

# Predicting Aqueous Free Energies of Solvation as Functions of Temperature

Adam C. Chamberlin, Christopher J. Cramer,\* and Donald G. Truhlar\*

Department of Chemistry and Supercomputing Institute, 207 Pleasant Street SE, University of Minnesota, Minneapolis, Minnesota 55455-0431

Received: December 12, 2005; In Final Form: January 25, 2006

This work introduces a model, solvation model 6 with temperature dependence (SM6T), to predict the temperature dependence of aqueous free energies of solvation for compounds containing H, C, and O in the range 273–373 K. In particular, we extend solvation model 6 (SM6), which was previously developed (Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133) for predicting aqueous free energies of solvation at 298 K, to predict the variation of the free energy of solvation relative to 298 K. Also, we describe the database of experimental aqueous free energies of solvation for compounds containing H, C, and O that was used to parametrize and test the new model. SM6T partitions the temperature dependence of the free energy of solvation into two components: the temperature dependence of the bulk electrostatic contribution to the free energy of solvation, which is computed using the generalized Born equation, and the temperature dependence of first-solvation-shell effects which is modeled using a parametrized solvent-exposed surface-area-dependent term. We found that SM6T predicts the temperature dependence of aqueous free energies of solvation with a mean unsigned error of 0.08 kcal/mol over our entire database, whereas using the experimental value at 298 K produces a mean unsigned error of 0.53 kcal/mol.

## 1. Introduction

Implicit solvation models<sup>1–4</sup> are widely used to predict a variety of solvent effects including free energies of solvation, solubilities, and vapor pressures. Often, as in industrial, environmental, and pharmacological applications, these quantities are needed over a range of temperatures; however, almost all implicit solvation models have been designed or tested only for predicting these quantities at room temperature. An exception is the work of Elcock and McCammon<sup>5</sup> who introduced an implicit solvation model for temperature-dependent hydration free energies of amino acids. There remains a need for models designed to treat a broad variety of solutes such as atmospheric pollutants,<sup>6–8</sup> drug leads,<sup>6,9</sup> high-energy molecules,<sup>10</sup> and reagents for organic synthesis.<sup>11</sup>

In the present paper, we present a temperature-dependent extension of our recent implicit solvation model, solvation model 6 (SM6).<sup>12</sup> We refer to this extension as solvation model 6 with temperature dependence (SM6T). SM6T is designed to predict changes in aqueous free energies of solvation as functions of temperature relative to the free energy of solvation at 298 K. In the present study, it was parametrized for compounds composed of H, C, and O in the temperature range 273–373 K (the range over which water is a liquid at 1 atm). SM6T is parametrized on the basis of the solute geometry and the atomic numbers of the atoms in the solute.

The standard-state temperature dependence of the free energy of solvation is given by<sup>13,14</sup>

$$\Delta G_S^\circ = \Delta H_S^\circ(T) - T\Delta S_S^\circ(T) \quad (1)$$

where  $T$  is temperature and  $\Delta H_S^\circ$  and  $\Delta S_S^\circ$  are, respectively, the standard-state enthalpy and entropy of solvation. Ignoring the

temperature dependence of  $\Delta H_S^\circ$  and  $\Delta S_S^\circ$  yields a linear dependence of  $\Delta G_S^\circ$  in  $T$ ; we call this the van't Hoff model. The temperature dependence of both  $\Delta H_S^\circ$  and  $\Delta S_S^\circ$  can be written in terms of the standard-state heat capacity of solvation,  $\Delta C_{P,S}^\circ$ , which yields<sup>14</sup>

$$\Delta G_S^\circ(T) = \Delta G_S^\circ(T_0) - \Delta S_S^\circ(T_0)[T - T_0] + \Delta C_{P,S}^\circ(T)[(T - T_0) - T \ln(T/T_0)] \quad (2)$$

where  $T_0$  is a reference temperature taken throughout as 298 K. Equation 2 with  $\Delta C_{P,S}^\circ$  taken as a constant will be the basis of our analysis. Experiment has shown that  $\Delta C_{P,S}^\circ$  is not necessarily a constant; for example, Bakk and Høye<sup>15</sup> found that it decreases by 12–35% for methane, ethane, and propane in water when  $T$  is increased from 273 to 323 K. One could of course include the temperature dependence of  $\Delta C_{P,S}^\circ$  in the model. However, such higher-order effects are probably beyond the resolution of our model. In the method presented here, we go beyond the van't Hoff model by including a nonzero heat capacity, but we do not include the temperature dependence of the heat capacity.

In section 2, we discuss the development of a database of temperature-dependent free energies of solvation in water. Section 3 presents the new model for predicting the temperature dependence of the free energies of solvation. Section 4 summarizes our software plans. The final section discusses the accuracy and applicability of the model.

## 2. Developing a Database

For a robust parametrization, it was necessary to obtain a database consisting of free energies of solvation as functions of temperature for a variety of compounds. First, we searched the literature for experimental measurements that could be used to obtain free energies of solvation. Next, we analyzed the data and discarded those that failed to meet certain criteria. Finally,

\* To whom correspondence should be addressed. E-mail: cramer@chem.umn.edu (C.J.C.); truhlar@umn.edu (D.G.T.).

we recorded the data as free energies of solvation relative to the experimental value at 298 K.

**2.A. Compilation of Free Energies of Solvation.** We obtained free energies of solvation from a variety of experimental sources including Henry's law constants,<sup>16–18</sup> saturated vapor pressures<sup>19–22</sup> of the solute over the pure liquid phase combined with aqueous solubilities,<sup>23–27</sup> and activity coefficients at infinite dilution<sup>28–31</sup> for the solutes in water. As with all of our prior models,<sup>32–34</sup> the free energies of solvation are tabulated for the standard-state choices of 1 M ideal gas and 1 M ideal solution, where the latter behaves as an infinitely dilute solution. As discussed by Ben-Naim,<sup>35</sup> such a choice of standard state, involving the same concentration in the two phases, removes the translational entropy of dilution.

To obtain the free energy of solvation, we considered the equilibrium between a solute (A) in the gas phase (g) and in aqueous solution (aq)



In the case of an infinitely dilute solution, the partition coefficient for this process is the Henry's law constant,  $K_H$ , which is related to the standard-state free energy of solvation by<sup>14,36</sup>

$$\Delta G_S^\circ = RT \ln(K_H) \quad (4)$$

where  $R$  is the gas constant<sup>14,37</sup> (1.985 cal K<sup>-1</sup> mol<sup>-1</sup>). Henry's law constants are reported in a variety of units, and considerable care must be exercised in converting the experimental data into the appropriate units; Staudinger and Roberts<sup>38</sup> provide a useful review on this subject. The Henry's law constant corresponding to our choice of standard states has units of (M of solute in gas phase/M of solute in solution).

Another widely measured quantity is the activity coefficient at infinite dilution. The activity coefficient is the constant of proportionality between the Henry's law constant and the Raoult's law<sup>14,37</sup> constant and is written as<sup>28,29,31,36</sup>

$$\gamma_A^\infty = \frac{K_H^X}{P_A^\bullet} \quad (5)$$

where  $K_H^X$  is the Henry's law constant in units of mole fractions in the gas and liquid phases and  $P_A^\bullet$  is the saturated vapor pressure of solute A over pure liquid solute A. The standard-state free energy of solvation can be calculated from the activity coefficient by

$$\Delta G_S^\circ = RT \ln\left(\frac{\gamma_A^\infty P_A^\bullet}{RTM_w^\bullet}\right) \quad (6)$$

where  $M_w^\bullet$  is the molarity of pure water, which serves as the unit conversion between mole fractions and molarity for the solute concentration in dilute aqueous solution. Equation 6 is derived in the appendix, which corrects an error in ref 36.

Alternatively, the solvation process can be considered to occur in two steps. The first is condensing the solute



where l denotes the pure liquid solute. The standard-state free energy associated with moving a solute from a 1 M concentration of solute A in the gas phase to a liquid solute A with ideal behavior in both phases is<sup>33,34</sup>

$$\Delta G_C^\circ = RT \ln\left(\frac{P_A^\bullet}{RTM_A^1}\right) \quad (8)$$

where  $M_A^1$  is the molarity of the solute in its pure liquid state. The second step of the process is transferring solute A from its pure liquid phase to aqueous solution



The standard-state free energy associated with this step is<sup>34</sup>

$$\Delta G_T^\circ = -RT \ln\frac{M_A^{\text{aq}}}{M_A^1} \quad (10)$$

where  $M_A^{\text{aq}}$  is the aqueous molar concentration of solute A in equilibrium with pure liquid; that is, it is the molarity of a saturated solution, which is easily calculated from the solubility. When eqs 8 and 10 are combined, the standard-state free energy of solvation is given by

$$\Delta G_S^\circ = RT \ln\frac{P_A^\bullet}{RTM_A^{\text{aq}}} \quad (11)$$

where eq 11 assumes that the solute is an ideal gas in the vapor phase and that the solution of solute A behaves as though it is very dilute so that a solute molecule only interacts with the solvent.

In practice, several free energies of solvation were computed from eq 11 for highly soluble compounds, for example, benzyl alcohol, which has a solubility of 0.41 mol/L at 293 K.<sup>39</sup> Equation 11 is not necessarily appropriate for such compounds; however, we have shown in a previous paper<sup>34</sup> that the error introduced by assuming ideal behavior for concentrated solutions is smaller than the inherent error of our solvation models.

A separate issue is that experimental solute vapor pressures and solubilities are rarely available at precisely the same temperature; thus, to calculate free energies of solvation, it was necessary to interpolate experimental vapor pressures to estimate values at the temperatures for which experimental solubilities were available. Fortunately, empirical fits of vapor pressures with accuracies comparable to the experimental errors of solubilities<sup>39–42</sup> are readily available in the literature.<sup>43,44</sup>

**2.B. Analysis and Removal of Data.** Once we obtained an extensive body of experimental free energies of solvation, we proceeded to analyze the data for each compound. We discarded experimental data on the basis of five criteria:

(1) We discarded any source that provided experimental data at only a single temperature for a given compound, except for the case of the free energies of solvation at 298 K obtained from the database used to parametrize SM6.<sup>12</sup> We used this particular criterion because combining such data with data at other temperatures from other sources may lead to spurious temperature dependence. We retained the data from the database used to parametrize the SM6 model because we found these to be reliable values of free energies of solvation at 298 K.

(2) We eliminated all experimental data from sources whose data points systematically deviated by 0.4 kcal/mol or more from values obtained from two or more other sources.

(3) We discarded any compound for which less than four experimental points were available. To estimate the curvature of the temperature dependence of the free energy, at least three data points are necessary. In most cases, the number of data

points is larger, but for a few compounds, such as sparingly soluble alkynes, we did actually use only four data points.

(4) We required that no data point deviate by more than two standard deviations from a fit, as described in section 2.C, to all of the retained experimental data for a given compound. Once a data point was removed, the fit was recalculated using the remaining data. This process was iterated until none of the remaining points deviated by more than two standard deviations from the fit. This criterion was applied to eliminate nonsystematic errors in the experimental data.

(5) In the process of assessing the data, we came to the conclusion that some sources were unreliable either because those sources often disagreed with other workers or because they did not provide complete enough experimental details; data from those sources were not used.

**2.C. Evaluation of Experimental Temperature Dependence.** With experimental free energies of solvation for a given compound in hand, we calculated the temperature dependence using the following equation

$$\Delta\Delta G_S^\circ = \Delta G_S^\circ(T) - \Delta G_S^\circ(T_0) \quad (12)$$

where  $\Delta G_S^\circ$  is the experimental free energy of solvation at a given temperature,  $T$ . To tabulate  $\Delta\Delta G_S^\circ$ , one requires a value of  $\Delta G_S^\circ(T_0)$  that is consistent with the  $\Delta G_S^\circ(T)$  data. In some cases, however, 298 K was not among the available temperatures in the data set produced by the steps in sections 2.A and 2.B. Furthermore, even when available,  $\Delta G_S^\circ$  at any one temperature such as  $T_0$  is subject to experimental noise. To make the data as consistent as possible, we always obtained  $\Delta G_S^\circ(T_0)$  by fitting all of the retained experimental data for a given compound,  $i$ , to

$$\Delta G_S^\circ(T; i) = a_i + b_i(T - T_0) + c_i[(T - T_0) - T \ln(T/T_0)] \quad (13)$$

where  $a_i$ ,  $b_i$ , and  $c_i$  are fitting parameters. The form of eq 13 is motivated by eq 2, which yields

$$a_i = \Delta G_S^\circ(T_0) \quad (14)$$

During the development of SM6T, we considered using a number of different functions instead of eq 13 including polynomial fits, the van't Hoff model, and a commonly used variant of the thermodynamic equation for the free energy of solvation proposed by Clarke and Glew.<sup>45</sup> We found that both the equation proposed by Clarke and Glew<sup>45</sup> and a second-order polynomial produced fits of equal accuracy to eq 13, but as neither was *more* accurate, we retained the form of eq 13.

The overall product of this analysis is a database of evaluated data. A listing of the individual compounds, the number of experimental points used for each compound, the temperature range covered, and the sources for the data used in the database is available in the Supporting Information.

The database is composed of 2364 aqueous free energies of solvation with temperatures ranging from 273 to 373 K for 182 compounds composed of H, C, and O. The database includes a variety of types of compounds including alkanes, alkenes, alkynes, cyclic alkanes, aromatics, alcohols, alkenols, cyclic alcohols, phenols, ethers, cyclic ethers, esters, and carboxylic acids. Table 1 shows that the database has a disproportionately large number of alcohols in comparison to the number of alkynes. Due to the experimental difficulties associated with measuring the equilibria of some of these solutes, particularly branched alkanes and unsaturated alkanes in aqueous solution, we often had to find compounds with combinations of func-

**TABLE 1: Number of Free Energies of Solvation in Database**

class of compound	no. of points	no. of compounds
unbranched alkanes	123	8
branched alkanes	49	6
cyclic alkanes	96	11
alkenes and alkynes	44	5
aromatics	459	22
alcohols, phenols, and water	702	47
aldehydes	142	13
ketones	236	18
ethers	115	12
esters	332	32
carboxylic acids	58	7
total	2356	181

tionalities that were more soluble. To ensure that the fit does not suffer from this uneven availability of data, the database was divided into 11 classes, based on functional groups, and a weighted error function was used in the parametrization. This will be described further in section 3.C.

### 3. Model Development

We must first describe certain key aspects of SM6 so that a description of SM6T will be clear. Then, we will introduce the functional form used to predict the temperature dependence of the free energy of solvation. We will also discuss tentative parametrizations of the model. Finally, we will present the final choice of functional form and the parameters used to develop the model.

**3.A. SM6.** SM6 is the latest in a sequence of closely related solvation models, with the original called SM1<sup>32,46</sup> and the penultimate called SM5.43R;<sup>47,48</sup> collectively, the models are referred to as SMx. SM6 is an implicit solvation model<sup>1,2,49,50</sup> used to calculate the free energy of solvation by approximating the solvent medium as a dielectric continuum, approximating the solute charge density as a collection of atom-centered partial atomic charges and approximating the solute cavity as a set of overlapping atom-centered spheres. In SM6, the free energy of solvation is partitioned into two contributions

$$\Delta G_S^\circ = \Delta G_{\text{ENP}} + G_{\text{CDS}} \quad (15)$$

where  $\Delta G_{\text{ENP}}$  is the contribution due to electrostatic interactions between the solute and the bulk solution and  $G_{\text{CDS}}$  is a parametrized term designed to account for first-solvation-shell effects and for approximations used in the calculation of  $\Delta G_{\text{ENP}}$ . As discussed elsewhere,<sup>2,50,51</sup> the decomposition of eq 15 is not well defined. However, a key working principle of the SMx models is that we define  $\Delta G_{\text{ENP}}$  as the Coulombic contribution calculated by bulk electrostatics for a given choice of intrinsic atomic radii called the Coulomb radii, which are assumed to be independent of atomic charge, hybridization state, and solvent. The usefulness of this principle has been validated by the success of the SMx models. To anticipate section 3.B, we note that a key element of SM6T is that these Coulomb radii are independent of temperature.

The bulk electrostatic component of the free energy of solvation is calculated as the sum of two contributions

$$\Delta G_{\text{ENP}} = \Delta E_{\text{EN}} + G_{\text{P}} \quad (16)$$

where  $\Delta E_{\text{EN}}$  is the change in the solute's internal free energy upon insertion into solution from the gas phase and  $G_{\text{P}}$  is the free energy of polarization. In practice, we usually retain only

271 the electronic contribution to  $\Delta E_{\text{EN}}$ , assuming that the vibrational–  
 272 librational–rotational contributions are small enough to be  
 273 neglected.

274 The free energy of polarization is the free energy change  
 275 associated with a bulk electrostatic treatment of the favorable  
 276 solute–solvent electrostatic interactions that form upon insertion  
 277 of the solute into the solvent and the accompanying distortion  
 278 of the solvent molecules and the solvent structure from the pure  
 279 solvent state. The electrostatic interactions between the partial  
 280 atomic charges of the solute and the surrounding dielectric  
 281 medium are calculated using the generalized Born equation<sup>32,52–54</sup>

$$G_{\text{P}} = -\frac{1}{2} \left( 1 - \frac{1}{\epsilon} \right) \sum_{k,k'} q_k q_{k'} \gamma_{kk'} \quad (17)$$

282 where the right-hand side of the equation is a double sum over  
 283 all the atoms  $k$  and  $k'$  in the solute,  $q_k$  is the partial atomic charge  
 284 of atom  $k$  computed using charge model 4 (CM4),<sup>12</sup>  $\epsilon$  is the  
 285 temperature-dependent dielectric constant of the solvent, and  
 286  $\gamma_{kk'}$  is the Coulomb integral between atoms  $k$  and  $k'$ . The  
 287 Coulomb integral is calculated as<sup>54</sup>

$$\gamma_{kk'} = \frac{1}{\sqrt{R_{kk'}^2 + \alpha_k \alpha_{k'} \exp(r_{kk'}/d_{kk'}) \alpha_k \alpha_{k'}}} \quad (18)$$

288 where  $r_{kk'}$  is the distance between atoms  $k$  and  $k'$ ,  $d_{kk'}$  is a  
 289 parameter,<sup>12</sup> and  $\alpha_k$  is an effective Born radius for atom  $k$  that  
 290 depends on the solute geometry and the Coulomb radius of each  
 291 atom in the solute. The Coulomb radii were determined in a  
 292 prior paper.<sup>12</sup>

293 The change in internal energy upon insertion of the solute  
 294 into liquid solution from the gas phase,  $\Delta E_{\text{EN}}$ , has two  
 295 contributions:  $\Delta E_{\text{E}}$  which is the energy required to distort the  
 296 solute electronic structure from its gas-phase optimum and  $\Delta E_{\text{N}}$   
 297 which is the equivalent cost associated with distortion of the  
 298 solute geometry from its gas-phase optimum. The total  $\Delta G_{\text{ENP}}$   
 299 is minimized in a self-consistent reaction field (SCRf)<sup>2,55</sup>  
 300 calculation. The energy change associated with distortion of the  
 301 solute geometry,  $\Delta E_{\text{N}}$ , can be calculated by optimizing the solute  
 302 geometry for the relaxed solute electronic structure. This latter  
 303 process can be relatively time-consuming, since the solute  
 304 electronic structure is iteratively computed for each step in the  
 305 optimization of the solute geometry in solution. One option that  
 306 has been explored in previous work<sup>56</sup> is to use the solute gas-  
 307 phase geometry and to allow only the electronic structure to  
 308 relax in solution. We have found that this approximation yields  
 309 reliable results for most cases; in particular, the effect of  
 310 optimization of the solute structure in solution on the overall  
 311 free energy of solvation is relatively small for molecules used  
 312 in our parametrization databases.<sup>56</sup>

313 As mentioned above,  $G_{\text{CDS}}$  is an empirically derived term  
 314 designed to minimize the deviation between experiment and  
 315  $\Delta G_{\text{ENP}}$ . It is intended to account for solvation phenomena that  
 316 are not included in the bulk electrostatic term. It includes the  
 317 free energy cost associated with forming a solute-shaped cavity  
 318 in the solvent, changes in solvent structure beyond those  
 319 included in  $\Delta E_{\text{EN}}$ , nonelectrostatic phenomena such as dispersion  
 320 and Pauli exchange–repulsion, and any approximations used  
 321 in the calculation of  $\Delta G_{\text{ENP}}$ . Previous studies<sup>32,49,57–60</sup> have  
 322 shown that these phenomena are proportional to the solvent  
 323 accessible surface area<sup>61–64</sup> (SASA) of the solute cavity. Note  
 324 that the atomic radii used to define the SASA for the  $G_{\text{CDS}}$  term  
 325 are *not* in general the same as the Coulomb radii. In particular,  
 326 we use the radii of Bondi<sup>65</sup> for this purpose.

The  $G_{\text{CDS}}$  term has the following form

$$G_{\text{CDS}} = \sum_k \sigma_k A_k \quad (19)$$

where the sum is over all atoms  $k$  in the solute,  $A_k$  is the solvent  
 accessible surface area of atom  $k$ , and  $\sigma_k$  is a parametrized  
 functional which is written as

$$\sigma_k = \tilde{\sigma}_{Z_k} + \sum_{m \neq k} \sum_{\tau} \tilde{\sigma}_{Z_k Z_m \tau} F_{Z_k Z_m \tau}(\{r_{km}\}) \quad (20)$$

where the sum is over all the atoms  $m$  in the molecule except  
 $k$ , the sum over  $\tau$  is over one or two types of  $F_{Z_k Z_m \tau}$ ,  $\tilde{\sigma}_{Z_k}$  is a  
 coefficient that depends on the atomic number,  $Z_k$ , of atom  $k$ ,  
 $\tilde{\sigma}_{Z_k Z_m \tau}$  is a coefficient that depends on the atomic numbers,  $Z_k$   
 and  $Z_m$ , of atoms  $k$  and  $m$  and the type,  $\tau$ , of geometric factor  
 $F_{Z_k Z_m \tau}$ . The function  $F_{Z_k Z_m \tau}$  is designed to account for the  
 dependence of atomic properties on the nature of the functional  
 group in which the atom appears.  $F_{Z_k Z_m \tau}$  depends on the set  $\{r_{km}\}$   
 of all of the internuclear distances in the molecule. For  
 compounds containing H, C, and O, SM6 has three  $\sigma_Z$   
 coefficients,  $\tilde{\sigma}_1$ ,  $\tilde{\sigma}_6$ , and  $\tilde{\sigma}_8$  corresponding respectively to H, C,  
 and O, and six  $\tilde{\sigma}_{ZZ\tau}$  coefficients,  $\tilde{\sigma}_{161}$ ,  $\tilde{\sigma}_{181}$ ,  $\tilde{\sigma}_{661}$ ,  $\tilde{\sigma}_{662}$ ,  $\tilde{\sigma}_{861}$ , and  
 $\tilde{\sigma}_{881}$ .

An important point to note is that calculated solute properties  
 depend on the choice of electronic structure theory and basis  
 set. It has been found<sup>12</sup> that the coefficients used to compute  
 $G_{\text{CDS}}$  vary somewhat depending on the choice of basis set, but  
 in SM6 and SM6T, we take them to be independent of other  
 aspects of the electronic structure theory chosen (for example,  
 they are independent of which density functional is used).

**3.B. SM6T.** Using the definition in eq 12 and motivated by  
 eq 15, we partition the temperature dependence of the free  
 energy of solvation as

$$\Delta \Delta G_{\text{S}}^{\circ} = \Delta \Delta G_{\text{ENP}}(T) + \Delta G_{\text{CDS}}(T) \quad (21)$$

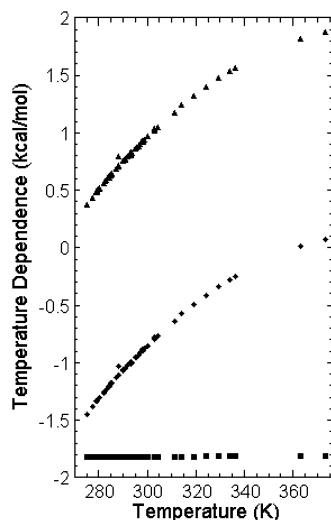
where  $\Delta \Delta G_{\text{ENP}}$  is the temperature dependence associated with  
 $\Delta G_{\text{ENP}}$  and  $\Delta G_{\text{CDS}}$  is an empirically derived term designed to  
 account for the temperature dependence of first-solvation-shell  
 phenomena and other approximations associated with use of  
 the generalized Born equation and the way we use it.

The first term on the right-hand side of eq 21 is directly  
 calculated as

$$\Delta \Delta G_{\text{ENP}}(T) = \Delta G_{\text{ENP}}(T) - \Delta G_{\text{ENP}}(T_0) \quad (22)$$

As described in section 3.A, the bulk electrostatic contribution  
 to the free energy of solvation depends on the partial atomic  
 charges of the solute, the Coulomb radii of the solute atoms,  
 the solute geometry, and the dielectric constant of the solvent,  
 all of which might, in principle, depend on temperature. Thus,  
 a strategic decision is required for how to model  $\Delta G_{\text{ENP}}$ .

There has been some discussion<sup>5,36,66–70</sup> concerning the use  
 of temperature-dependent radii to describe the electrostatics of  
 the solute cavity. An argument<sup>69</sup> can be made in favor of such  
 an approach based on the observation from statistical mechanical  
 studies that the radial distribution of water around a solute  
 changes as a function of temperature.<sup>66</sup> There have also been  
 studies<sup>70,71</sup> that used the isothermal compressibility of the solvent  
 to model the temperature dependence of the solute radii in  
 solution. However, the idea that the radii used to define the  
 solute cavity describe a physical boundary is oversimplified.<sup>1–3,50</sup>  
 Therefore, on the basis of the principle discussed below eq 15,  
 we treat the Coulomb radii as constant (i.e., independent of  $T$ ).



**Figure 1.** Experimental (◆) free energy of solvation for benzene in water as a function of temperature compared to the corresponding computed values of  $\Delta G_{\text{ENP}}(T)$  (■) and  $G_{\text{CDS}}(T)$  (▲).

379 Furthermore, in keeping with the discussion above of  $\Delta E_N$ , we  
380 neglect, at least in the parametrization step, the temperature  
381 dependence of the geometry of the solute because it is small.  
382 (It can be included as desired but will have an effect smaller  
383 than the reliability of the aqueous SM6T model.) The temper-  
384 ature dependence of the dielectric constant reflects the ability  
385 of the bulk solvent to reorient and polarize in response to the  
386 solute charge density. As such, it represents a clear physical  
387 phenomenon which can be accurately simulated by incorporating  
388 a temperature-dependent dielectric constant into the model. For  
389 water, the temperature dependence of the dielectric constant is  
390 well approximated by a cubic polynomial.<sup>72</sup>

391 The dielectric constant of water changes from 87.90 at 273  
392 K to 55.58 at 373 K; however,  $G_P$  and consequently  $\Delta G_{\text{ENP}}$   
393 depend on  $(1 - (1/\epsilon))$ , which changes much less than  $\epsilon$ . For  
394 example,  $\Delta G_{\text{ENP}}$  for ethyl benzene in water changes from  $-1.68$   
395 kcal/mol at 273 K to  $-1.66$  kcal/mol at 373 K, while its  
396 experimental free energy of solvation varies by 2.24 kcal/mol  
397 (see Figure 1) over this range. The temperature dependence of  
398 the bulk electrostatic contribution to the aqueous free energy  
399 of solvation remains small even for solutes with larger dipole  
400 moments such as furfural (dipole moment of 3.97 D),<sup>73</sup> for  
401 which the bulk electrostatic contribution to the aqueous free  
402 energy of solvation changes from  $-5.79$  at 277 K to  $-5.74$  at  
403 373 K. (The temperature dependence of the electrostatic  
404 interaction between the solute and bulk solvent becomes  
405 considerably larger for less polar solvents such as octanol. For  
406 example, octanol has a dielectric constant that varies from 11.92  
407 to 5.32 over the temperature range 273–373 K,<sup>72</sup> and the  
408 computed value of  $\Delta G_{\text{ENP}}$  for ethyl benzene in octanol varies  
409 by 0.19 kcal/mol. Thus, while  $\Delta\Delta G$  does not contribute  
410 significantly to the temperature dependence of the free energy  
411 of solvation in aqueous solutions, it does in nonpolar solvents.)

412 Throughout the course of this project, the bulk electrostatic  
413 contribution to the free energy of solvation,  $\Delta G_{\text{ENP}}$ , both at the  
414 reference temperature,  $T_0$ , and at the temperature of interest,  $T$ ,  
415 was computed using SM6 as it is implemented in MN-GSM<sup>74</sup>  
416 (a locally developed code that we cannot distribute due to  
417 Gaussian license restrictions), in SMXGAUSS<sup>75</sup> (a code that is  
418 freely available to the general scientific community and that  
419 was developed within our group to integrate our solvation  
420 models with Gaussian 03<sup>76</sup>), and in HONDOPLUS.<sup>77</sup>

421 Since the temperature dependence of the bulk electrostatic  
422 contribution to the free energy of solvation in water is negligible,  
423 the  $\Delta G$  term incorporates almost all of the temperature  
424 dependence of the aqueous free energy of solvation in the SM6T  
425 model. This term may be written as

$$\Delta G_{\text{CDS}} = G_{\text{CDS}}(T) - G_{\text{CDS}}(T_0) \quad (23)$$

426 Comparing eq 23 to eq 2 suggests that we model this as

$$\Delta G_{\text{CDS}} = B(T - T_0) + C[(T - T_0) - T \ln(T/T_0)] \quad (24)$$

427 where  $B$  and  $C$  are parameters of the model. As in the SM6  
428 model for  $G_{\text{CDS}}$ , we model  $\Delta G_{\text{CDS}}$  as being proportional to  
429 solvent-exposed surface areas.<sup>2,32,49–51,78</sup> We applied the same  
430 functions and solvent accessible surface areas used in eqs 19  
431 and 20 to calculate  $B$  and  $C$

$$B = \sum_k \sigma_k^B A_k \quad (25)$$

$$C = \sum_k \sigma_k^C A_k \quad (26)$$

432 where  $\sigma_k^B$  and  $\sigma_k^C$  are parametrized as

$$\sigma_k^B = \tilde{\sigma}_{Z_k}^B + \sum_{m \neq k} \sum_{\tau} \tilde{\sigma}_{Z_k Z_m \tau}^B F_{Z_k Z_m \tau}(\{r_{km}\}) \quad (27)$$

$$\sigma_k^C = \tilde{\sigma}_{Z_k}^C + \sum_{m \neq k} \sum_{\tau} \tilde{\sigma}_{Z_k Z_m \tau}^C F_{Z_k Z_m \tau}(\{r_{km}\}) \quad (28)$$

433 where  $\tilde{\sigma}_{Z_k}^B$ ,  $\tilde{\sigma}_{Z_k}^C$ ,  $\tilde{\sigma}_{Z_k Z_m \tau}^B$ , and  $\tilde{\sigma}_{Z_k Z_m \tau}^C$  are coefficients that depend on  
434 the atomic numbers of atoms  $k$  and  $m$ . Since  $F_{881}$  primarily  
435 distinguishes peroxides from other oxygen-containing solutes  
436 and since the database used here for SM6T has no peroxides,  
437  $\tilde{\sigma}_{881}^B$  and  $\tilde{\sigma}_{881}^C$  were not used. Thus, there are potentially as  
438 many as 16 parameters in SM6T, 8 for  $B$  and 8 for  $C$ . However,  
439 we will reduce the number of independent parameters during  
440 the fit, which is discussed in section 3.D.

441 **3.C. Error Function.** Several key classes of compounds are  
442 under-represented in the database used to develop parameters  
443 for the coefficients  $B$  and  $C$ . The disproportionate number of  
444 alcohols versus unsaturated alkanes arose due to the scarcity of  
445 experimental solubility data for the latter. To account for this,  
446 a weighted error function was used. The error function is defined  
447 as

$$D = \sum_{t=1}^{11} \sum_{m=1}^{s_t} \sqrt{\frac{\sum_i^{n_m} (\Delta G_{\text{CDS}}(\text{exptl};i) - \Delta G_{\text{CDS}}(\text{calcd};i))^2}{n_m s_t^2 11^2}} \quad (29)$$

448 where the first sum is over all 11 classes,  $t$ , of compounds in  
449 the database, the second sum is over all  $s_t$  compounds,  $m$ , in a  
450 given class  $t$ , and the third sum is over all  $n_m$  experimental data,  
451  $i$ , for a given compound,  $m$ .  $\Delta G_{\text{CDS}}(\text{calcd};i)$  is calculated using  
452 eqs 25–28, and  $\Delta G_{\text{CDS}}(\text{exptl};i)$  is the experimental value  
453 obtained by

$$\Delta G_{\text{CDS}}(\text{exptl}) = \Delta \Delta G_{\text{S}}^{\circ}(T) - \Delta \Delta G_{\text{ENP}}(T) \quad (30)$$

454 where the first term on the right-hand side is obtained from eqs  
455 12 and 13 and the second term is calculated by the generalized  
456 Born approximation with temperature-independent radii.

457 The error function in eq 29 ensures that each class is weighted  
 458 equally and that each compound within the class is weighted  
 459 equally with other compounds in the same class; this counteracts  
 460 the disproportionate number of alcohols and esters in the  
 461 database. The danger of this approach is that classes with fewer  
 462 compounds are more subject to the effects of random errors in  
 463 the experimental measurements for a given compound; this  
 464 concern has been addressed through the use of strict criteria in  
 465 selecting experimental data, as described in section 2.B.

466 The error function,  $D$ , was minimized using the generalized  
 467 reduced gradient code (GRG2) of Lasdon et al.<sup>79</sup> as implemented  
 468 in Microsoft Excel. The results produced by this code were  
 469 compared to those produced by the NAG linear optimizer, and  
 470 good agreement was found between the two.

471 **3.D. Parametrizations.** In this paper, all electronic structure  
 472 calculations were performed using the 6-31+G(d,p)<sup>80,81</sup> basis  
 473 set. Having chosen a basis set, a functional form for the  
 474 predictive model, and an appropriate error function, we pro-  
 475 ceeded to determine which parameters,  $\tilde{\sigma}_{ZZ\tau}^B$  and  $\tilde{\sigma}_{ZZ\tau}^C$ , were  
 476 necessary to accurately reproduce the experimental temperature  
 477 dependence of the aqueous free energies of solvation.

478 As a first step, we considered the case where a unique pair  
 479 of values for  $B$  and  $C$  is determined for each individual  
 480 compound (we refer to this case as “molecular coefficients”)  
 481 and the case where all of the parameters were set equal to zero  
 482 (we refer to this case as the “null hypothesis”).

483 The use of molecular coefficients corresponds to the best case  
 484 scenario for any model that uses eq 24 to calculate  $\Delta G_{\text{CDS}}$ , and  
 485 it can also be considered to give a rough estimate of the  
 486 nonsystematic experimental error in the temperature dependence  
 487 of the free energies of solvation in our database. Table 2 shows  
 488 that such nonsystematic errors produce a mean unsigned error  
 489 of only 0.02 kcal/mol, on average.

490 The null hypothesis, in contrast, corresponds to assuming that  
 491  $G_{\text{CDS}}$  does not change as a function of temperature. Table 2  
 492 shows that this produces a mean unsigned error of 0.53 kcal/  
 493 mol, which is considerably larger than 0.02 kcal/mol. Clearly  
 494 then,  $G_{\text{CDS}}$  *does* exhibit temperature dependence. Since SM6  
 495 can reliably reproduce free energies of solvation to within an  
 496 average error of  $\sim 0.5$  kcal/mol,<sup>12</sup> a variation of the free energy  
 497 of solvation by 0.53 kcal/mol is quite significant.

498 The error in the temperature dependence of free energies of  
 499 solvation and the error in free energies of solvation, which we  
 500 estimate<sup>47</sup> to be around 0.2 kcal/mol, are two distinct properties.  
 501 The estimated experimental error in the temperature dependence  
 502 of the free energies of solvation using the molecular coefficients  
 503 (0.02 kcal/mol) is considerably smaller than the typical estimated  
 504 error in the experimental determination of absolute free energies  
 505 of solvation.

506 Next, we performed a parametrization against the entire  
 507 database using the full set of 16 coefficients labeled  $\tilde{\sigma}_1^B, \tilde{\sigma}_6^B, \tilde{\sigma}_8^B,$   
 508  $\tilde{\sigma}_{161}^B, \tilde{\sigma}_{181}^B, \tilde{\sigma}_{661}^B, \tilde{\sigma}_{662}^B, \tilde{\sigma}_{861}^B, \tilde{\sigma}_1^C, \tilde{\sigma}_6^C, \tilde{\sigma}_8^C, \tilde{\sigma}_{161}^C, \tilde{\sigma}_{181}^C, \tilde{\sigma}_{661}^C,$  and  
 509  $\tilde{\sigma}_{861}^C$ . Analysis of the covariance between the various param-  
 510 eters indicated considerable redundancy in the parameters. The  
 511 covariance between the parameters in the 16-parameter fit  
 512 showed that  $\tilde{\sigma}_H^B$  and  $\tilde{\sigma}_H^C$  were highly correlated with the  
 513 remaining parameters and thus could be discarded. This reduces  
 514 the number of parameters to 14; the resulting fit with 14  
 515 parameters is called the “unrestricted” fit, and Table 2 shows  
 516 that the unrestricted fit produces a mean unsigned error of 0.07  
 517 kcal/mol. Table 3 shows the value for each coefficient obtained  
 518 by parametrizing the unrestricted fit against the entire database.  
 519 Our next step was to determine how to further reduce the

**TABLE 2: Mean Signed Error (MSE), Mean Unsigned Error (MUE), and Root Mean Squared Error (RMSE) for Each Class of Compounds (kcal/mol) for Various Parameterizations of SM6T**

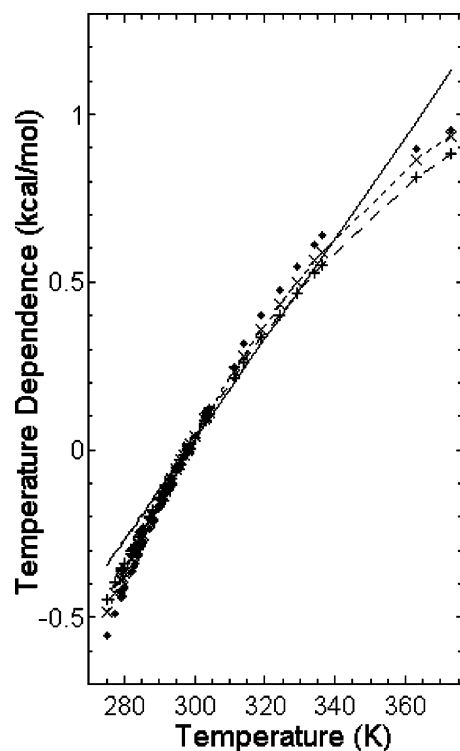
	MSE	MUE	RMSE	MSE	MUE	RMSE
	Molecular Coefficients			Null Hypothesis		
unbranched alkanes	0.00	0.05	0.06	-0.39	0.61	0.80
branched alkanes	0.00	0.02	0.02	-0.26	0.62	0.79
cyclic alkanes	0.00	0.01	0.01	-0.04	0.38	0.47
alkenes and alkynes	0.00	0.01	0.01	-0.19	0.28	0.37
aromatics	0.00	0.02	0.03	-0.03	0.31	0.39
alcohols, phenols, and water	0.00	0.03	0.03	-0.58	0.83	1.00
aldehydes	0.00	0.03	0.03	-0.16	0.38	0.46
ketones	0.00	0.03	0.04	-0.36	0.63	0.76
ethers	0.00	0.02	0.03	-0.23	0.58	0.70
esters	0.00	0.02	0.03	-0.39	0.69	0.81
carboxylic acids	0.00	0.02	0.03	-0.22	0.50	0.59
average	0.00	0.02	0.03	-0.26	0.53	0.65
	Unrestricted			van't Hoff		
unbranched alkanes	-0.01	0.07	0.08	0.03	0.10	0.13
branched alkanes	-0.04	0.13	0.15	0.01	0.16	0.21
cyclic alkanes	-0.01	0.03	0.04	0.02	0.06	0.08
alkenes and alkynes	0.03	0.06	0.08	0.01	0.06	0.09
aromatics	0.01	0.06	0.07	0.04	0.07	0.10
alcohols, phenols, and water	0.03	0.09	0.11	0.04	0.11	0.14
aldehydes	-0.01	0.07	0.08	0.00	0.10	0.12
ketones	-0.01	0.06	0.07	0.02	0.09	0.11
ethers	-0.04	0.08	0.09	0.00	0.09	0.11
esters	0.03	0.09	0.11	0.07	0.11	0.15
carboxylic acids	-0.01	0.06	0.08	0.00	0.07	0.09
average	0.00	0.07	0.09	0.02	0.09	0.12
	Final					
unbranched alkanes	-0.01	0.08	0.09			
branched alkanes	-0.04	0.13	0.15			
cyclic alkanes	-0.02	0.04	0.05			
alkenes and alkynes	0.02	0.07	0.09			
aromatics	0.00	0.07	0.09			
alcohols, phenols, and water	0.02	0.10	0.11			
aldehydes	-0.01	0.08	0.09			
ketones	-0.01	0.06	0.07			
ethers	-0.04	0.08	0.10			
esters	0.02	0.08	0.10			
carboxylic acids	-0.01	0.06	0.07			
average	-0.01	0.08	0.09			

**TABLE 3: Coefficients Used for Each Parametrization (cal  $\text{\AA}^{-2} \text{mol}^{-1} \text{K}^{-1}$ )**

	6	8	161	181	661	662	861
	Parameters Used in Calculation of $B$ Coefficient						
unrestricted	38.52	1.96	25.85	73.26	-24.36	-12.51	-8.54
van't Hoff	22.49	17.53	22.34	54.76	-13.07	4.25	-26.23
final SM6T	2.18	2.18	29.47	71.95	2.18	2.18	2.18
	Parameters Used in Calculation of $C$ Coefficient						
unrestricted	247.79	-200.07	43.70	294.55	-155.43	-370.03	246.44
final SM6T	-84.45	-84.45	78.59	166.23	78.59	-266.91	166.23

number of parameters without appreciably increasing the error  
 of the model.

For the purpose of testing whether the coefficient  $C$  was  
 necessary, we parametrized the model using just the coefficient  
 $B$ . Calculations employing only the coefficient  $B$  involved seven  
 parameters, five corresponding to the C, O, H–O, H–C, and  
 O–C bonds and two corresponding to C–C bonds (these  
 parameters are also given in Table 3). Table 2 shows that this  
 “van't Hoff” fit is reasonably accurate, with a mean unsigned  
 error of 0.09 kcal/mol. However, if one inspects the predicted  
 temperature dependence of the aqueous free energies of solva-



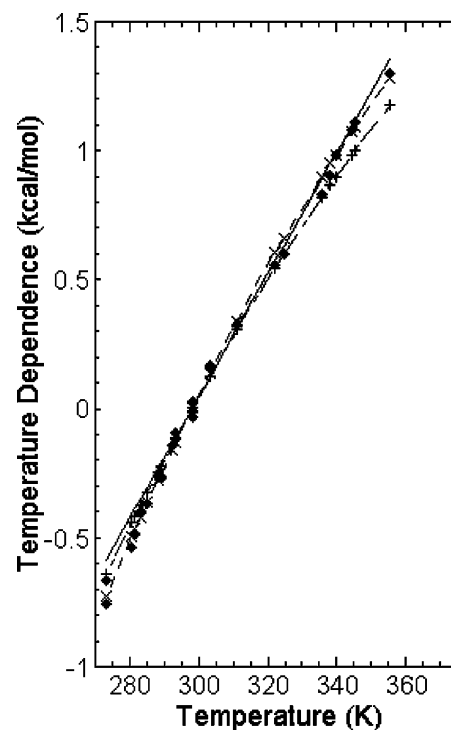
**Figure 2.** Experimental ( $\blacklozenge$ ) value of  $\Delta G_{\text{CDS}}$  for benzene compared to various possible parametrizations of SM6T: unrestricted (- + -); van't Hoff (-); final (- x -)

tion and compares those values to the experimental data for some representative compounds such as benzene (see Figure 2), one immediately sees that the van't Hoff fit fails to capture the obvious curvature in the experimental data.

We also considered combining parameters into a single general parameter. We found after a careful analysis of the covariance matrices, observing the values of the different parameters, and through trial and error that  $\tilde{\sigma}_6^B$ ,  $\tilde{\sigma}_8^B$ ,  $\tilde{\sigma}_{661}^B$ ,  $\tilde{\sigma}_{662}^B$ , and  $\tilde{\sigma}_{861}^B$  could be set equal to a single "general" parameter,  $\tilde{\sigma}_{b1}^B$ , with relatively little effect on the accuracy or apparent physicality of the model. However, we found that combination of either  $\tilde{\sigma}_{181}^B$  or  $\tilde{\sigma}_{161}^B$  with  $\tilde{\sigma}_{b1}^B$  or removal of either of these coefficients significantly reduced the accuracy of the model. We also observed that  $\tilde{\sigma}_{b1}^B$  was quite small and that removal of this parameter introduced relatively little additional error into the model.

We also considered combining parameters for the  $C$  coefficient. The approach which sacrificed the least accuracy involved setting  $\tilde{\sigma}_6^C$  and  $\tilde{\sigma}_8^C$  equal to a single parameter,  $\tilde{\sigma}_{c1}^C$ ,  $\tilde{\sigma}_{161}^C$  and  $\tilde{\sigma}_{661}^C$  equal to a single parameter,  $\tilde{\sigma}_{c2}^C$ , and  $\tilde{\sigma}_{181}^C$  and  $\tilde{\sigma}_{861}^C$  equal to a single parameter,  $\tilde{\sigma}_{c3}^C$  (see "final" parametrization in Table 2). With these restrictions, the mean unsigned error of the model changed by less than 0.02 kcal/mol relative to the unrestricted parametrization.

Our final model uses a total of seven parameters to predict the temperature dependence of the free energy of solvation; three parameters,  $\tilde{\sigma}_{161}^B$ ,  $\tilde{\sigma}_{181}^B$ , and  $\tilde{\sigma}_{b1}^B$ , are used to predict the  $B$  coefficient, and four parameters,  $\tilde{\sigma}_{662}^C$ ,  $\tilde{\sigma}_{c1}^C$ ,  $\tilde{\sigma}_{c2}^C$ , and  $\tilde{\sigma}_{c3}^C$ , are used to predict the  $C$  coefficient. This model is the most attractive because it introduces parameters with relatively small values as opposed to those for the unrestricted parametrization (as is shown in Table 3). Parametrizations with smaller values of the parameters are less likely to predict unphysical values when applied to problems outside its applicable range, for example, if the model were to be applied to functional groups



**Figure 3.** Experimental ( $\blacklozenge$ ) value of  $\Delta G_{\text{CDS}}$  for ethoxyethane compared to various possible parametrizations of SM6T: unrestricted (- + -); van't Hoff (-); final (- x -).

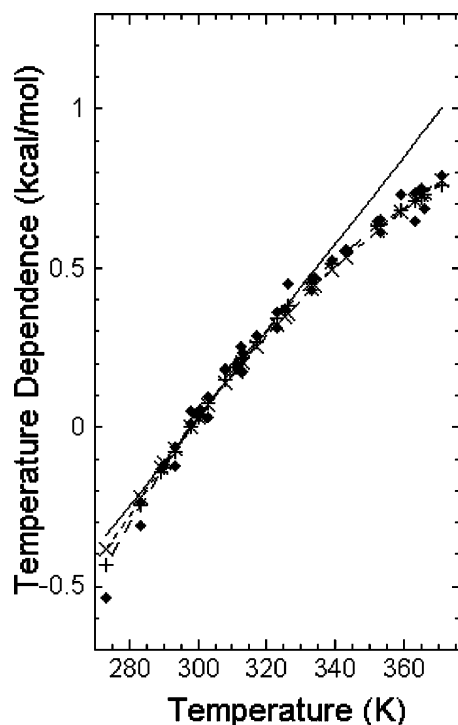
that were not included in the data set. The choice of four parameters,  $\tilde{\sigma}_{662}^C$ ,  $\tilde{\sigma}_{c1}^C$ ,  $\tilde{\sigma}_{c2}^C$ , and  $\tilde{\sigma}_{c3}^C$ , for the prediction of the  $C$  coefficient is based on the empirical observation that the errors in the model do not change significantly relative to the unrestricted model.

The predictions of the final parametrization do not vary significantly from those of the unrestricted parametrization. Figures 2–4 show each parametrization compared to the experimental temperature dependence of the free energy of solvation for benzene, ethoxyethane, and furfural, and one sees relatively good agreement between the results of the various parametrizations. Even at high temperatures around 373 K, both the final parametrization and the unrestricted parametrization predict very similar values for  $\Delta \Delta G_{\text{S}}^{\circ}$ . The final parametrization reproduces the temperature dependence of the free energy of solvation with more accuracy than the fit using just the  $B$  coefficient with an equal number of parameters (seven parameters).

The final parametrization uses unique values for the  $\tilde{\sigma}_{161}^B$  and the  $\tilde{\sigma}_{181}^B$  parameters. This is physically reasonable because it suggests that the hydrophobic effect, which is proportional to the number of CH groups exposed to the surrounding water, and the formation of hydrogen bonds with OH groups contribute in different ways to the entropy of solvation. The single general parameter,  $\tilde{\sigma}_{b1}^B$ , is retained; while the main contributions to the entropy of solvation in water may arise from the OH and CH groups, this may not necessarily be the case in nonaqueous solvents in which case the general parameter may contribute significantly to the accuracy of SM6T.

**3.E. Predicting Thermodynamic Properties.** In section 3.D, we showed that a parametrized model based on solvent accessible surface areas can reproduce the temperature dependence of aqueous free energies of solvation quite well. This model predicts two coefficients,  $B$  and  $C$ , which for the case of aqueous solutions correspond respectively to the entropy and heat capacity of solvation (compare eq 24 to eq 2). Table 4

566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601



**Figure 4.** Experimental (◆) value of  $\Delta G_{\text{CDS}}$  for furfural compared to various possible parametrizations of SM6T: unrestricted (- + -); van't Hoff (-); final (- × -).

compares the predicted values of  $B$  using the various parametrizations discussed above to experimental<sup>82–85</sup> and theoretical estimates of the entropy of solvation, such as those based on explicit solvent models by Gallichio et al.,<sup>86</sup> Cui et al.,<sup>87,88</sup> Schravendijk et al.,<sup>89</sup> and Rick.<sup>90</sup> Table 5 compares the predicted values of  $C$  to estimates of the heat capacity of solvation obtained from experimental work<sup>15,91–101</sup> and from the theoretical work of Graziano.<sup>102</sup> Previous estimates of these thermodynamic properties vary quite considerably between authors. For instance, note the large difference between the values for the entropy of solvation reported by Abraham et al.<sup>84</sup> and Wilhelm et al.<sup>85</sup> in Table 4. Furthermore, the values we obtained from a direct fit of the experimental data for each compound in our database (in the column labeled “molecular coefficients” in Tables 4 and 5) can vary quite considerably from other experimental sources. Franks et al.<sup>92</sup> pointed out that the entropy and heat capacity of solvation are difficult to obtain experimentally and require a considerable number of measurements—they estimated at least 40—over a relatively broad temperature range to accurately obtain such information. In some cases, our final choice of parameters for SM6T produces estimates of the entropy and heat capacity of solvation that agree better with the literature than our molecular coefficients. This may indicate that the final fit is more reliable than fitting each individual molecule because the considerable noise due to the sparsity of data for some compounds is averaged out to some extent when fitting the whole data set.

We emphasized in sections 3.A and 3.B that the partitioning of solvent effects between the two terms of eq 15,  $\Delta G_{\text{ENP}}$  and  $G_{\text{CDS}}$ , or the two terms of eq 21,  $\Delta \Delta G_{\text{ENP}}$  and  $\Delta G_{\text{CDS}}$ , is somewhat arbitrary. We use a fixed cavity size (i.e., our radii are independent of temperature). The extent that the electrostatic terms would change in either of these equations if the temperature dependence of the “true” Coulomb radii were included is, to first order, associated with the region at the solute–solvent boundary, and hence, it is equally well treated by the first-solvation-shell terms. However, changing the Coulomb radii has

the second-order effect that it would also change the calculated partial atomic charges of the solute, which would affect the long-range electrostatic interactions. This effect is not as obviously related to the solvent accessible surface area, but the success of our model argues that either the effect is small or it *can* be modeled well using surface tensions. We have observed this effect numerous times with our prior models.<sup>12,51,103</sup> That is, the choice of Coulomb radii has little effect on the accuracy of the SMx models for neutral compounds provided that the surface tensions are reoptimized as well, indicating that surface tensions can account for the differences in  $\Delta G_{\text{ENP}}$  due to different Coulomb radii. In our model, since  $\Delta \Delta G_{\text{ENP}}$  is small, it was not necessary to partition it into enthalpic and entropic contributions (it is essentially all enthalpic, with negligible entropy and heat capacity). An alternative approach, first implemented by Bonnacorsi et al.,<sup>71</sup> is to compute the change in  $\Delta G_{\text{ENP}}$  due to the temperature dependence of the electrostatic cavity using the coefficient of thermal expansion.<sup>70,71</sup>

SM6T does not account for the temperature dependence of the heat capacity of solvation, which can be significant.<sup>15,102</sup> To test the effect of the variation of the heat capacity on the free energy of solvation we compared the free energy of solvation for benzene in water at 373 K computed using a temperature-independent heat capacity obtained from the fit of our experimental data using eq 13 to a temperature-dependent heat capacity predicted using a model reported by Bakk and Hoye.<sup>15</sup> We used the same SM6 free energies and SM6T entropies of solvation at 298 K for both types of heat capacities. First, we compared the temperature-independent heat capacity of solvation ( $\Delta C_{\text{P,S}}^{\circ} \approx 70.3 \text{ cal mol}^{-1} \text{ K}^{-1}$ ) to the heat capacity at 298 K obtained from the model of Bakk and Hoye<sup>15</sup> ( $\Delta C_{\text{P,S}}^{\circ}(298 \text{ K}) \approx 68.8 \text{ cal mol}^{-1} \text{ K}^{-1}$ ). We found that a difference of  $1.5 \text{ cal mol}^{-1} \text{ K}^{-1}$  between the two heat capacities produces a difference of  $0.01 \text{ kcal/mol}$  in the aqueous free energy of solvation for benzene at 373 K, which is well within the error of SM6T. Then, we compared free energies of solvation computed using the temperature-independent heat capacity and the temperature-dependent heat capacity at 373 K ( $\Delta C_{\text{P,S}}^{\circ}(373 \text{ K}) \approx 54.4 \text{ cal mol}^{-1} \text{ K}^{-1}$ ), and we found a difference of  $0.14 \text{ kcal/mol}$  between the two calculations. Such an error is comparable to the deviation of SM6T from the experimental temperature dependence of the free energy of solvation of benzene at 373 K, but it is much smaller than the absolute error of SM6. In the interest of minimizing the number of parameters in our model and to avoid including poorly determined small terms, we do not include the temperature dependence of the heat capacity of solvation, but it is evident that the assumption of a temperature-independent heat capacity of solvation does introduce some error.

From the standpoint of our model, which was simultaneously parametrized to predict both linear and logarithmic terms in the temperature dependence of the free energy of solvation, the computation of the entropy and heat capacity of solvation is further complicated because contributions from the linear and logarithmic terms may partially cancel one another to produce reliable and accurate free energies of solvation without necessarily producing a model suitable for predicting the entropy and heat capacity of solvation. In particular, one should note that in Table 3 some of the coefficients for the heat capacity of solvation are negative while experimental heat capacities of solvation are positive. This in itself need not be meaningful because one adds environmentally dependent contributions from all of the atoms in a molecule to obtain the net  $C$  coefficient for the molecule. Application of SM6T to the database yields

639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702



**TABLE 4: Theoretical and Experimental Entropies of Solvation (cal mol<sup>-1</sup> K<sup>-1</sup>) from the Literature Compared to Predictions Based on Various Parametrizations**

	SM6T parametrizations				experimental literature	theoretical literature
	molecular coefficients	unrestricted	van't Hoff	final		
methane	-18	-16	-12	-11	-32, <sup>a</sup> -16 <sup>b</sup>	-15, <sup>c</sup> -13 <sup>d</sup>
ethane	-22	-18	-15	-16	-36, <sup>a</sup> -20 <sup>b</sup>	-19, <sup>c</sup> -25 <sup>d</sup>
propane	-24	-22	-18	-21	-39, <sup>a</sup> -23 <sup>b</sup>	-40 <sup>d</sup>
butane	-28	-26	-22	-25	-42 <sup>a</sup>	-26 <sup>c</sup>
hexane	-35	-34	-28	-34	-11 <sup>e</sup>	-32 <sup>c</sup>
cyclohexane	-28	-27	-23	-30	-11 <sup>e</sup>	-30 <sup>c</sup>
benzene	-21	-18	-15	-19	-21 <sup>f</sup>	-19, <sup>g</sup> -15, <sup>h</sup> -8.7 <sup>h</sup>

<sup>a</sup> Abraham et al.<sup>84</sup> <sup>b</sup> Wilhelm et al.<sup>85</sup> <sup>c</sup> Gallichio et al.<sup>86</sup> <sup>d</sup> Cui et al.<sup>87,88</sup> <sup>e</sup> Ben-Naim et al.<sup>82</sup> <sup>f</sup> Makhatadze and Privalov.<sup>83</sup> <sup>g</sup> Schravendijk et al.<sup>89</sup> <sup>h</sup> Rick.<sup>90</sup>

**TABLE 5: Experimental Heat Capacities of Solvation (cal mol<sup>-1</sup> K<sup>-1</sup>) from the Literature Compared to Predictions Based on Various Parameterizations**

	SM6T parameterizations			literature
	molecular coefficients	unrestricted	final	
ethane	68	54	34	70, <sup>a</sup> 67, <sup>b</sup> 95, <sup>c</sup> 65 <sup>d</sup>
propane	73	60	47	81, <sup>a</sup> 78, <sup>b</sup> 76 <sup>d</sup>
butane	84	68	59	102, <sup>a</sup> 95 <sup>b</sup>
pentane	70	76	70	108 <sup>a</sup>
hexane	104	84	82	119 <sup>a</sup>
cyclohexane	42	56	79	100 <sup>e</sup>
benzene	70	49	55	69, <sup>d</sup> 67, <sup>e</sup> 72, <sup>f</sup> 84, <sup>g</sup> 70 <sup>h</sup>
methanol	21	30	11	32 <sup>i</sup>
ethanol	117	33	30	48 <sup>i</sup>
<i>n</i> -propanol	2	40	42	64 <sup>i</sup>
butanol	70	48	54	79 <sup>i</sup>
2-butanol	92	56	53	83 <sup>i</sup>
pentanol	89	52	63	89 <sup>i</sup>

<sup>a</sup> Olofsson et al.<sup>96</sup> <sup>b</sup> Dec et al.<sup>95,97,98</sup> <sup>c</sup> Naghibi et al.<sup>99,100</sup> <sup>d</sup> Bakk and Hoye.<sup>15</sup> <sup>e</sup> Gill et al.<sup>94</sup> <sup>f</sup> Arnold et al.<sup>101</sup> <sup>g</sup> Franks et al.<sup>92</sup> <sup>h</sup> Graziano.<sup>102</sup> <sup>i</sup> Arnett et al.<sup>93</sup>

two instances (namely, ethyne and propyne) within the database where the *C* coefficient for a molecule is negative. Also, there are numerous instances within the set where the sum of contributions to the *C* coefficient from all carbon atoms is negative but is counterbalanced by a positive contribution from the H atoms. This reflects the fact that the coefficients used to compute *C* were chosen to best reproduce experimental free energies of solvation with a minimum of parameters and that independent physical meaning cannot be assigned to the value of each coefficient. This is confirmed by a study of the covariance matrices of these parameters, which shows that they are not entirely independent. Noting this, and that we neglected the temperature dependence of the heat capacity of solvation, that we assumed temperature-independent radii, and that an average of only 12 experimental points per compound was used to develop the model, we conclude that, despite the encouraging agreement with data from other groups, the division of the free energy of solvation predicted by SM6T into entropy and heat capacity should be interpreted with caution.

#### 4. Software

New versions of HONDOPLUS,<sup>77,104</sup> GAMESSPLUS,<sup>105,106</sup> SMXGAUSS,<sup>75</sup> NWChem,<sup>107,108</sup> and Jaguar<sup>109</sup> containing SM6T are planned for the near future.

#### 5. Concluding Remarks

We have developed a model for calculating the temperature dependence of the aqueous free energy of solvation of compounds composed of H, C, and O over the temperature range

273–373 K. This involved creating an extensive database of experimental free energies of solvation, selecting an appropriate functional form for the model, and parametrizing the model against experiment. We found that an accurate model requires a heat-capacity-like term to capture the curvature of the temperature dependence of the free energy of solvation, that a parametrized model based on solvent accessible surface areas produces quite accurate results, and that the temperature dependence of the dielectric constant produces relatively little effect in aqueous solutions, although it is much more significant in organic solutions. The model reduces the mean unsigned error in the temperature dependence of the solvation free energy by a factor of 6.5 relative to the null hypothesis. The next step will be to extend the model to broader classes of compounds and to organic solvents. In future work, it would also be interesting to extend the SM6T model to predict free energies of solvation in near-critical and supercritical water which is the subject of considerable theoretical research.<sup>110–121</sup>

**Acknowledgment.** The authors thank Casey Kelly for his help throughout this project. This work was supported by the U.S. Army Research Office under Multidisciplinary Research Program of the University Research Initiative (MURI) under Grant No. DAAD19-02-1-0176 and by the Office of Naval Research under Grant No. N00012-05-01-0538.

#### Appendix 1. Derivation of eq 6

We start with the standard thermodynamic equation relating the concentration of the solute in the gas phase to the concentration of the solute in solution at equilibrium<sup>28,29,37</sup>

$$\gamma_a^{g,x} y_a P = \gamma_a^{\infty,x} x_a P_a^* \quad (31)$$

where  $y_a$  is the mole fraction of compound *a* in the gas phase,  $P$  is the total vapor pressure,  $\gamma_a^{g,x}$  is the activity coefficient of *a* in the gas phase,  $x_a$  is the mole fraction of *a* in solution,  $P_a^*$  is the vapor pressure of the pure liquid solute, and  $\gamma_a^{\infty,x}$  is the infinite dilution activity coefficient. Assuming that the solute behaves as an ideal gas, eq 31 can be rewritten as

$$\frac{\gamma_a^{\infty,x} P_a^*}{P} = \frac{y_a}{x_a} \quad (32)$$

The free energy of solvation for a standard state of 1 M in the gas phase and 1 M in solution equates to

$$\Delta G_S^{\circ} = RT \ln(K_H) = RT \ln\left(\frac{c_a}{S_a}\right) \quad (33)$$

where  $K_H$  is the Henry's law constant,  $c_a$  is the molar

730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766

767 concentration of the solute in the gas phase, and  $S_a$  is the molar  
768 concentration of the solute in solution. Assuming ideality of  
769 the solute in the gas phase, we find

$$y_a = c_a \frac{RT}{P} \quad (34)$$

770 Assuming that the solute is infinitely dilute in the solution yields

$$x_a = \frac{S_a}{M_w} \quad (35)$$

771 where  $M_w$  is the molarity of the solvent, which in this case is  
772 water. Combining eqs 34 and 35 with eq 32 and rearranging  
773 yields

$$\frac{\gamma_a^{\infty,x} P_a^*}{RTM_w} = \frac{c_a}{S_a} \quad (36)$$

774 which, when combined with eq 33, results in eq 6.

775 **Supporting Information Available:** Tables listing the  
776 individual compounds, the number of experimental points used  
777 for each compound, and the references for the experimental data.  
778 This material is available free of charge via the Internet at [http://](http://pubs.acs.org)  
779 [pubs.acs.org](http://pubs.acs.org).

## 780 References and Notes

- 781 (1) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.  
782 (2) Cramer, C. J.; Truhlar, D. G. In *Solvent Effects and Chemical*  
783 *Reactivity*; Tapia, O., Bertrán, J., Eds.; Kluwer: Boston, MA, 1996; p 1.  
784 (3) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.  
785 (4) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.  
786 (5) Elcock, A. H.; McCammon, J. A. *J. Phys. Chem. B* **1997**, *101*,  
787 9624.  
788 (6) Grant, D. J. W.; Higuchi, T. *Solubility Behavior of Organic*  
789 *Compounds*; Wiley: New York, 1990.  
790 (7) Hamad, S.; Lago, S.; Mejias, J. A. *J. Phys. Chem. A* **2002**, *106*,  
791 9104.  
792 (8) Mirbagheri, S. A. *Int. J. Environ. Sci. Technol.* **2004**, *1*, 141.  
793 (9) Silverman, R. B. *The Organic Chemistry of Drug Design and Drug*  
794 *Action*; Academic: San Diego, CA, 1992.  
795 (10) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **2005**,  
796 *113*, 133.  
797 (11) Li, C.-J.; Chan, T.-H. *Organic Reactions in Aqueous Media*;  
798 Wiley: New York, 1997.  
799 (12) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.*  
800 **2005**, *1*, 1133.  
801 (13) Addison, C. C. *J. Chem. Soc.* **1945**, 1945, 98.  
802 (14) Klotz, I. M.; Rosenberg, R. M. *Chemical Thermodynamics*, 5th ed.;  
803 Wiley: New York, 1994.  
804 (15) Bakk, A.; Hoye, J. S. *Physica A* **2002**, *303*, 286.  
805 (16) Sander, R. *Compilation of Henry's Law Constants for Inorganic*  
806 *and Organic Species of Potential Importance in Environmental Chemistry*,  
807 3rd ed.; 1999; (<http://www.mpch-mainz.mpg.de/~sander/res/henry.html>).  
808 (17) Shiu, W.; Ma, K. *J. Phys. Chem. Ref. Data* **2000**, *29*, 41.  
809 (18) Dohanyosova, P.; Sarraute, S.; Dohnal, V.; Majer, V.; Gomes, M.  
810 *C. Ind. Eng. Chem. Res.* **2004**, *43*, 2805.  
811 (19) Stephenson, R. M.; Malanowski, S. *Handbook of the Thermody-*  
812 *namics of Organic Compounds*; Elsevier: New York, 1987.  
813 (20) Yaws, C. L. *Handbook of Vapor Pressures*; Gulf Publishing  
814 Company: Houston, TX, 1994; Vols. 1–4.  
815 (21) Yaws, C. L. *Yaws' Handbook of Thermodynamic and Physical*  
816 *Properties of Chemical Compounds*; Knovel: Norwich, NY, 2003.  
817 (22) Plyasunov, A. V.; Plyasunova, N. V.; Shock, E. L. *The Database*  
818 *for the Thermodynamic Properties of Neutral Compounds in the State of*  
819 *Aqueous Solution*; Department of Geological Sciences, Arizona State  
820 University: Tempe, AZ, 2005 (<http://orchyd.asu.edu/>).  
821 (23) Yalkowsky, S. H.; He, Y. *Handbook of Aqueous Solubility Data*;  
822 CRC Press: New York, 2003.  
823 (24) Maczynski, A.; Wisniewska-Gocłowska, B.; Góral, M. *J. Phys.*  
824 *Chem. Ref. Data* **2004**, *33*, 549.  
825 (25) Góral, M.; Maczynski, A.; Wisniewska-Gocłowska, B. *J. Phys.*  
826 *Chem. Ref. Data* **2004**, *33*, 579.  
(26) Góral, M.; Wisniewska-Gocłowska, B.; Maczynski, A. *J. Phys.* 827  
*Chem. Ref. Data* **2004**, *33*, 1159. 828  
(27) IUPAC–NIST Solubility Database; National Institute of Standards 829  
and Technology: Boulder, CO, 2005 (<http://srdata.nist.gov/solubility/>). 830  
(28) Gautreaux, M. F.; Coates, J. *AIChE J.* **1955**, *1*, 496. 831  
(29) Slocum, E. W.; Dodge, B. F. *AIChE J.* **1964**, *10*, 364. 832  
(30) Gmehling, J.; Menke, J.; Schiller, M. *Activity Coefficients at Infinite* 833  
*Dilution: C<sub>10</sub>–C<sub>36</sub> with O<sub>2</sub>S and H<sub>2</sub>O*; Dechema: Frankfurt, Germany, 834  
1994; Vol. 9, Part 4. 835  
(31) Hovorka, S.; Dohnal, V.; Roux, A. H.; Roux-Desgranges, G. *Fluid* 836  
*Phase Equilib.* **2002**, *201*, 135. 837  
(32) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305. 838  
(33) Winget, P.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys.* 839  
*Chem. B* **2000**, *104*, 4726. 840  
(34) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* 841  
**2003**, *119*, 1661. 842  
(35) Ben-Naim, A. *Solvation Thermodynamics*; Plenum: New York, 843  
1987. 844  
(36) Cramer, C. J.; Truhlar, D. G. In *Free Energy Calculations in* 845  
*Rational Drug Design*; Reddy, M. R., Erion, M. D., Eds.; Kluwer: New 846  
York, 2001; p 63. 847  
(37) Atkins, P. *Physical Chemistry*, 6th ed.; Freeman: New York, 1998. 848  
(38) Staudinger, J.; Roberts, P. V. *Chemosphere* **2001**, *44*, 561. 849  
(39) Stephenson, R.; Stuart, J. *J. Chem. Eng. Data* **1986**, *31*, 56. 850  
(40) Liabastre, A. A. Ph.D. Thesis, Georgia Institute of Technology, 851  
1974. 852  
(41) Price, L. C. *Am. Assoc. Pet. Geo. Bull.* **1976**, *60*, 213. 853  
(42) Stephenson, R.; Stuart, J.; Tabak, M. *J. Chem. Eng. Data* **1984**, 854  
*29*, 287. 855  
(43) Lemmon, E. W.; Goodwin, A. R. H. *J. Phys. Chem. Ref. Data* 856  
**2000**, *29*, 1. 857  
(44) Poling, B. E.; Prausnitz, J. M.; O'Connell, J. P. *The Properties of* 858  
*Gases and Liquids*, 5th ed.; McGraw-Hill: New York, 2001. 859  
(45) Clarke, E. C. W.; Glew, D. N. *Trans. Faraday Soc.* **1966**, *62*, 539. 860  
(46) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8552. 861  
(47) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* 862  
**2004**, *108*, 6532. 863  
(48) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* 864  
**2005**, *113*, 107. 865  
(49) Cramer, C. J.; Truhlar, D. G. In *Reviews in Computational* 866  
*Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VHC: New York, 1995; 867  
Vol. 6, p 7. 868  
(50) Cramer, C. J. *Essentials of Computational Chemistry: Theories* 869  
*and Models*, 2nd ed.; Wiley: New York, 2002. 870  
(51) Hawkins, G.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, 871  
*101*, 7147. 872  
(52) Daudel, R. *Quantum Theory of Chemical Reactivity*; Reidel: 873  
Dordrecht, The Netherlands, 1973. 874  
(53) Tucker, S. C.; Truhlar, D. G. *Chem. Phys. Lett.* **1989**, *157*, 164. 875  
(54) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am.* 876  
*Chem. Soc.* **1990**, *112*, 6127. 877  
(55) Tapia, O. In *Quantum Theory of Chemical Reactions, Vol. 2. Solvent* 878  
*Effect, Reaction Mechanisms, Photochemical Processes*; Daudel, R., Pull- 879  
man, A., Salem, L., Veillard, A., Eds.; Kluwer: Boston, MA, 1981; Vol. 880  
2, p 25. 881  
(56) Li, J.; Zhu, T.; Hawkins, G.; Winget, P.; Liotard, D. A.; Cramer, 882  
C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1999**, *103*, 9. 883  
(57) Giesen, D. J.; Storer, J. W.; Cramer, C. J.; Truhlar, D. G. *J. Am.* 884  
*Chem. Soc.* **1995**, *117*, 1057. 885  
(58) Makhatadze, G. I.; Privalov, P. L. *Adv. Protein Chem.* **1995**, *47*, 886  
307. 887  
(59) Lazaridis, T.; Karplus, M. *Proteins* **1999**, *35*, 133. 888  
(60) Shimizu, S.; Chan, H. S. *J. Chem. Phys.* **2000**, *113*, 4683. 889  
(61) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379. 890  
(62) Hermann, R. B. *J. Phys. Chem.* **1972**, *76*, 2754. 891  
(63) Hermann, R. B. *J. Phys. Chem.* **1975**, *79*, 163. 892  
(64) Liotard, D. A.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; 893  
Truhlar, D. G. *J. Comput. Chem.* **1995**, *16*, 422. 894  
(65) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441. 895  
(66) Yu, H.-A.; Roux, B.; Karplus, M. *J. Chem. Phys.* **1990**, *92*, 5020. 896  
(67) Shock, E. L.; Oelkers, E. H.; Johnson, J. W.; Sverjensky, D. A.; 897  
Helgeson, H. C. *J. Chem. Soc., Faraday Trans. 1* **1992**, *88*, 803. 898  
(68) Rashin, A. A.; Bukatin, M. A. *Biophys. Chem.* **1994**, *51*, 167. 899  
(69) Kono, H.; Ohtsuki, Y.; Abe, T. *J. Phys. Chem.* **1996**, *100*, 9935. 900  
(70) Cubero, E.; Orozco, M.; Luque, F. J. *J. Org. Chem.* **1998**, *63*, 2354. 901  
(71) Bonaccorsi, R.; Palla, P.; Tomasi, J. *J. Am. Chem. Soc.* **1984**, *106*, 902  
1945. 903  
(72) *CRC Handbook of Chemistry and Physics*; Lide, D. R., Ed.; CRC: 904  
Boca Raton, FL, 1995; Vol. 75. 905  
(73) Abraham, R. J.; Bretschneider, E.; In *Internal Rotation in Molecules*; 906  
Orville-Thomas, W. J., Ed.; Wiley: New York, 1974; p 481. 907  
(74) Kelly, C. P.; Xidos, J. D.; Li, J.; Thompson, J. D.; Hawkins, G. 908  
D.; Winget, P. D.; Zhu, T.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; 909

- 910 Truhlar, D. G.; Frisch, M. J. *MN-GSM 5.2*; Univeristy of Minnesota: Minneapolis, MN, 2005.
- 911 (75) Kelly, C. P.; Thompson, D. J.; Lynch, B. J.; Xidos, J. D.; Li, J.;
- 912 Hawkins, G. D.; Zhu, T.; Volobuev, Y.; Dupuis, M.; Rinaldi, D.; Liotard,
- 913 D. A.; Cramer, C. J.; Truhlar, D. G. *SMXGAUSS*, version 2.0.2; University
- 914 of Minnesota: Minneapolis, MN, 2003.
- 915 (76) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb,
- 916 M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K.
- 917 N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.;
- 918 Mennucci, B.; Cossi, M.; Scalamani, G.; Rega, N.; Petersson, G. A.;
- 919 Nagatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.;
- 920 Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li,
- 921 X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.;
- 922 Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.;
- 923 Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.;
- 924 Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels,
- 925 A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.;
- 926 Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.;
- 927 Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz,
- 928 P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.;
- 929 Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson,
- 930 B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*,
- 931 revision C.01; Gaussian: Pittsburgh, PA, 2003.
- 932 (77) Kelly, C. P.; Nakamura, H.; Xidos, J. D.; Thompson, D. J.; Li, J.;
- 933 Hawkins, G. D.; Zhu, T.; Lynch, B. J.; Volobuev, Y.; Rinaldi, D.; Liotard,
- 934 D. A.; Cramer, C. J.; Truhlar, D. G. *HONDOPLUS*, version 4.7; University
- 935 of Minnesota: Minneapolis, MN, 2005 (based on HONDO v. 99.6).
- 936 (78) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J.*
- 937 *Phys. Chem.* **1996**, *100*, 16385.
- 938 (79) Lasdon, L. S.; Waren, A. D.; Jain, A.; Ratner, M. *ACM Trans.*
- 939 *Math. Software* **1978**, *4*, 34.
- 940 (80) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
- 941 (81) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio*
- 942 *Molecular Orbital Theory*; Wiley: New York, 1986.
- 943 (82) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016.
- 944 (83) Makhatazde, G. I.; Privalov, P. L. *Biophys. Chem.* **1994**, *50*, 285.
- 945 (84) Abraham, M. H.; Matteoli, E. *J. Chem. Soc., Faraday Trans. 1*
- 946 **1988**, *84*, 1985.
- 947 (85) Wilhelm, E.; Battino, R.; Wilcock, R. J. *Chem. Rev.* **1977**, *77*, 219.
- 948 (86) Gallicchio, E.; Kubo, M. M.; Levy, R. M. *J. Phys. Chem. B* **2000**,
- 949 *104*, 6271.
- 950 (87) Cui, Q.; Smith, V. H. *J. Chem. Phys.* **2000**, *113*, 10240.
- 951 (88) Cui, Q.; Smith, V. H. *J. Chem. Phys.* **2001**, *115*, 2228.
- 952 (89) Schravendijk, P.; van der Vegt, N. F. A. *J. Chem. Theory Comput.*
- 953 **2005**, *1*, 643.
- 954 (90) Rick, S. J. *J. Phys. Chem. B* **2003**, *107*, 9853.
- 955 (91) Arnold, V. W.; Washburn, E. R. *J. Phys. Chem.* **1958**, *62*, 1088.
- 956 (92) Franks, F.; Gent, M.; Johnson, H. H. *J. Chem. Soc.* **1963**, 2716.
- 957 (93) Arnett, E. M.; Kover, W. B.; Carter, J. V. *J. Am. Chem. Soc.* **1969**,
- 958 *91*, 4028.
- 959 (94) Gill, S. J.; Nichols, N. F.; Wadso, I. *J. Chem. Thermodyn.* **1976**, *8*,
- 960 *445*.
- 961 (95) Dec, S. F.; Gill, S. J. *J. Solution Chem.* **1984**, *13*, 27.
- 962 (96) Olofsson, G.; Oshodj, A. A.; Warnstrom, E.; Wadso, I. *J. Chem.*
- 963 *Thermodyn.* **1984**, *16*, 1041.
- 964 (97) Dec, S. F.; Gill, S. J. *J. Solution Chem.* **1985**, *14*, 417.
- 965 (98) Dec, S. F.; Gill, S. J. *J. Solution Chem.* **1985**, *14*, 827.
- 966 (99) Naghibi, H.; Dec, S. F.; Gill, S. J. *J. Phys. Chem.* **1986**, *90*, 4261.
- 967 (100) Naghibi, H.; Dec, S. F.; Gill, S. J. *J. Phys. Chem.* **1987**, *91*, 245.
- (101) Arnold, D. S.; Plank, C. A.; Erickson, E. E.; Pike, F. P. *Chem.*
- Eng. Data Ser.* **1958**, *3*, 253.
- (102) Graziano, G. *Biophys. Chem.* **2005**, *116*, 137.
- (103) Hawkins, G.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *J. Org.*
- Chem.* **1998**, *63*, 4305.
- (104) Dupuis, M.; Marquez, A.; Davidson, E. R. *HONDO 99.6*; Quantum
- Chemistry Program Exchange (QCPE), Indiana University: Bloomington,
- IN, 1999 (based on HONDO 95.3).
- (105) Kelly, C. P.; Pu, J.; Thompson, D. J.; Xidos, J. D.; Li, J.; Zhu,
- T.; Hawkins, G. D.; Chuang, Y.-Y.; Fast, P. L.; Lynch, B. J.; Liotard, D.
- A.; Rinaldi, D.; Gao, J.; Cramer, C. J.; Truhlar, D. G. *GAMESSPLUS*,
- version 4.7; University of Minnesota: Minneapolis, MN, 2005 (based on
- the General Atomic and Molecular Electronic Structure System (GAMESS)).
- (106) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.;
- Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.;
- Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.*
- 1993**, *14*, 1347.
- (107) Aprà, E.; Windus, T. L.; Straatsma, T. P.; Bylaska, E. J.; de Jong,
- W.; Hirata, S.; Valiev, M.; Hackler, M.; Pollack, L.; Kowalski, K.; Harrison,
- R.; Dupuis, M.; Smith, D. M. A.; Nieplocha, J.; V. T.; Krishnan, M.; Auer,
- A. A.; Brown, E.; Cisneros, G.; Fann, G.; Fruchtl, H.; Garza, J.; Hirao, K.;
- Kendall, R.; Nichols, J.; Tsemekhman, K.; Wolinski, K.; Anshell, J.;
- Bernholdt, D.; Borowski, P.; Clark, T.; Clerc, D.; Dachsel, H.; Deegan,
- M.; Dyall, K.; Elwood, D.; Glendening, E.; Gutowski, M.; Hess, A.; Jaffe,
- J.; Johnson, B.; Ju, J.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.;
- Long, X.; Meng, B.; Nakajima, T.; Niu, S.; Rosing, M.; Sandrone, G.; Stave,
- M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wong, A.; Zhang, Z. *NWChem,*
- A Computational Chemistry Package for Parallel Computers*, version 4.7;
- Pacific Northwest National Laboratory: Richland, WA, 2005.
- (108) Kendall, R. A.; Apra, E.; Bernholdt, D. E.; Bylaska, E. J.; Dupuis,
- M.; Fann, G. I.; Harrison, R. J.; Ju, J.; Nichols, J. A.; Nieplocha, J.;
- Straatsma, T. P.; Windus, T. L.; Wong, A. T. *Comput. Phys. Commun.*
- 2000**, *128*, 260.
- (109) *Jaguar 6.5*; Schrodinger LLC: Portland, OR, 2005.
- (110) Gilson, M.; Sharp, K.; Honig, B. *J. Comput. Chem.* **1988**, *9*, 327.
- (111) Tucker, S. C.; Gibbons, E. M. In *Structure and Reactivity in*
- Aqueous Solution*; Cramer, C. J., Truhlar, D. G., Eds.; ACS Symposium
- Series, Vol. 528; The American Chemical Society: Washington, DC, 1994;
- p 196.
- (112) Bennett, G. E.; Rossky, P. J.; Johnston, K. P. *J. Phys. Chem.* **1995**,
- 99*, 16136.
- (113) Luo, H.; Tucker, S. C. *J. Am. Chem. Soc.* **1995**, *117*, 11359.
- (114) Pomelli, C. S.; Tomasi, J. *J. Phys. Chem. A* **1997**, *101*, 3561.
- (115) Plyasunov, A. V.; O'Connell, J. P.; Wood, R. H.; Shock, E. L. *Geochim. Cosmochim. Acta* **2000**, *64*, 2779.
- (116) Yezdimer, E. M.; Sedlbauer, J.; Wood, R. H. *Chem. Geol.* **2000**,
- 164*, 259.
- (117) Plyasunov, A. V.; O'Connell, J. P.; Wood, R. H.; Shock, E. L. *Fluid Phase Equilib.* **2001**, *183-184*, 133.
- (118) Hori, T.; Takahashi, H.; Nitta, T. *J. Comput. Chem.* **2002**, *24*,
- 209*.
- (119) Akinfiev, N. N.; Diamond, L. W. *Geochim. Cosmochim. Acta*
- 2003**, *67*, 613.
- (120) Anisimov, M. A.; Sengers, J. V.; Sengers, J. M. H. L. In *Aqueous*
- Systems at Elevated Temperatures and Pressures*; Palmer, D. A., Fernandez-
- Prini, R., Harvey, A. H., Eds.; Elsevier: London, 2004; p 29.
- (121) Majer, V.; Sedlbauer, J.; Wood, R. H. In *Aqueous Systems at*
- Elevated Temperatures and Pressures*; Palmer, D. A., Fernandez-Prini, R.,
- Harvey, A. H., Eds.; Elsevier: London, 2004; p 99.