Thermodynamic Decomposition of Hydration Free Energies by Computer Simulation: Application to Amines, Oxides, and Sulfides

Masahito Mogami Kubo, Emilio Gallicchio, and Ronald M. Levy*

Department of Chemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08855-0939

Received: August 8, 1997; In Final Form: October 2, 1997[⊗]

The calculation of the Gibbs free energy, enthalpy, and entropy of hydration of ammonia, methylamine, dimethylamine, trimethylamine, water, methanol, dimethyl ether, hydrogen sulfide, methanthiol, and dimethylsulfide is presented to illustrate the usefulness of the enthalpy and entropy of solvation in studying microscopic phenomena affecting the thermodynamics of the hydration of simple organic molecules. The free energy perturbation (FEP) method is used in conjunction with constant temperature and constant pressure molecular dynamics (MD) configurational sampling. The hydration free energies are studied as a function of the temperature in order to evaluate the hydration entropy by finite differences (FD). The TIP3P water model is used for the solvent water and revised AMBER parameters for the solutes. Partial charges of the solutes are obtained from fitting the electrostatic potential obtained from electronic structure calculations. Discrepancies with the experiments, especially noticeable for the amines, are observed for the hydration enthalpies and entropies even in cases where the hydration free energies are in agreement with the experiments. We conclude that this molecular force field requires additional parametrization against experimental entropies and enthalpies of hydration. Other molecular force fields may also need reparametrization.

Downloaded by UNIV OF TEXAS AUSTIN on July 7, 2009
Published on December 4, 1997 on http://pubs.acs.org | doi: 10.1021/jp972598b

1. Introduction

Computer simulations have been applied to the calculation of solvation free energies for a wide range of molecules, from small ions and inorganics to larger organic and biological molecules.[1−5] On the other hand, entropies and enthalpies of solvation have not received the same attention partly because of the more stringent computer requirements.

Several studies have shown that existing molecular force fields such as AMBER[6] and OPLS[7] have been often successfully used in predicting the free energy of hydration of simple organic molecules at room temperature and standard pressure.[8−10] The question arises if these force fields are also able to reproduce the experimental enthalpies and entropies of solvation for a large variety of solutes. Accurate reproduction of entropies and enthalpies of solvation is, in our opinion, essential if computer simulations are to become a more useful tool to obtain molecular insights into solvation phenomena.

The entropy and enthalpy of solvation are an important complement to the free energy of solvation because they provide additional information to help understand and interpret the physics of the solvation process. The entropies and enthalpies of solvation are also thermodynamic state functions and experimentally measurable quantities and therefore provide additional benchmarks to be used in developing force fields for condensed-phase simulations.

The decomposition of the free energy into entropic and enthalpic contributions is helpful in determining which microscopic phenomena are important for the solvation process. One can envision, for example, two solutes having similar and favorable solvation free energies for completely distinct reasons at a molecular level. For instance, the solvation of one solute relative to another could be mostly driven by a favorable entropy, that is, by an increase of the configurational space spanned by the solvent and the solute with little or even positive internal energy change. For another solute, instead, the dominant effect could be a large and favorable interaction energy with the solvent that, however, causes a loss of entropy due to the system locking in a particular low-energy configuration. The hydrophobic effect, assumed to be entropically dominated, is a concrete example of how the entropy can be used to classify solvation processes.

It is observed that in many molecular series the trends in the enthalpies of solvation are more predictable than the trends in the free energies of solvation. For the dissolution of methyl-substituted amines in water, for example, it is observed experimentally[11] that, whereas the free energies of hydration often show small variations and follow a complex trend that cannot be easily interpreted, the enthalpies and entropies of hydration show large, predictable, and distinctive changes. The study of the entropies and enthalpies of solvation, therefore, can more readily provide insights into the dominant elements that contribute to the general trends, while the study of the free energies often involves the calculation and analysis of small differences between large opposing effects (ΔH and $-T\Delta S$).

Computationally, the study of solvation enthalpies can provide more direct insights because of their higher sensitivity, compared to the free energies, to the interaction potentials. Indeed, often a variation in the force field parameters results in a substantial change in the enthalpy that is accompanied by a compensating change in the entropy in such a way that negligible changes of the free energy are observed. This effect is sometimes referred to as entropy−enthalpy compensation.

Often researchers seek to obtain insights into the features of the solvent−solute interactions from a series of single-point electronic structure calculations. They then tend to discuss trends in solvation free energies on the basis of such features.[12] However, such arguments are more predictive of the enthalpies of solvation rather than the free energies because it is difficult to discuss entropy variations without extensively exploring the potential landscape of the solute−solvent system. In light of this, more robust computational methods are needed to bridge ab initio calculations and simulations of enthalpies of solvation.

The entropies and enthalpies of solvation, when compared to available experimental data, can also provide a means to

* Corresponding author.
[⊗] Abstract published in *Advance ACS Abstracts,* November 15, 1997.

S1089-5647(97)02598-4 CCC: $14.00 © 1997 American Chemical Society

improve existing empirical molecular force fields. Whereas the free energy is a quantity directly related to the partition function, the entropy is related to its temperature derivative. A molecular force field able to reproduce entropy-related quantities thus provides a molecular model that correctly describes the changes in the system while its temperature is varied. This force field, therefore, in addition to being applicable for a wider range of temperatures, would also provide better estimates of quantities such as heat capacities related to temperature derivatives of molecular distribution functions. The optimization of force field parameters against entropies and enthalpies of solvation can also lead to a model of molecular interactions that is more realistic and therefore more transferable from one molecular series to another. Although the high computational cost makes it difficult at present to systematically optimize force field parameters with respect to entropies and enthalpies of solvation, these thermodynamic quantities can still be used to analyze deficiencies in current force fields and provide insights into the means for improving them.

The slow convergence has always been a challenge intrinsic to the calculation of derivative quantities such as entropy and enthalpy changes. In one of the few such attempts, Fleischman et al. investigated the decomposition of relative Helmholtz free energies for several organic molecules.[13] Although their approach was innovative, the statistical uncertainties associated with the results were too large to achieve any significant molecular insights. The ensemble dependent effects also made the comparison with experiments more difficult. Their simulations, in fact, were performed at constant volume rather than constant pressure, and they did not properly correct for this when comparing calculated and experimental solvation enthalpies and entropies. The need for extra care when dealing with enthalpies and entropies is due to the fact that the free energy changes at the same thermodynamic conditions but under different environmental constraints are identical, whereas their temperature derivatives are not (see section 2.1).

More recently, Smith et al.[14] reported a similar approach to calculate the entropy changes. Their finite-difference (FD-E) method approximates the entropies by calculating the change of free energy of solvation between two temperatures. The FD-E formula is based on the assumption that the heat capacity is constant over a range of temperatures near the target. According to the authors, this assumption is typically valid around room temperature with a temperature perturbation ΔT as large as 50 K.[14] Yet the choice of ΔT is critical for reasons of both accuracy and precision. If ΔT is too large, the heat capacity cannot be considered constant in the $T \pm \Delta$T range. If ΔT is too small, on the other hand, significantly more expensive computations are required to overcome the increase in the statistical fluctuations.

The present study demonstrates that the decomposition of the free energy of solvation into entropic and enthalpic components with reasonable precision is now computationally feasible. We present the thermodynamic decomposition of the solvation free energies using the FD-E approach. Three sets of molecular series, namely, the amine (N), oxide (O), and sulfide (S) series, are considered. Starting from the totally hydrogenated molecules (NH_3, H_2O, SH_2), hydrogens are successively substituted by methyl groups.

The amine series is one of the few cases that exhibit significant discrepancy between the calculated and experimental free energies.[15,16] Insights into the molecular origin of the discrepancy in this series can be obtained by comparing the experimental and theoretical trends of the entropies and enthalpies of solvation.

The O series exhibits closer agreement between the calculated and experimental free energies of solvation. For these solutes, we examine the variations in the entropies and enthalpies of solvation upon methylation and we compare our results with the available experimental measurements. For the S series, we also decompose the free energy of solvation, although experimental measurements for this series are not extensively available.

In studying the entropy and enthalpy of hydration of these molecular series, we also hope to address the issue of entropy−enthalpy compensation in solvation processes. By entropy−enthalpy compensation we mean the empirical observation that the difference of free energy of solvation of two solutes is often smaller than the difference observed for the entropy and enthalpy. Since the free energy is given by the sum of ΔH and $-T\Delta S$, it must be for these cases that, by mutation of one solute into the other, the change in enthalpy is partially compensated by the change in entropy and vice versa. Because exceptions to this rule can be easily found, for instance in the methylation of methanol to obtain dimethyl ether,[17,18] with our study we would like to develop an understanding of the microscopic phenomena leading to entropy−enthalpy compensation. It is important to note that this phenomenon cannot be fully predicted on the basis of thermodynamic laws alone, since it is a consequence also of the particular interactions between the species in solution. This is in contrast to forms of exact entropy−enthalpy compensation, which are necessary on the basis of thermodynamic laws. For instance, it is known that the free energy of solvation is independent of the ensemble (constant pressure or constant volume) while entropy and enthalpy are ensemble dependent quantities so that the difference between the enthalpy of solvation at constant pressure and that at constant volume must be exactly canceled by the corresponding difference in the entropy.[19]

It is computationally more expensive to obtain the entropy and enthalpy of solvation than to obtain the corresponding free energy. It is important, therefore, to develop theoretical and computational methods to efficiently evaluate the decomposition of the free energy. In this work, we use molecular dynamics (MD) simulations and we apply the FD-E method for the calculation of entropies of solvation. In a subsequent paper we will report on the precision and accuracy of other methodologies as well.

2. Method

2.1. Ensemble Dependence. The free energy perturbation (FEP) method is a well-established computational tool to compute the difference of solvation free energy between two solutes.[1−5] In the following we will express the free energy and related quantities according to the Ben-Naim standard state convention.[18] The thermodynamic quantities reported in the Ben-Naim convention will be marked by an asterisk.

In our simulations a parameter λ is introduced in the potential function so that in going from $\lambda = 0$ to $\lambda = 1$ the system is removed from the initial state and transformed into the final state. This is accomplished in several steps, or windows. In the *i*th window the system configurations are sampled according to a MD algorithm setting $\lambda = \lambda_i$. The Gibbs free energy change in varying λ_i to λ_{i+1} is then expressed as

$$\Delta G^*_{i+1,i} = -kT \ln\langle \exp(-\Delta U_{i+1,i}/kT)\rangle_i \qquad (1)$$

where $\langle ... \rangle$ denotes an isobaric ensemble average for $\lambda = \lambda_i$ and $\Delta U_{i+1,i} = U_{i+1} - U_i$ is the potential energy difference between the $\lambda = \lambda_{i+1}$ and $\lambda = \lambda_i$ states.

Downloaded by UNIV OF TEXAS AUSTIN on July 7, 2009
Published on December 4, 1997 on http://pubs.acs.org | doi: 10.1021/jp972598b

Downloaded by UNIV OF TEXAS AUSTIN on July 7, 2009
Published on December 4, 1997 on http://pubs.acs.org | doi: 10.1021/jp972598b

It should be noted that the choice of the ensemble (canonical or isobaric) in eq 1 is somewhat arbitrary. In a constant volume ensemble, eq 1 would correspond to the change in Helmholtz free energy at a specific volume. Not considering finite size effects, this Helmholtz free energy change would be equal to the corresponding Gibbs free energy change at the pressure determined by the equation of state of the system. Both quantities correspond to the change in chemical potential of the solute.[19]

The choice of the ensemble is, however, critical when considering the entropy and enthalpy changes.[20] For instance, by using standard thermodynamics arguments, we find that the difference between the enthalpy of solvation at constant pressure $(\Delta H)_P$ and the internal energy of solvation at constant volume $(\Delta E)_V$ (P and V being related by the equation of state of the pure solvent at temperature T) is given by

$$(\Delta H)_P - (\Delta E)_V = \frac{T\alpha}{\kappa}(\Delta V)_P \qquad (2)$$

where $(\Delta V)_P$ is the partial molar volume of the solute and α and κ are, respectively, the thermal expansion coefficient and the isothermal compressibility of the pure solvent. Explicit derivations of this and related thermodynamic relation are given in the appendix. In general, the right-hand side of eq 2 is not negligible (for the dissolution of methane in water, it is about 1.5 kcal/mol) and not related to $P(\Delta V)_P$ as is sometimes assumed.

The simulations described in this paper have been performed in the isobaric ensemble to connect to existing experimental measurements of enthalpies and entropies of solvation without having to estimate the thermal expansion coefficient and the isothermal compressibility of the water model used.

2.2. Finite-Difference Method. The entropy change between states i and j at temperature T is given by the relationship

$$\Delta S^*_{ji}(T) = -\left(\frac{\partial \Delta G^*_{ji}}{\partial T}\right)_{P,N} \qquad (3)$$

According to the finite-difference (FD-E) method, the above relationship is approximated as[14,21]

$$\Delta S^*_{ji}(T) \simeq -\frac{\Delta G^*_{ji}(T+\Delta T) - G^*_{ji}(T-\Delta T)}{2\Delta T} \qquad (4)$$

Therefore, to obtain the complete set of thermodynamic quantities (i.e., ΔG^*, $-T\Delta S^*$, ΔH^*), this method requires an independent simulation at each of three different temperatures T, $T + \Delta T$, and $T - \Delta T$.

The FD-E approximation is based on the assumption that the heat capacity is constant over a certain range of temperatures near the target temperature T. For aqueous solutions, this assumption normally holds valid near room temperature with ΔT as large as 50 K,[22,14] making FD-E an attractive way to compute the temperature derivative of the free energy. The choice of ΔT is, however, important to minimize the statistical uncertainty of the FD-E formula. The statistical uncertainty of the entropy is inversely proportional to ΔT as given by the relationship

$$\sigma_{[T\Delta S^*_{ji}(T)]} = \frac{T}{2\Delta T}\{\sigma_{[\Delta G^*_{ji}(T+\Delta T)]} + \sigma_{[\Delta G^*_{ji}(T-\Delta T)]}\} \qquad (5)$$

where $\sigma_{[\Delta G^*_{ji}(T)]}$ is the standard deviation associated with the Gibbs free energy of solvation at temperature T. Thus, as ΔT becomes smaller, significantly more expensive computations are required to achieve converged results for the entropy of solvation. If ΔT is too large, on the other hand, the accuracy of the FD-E formula grows worse because the linearity assumption of the free energy change over temperature does not hold any longer. For a typical simulation at room temperature, if ΔT is chosen to be 30 K, the statistical uncertainty associated with $T\Delta S$ is roughly 10 times larger than that of the free energy (298 K/30 K $\simeq$ 10). Assuming a good quality free energy calculation bears an uncertainty of 0.1 kcal/mol, the corresponding uncertainty for $T\Delta S^*$ would roughly be about 1 kcal/mol.

The enthalpy change $\Delta H^*(T)$ is reported as the difference between $\Delta G^*(T)$ and $-T\Delta S^*(T)$. The associated error is, therefore, expressed as the sum of the statistical uncertainties of the two quantities.

TABLE 1: Partial Atomic Charges Used for Simulations

molecule	atom	ESP charge
NH_3	N	−1.076
	H	0.359
NH_2CH_3	N	−1.016
	H (amine)	0.369
	C	0.441
	H (methyl)	−0.054
$NH(CH_3)_2$	N	−0.768
	H (amine)	0.366
	C	0.191
	H (methyl)	0.004
$N(CH_3)_3$	N	−0.420
	C	−0.015
	H	0.052
H_2O (TIP3P)	O	−0.834
	H	0.417
CH_3OH	O	−0.685
	H (alcohol)	0.426
	C	0.278
	H (methyl)	−0.006
$(CH_3)_2O$	O	−0.420
	C	0.124
	H	0.064
SH_2	S	−0.346
	H	0.173
CH_3 SH	S	−0.371
	H (thiol)	0.202
	C	0.031
	H (methyl)	0.019
$(CH_3)_2S$	S	−0.284
	C	−0.049
	H	0.064

2.3. Computational Details. Solute geometries are generated by ab initio quantum chemistry calculations at the HF/6-31G* level for the N series and at the HF/6-31G** level for the other series using the GAUSSIAN 92 or GAUSSIAN 94 programs.[23,24] Charges at each atom site are fitted to the molecular electrostatic potentials generated from CHELPG calculations[25] along with the geometry optimization. As expected, generated dipole moments are roughly 20% larger than the experimental gas-phase values. This exaggerated polarization is an artifact of the relative low level of accuracy of the quantum calculations that yield partial charges, used in nonpolarizable force fields, that fortuitously account for some of the condensed-phase polarization effects in a mean field way. Partial charges are listed in Table 1.

A series of MD simulations of simple organic solutes (amines, oxides, and sulfides) in water have been performed. The free energy changes are computed by the FEP method using the program IMPACT.[26] The AMBER pairwise additive potential[27] is used to model nonelectrostatic molecular interactions. The all-atom model is used throughout this study. The TIP3P model is used for water both as solute and as solvent.[28]

One solute molecule is placed in a cubic box containing about 216 water molecules. Periodic boundary conditions and a cutoff of 8.5 Å are applied for the nonbonded interactions. The radial distribution function is assumed to be constant beyond the cutoff distance for Lennard-Jones (LJ) interaction computation. Electrostatic interaction is neglected beyond the cutoff. The temperature and the pressure are kept constant in a simulation using Berendsen's scaling method with the relaxation time of 0.2 ps.[29] The reference (central) temperature is 298.15 K and the pressure is 1.0 atm. The equations of motion are integrated using Andersen's RATTLE algorithm with a time step of 1 fs.[30] All bond lengths and the bond angle of the solvent are constrained using the SHAKE algorithm.[31]

The free energy perturbation calculation is performed starting with the fully formed solute. The charges on the solute sites are then annihilated. This is followed by the annihilation of the Lennard-Jones nonelectrostatic interactions. The LJ annihilation process is further divided into two steps depending on the solute. If the LJ term for all the solute atoms are annihilated simultaneously, unrealistic solute−solvent interactions may be produced. Following the guideline that only atoms located inside the LJ radii of remaining atoms are annihilated in one step, hydrogens in methyl groups are initially annihilated. In the second step, all the remaining atoms are transformed into a single spherical cavity with Lennard-Jones parameters corresponding to a neutral nitrogen atom.[21]

The following procedure is used for the free energy simulations. The system is first energy minimized for 2000 steps, then equilibrated for 10 ps. FEP is carried out with the double-wide window method with a total number of windows ranging from 100 to 250, depending on the size and the partial charges of the solute. Within each window, the system is equilibrated for 1 ps and data are collected for 5 ps or more until the desired convergence is achieved. To obtain the desired convergence for the entropy, the free energy perturbation calculations at the lower and higher temperatures require particularly long collection times. Forward and backward perturbations are performed both in double-wide sampling windows so that there are two forward and two backward results. Those four results are averaged, and statistical uncertainty is reported simply by taking the hysteresis. Error analysis is not performed explicitly otherwise. The hysteresis can be considered as the lower limit of the error. A detailed analysis of the precision inherent in the different methods for computing solvation enthalpies and entropies will be presented elsewhere.

For the N series, the original AMBER parameters for LJ interactions are used in order to make direct comparison with previously reported results.[15] For the O and S series, on the other hand, the LJ parameters for hydrogens in methyl groups are adopted from the result of Hägler et al.[32] These hydrogen parameters mixed with original AMBER nonbonded parameters for carbons generally result in better agreement with the experimental solvation free energy,[33] although they often overestimate the LJ interaction energy.[34] Revised AMBER parameters are used for the sulfur atom.[6] The LJ parameters used in this study are listed in Table 2.

To avoid the complete annihilation or creation of the solute,[35−37] the thermodynamic potentials are calculated relative to a neutral sp^3 nitrogen atom for all FEP simulations.[15] The solvation of the nitrogen atom is separately studied using similar simulation conditions with the exception of the adoption of a time step of 2 fs. This result is added to the relative free energy changes in order to report the absolute solvation free energy of the solute. The solvation process of the nitrogen atom is further divided into three steps. The nitrogen atom is annihilated

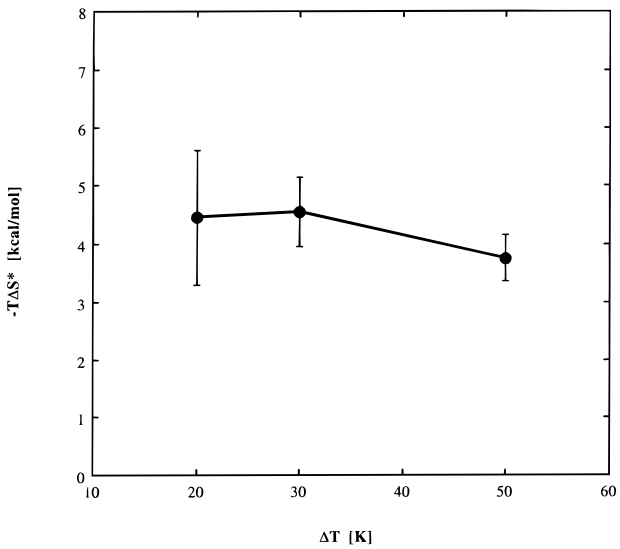


Figure 1. Calculated entropy of hydration of methanol and the associated error as a function of temperature perturbation.

TABLE 2: Nonbonded Lennard-Jones Parameters Used for Simulations

molecular series	atom type	σ [Å]	ϵ [kcal/mol]
amine (N)	C	1.6036	0.0600
	H (methyl)	1.3720	0.0100
	N	1.6482	0.1200
	H (amine)	0.8909	0.0200
O and S	C	1.6036	0.0600
	H (methyl)	1.2250	0.0380
	O	1.4700	0.1500
	S	1.7818	0.2500
	H (alcohol/thiol)	0.8909	0.0200

introducing a coupling parameter λ. From $\lambda = 1$ to $\lambda = 0.005$, 199 double sampling windows have been used and from $\lambda = 0.005$ to $\lambda = 2.5 \times 10^{-5}$ another 199 double sampling windows were adopted. Finally, it is annihilated "completely" ($\sigma = 0.0$ Å, $\epsilon = 1 \times 10^{-30}$ kcal/mol) using the slow growth method with 40 ps of simulation time. In the first and second steps, the system is equilibrated for 1 ps, then data is collected for 10 ps in each window.

3. Results and Discussion

3.1. Choice of Temperature Perturbation. The solvation of methanol is chosen as the test case for optimization of ΔT. With other simulation conditions fixed, ΔT is changed from 20 to 50 K to estimate the effect of the finite difference approximation on the accuracy and precision of the calculated entropies. The total number of windows in this study is 100; we used 1 ps of equilibration and 5 ps of data collection for each window.

In Figure 1, we show the calculated entropy and the statistical error (hysteresis) as a function of the temperature perturbation ΔT. This figure clearly shows the inverse proportionality between the statistical error and ΔT. The calculated entropy does not show appreciable changes below $\Delta T = 30$ K, at least within this limited test case. In fact, even though the error approximately doubles, the mean value of the enthalpy change for $\Delta T = 20$ K is comparable to the one with $\Delta T = 30$ K. Considering the additional computational time necessary to obtain acceptable converged results for ΔT below 30 K, we decided that the temperature perturbation $\Delta T = 30$ K is the most suitable for the present study.

Downloaded by UNIV OF TEXAS AUSTIN on July 7, 2009
Published on December 4, 1997 on http://pubs.acs.org | doi: 10.1021/jp972598b

TABLE 3: Solvation Free Energy, Entropy, and Enthalpy of N Series [kcal/mol]

	ΔG^*		$-T\Delta S^*$		ΔH^*	
molecule	calcd	exptl[a]	calcd	exptl[a]	calcd	exptl[a]
NH_3	-4.19 ± 0.14	-4.31	6.00 ± 0.56	3.57	-10.19 ± 0.69	-7.87
NH_2CH_3	-3.03 ± 0.17	-4.57	5.18 ± 0.86	5.70	-8.13 ± 1.02	-10.27
$NH(CH_3)_2$	-0.62 ± 0.35	-4.30	4.46 ± 1.20	7.84	-5.08 ± 1.53	-12.14
$N(CH_3)_3$	2.83 ± 0.47	-3.24	3.21 ± 1.27	8.81	-0.38 ± 1.63	-12.05

[a] From ref 11, after standard states conversion. See Appendix A.

TABLE 4: Solvation Free Energy, Entropy, and Enthalpy of O Series [kcal/mol]

	ΔG^*		$-T\Delta S^*$		ΔH^*	
molecule	calcd	exptl	calcd	exptl	calcd	exptl
H_2O	-6.46 ± 0.08	-6.32[a]	4.65 ± 0.28	3.65[a]	-11.11 ± 0.36	-9.97[a]
CH_3OH	-4.19 ± 0.08	-5.10[a]	4.55 ± 0.59	5.15[a]	-8.74 ± 0.66	-10.25[a]
$(CH_3)_2O$	-0.39 ± 0.11	-1.90[b]	4.57 ± 0.58	5.77[c]	-4.97 ± 0.68	-7.66[c]

[a] From ref 18. [b] From ref 41. [c] From ref 17, after standard states conversion. See Appendix A.

TABLE 5: Solvation Free Energy, Entropy, and Enthalpy of S Series [kcal/mol]

	ΔG^*		$-T\Delta S^*$		ΔH^*	
molecule	calcd	exptl	calcd	exptl	calcd	exptl
H_2S	-0.39 ± 0.07	-0.54[a]	1.93 ± 0.41	3.04[a]	-2.32 ± 0.47	-3.58[a]
CH_3SH	-0.69 ± 0.11	-1.24[b]	2.40 ± 1.06		-3.09 ± 1.17	
$(CH_3)_2S$	0.15 ± 0.24	-1.54[b]	2.52 ± 1.13		-2.37 ± 1.36	

[a] From ref 42, after standard states conversion. See Appendix A. [b] From ref 41.

3.2. Thermodynamic Decomposition. We have calculated entropies and enthalpies for the three molecular series, i.e., the N, O, and S series, using our MD/FEP/FD-E method. Results are summarized in Tables 3−5 for each molecular series.

In all three molecular series, the calculations show that the free energy of hydration increases with the degree of methylation. Although this is in general qualitative agreement with the experiments, the change in free energy of hydration in each molecular series is quantitatively overestimated by the calculations.

For the amine series, for example, in going from ammonia to trimethylamine the calculations predict a more positive free energy of hydration of about 7 kcal/mol. Experimental solubility data show, instead, an increase of only about 1 kcal/mol. For the oxygen and sulfur series the discrepancy between calculated and experimental changes of free energies of hydration is smaller. In both cases, upon full methylation, the calculations predict an increase in free energy of hydration that is only about 1.5 kcal/mol more positive than the experiments.

The results of the present calculations for the amine series are not unique, since many other model studies of the free energies of hydration of nitrogen-containing organic compounds found a clear disagreement with the experimental results.[15,16,38,39] This issue is a topic of much debate, and several interpretations of this phenomenon can be found in the literature.[12,15,16] The present study, however, offers a unique perspective on the problem because it shows that the quantitative disagreement between calculated and experimental free energies of hydration for the amines is actually caused by a much larger qualitative disagreement between the calculated and experimental entropies and enthalpies of hydration.

From Table 3, we see that the calculations predict an increase of the entropy of hydration from ammonia to trimethylamine whereas the experiments show the opposite trend. The discrepancy between calculations and experiments is even more evident in the enthalpy changes. The calculations predict an increased (more positive) enthalpy of hydration from ammonia to trimethylamine of 9.81 kcal/mol in contrast to the experiments that show instead a decrease of 4.18 kcal/mol. It is amusing to note that the only reason the trend of the calculated free energy of hydration is in even qualitative agreement with the experiments is because, upon methylation, the disagreement between calculations and experiments grows faster for the enthalpy than for the entropy.

It is, therefore, our belief that any attempt to reconcile the present solvation model of the amines with the experiments should focus on resolving the qualitative disagreement found for the enthalpies and entropies of hydration rather than on the free energies alone. It would be pointless, for example, to change the model parameters to fit the experimental free energies of hydration in such a way as to cause an even larger disagreement between the calculated entropies and enthalpies and the experimental measurements.

It has been proposed by Marten et al.[12] that the incorrect modeling of the hydrogen bonding pattern in hydrated amines is responsible for the observed discrepancy between the calculated and experimental hydration free energies. Marten et al. constructed their argument by observing discrepancies between single-point ab initio energies of solutes interacting with one or two water molecules with energies calculated from commonly used molecular force fields (OPLS and AMBER). Although their extrapolation from the observed single-point potential energies to the hydration free energies is undermined by having ignored entropic effects, we observe that their conclusions are nonetheless confirmed by our study of hydration enthalpies. We believe, in fact, that the analysis of Marten et al. is more directly applicable to hydration enthalpies than to hydration free energies. The increase of the calculated hydration enthalpies upon methylation can be explained by the unrealistic loss of a hydrogen bond between the nitrogen and a water molecule (caused by the sharp decrease of the partial charge on the nitrogen upon methylation (see Table 1)) by overestimating the loss of hydrogen bonding between the hydrogens bonded to nitrogen and the oxygen atoms of water and by underestimating the interaction between the alkylic groups and the water molecules. Hence, the analysis of Marten et al., when applied to hydration enthalpies, acquires a stronger meaning and supports the findings of the present study.

Downloaded by UNIV OF TEXAS AUSTIN on July 7, 2009
Published on December 4, 1997 on http://pubs.acs.org | doi: 10.1021/jp972598b

Marten et al. observed that the OPLS force field proved superior to the AMBER force field, used in the present study, in reproducing ab initio energies. Preliminary results also show that the OPLS force field produces hydration enthalpies and entropies of the substituted amines in closer agreement with the experimental results.[40]

It has been found by Ding et al.[15] that the calculated hydration free energies of the substituted amines is noticeably affected by the introduction of many-body polarization into the energy function. In fact, the correct handling of the molecular polarizabilities drives the computed free energies of hydration to a closer agreement with the experimental results. The molecular dipole moments in nonpolarizable force fields are inferred from gas-phase calculations and empirically corrected to reproduce liquid-phase average dipole moments. Such dipole moments are unresponsive to changes of the surrounding solvent. Interestingly, Ding et al. found, using a polarizable model, that the free energies of hydration of the substituted amines in water is well correlated to their average dipole moments. The magnitude of the dipole moments, however, is not directly correlated to the molecular polarizability.

Some of the discrepancies between the single-point ab initio energies and the nonpolarizable molecular force field predictions observed by Marten et al. can be explained in terms of many-body polarization. In particular, the inclusion of polarization will cause a more favorable interaction between the highly polarizable methyl groups and the water molecules. We expect, therefore, that by taking into account polarization, the calculated hydration enthalpies of the amines would not increase as sharply as observed in this study.

In the oxygen series we observe the same trends discussed previously for the amines, although the discrepancies between calculations and experiments are on a smaller scale. With respect to the experiments, the calculations predict a steeper increase of the free energy of hydration, underestimate the decrease of entropy of hydration, and overestimate the decrease of enthalpy of hydration upon methylation of the oxygen atom. The consequence of this is that, although the calculated and experimental free energies of solvation are in good agreement for the O series, the calculated trend in free energy is dominated by the enthalpy term while the experimental free energy trend is dominated by the entropy term. Some of the phenomena responsible for the discrepancies between calculations and experiments in the amine series may be also of importance for the oxide series.

In the case of the sulfide series, the comparisons with the experiments are more difficult because experimental data are not readily available for the solvation enthalpy and entropy. The experimental results for the S series exhibit an "unexpected" free energy trend upon methylation similar to results for the amines. Experimentally, in fact, the free energy becomes more favorable as methylation occurs. The calculations somewhat reproduce this trend for the first methylation but predict an unfavorable change upon the second methylation. The calculations also show a nearly constant entropy and enthalpy of hydration upon methylation. This phenomenon has not been observed in the other two molecular series.

The phenomenon of entropy−enthalpy compensation is observed for the amines. However, the phenomenon is not universal. In the oxide series, for example, we observe variations in the enthalpy of hydration whereas the entropy is virtually constant. A similar phenomenon is observed for the sulfur series. In future work we plan to use computer simulations to elucidate the molecular origins of entropy−enthalpy compensation in the solvation of organic solutes in water.

4. Concluding Remarks

Thermodynamic decomposition computations of the hydration free energies for 10 organic solutes in the amine, oxide, and sulfide molecular series have been performed. These computations are intrinsically more expensive than the corresponding free energy computations.

We have shown that reasonable precision is attainable for these systems by adopting the finite difference (FD-E) approximation for the entropy change and using an efficient isobaric molecular dynamics sampling procedure. A key factor in determining the feasibility of such calculations is a proper selection of the temperature difference (ΔT) between the three independent calculations of the hydration free energy required by the FD-E method.[14] By using this procedure and with the help of today's computational capabilities, we were able to compute absolute enthalpies of hydration with an uncertainty on the order of 1 kcal/mol, a more than 10-fold improvement over previously reported studies of similar systems.[13]

Even though in several instances the computations are found to be in disagreement with the experiments, it is still possible to provide insightful molecular information that cannot be obtained by experiments alone. The experimental thermodynamic data for the amines vs degree of methylation can be interpreted in terms of the decreasing electrostatic and hydrogen-bonding solute−solvent interactions and increasing van der Waals solute−solvent interactions leading to more favorable solvation enthalpy overall, whereas the entropy becomes unfavorable owing to the ordering by methyl groups. In the calculations unrealistic treatment of the hydrogen bonding pattern is considered the dominant cause for the failure of the current molecular model to reproduce the experimental results. This interpretation is in agreement with ab initio quantum mechanical results.[12]

In forthcoming papers we plan to report on more accurate and more efficient methods for the calculation of the entropy and enthalpy of solvation that we are currently investigating. An extensive study to address the efficiency of these calculations is important to develop practical molecular force field optimization methods based on entropies and enthalpies of solvation. These are, in fact, slowly convergent properties, and force field optimization procedures require a long series of simulations.

We also plan to study a larger series of solutes in order to correlate deviations of the calculated enthalpies and entropies of solvation from the experiments to molecular properties. From this study we hope to address some of the pitfalls of the existing molecular force fields and issues related to their improvement.

Acknowledgment. This work has been supported by NIH grants GM 52210 and GM 30580, by a grant of the Center of Biomolecular Simulations at Columbia University NIH RR068R, and by a NSF Metacenter grant (MCA 935021P) of computer time. M.M.K. is grateful to Dr. Y. Ding for helpful discussions.

Apendix A. Sources of Experimental Quantities

In the preceding sections we were concerned with the calculation of the free energy, enthalpy and entropy in the Ben-Naim standard states convention[18] (eqs 1 and 3). Experimental solution thermodynamic quantities are, instead, tabulated in terms of a variety of standard states. The conversion between the two conventions is accomplished by the following relations

$$\Delta G^* = \Delta G^\circ - kT \ln \frac{\rho_u kT}{P} \tag{6}$$

Downloaded by UNIV OF TEXAS AUSTIN on July 7, 2009
Published on December 4, 1997 on http://pubs.acs.org | doi: 10.1021/jp972598b

$$-T\Delta S^* = -T\Delta S^\circ - kT(1 - T\alpha_w) - kT\ln\frac{\rho_u kT}{P} \quad (7)$$

and

$$\Delta H^* = \Delta H^\circ + kT(1 - T\alpha_w) \quad (8)$$

where the standard experimental quantities have been marked with a circle, ρ_u is the number density of the solute at the experimental solution standard state, α_w the thermal expansion coefficient of water at 298 K, and P is the atmospheric pressure. These transformations have been applied to the standard free energies, enthalpies, and entropies of hydration obtained from the tabulated data from refs 11, 41, 17, and 42 to obtain the free energies, enthalpies and entropies of hydration in the Ben-Naim convention reported in Tables 3−5.

Appendix B. Thermodynamic Relations

B.1. Internal Energy Change for the Dissolution Process. The change of internal energy of a solution by adding one molecule of species k at constant pressure is

$$(\delta E)_P = \left(\frac{\partial E}{\partial N_k}\right)_{P,T,N_i'} \delta N_k = \left(\frac{\partial E}{\partial N_k}\right)_{P,T,N_i'} \quad (9)$$

where $(\partial/\partial N_k)_{N_i'}$ denotes a derivative, keeping all the composition variables constant except N_k, and we have set the variation in number of molecules δN_k to 1.

The change of internal energy by adding one molecule of species k at constant volume is instead

$$(\delta E)_V = \left(\frac{\partial E}{\partial N_k}\right)_{V,T,N_i'} \delta N_k = \left(\frac{\partial E}{\partial N_k}\right)_{V,T,N_i'} \quad (10)$$

Using the chain rule for the derivatives of thermodynamic quantities, we rewrite the right-hand side of eq 9 as

$$\left(\frac{\partial E}{\partial N_k}\right)_{P,T,N_i'} = \left(\frac{\partial E}{\partial N_k}\right)_{V,T,N_i'} + \left(\frac{\partial E}{\partial V}\right)_{T,N_i}\left(\frac{\partial V}{\partial N_k}\right)_{P,T,N_i'}$$

$$= (\delta E)_V + (\delta_V)_P\left(\frac{\partial E}{\partial V}\right)_{T,N_i} \quad (11)$$

The last term on the right side of eq 11 can be further expanded to obtain

$$\left(\frac{\partial E}{\partial V}\right)_{T,N_i} = \left(\frac{\partial E}{\partial V}\right)_{S,N_i} + \left(\frac{\partial E}{\partial S}\right)_{V,N_i}\left(\frac{\partial S}{\partial V}\right)_{T,N_i}$$

$$= -P + T\left(\frac{\partial S}{\partial V}\right)_{T,N_i} \quad (12)$$

By using the following Maxwell relation

$$\left(\frac{\partial S}{\partial V}\right)_{T,N_i} = \left(\frac{\partial P}{\partial T}\right)_{V,N_i} \quad (13)$$

and the variable transformation

$$\left(\frac{\partial P}{\partial T}\right)_{V,N_i} = -\left(\frac{\partial V}{\partial T}\right)_{P,N_i}\left(\frac{\partial P}{\partial V}\right)_{T,N_i}$$

$$= (-V\alpha)\left(-\frac{1}{V\kappa}\right) = \frac{\alpha}{\kappa} \quad (14)$$

where

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_{P,N_i} \quad (15)$$

is the cubic thermal expansion coefficient and

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_{T,N_i} \quad (16)$$

is the isothermal compressibility, we finally obtain

$$(\delta E)_P - (\delta E)_V = (\delta V)_P\left(-P + T\frac{\alpha}{\kappa}\right) \quad (17)$$

For instance, for the dissolution of methane in water at $T = 298K$, setting $(\delta V)_P = 36.2$ cm^3/mol and $\alpha = 2.57 \times 10^{-4}\,K^{-1}$, $\kappa = 4.58 \times 10^{-5}$ atm^{-1}, we obtain

$$(\delta E)_P - (\delta E)_V \simeq 1.5 \text{ kcal/mol}$$

B.2. Enthalpy Change for the Dissolution Process. The enthalpy change by adding one molecule of species k at constant pressure is

$$(\delta H)_P = \left(\frac{\partial H}{\partial N_k}\right)_{P,T,N_i'} \delta N_k = \left(\frac{\partial H}{\partial N_k}\right)_{P,T,N_i'} \quad (18)$$

The enthalpy change by adding one molecule of species k at constant volume is, instead,

$$(\delta H)_V = \left(\frac{\partial H}{\partial N_k}\right)_{V,T,N_i'} \delta N_k = \left(\frac{\partial H}{\partial N_k}\right)_{V,T,N_i'} \quad (19)$$

Using the chain rule for the derivatives of thermodynamic quantities, we rewrite the right-hand side of eq 18 as

$$\left(\frac{\partial H}{\partial N_k}\right)_{P,T,N_i'} = \left(\frac{\partial H}{\partial N_k}\right)_{V,T,N_i'} + \left(\frac{\partial H}{\partial V}\right)_{T,N_i}\left(\frac{\partial V}{\partial N_k}\right)_{P,T,N_i'}$$

$$= (\delta H)_V + (\delta V)_P\left(\frac{\partial H}{\partial V}\right)_{T,N_i} \quad (20)$$

The last term on the right side of eq 20 can be further expanded to obtain

$$\left(\frac{\partial H}{\partial V}\right)_{T,N_i} = \left(\frac{\partial H}{\partial P}\right)_{T,N_i}\left(\frac{\partial P}{\partial V}\right)_{T,N_i}$$

$$= \left[\left(\frac{\partial H}{\partial P}\right)_{S,N_i} + \left(\frac{\partial H}{\partial S}\right)_{P,N_i}\left(\frac{\partial S}{\partial P}\right)_{T,N_i}\right]\frac{-1}{V\kappa}$$

$$= -\left[V + T\left(\frac{\partial S}{\partial P}\right)_{T,N_i}\right]\frac{1}{V\kappa} \quad (21)$$

By using the following Maxwell relation

$$\left(\frac{\partial S}{\partial P}\right)_{T,N_i} = -\left(\frac{\partial V}{\partial T}\right)_{P,N_i} = -V\alpha \quad (22)$$

we finally obtain

$$(\delta H)_P - (\delta H)_V = (\delta V)_P\frac{T\alpha - 1}{\kappa} \quad (23)$$

Downloaded by UNIV OF TEXAS AUSTIN on July 7, 2009
Published on December 4, 1997 on http://pubs.acs.org | doi: 10.1021/jp972598b

For the dissolution of methane in water at $T = 298$ *K* we obtain

$$(\delta H)_P - (\delta H)_V \simeq -19.5 \text{ kcal/mol}$$

B.3. Comparison between $(\delta H)_P$ and $(\delta E)_V$. Since we have $H = E + PV$, we find

$$(\delta H)_P = \left(\frac{\partial H}{\partial N_k}\right)_{P,T,N'_i} = \left(\frac{\partial E}{\partial N_k}\right)_{P,T,N'_i} + P\left(\frac{\partial V}{\partial N_k}\right)_{P,T,N'_i} = (\delta E)_P + P(\delta V)_P \quad (24)$$

However, from eq 17,

$$(\delta E)_P = (\delta E)_V + (\delta V)_P\left(-P + T\frac{\alpha}{\kappa}\right) \quad (25)$$

Therefore,

$$(\delta H)_P - (\delta E)_V = P(\delta V)_P + \left(-P + T\frac{\alpha}{\kappa}\right)(\delta V)_P = \frac{T\alpha}{\kappa}(\delta V)_P \quad (26)$$

For the dissolution of methane in water at $T = 298$ *K* we obtain

$$(\delta H)_P - (\delta E)_V \simeq 1.5 \text{ kcal/mol}$$

References and Notes

(1) Tembe, B.; McCammon, J. *Comput. Chem.* **1984**, *8*, 281.
(2) Beveridge, D.; Di Capua, F. *Annu. Rev. Biophys. Chem.* **1989**, *18*, 431.
(3) Jorgensen, W. *Acc. Chem. Res.* **1989**, *22*, 184.
(4) Warshel, A.; Sussman, F.; King, G. *Biochemistry* **1986**, *25*, 8368.
(5) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.
(6) Cornell, W.; Cieplak, P.; Bayly, C.; Gould, I.; Merz, K.; Ferguson, D.; Spellmeyer, D.; Fox, T.; Caldwell, J.; Kollman, P. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(7) Jorgensen, W.; Maxwell, D.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
(8) Carlson, H.; Nguyen, T.; Orozco, M.; Jorgensen, W. *J. Comput. Chem.* **1993**, *14*, 1240.
(9) Kaminski, G.; Duffy, E.; Matsui, T.; Jorgensen, W. *J. Phys. Chem.* **1994**, *98*, 13077.
(10) Sun, Y.; Spellmeyer, D.; Pearlman, D.; Kollman, P. *J. Am. Chem. Soc.* **1992**, *114*, 6798.
(11) Jones, F.; Arnett, E. *Prog. Phys. Org. Chem.* **1974**, *11*, 263.
(12) Marten, B.; Kim, K.; Cortis, C.; Friesner, R.; Murphy, R.; Ringnalda, M.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 11775.
(13) Fleischman, S.; Brooks, C. *J. Chem. Phys.* **1987**, *87*, 3029.
(14) Smith, D.; Haymet, A. *J. Chem. Phys.* **1993**, *98*, 6445.
(15) Ding, Y.; Bernardo, D.; Krogh-Jespersen, K.; Levy, R. *J. Phys. Chem.* **1995**, *99*, 11575.
(16) Morgantini, P.; Kollman, P. *J. Am. Chem. Soc.* **1995**, *117*, 6057.
(17) Wagman, D.; Evans, W.; Parker, V.; Schumm, R.; Halow, I.; Bailey, S.; Churney, K.; Nuttall, R. *J. Phys. Chem. Ref. Data: NBS Tables of Chemical Thermodynamic Properties*; American Institute of Physics, Inc.: New York, 1982.
(18) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016.
(19) Qian, H.; Hopfield, J. *J. Chem. Phys.* **1996**, *105*, 9292.
(20) Matubayasi, N.; Reed, L.; Levy, R. *J. Phys. Chem.* **1994**, *98*, 10640.
(21) Ding, Y. Ph.D. Thesis, Rutgers, The State University of New Jersey, New Brunswick, NJ, 1996.
(22) Privalov, P.; Gill, S. *Adv. Protein Chem.* **1988**, *39*, 191.
(23) Frisch, M.; et al. *Gaussian 92*; Gaussian Inc.: Pittsburgh, 1992.
(24) Frisch, M.; et al. *Gaussian 94*; Gaussian Inc.: Pittsburgh, 1995.
(25) Breneman, C.; Wiberg, K. *J. Comput. Chem.* **1990**, *5*, 129.
(26) Kitchen, D.; Hirata, F.; Westbrook, J.; Levy, R.; Kofke, D.; Yarmush, M. *J. Comput. Chem.* **1990**, *11*, 1169.
(27) Weiner, S.; Kollman, P.; Nguyen, D.; Case, D. *J. Comput. Chem.* **1986**, *7*, 230.
(28) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926.
(29) Berendsen, H.; Postma, J.; van Gunsteren, W.; Di Nola, A.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684.
(30) Andersen, H. *J. Comput. Phys.* **1983**, *52*, 24.
(31) van Gunsteren, W.; Berendsen, H. *Mol. Phys.* **1977**, *34*, 1311.
(32) Hägler, A.; Euler, E.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5319.
(33) Chipot, C.; Millot, C.; Maigret, B.; Kollman, P. *J. Phys. Chem.* **1994**, *98*, 11362.
(34) Fukunishi, Y. Unpublished data.
(35) Mezei, M.; Beveridge, D. *Ann. N. Y. Acad. Sci.* **1986**, *482*, 1.
(36) Cross, A. *J. Chem. Phys. Lett.* **1986**, *128*, 198.
(37) Lin, C.; Wood, R. *J. Comput. Chem.* **1994**, *15*, 149.
(38) Rao, B.; Singh, U. *J. Am. Chem. Soc.* **1989**, *111*, 3125.
(39) Orozco, M.; Jorgensen, W.; Luque, F. *J. Comput. Chem.* **1993**, *14*, 1498.
(40) Gallicchio, E.; Kubo, M.; Levy, R. Unpublished data.
(41) Cavani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563.
(42) Wilhelm, E.; Battino, R.; Wilcock, R. *Chem. Rev.* **1977**, *77*, 219.

Downloaded by UNIV OF TEXAS AUSTIN on July 7, 2009
Published on December 4, 1997 on http://pubs.acs.org | doi: 10.1021/jp972598b