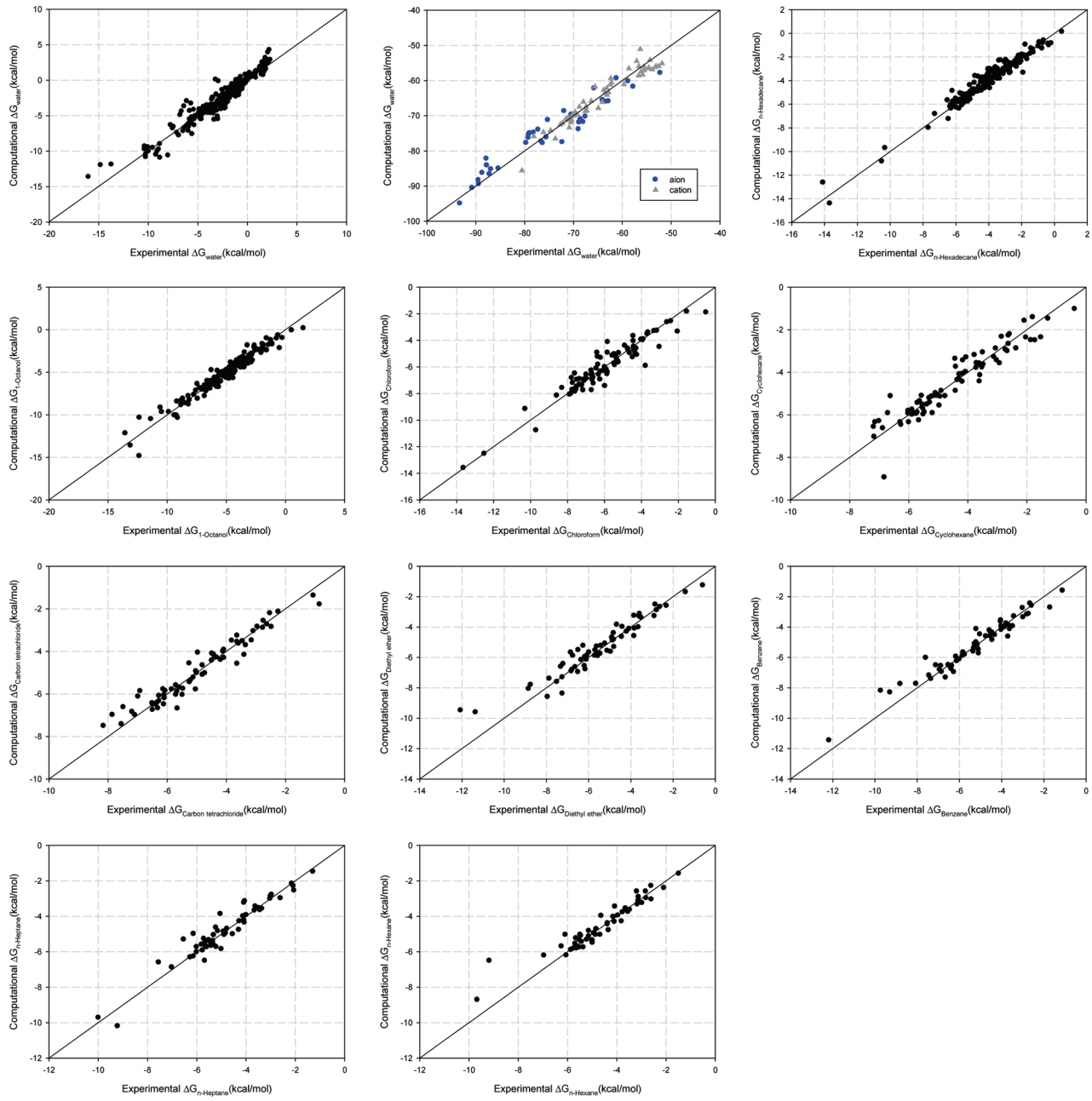


Figure 2. Comparison of calculated and experimental solvation free energies in each solvent model.

Table 11. Solvation Free Energies of the N-Methylated Nucleic Acid Bases[a] in Water and Chloroform

	water					chloroform	
DNA bases	SFED	SM5.4[b]	Miller et al.[c]	Galesa et al.[d]	exp	SFED	SM5.4[b]
9-methyladenine	−16.05	−15.2	−11.65	−11.25	−13.6	−12.38	−13.3
9-methylguanine	−20.38	−20.1	−22.65	−25.34		−14.04	−15.9
1-methylthymine	−10.24	−9.6	−12.58		−10.4	−10.63	−9.2
1-methylcytosine	−12.87	−19.9	−8.09			−10.91	−15.5

[a] Energies in kcal/mol. [b] Ref 58. [c] Ref 59. [d] Ref 60.

compounds are much improved or comparable despite of the inclusion of perhalogenated compounds. Despite of the appreciable improvement in amine class, surface-SFED model is still need more elaborate description of the dependency for the order.

The surface-SFED model was applied to calculate the solvation free energies of the N-methylated nucleic acid bases in water and chloroform. Table 11 compares our results to experimental data and other works which include semiempirical,[58] explicit solvation[59] and ab initio models.[60] We find that surface-SFED model slightly overestimate the solvation free energy of 9-methyladenine and are well within the experimental range for 1-methylthymine in water. In general, the surface-SFED model agrees

as well with the semiempirical model except that the solvation free energy of 1-methylcytosine is underestimated.

CONCLUSION

The previous SFED model was successful in calculating the solvation free energy of limited cases of neutral solutes. However, this model had limited applicability and restrictions, such as bumping of the grid between molecules occupying a short distance. In this study, we modified the SFED model to overcome these shortcomings. First, the surface model was introduced to improve computational efficiency and increase applicability. The basis function for electrostatic potential was modified to calculate more accurate solvation free energies of ionic molecules. Two basis functions were introduced to express contributions from hydrogen bonding, and the hydrogen bond coefficients were determined with various solvents. Finally, the parametrization process was performed with a large, diverse set of solvation free energies that included highly polar and ionic molecules.

Because of the limited availability of experimental data, only 10 solvents were investigated. The calculated solvation free energies of a wide range of solute functional groups in diverse solvents were consistent with experimental data. MAE for 1200 neutral solutes in these solvents was 0.40 kcal/mol, and MAE for 90 hydration free energies of ions was 1.7 kcal/mol.

Although the linear combination coefficients for each solvent model were determined independently, some consistency was observed depending on the solvent properties. Since the surface-SFED model produced reliable results for predicting experimental solvation free energies for diverse solvents, this model can be further developed into a more general model if the linear expansion coefficients can be calculated as a function of solvent properties.

By applying the surface cavity, the molecular interactions between molecules in a short distance can be described accurately with shorter computational time compared to the shell-SFED model. The surface-SFED model can calculate the physical properties of small and bio molecules with greater accuracy and can be used to develop three-dimensional descriptors. Moreover, the surface-SFED model could be applied to even larger molecules, such as proteins and their interaction with organic molecules and solvents. Therefore the surface-SFED model combined with other method such like docking and QSAR can be a powerful and practical tool for rational drug design and protein engineering.

ACKNOWLEDGMENT

This work was partly supported by the IT R&D program of MKE/IITA [KI001826, Development of e-Organ system based on Cyber Computing] and Korea Science and Engineering Foundation(KOSFF) grant funded by Korea government (MEST)[2010-0001917, Computer-aided molecular modeling and optimization for function-controlled materials].

Note Added after ASAP Publication. This paper was published ASAP on December 6, 2010, with a spelling error in an author's name. The corrected version was published ASAP on December 9, 2010.

Supporting Information Available: Experimental and calculated solvation free energies for neutral solutes used in the parametrizations of the SFED model (Table S1) and experimental and calculated solvation free energy of anions and cations in aqueous solution used in the parametrizations of the SFED model (Tables S2 and S3). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Tapia, O. Solvent effect theories: Quantum and classical formalisms and their applications in chemistry and biochemistry. *J. Math. Chem.* **1992**, *98*, 139–181.
(2) Kollman, P. Free energy calculations: Applications to chemical and biological phenomena. *Chem. Rev.* **1993**, *93*, 2395–2417.
(3) Hine, J.; Mookerjee, P. K. The intrinsic hydrophilic character of organic compound. Correlations in terms of structural contributions. *J. Org. Chem.* **1975**, *40*, 292–298.
(4) Viswanadhan, V. N.; Ghose, A. K.; Singh, U. C.; Wendoloski, J. J. Prediction of solvation free energies of small organic molecules: Additive-constitutive models based on molecular fingerprints and atomic constants. *J. Chem. Inf. Comput. Sic.* **1999**, *39*, 405–412.
(5) Ooi, T.; Oobatake, M.; Némethy, G.; Scheraga, H. A. Accessible surface areas as a measure of the thermodynamics parameters of hydration of peptides. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086–3090.
(6) Eisenberg, D.; McLachlan, A. D. Solvation energy in protein folding and binding. *Nature* **1986**, *319*, 199–203.
(7) Jorgensen, W. L. Free energy calculations: A breakthrough for modeling organic chemistry in solution. *Acc. Chem. Res.* **1989**, *22*, 184–189.
(8) Smith, P. E.; Pettitt, B. M. Modeling solvent in biomolecular systems. *J. Phys. Chem.* **1994**, *98*, 9700–9711.
(9) Nicholas, M.; Arianna, W. R.; Marshall, N. R.; Augusta, H. T.; Edward, T. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
(10) Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* **1957**, *27*, 1208.
(11) Aneesur, R.; Frank, H. S. Molecular dynamics study of liquid water. *J. Chem. Phys.* **1971**, *55*, 3336–3359.
(12) Born, M. Volumen und hydratationswrme der ionen. *Z. Phys.* **1920**, *1*, 45–48.
(13) Bell, R. P. The electrostatic energy of dipole molecules in different media. *Trans. Faraday. Soc.* **1931**, *27*, 797–802.
(14) Kirkwood, J. G. Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *J. Chem. Phys.* **1934**, *2*, 351–361.
(15) Onsager, L. Electric moments of molecules in liquids. *J. Am. Chem. Soc.* **1936**, *58*, 1486–1493.
(16) Tomasi, J.; Persico, M. Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent. *Chem. Rev.* **1994**, *94*, 2027–2094.
(17) Cramer, C. J.; Truhlar, D. G. Implicit solvation models: Equilibria, structure, spectra, and dynamics. *Chem. Rev.* **1999**, *99*, 2161–2200.
(18) Bren, U.; Martínek, V.; Florián, J. Decomposition of the solvation free energies of deoxyribonulceoside using the free energy perturbation method. *J. Phys. Chem. B* **2006**, *110*, 12782–12788.
(19) Bren, M.; Florián, J.; Mavri, J.; Bren, U. Do all pieces make a whole? Thiele cumulants and the free energy decomposition. *Theor. Chem. Acc.* **2007**, *117*, 535–540.
(20) Still, W. C.; Tempczyk, A.; Haawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
(21) Lazaridis, T.; Karplus, M. Effective energy function for proteins in solution. *Proteins* **1999**, *35*, 133–152.
(22) Tannor, D. J.; Marten, B.; Murphy, R.; Friesener, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A., III; Honig, B. Accurate first principles calculation of molecular charge distributions and solvation energies from ab initio quantum mechanics and continuum dielectric theory. *J. Am. Chem. Soc.* **1994**, *116*, 11845–11882.
(23) Sharp, K. A.; Honig, B. Electrostatic interactions in macromolecules: Theory and applications. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301–332.
(24) Honig, B.; Nicholls, A. Classical electrostatics in biology and chemistry. *Science* **1995**, *268*, 1144–1149.

(25) Frank, N. H.; Tobocman, W. *Fundamental Formulas of Physics*; Menzel, D. H., Ed.; Dover: New York, 1960; Vol. 1, pp 307−649.
(26) Malcolm, N. O. J.; McDouall, J. J. W. Assessment of the Langevin dipoles solvation model for Hartree-Fock wavefunctions. *J. Mol. Struct. (THEOCHEM)* **1996**, *366*, 1–9.
(27) Florián, J.; Warshel, A. Calculations of hydration entropies of hydrophobic, polar, and ionic solutes in the framework of the Langevin dipoles solvation model. *J. Phys. Chem. B* **1999**, *103*, 10282–10288.
(28) No, K. T.; Kim, S. G.; Cho, K.-H.; Scheraga, H. A. Description of hydration free energy density as a function of molecular physical properties. *Biophys. Chem.* **1999**, *78*, 127–145.
(29) In, Y.; Chai, H. H.; No, K. T. A partition coefficient calculation method with the SFED model. *J. Chem. Inf. Model.* **2005**, *45*, 254–263.
(30) No, K. T.; Grant, J. A.; Scheraga, H. A. Determination of net atomic charges using a modified partial equalization of orbital electronegativity method. 1. Application to neutral molecules as models for polypeptides. *J. Phys. Chem.* **1990**, *94*, 4732–4739.
(31) No, K. T.; Grant, J. A.; Jhon, M. S.; Scheraga, H. A. Determination of net atomic charges using a modified partial equalization of orbital electronegativity method. 1. Application to ionic and aromatic molecules as models for polypeptides. *J. Phys. Chem.* **1990**, *94*, 4140–4746.
(32) No, K. T.; Cho, K. H.; Jhon, M. S.; Scheraga, H. A. An empirical method to calculate average molecular polarizabilities from the dependence of effective atomic polarizabilities on net atomic charge. *J. Am. Chem. Soc.* **1993**, *115*, 2005–2014.
(33) Raevsky, O. A.; Grigori've, V. Y.; Kireev, D. B.; Zefirov, N. S. Complete thermodynamic description of H-bonding in the framework of multiplicative approach. *Quant. Struct.−Act. Relat.* **1992**, *11*, 49–63.
(34) Raevsky, O. A. Molecular structure descriptors in the computer-aided design of biologically active compounds. *Russ. Chem. Rev.* **1999**, *68*, 505–524.
(35) Abraham, M. H. Scales of solute hydrogen-bonding: Their construction and application to physicochemical processes. *Chem. Soc. Rev.* **1993**, *22*, 73–83.
(36) Abraham, M. H. Hydrogen bonding. 31. Construction of a scale of solute effective or summation hydrogen-bond basicity. *J. Phys. Org. Chem.* **1993**, *6*, 660–684.
(37) Abraham, M. H. Hydrogen bonding. 32. An analysis of water−octanol and water−alkane partitioning and the logp parameter of seiler. *J. Pharm. Sci.* **1994**, *83*, 1085–1100.
(38) Meot-Ner (Mautner), M. The ionic hydrogen bond and solvation. 1. $NH^+\cdots O$, $NH^+\cdots N$, and $OH^+\cdots O$ bonds. Correlations with proton affinity. Deviations due to structure effects. *J. Am. Chem. Soc.* **1984**, *106*, 1257–1264.
(39) Meot-Ner (Mautner), M. The ionic hydrogen bond and ion solvation. 2. Solvation of onium ions by one to seven H_2O molecules. Relations between monomolecular, specific, and bulk hydration. *J. Am. Chem. Soc.* **1984**, *106*, 1265–1272.
(40) Speller, C. V.; Meot-Ner (Mautner), M. The ionic hydrogen bond and ion solvation. 3. Bonds involving cyanides. Correlation with proton affinities. *J. Phys. Chem.* **1985**, *89*, 5217–5222.
(41) Sieck, L. W.; Meot-Ner (Mautner), M. The ionic hydrogen bond and ion solvation. 4. $SF^+\cdots O$ and NH^+-S bonds. Correlations with proton affinity. Mutual effects of weak and strong ligands in mixed clusters. *J. Phys. Chem.* **1985**, *89*, 5222–5225.
(42) Sieck, L. W.; Meot-Ner (Mautner), M. The ionic hydrogen bond and ion solvation. 5. OH...O^- bonds. Gas-phase solvation and clustering of alkoxide and carboxylate anions. *J. Am. Chem. Soc.* **1986**, *108*, 7525–7529.
(43) Meot-Ner (Mautner), M. Ionic hydrogen bond and ion solvation. 6. Interaction energies of the acetate ion with organic molecules. Comparison of CH_3COO^- with Cl^-, CN^-, and SH^-. *J. Am. Chem. Soc.* **1988**, *110*, 3854–3858.
(44) Meot-Ner (Mautner), M. The ionic hydrogen bond and ion solvation. 7. Interaction energies of carbanions with solvent molecules. *J. Am. Chem. Soc.* **1988**, *110*, 3858–3862.
(45) Sieck, L. W.; Meot-Ner (Mautner), M. Ionic hydrogen bond and ion solvation. 8. $RS^-\cdots HOR$ bond strengths. Correlation with acidities. *J. Am. Chem. Soc.* **1989**, *93*, 1586–1588.
(46) *Masterfile*; MedChem Software, BioByte Corp: Claremont, CA, 1994.
(47) Wang, J.; Wang, W.; Huo, S.; Lee, M.; Kollman, P. A. Solvation model based on weighted solvent accessible surface area. *J. Phys. Chem. B* **2001**, *105*, 5055–5067.
(48) Pearson, R. G. Ionization potentials and electron affinities in aqueous solution. *J. Am. Chem. Soc.* **1986**, *108*, 6109–6114.
(49) Zhan, C.-G.; Dixon, D. A. Absolute hydration free energy of the proton from first-principles electronic structure calculations. *J. Phys. Chem. A* **2001**, *105*, 11534–11540.
(50) Lias, S. G.; Bartness, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G. 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; http://webbook.nist.gov (accessed in 2008).
(51) Hunter, E. P. L.; Lias, S. G. Evaluated gas phase basicities and proton affinities of molecules: An update. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413–656.
(52) Tehan, B. G.; Lloyd, E. J.; Wong, M. G.; Pitt, W. R.; Montana, J. G.; Manallack, D. T.; Gancia, E. Estimation of pK_a using semiempirical molecular orbital method. Part 1: Application to phenols and carboxylic acids. *Quant. Struct.−Act. Relat.* **2002**, *21*, 457–472.
(53) Tehan, B. G.; Lloyd, E. J.; Wong, M. G.; Pitt, W. R.; Gancia, E.; Manallack, D. T. Estimation of pK_a using semiempirical molecular orbital method. Part 2: Application to amines, anilines and various nitrogen containing heterocyclic compounds. *Quant. Struct.−Act. Relat.* **2002**, *21*, 473–485.
(54) *Physical/Chemical Property DataBase (PHYSPROP)*; SRC Environmental Science Center: Syracuse, NY, 1994.
(55) Albert, A.; Serjeant. E. P. *The Determination of Ionization Constants: A Laboratory Manual*; Chapman and Hall: New York, 1984.
(56) Tapia, O. In *Molecular Interactions*; Ratajczak, H., Orville-Thomas, W. J., Eds.; John Wiley & Sons: Chichester, U.K., 1982; Vol. 3, pp 47−118.
(57) Ashbaugh, H. S.; Paulaitis, M. E. Effect of solute size and solute−water attractive interactions on hydration water structure around hydrophobic solutes. *J. Am. Chem. Soc.* **2001**, *123*, 10721–10728.
(58) Giesen, D. J.; Chambers, C. C.; Cramer, C. J.; Truhlar, D. G. What controls partitioning of the nucleic acid bases between chloroform and water. *J. Phys. Chem. B* **1997**, *101*, 5084–5088.
(59) Miller, J. L.; Kollman, P. A. Solvation free energies of the nucleic acid bases. *J. Phys. Chem.* **1996**, *100*, 8587–8594.
(60) Galeă, K.; Bren, U.; Kranjc, A.; Mavri, J. Carcinogenicity of acrylamide: A computational study. *J. Agric. Food Chem.* **2008**, *56*, 8720–8727.

CI100299M