Explicitly Representing the Solvation Shell in Continuum Solvent Calculations

Eirik F. da Silva,*,†,‡ Hallvard F. Svendsen,† and Kenneth M. Merz‡

Department of Chemical Engineering, Norwegian University of Science and Technology, N-7491 Trondheim, Norway, SINTEF Materials and Chemistry, R. Birkelands Vei 2B, N-7465 Trondheim, Norway and Quantum Theory Project, University of Florida, Gainesville, Florida 32611

A method is presented to explicitly represent the first solvation shell in continuum solvation calculations. Initial solvation shell geometries were generated with classical molecular dynamics simulations. Clusters consisting of solute and 5 solvent molecules were fully relaxed in quantum mechanical calculations. The free energy of solvation of the solute was calculated from the free energy of formation of the cluster, and the solvation free energy of the cluster was calculated with continuum solvation models. The method has been implemented with two continuum solvation models, a Poisson−Boltzmann model and the IEF-PCM model. Calculations were carried out for a set of 60 ionic species. Implemented with the Poisson−Boltzmann model the method gave an unsigned average error of 2.1 kcal/mol and a rmsd of 2.6 kcal/mol for anions; for cations the unsigned average error was 2.8 kcal/mol and the rmsd 3.9 kcal/mol. Similar results were obtained with the IEF-PCM model.

1. Introduction

Solvation phenomena play an important part in many areas of chemistry, and there has for some time been significant interest among chemists in predicting solvation effects. One property of particular importance is the free energy of solvation. Accurate calculation of this property is essential in predicting chemical reactions in solution.

In the last decades theoretical and computational chemists have put considerable efforts into the development of methods to calculate the free energy of solvation. Much of this work is covered in reviews by Cramer and Truhlar,[1] Orozco and Javier Luque,[2] and Tomasi, Menucci, and Cammi.[3] Present solvation models tend to perform well in predicting the solvation energy of small neutral species but are less successful for ionic species.

Continuum solvation models represent an appealing approach to calculating solvation free energies. Compared to free energy calculations with explicit solvent models, continuum model calculations are both easy to set up and computationally inexpensive. The main shortcoming of continuum models is probably their failure to account for chemical interactions between a solute and solvent molecules in the first solvation shell. Such interactions are expected to be of particular importance for ionic solutes.[4−7] Continuum models are also not able to capture effects of charge transfer between solute and solvent, an effect that may play a significant role in solvation.[8,9]

It has repeatedly been proposed that this shortcoming can be overcome by adding explicit solvent molecules in the continuum model calculations.[1,4,5,9−19] One approach is quasi-chemical theory, where the closest solute molecules are regarded as ligands binding to the solute.[11] Other approaches involve including a limited number of solvent molecules in continuum solvation calculation, with varying thermodynamic cycles.[4−7]

* To whom correspondence should be addressed. E-mail: Eirik.Silva@ sintef.no.
† Norwegian University of Science and Technology and SINTEF Materials and Chemistry.
‡ University of Florida.

SCHEME 1
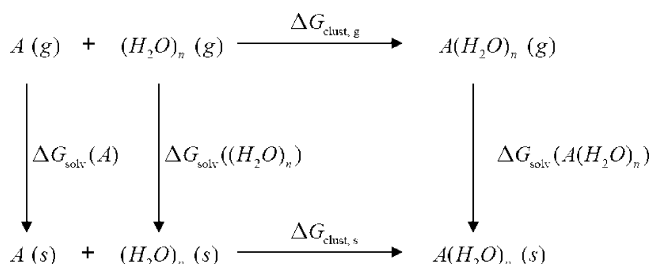
Such models have been referred to by varying names such as the supermolecule method or cluster-continuum method.[4]

In the present work a new method is proposed for adding explicit solvent molecules in continuum model calculations. The focus of the present work is to produce a robust method that may be applied to a wide variety of ionic species. The method is fully scripted, only requiring solute geometry and charge as input.

2. Theory

In the thermodynamic cycle in Scheme 1 we show that the free energy of solvation of a given solute can be calculated from the solvation free energies of solute−solvent clusters and the free energy of solute−solvent cluster formation. If the solvent and solute−solvent clusters have the structure of bulk solvent, the free energy of cluster formation in solution is 0. The reaction in solution can in this case be regarded as exchange of implicit and explicit solvent molecules.

It follows that the free energy of solvation of the solute can be calculated from the free energy of cluster formation in the gas phase and the free energy of solvation of the clusters:

$$\Delta G_{solv}(A) = \Delta G_{clust,g}(A(H_2O)_n) + \Delta G_{solv}(A(H_2O)_n) - \Delta G_{solv}((H_2O)_n) \quad (1)$$

Experimental and calculated solvation energies are usually reported with a standard state of 1 mol/L (indicated with an

Continuum Solvent Calculations

J. Phys. Chem. A, Vol. 113, No. 22, 2009 **6405**

asterisk (*) in the present work) for both the gas and liquid state. Adopting this standard state the following equation is obtained:

$$\Delta G^*_{solv}(A) = \Delta G^*_{clust, g}(A(H_2O)_n) + \Delta G^*_{solv}(A(H_2O)_n) - \Delta G^*_{solv}((H_2O)_n) - RT \ln([H_2O]/n) \quad (2)$$

The last term in eq 2 is a correction term for the free energy change of 1 mol of $(H_2O)_n$ gas from 55.34/n M liquid to 1 mol/ L. The correction from ΔG_{clust} at a standard state of 1 atm to ΔG^*_{clust} is $-RT \ln(R'T)$, where R' is 0.082 053 K^{-1}.[4] A discussion of the standard state corrections is given by Bryantsev et al.[7] Equation 2 is the same as that applied by Bryantsev et al.[7] in their recent work. This equation also gives the same relative solvation energies as the method Pliego and Riveros[4] proposed in their cluster-continuum model. We however chose to calculate the energies relative to pure solvent clusters, while Pliego and Riveros calculated energies relative to the experimentally determined vaporization free energy of the solvent.

It should be noted that the accuracy of calculations based on eq 1 will depend on how close the utilized cluster geometries are to the real structure of solvation shells in bulk water.

3. Methods

In this work we extract the geometry of the solvation shell from molecular simulations of the solute in bulk solvent and fully optimize this solvation shell geometry with quantum mechanical calculations. One could also do single-point calculations on the simulation geometries, but our initial work suggested that the present molecular dynamics geometries (from classical simulations) are not good enough for such an approach to yield accurate solvation free energies.

The clusters consisted of the solute and the closest 5 water molecules. In selecting the number of solvent molecules to include in the solvation shell, there are several considerations. We wish to include enough solvent molecules to describe all strong solute−solvent interactions. On the other hand, adding more solvent molecules adds to the cost of the calculations. Since we are fully relaxing the solvation shell in vacuum there is also a risk that the structure will deviate more from the structure of bulk solvent as more solvent molecules are added. The present method may have systematic errors that depend on the number of solvent molecules in the calculation (an issue to which we will return in the discussion); we therefore believe that the most reliable results are obtained by using a constant number of solvent molecules for all solutes. For the small solutes studied in the present work, we believe 5 solvent molecules to be enough to represent strong solute−solvent interactions.

We wished to include the solvent molecules that are bound most strongly to the solute in the cluster. The solvent molecules were therefore selected based on their distance to hydrophile atoms in the solute. Hydrophile atoms are in the present work defined as any solute atoms that are not carbon atoms or hydrogen atom bound to a carbon atom.

Full optimization of solute and the solute−solvent clusters were carried out in vacuum at the HF/6-31+G(d) level. Pure solvent clusters were optimized at the same level of theory. The entropy was determined from vibrational frequency calculations on the optimized clusters at the same level of theory at a temperature of 298 K.

The interactions between a solute and solvent molecules are in general not well represented by a single geometry. In the present work we therefore chose to carry out calculations with an ensemble of clusters. One hundred different cluster geometries were extracted from molecular dynamics simulations for
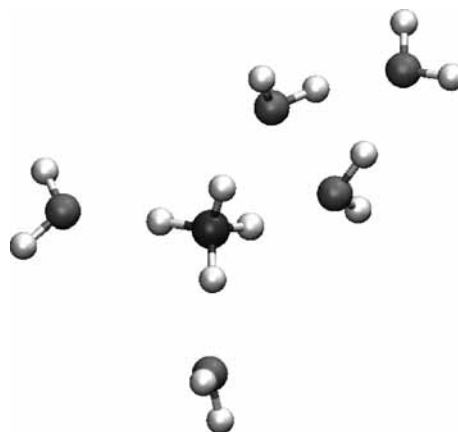

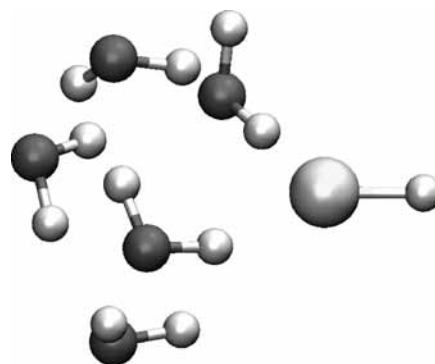**Figure 1.** Optimized cluster of NH_4^+.

Figure 2. Optimized cluster of HS^-.

each solute. The free energies we report are linear averages from calculations on the full set of clusters.

We chose to implement the method with two continuum solvation models; this is in part done to test to what extent the method performance depends on the continuum model. The continuum models employed for calculations of cluster solvation energies were the Poisson−Boltzmann-based model in the DIVCON code[20] and the IEF-PCM[21] model in Gaussian 03.[22] The continuum solvation calculations were carried out as single-point calculations on the optimized HF/6-31+G(d) clusters. The Poisson−Boltzmann model calculations were carried out at the AM1 level (the model is not implemented at HF level), while the IEF-PCM calculations were carried out at the HF/6-31+G(d) level. The IEF-PCM calculations were carried out with default settings in Gaussian 03.[22]

A summary of the Poisson−Boltzmann model is given in the Supporting Information. All quantum mechanical calculations were carried out in Gaussian 03,[22] and all simulations were carried out using Sander from the AMBER 9 suite.[23] Details of the molecular dynamics simulations are given in the Supporting Information.

4. Results

In Figures 1 and 2 are shown examples of optimized clusters produced with the present method. Figure 1 shows a case of the solvent molecules interacting with all parts of the solute, creating a complete solvation shell. In other cases (such as that shown in Figure 2) the solvent molecules will cluster around hydrophilic groups in the solute, leaving hydrophobic groups exposed.

Table 1 shows the calculated solvation energies, cluster formation energies, entropies, and cluster solvation energies for

TABLE 1: Experimental and Calculated Free Energies of Solvation for Anions (all data in kcal/mol)

A^-	AH	ΔG_{solv} (exp)[a]	ΔG_{solv} (calcd)[b]	error	$\Delta E^*_{cluster}$[c]	$-T\Delta S^*_{cluster}$[d]	$\Delta G_s(A(S)_n)$[e]
Cl^-		−74.5	−77.0	−2.5	−31.6	5.9	−61.2
OH^-	water	−104.7	−106.3	−1.6	−66.4	11.2	−61.1
CH_3O^-	methanol	−95	−94.6	0.4	−57.2	10.8	−58.2
$CH_3CH_2O^-$	ethanol	−90.7	−90.5	0.2	−56.3	11.7	−55.8
$C_6H_5O^-$	phenol	−71.9	−68.8	3.1	−34.6	9.7	−53.7
HCO_2^-	formic acid	−76.2	−78.4	−2.2	−41.4	10.5	−57.3
NH_2^-	ammonia	−92.2[f]	−97.6[g]	−5.4	−52.6[g]	16.4[g]	−71.2[g]
HS^-	hydrogen sulfide	−72.1	−77.5	−5.4	−31.4	7.7	−63.9
CH_3S^-	methanethiol	−73.8	−75.9	−2.1	−32.1	8.4	−62.1
$CH_3CH_2CH_2O^-$	1-propanol	−88.3	−88.5	−0.2	−55.1	12.0	−55.4
$(CH_3)_2CHO^-$	2-propanol	−86.3	−85.9	0.4	−53.7	12.2	−54.3
$C(CH_3)_3O^-$	*tert*-butyl alcohol	−82.3	−82.4	−0.1	−52.5	12.9	−52.7
CH_3OO^-	methyl peroxide	−93.2	−89.5	3.7	−52.8	11.6	−58.3
CN^-	hydrogen cyanide	−70.2	−73.7	−3.5	−30.9	8.4	−61.1
$CH_2NO_2^-$	nitromethane	−76.5	−77.6	−1.1	−39.7	11.6	−59.4
HO_2^-	hydrogen peroxide	−97.3	−95.6	1.7	−55.4	11.2	−61.4
F^-		−104.4	−106.7	−2.3	−62.7	7.2	−61.2
Br^-		−68.3	−72.5	−4.2	−30.7	9.5	−61.2
HCC^-	acetylene	−76.5	−75.4	1.1	−32.0	8.2	−61.5
$CH_3CO_2^-$	acetic acid	−77.6	−79.7	−2.1	−42.0	9.7	−57.3
$CH_3CH_2CO_2^-$	propanoic acid	−76.2	−75.9	0.3	−41.6	11.6	−55.8
$H_2C{=}CHCO_2^-$	acrylic acid	−74	−75.3	−1.3	−39.9	11.0	−56.4
$C_6H_5CO_2^-$	benzoic acid	−71.2	−71.9	−0.7	−37.8	11.5	−55.6
$H_2C{=}CHCH_2O^-$	allyl alcohol	−86.6	−85.6	1.0	−52.8	12.2	−54.9
$C_6H_5CH_2O^-$	benzyl alcohol	−85.1	−80.6	4.5	−48.6	13.0	−55.0
$CH_3OCH_2CH_2O^-$	2-methoxyethanol	−89.4	−88.0	1.4	−49.6	10.7	−59.0
$C_6H_5CO_2^-$	benzoic acid	−71.2	−71.9	−0.7	−37.8	11.5	−55.6
o-$CH_3C_6H_4O^-$	2-methylphenol	−70.2	−66.5	3.7	−34.7	12.1	−53.8
m-$CH_3C_6H_4O^-$	3-methylphenol	−71.1	−68.7	2.4	−35.0	10.8	−54.4
p-$CH_3C_6H_4O^-$	4-methylphenol	−72.0	−68.4	3.6	−37.3	12.8	−53.8

[a] Experimental free energies of solvation at 298 K as reported by Kelly, Cramer, and Truhlar.[24] [b] Calculated free energy of solvation; all values shifted by −2.41 kcal/mol to remove systematic error relative to experimental values. Estimated sampling standard deviation is 1 kcal/mol. [c] Energy of formation of the cluster at the HF/6-31+G(d) level, converted from a standard state of 1 atm to 1 mol/L. Thermal corrections to the energy and zero-point energies not included. [d] Temperature (298 K) multiplied by the entropy of formation of the cluster at the HF/6-31+G(d) level. [e] Free energy of solvation of the cluster calculated with the Poisson−Boltzmann continuum model. [f] Experimental data from Pliego and Riveros.[18] [g] All solvent molecule O−H bonds constrained in calculations.

a set of 30 anionic compounds. In Table 2 are shown data for cations presented in the same form. The set of ions covers many important functional groups in chemistry such as alcohols, amines, and carboxylic acids. This set also includes all ions studied by Pliego and Riveros.[4] The cluster solvation energies in the tables are calculated with the Poisson−Boltzmann continuum solvation model; results with the IEF-PCM solvation model are given in the Supporting Information.

All calculated solvation energies were shifted to remove the average error relative to experimental data. For anionic species the shift is −2.41 kcal/mol, while for cationic species the shift is −5.60 kcal/mol. This means that the model is only used to predict relative free energies of solvation and not their absolute values.

The unsigned average error in the calculated solvation free energies for anions is 2.1 kcal/mol, while the root-mean-square deviation (rmsd) is 2.6 kcal/mol. For cations the unsigned average error in the calculated solvation free energies is 2.8 kcal/mol and the rmsd is 3.9 kcal/mol. It can be seen from the data in Tables 1 and 2 that the energy of cluster formation is the most important term in determining differences in solvation energies. The clusters solvation energies do also vary considerably.

Calculations with the IEF-PCM continuum model gave similar results. For anions the unsigned average error is 1.6 kcal/mol and the rmsd 2.1 kcal/mol. For cations the unsigned average error is 3.2 kcal/mol and the rmsd 4.5 kcal/mol. The average shift of calculated relative to experimental data for anions is −12.0 kcal/mol. For cations this average shift is −5.5 kcal/

mol. The two continuum models predict similar absolute solvation energies for anions; for cations the IEF-PCM solvation energies are however on average 7.7 kcal/mol lower than the solvation energies from the Poisson−Boltzmann model. The two solvation models also give substantially different solvation energies for the pure solvent clusters. The Poisson−Boltzmann model gives an average value of −15.7 kcal/mol, while the IEF-PCM model gives an average value of −23.6 kcal/mol. Similar differences between continuum models in solvent cluster calculations were reported by Bryantsev et al.[7]

For NH_2^- proton exchange occurred between the solute and solvent molecules. This problem was overcome by constraining all solvent O−H bond lengths. A similar issue with proton exchange was found for protonated acetone and dimethyl ether. For clusters of these solutes the O−H bond in the solute was constrained. These three solutes were not included in calculations of systematic errors between calculated and experimental data.

In the presented results thermal corrections to the energy and zero-point energy calculations were not included. We found that including these effects would result in somewhat worse results. Inclusion of the thermal correction and the zero-point energy would have formally been the more correct choice. The omission of these corrections can be regarded as an implicit form of parameter fitting in the present method. Results with these terms included are shown in the Supporting Information.

The statistical uncertainty of the method was estimated by carrying out calculations at the semiempirical AM1 level of

Continuum Solvent Calculations

J. Phys. Chem. A, Vol. 113, No. 22, 2009 **6407**

TABLE 2: Experimental and Calculated Free Energies of Solvation for Cations (all data in kcal/mol)

BH^+	B	ΔG_{solv} (exp)[a]	ΔG_{solv} (calcd)[b]	error	ΔE^*_{clust}[c]	$-T\Delta S^*_{clust}$[d]	$\Delta G_s(A(S)_n)$[e]
H^+		−265.9	−272.0	−6.1	−229.8	8.0	−57.0
NH_4^+	ammonia	−85.2	−91.5	−6.3	−47.7	6.7	−57.3
H_3O^+	water	−110.3	−110.8	−0.5	−71.6	9.6	−55.5
K^+		−86.0	−89.6	−3.6	−45.5	3.0	−53.8
$CH_3OH_2^+$	methanol	−93.0	−91.4	1.6	−54.4	8.9	−52.6
$CH_3CH_2OH_2^+$	ethanol	−88.4	−85.6	2.8	−49.9	8.3	−50.7
$CH_3NH_3^+$	methyalmine	−76.4	−80.4	−4.0	−41.1	8.0	−54.1
$C_6H_5NH_3^+$	aniline	−72.4	−71.9	0.5	−38.6	9.5	−49.6
$C_4H_5NH^+$	pyrrole	−61.4	−67.1	−5.7	−33.7	10.9	−51.0
$CH_3(CH_2)_2NH_3^+$	*n*-propylamine	−71.5	−70.9	0.6	−36.6	10.5	−51.6
$C(CH_3)_3NH_3^+$	*tert*-butylamine	−67.3	−63.3	4.0	−32.3	10.5	−48.3
$(CH_3)_2NH_2^+$	dimethylamine	−68.6	−69.8	−1.2	−34.7	9.4	−51.3
$(CH_3)_3NH^+$	trimethylamine	−61.1	−61.2	−0.1	−25.8	9.1	−51.2
$C_3H_6NH_2^+$	azetidine	−67.7	−64.6	3.1	−30.3	10.6	−51.7
$C_4H_8NH_2^+$	pyrrolidine	−66.0	−64.8	1.2	−32.2	9.8	−49.2
$C_5H_{10}NH_2^+$	piperidine	−64.2	−62.5	1.7	−31.1	10.0	−48.2
CH_3CNH^+	acetonitrile	−75.3	−72.6	2.7	−33.6	9.6	−55.4
$(CH_3)_2OH^+$	dimethyl ether	−79.7	−69.2[f]	10.5	−35.3[f]	12.2[f]	−52.8[f]
$CH_3C(OH)CH_3^+$	acetone	−77.1	−67.0[f]	10.1	−34.0[f]	11.9[f]	−51.6[f]
$CH_3C(OH)C_6H_5^+$	acetophenone	−64.5	−58.5	6.0	−27.5	10.3	−48.0
$(CH_3)_2CHNH_3^+$	isopropylamine	−69.6	−69.3	0.3	−35.6	9.5	−49.9
c-$C_6H_{11}NH_3^+$	cyclohexanamine	−68.7	−66.2	2.5	−35.6	10.4	−47.7
$H_2C{=}CHCH_2NH_3^+$	allyl amine	−72.0	−72.4	−0.4	−37.3	9.5	−51.3
$C_4H_8NHNH_2^+$	piperazine	−66.0	−65.5	0.5	−27.1	8.0	−53.2
$C_5H_5NH^+$	pyridine	−61.1	−60.8	0.3	−26.6	10.5	−51.5
$H_2NNH_3^+$	hydrazine	−84.6	−88.3	−3.7	−46.3	8.1	−57.0
$C_4H_8ONH_2^+$	morpholine	−69.6	−66.6	3.0	−30.3	9.9	−53.0
$C_2H_4NH_2^+$	aziridine	−70.9	−71.4	−0.5	−34.3	9.5	−53.3
$(C_2H_5)_2NH_2^+$	diethylamine	−63.4	−62.6	0.8	−31.6	10.1	−47.9
$(n$-$C_3H_7)_2NH_2^+$	di-*n*-propylamine	−60.5	−60.0	0.5	−31.3	10.4	−45.9

[a] Experimental free energies of solvation at 298 K as reported by Kelly, Cramer, and Truhlar.[24] [b] Calculated free energy of solvation; all values shifted by −5.60 kcal/mol to remove systematic error relative to experimental values. Estimated sampling standard deviation is 1 kcal/mol. [c] Energy of formation of the cluster at the HF/6-31+G(d) level, converted from a standard state of 1 atm to 1 mol/L. Thermal corrections to the energy and zero-point energies not included. [d] Temperature (298 K) multiplied by the entropy of formation of the cluster at the HF/6-31+G(d) level. [e] Free energy of solvation of the cluster calculated with the Poisson−Boltzmann continuum model. [f] Solute O−H bonds constrained in calculations to its length in vacuum.

theory.[27] Ten sets of calculations were performed for the cluster formation energies (based on a total of 1000 different cluster geometries). This gave a standard deviation in calculated cluster formation energies of around 1 kcal/mol.

Some of the cluster geometry optimizations failed, resulting in the breakdown of the geometry of the solute or solvent molecules. In some cases the vibration frequency calculations also failed. These clusters were omitted in the calculations of solvation energies. For most solutes there were 10−20 failed calculations. A likely cause of such failed calculations is poor initial geometries obtained from the molecular dynamics simulations.

5. Discussion

5.1. Method Performance. The present results are to our knowledge among the best reported relative solvation energies for ions. The error bars are not too far from the error bars in the experimental data suggested by Kelly, Cramer, and Truhlar (±2−3 kcal/mol).[25] The size of the study set also gives confidence that the method is robust. Much of the work with models combining continuum models with explicit solvent representation has been focused on solutes with no hydrophobic groups.[7,13,14,19] For such solutes optimized clusters will form complete solvation shells, and such systems are likely to be ideally suited for models with explicit representations of the solvation shell in continuum. The present work does however show that even for systems with hydrophobic groups where a complete solvation shell is not formed, good results can be obtained.

It is also noteworthy that methods combining continuum models with explicit solvent representation, unlike a conventional continuum model or classical simulation, to some extent take account of all contributions to the solvation energy. Long-range electrostatics, solute polarization, and cavitation energy contributions are accounted for in the continuum model, while solute−solvent hydrogen bonding and entropy contributions are accounted for in the calculation of cluster formation energy.

The quality of the results obtained is likely to depend strongly on the quality of the quantum mechanical method employed. The continuum solvation model is also likely to have a significant effect on the quality of the results. The entropy term contributes less to relative solvation energies but may also have a significant effect on the results. We do make the approximation that the vibration entropy can be calculated as if the cluster was a single molecule. Since the clusters are completely relaxed in quantum mechanical calculations, the initial molecular dynamics geometry probably does not affect the final results. As noted in the Theory section, the results also depend on how closely the optimized cluster resembles the solvation shell geometry in bulk liquid. This may especially be an issue for solutes with hydrophobic groups that are not forming complete solvation shells.

For NH_2^-, acetone, and dimethyl ether calculations were carried out with constrained bonds. The errors in terms of relative solvation free energies for these species are all somewhat large. While this approach seems to work reasonably well for solutes that might otherwise undergo some form of proton transfer, it does appear to lead to somewhat less accurate results.

If the three solutes with constrained bonds were to be removed from the data set the unsigned average error would fall to 2.0 kcal/mol for anions and 2.2 kcal/mol for cations (with the Poisson−Boltzmann continuum model). With the IEF-PCM continuum model the unsigned average errors improve to 1.5 (anions) and 2.6 kcal/mol (cations).

The calculated solvation free energies for purely hydrophilic solutes also have relatively large errors. For these solutes the calculated solvation free energies are more negative than the experimental values. These solutes such as H^+ and NH_4^+ form ideal solvation shells. The relative differences in calculated solvation free energies between entirely hydrophilic solutes and solutes with hydrophobic groups suggests that the model performance to some extent depends on the structure of the solvation shell formed.

In the present model hydrophobic groups will not interact with the explicit solvation shell. It is the continuum model that accounts for the hydrophobic group's contribution of the free energy of solvation. Relatively large errors for species with large hydrophobic groups such as benzyl alcohol are probably due to the performance of the continuum model.

5.2. Absolute Solvation Energies. The present work is focused on obtaining relative solvation energies. Predictions of absolute solvation energies can be made by correcting with the correction factors reported in the present work.

It is however also of interest to look at how close the absolute solvation energies produced by the model are to experimental data. The thermodynamic cycle we employ should in principle yield absolute values of the solvation energies. Models similar in nature to the present work have also been used to predict absolute solvation energies.[7,13,14,17−19]

In looking at the absolute values predicted by the present model, we should look at the formally correct calculations including the zero-point energy and thermal corrections to the energy. If this term is added the average shift for anions is −5.7 kcal/mol and for cations it is −7.3 kcal/mol with the Poisson−Boltzmann continuum model. If we looked at the average shift for hydrophilic species only, the shift would be −2.6 kcal/mol for anions and −3.3 kcal/mol for cations. With the IEF-PCM continuum solvation model the average shift for anions is −15.3 kcal/mol and for cations −7.2 kcal/mol. For hydrophilic species only, the shift was in this case −12.6 kcal/mol for anions and −2.6 kcal/mol for cations.

These results suggest that part of the difference between absolute values of experimental and calculated solvation free energies are systematic errors in the calculation of hydrophobic species. It may be that the continuum solvation models underestimate the contribution to the free energy of solvation of hydrophobic groups.

The absolute solvation energies do depend to a large extent on the continuum solvation model, and the two continuum models we utilize produce quite different results. Asthagiri et al.[13] also reported that the absolute values of long-range electrostatic terms depend significantly on the calculation method. Large differences between continuum solvation in calculations of solvation free energies of solute−solvent clusters have also been observed by Bryantsev et al.[7] The large uncertainty in the absolute energy values produced by continuum models suggests that contributions from these calculations may be the main cause of the systematic difference observed between calculated and experimental values. Hartree−Fock level calculations can also not be expected to yield accurate absolute values of binding energies. The full relaxation of the solvation shell in vacuum may also lead to bonding that deviates systematically

from that of bulk solvent structure. There may also be systematic errors from the approach taken to calculate the entropy terms. It should furthermore be noted that there is significant uncertainty in the absolute value of the experimental solvation free energies; this is due to the uncertainty in the solvation energy of the proton itself. Kelly, Cramer, and Truhlar suggest an uncertainty of no less than 2 kcal/mol for the free energy of solvation of the proton.[16]

As noted in the Methods section, we believe any such systematic error in the present method is likely to depend on the number of solvent molecules included in the clusters. Since we are using a constant number of solvent molecules, any such systematic error should not affect the calculated relative solvation free energies.

5.3. Number of Solvent Molecules. A key issue in a method to explicitly represent solute−solvent interactions is the number of solvent molecules which to include. The number we employ in the present work is higher than in some other studies; it can, for example, be noted that Kelly, Cramer, and Truhlar[28] suggested adding a single solvent molecule. We believe solvent molecules should be added to represent all strong solute−solvent interactions. Some ions such as Cl^- will interact strongly with at least 5 water molecules. In calculations with a similar method Bryantsev et al.[7] reported convergence for H^+ with 14 water molecules. We did however find that not all solutes would bind directly to 5 solvent molecules, suggesting that the optimal number of solvent molecules will vary from ion to ion. The risk with adding more solvent molecules than binds to the solute is that when the cluster is optimized in vacuum the geometry may come to deviate more from that of a solvation shell in bulk liquid. The quality of the present results are however very encouraging, suggesting that 5 solvent molecules is a reasonable choice for small ionic species. While there is some uncertainty as to the optimal number of solvent molecules to include in the calculations, it would also seem clear that for ions including solvent molecules produces significantly better results than a continuum model by itself.

Schemes to automatically select the number of solvent molecules to include in the cluster could be based on binding energies or geometry. A series of calculations could, for example, be made to determine how many solvent molecules bind strongly to the solute. Geometry criteria could, on the other hand, be based on the distance between solute and solvent molecules. Such schemes would result in a more elaborate method, but the calculations could probably be carried out at a low level of theory and need not be time consuming.

Bryantsev et al.[7] in recent work looked at adding solvent molecules until the calculated solvation energies converge, i.e., until results are not affected by the addition of further solvent molecules. This would appear to be a rigorous approach, but it should be noted that the method involves rather large calculations.

5.4. Comparison with Other Explicit Solvation Shell Methods. The present work was greatly influenced by the work of Pliego and Riveros,[4] and the present method is similar to their cluster-continuum method. The overall quality of their results appears to be similar to that of the present method. Using the same procedure as we employ in Tables 1 and 2 and comparing against the same experimental data unsigned average errors of 1.8 (anions) and 2.4 kcal/mol (cations) are obtained. The Pliego and Riveros data set is however small (14 ions), and the statistics is therefore more uncertain than in the present work.

The cluster-continuum method and the present method differ in the number of clusters in the calculation, number of solvent molecules included, and the level of quantum mechanical calculations. The work also differs in the continuum solvation

Continuum Solvent Calculations

J. Phys. Chem. A, Vol. 113, No. 22, 2009 **6409**

models employed. Pliego and Riveros[4] propose a kind of variational principle to determine the number of solvent molecules included in their clusters. As a result of this rule, Pliego and Riveros employ different number of solvent molecules for different solutes. We suspect that this may introduce errors in the calculations. In the same way as the present method, the cluster-continuum method may have systematic errors depending on the number of solvent molecules in the clusters. The quality of Pliego and Riveros results does however suggest that any such systematic error is likely to be small. We are also not confident that their proposed variational principle will yield the best results. We believe that when enough solvent molecules are added to saturate the solvation shell, the calculated solvation energies should converge. In other words, once enough solvent molecules have been added to represent all strong solute−solvent interactions, addition of further solvent molecules should not significantly affect the results.

The level of theory and basis set employed in the work of Pliego and Riveros[4] are better than in the present work. The present method can also be implemented at higher levels of theory, and we believe results are likely to improve with the quality of the quantum mechanical method.

The present method is based on ensemble averages, while Pliego and Riveros[4] did calculations on a single cluster. We have done some initial calculations, suggesting that comparable results are obtained with ensemble averages and calculations on the single most stable cluster in the ensemble. This would imply that similar results can be obtained from methods based on ensemble average and procedures to find the global energy minimum.

5.5. Comparison with Continuum Solvation Models. The recently developed SM6 continuum model has reported unsigned average errors for ions of 4.19 kcal/mol for a data set of 112 ions.[25] While our data set is smaller, it does cover all important classes of ions, and we believe the reported performances are comparable. The performance of the SM6 model is achieved by direct fitting of the model to experimental free energy of solvation data. SM6 is also the latest generation of a series of solvation models that have been developed for over 15 years. That the present model performs better than SM6 is very encouraging. It is also noteworthy that this is achieved with a very limited degree of parametrization.

Kelly, Cramer, and Truhlar[24] also report performance data for a number of other continuum solvation models such as SM5 and IEF-PCM. The present method is substantially better than any of these in predicting relative free energies of solvation for ions. The IEF-PCM method at the HF/6-31G(d) level of theory is, for example, reported to have unsigned average errors of 7.2 kcal/mol.

6. Conclusions

The present solvation model provides relative solvation energies in good agreement with experimental data with very limited parameter fitting. The results suggest that methods based on adding explicit solvent molecules in a continuum model perform significantly better than a continuum model by itself in calculations of free energy of solvation for ionic species.

Supporting Information Available: Details of molecular dynamics simulations, results with the IEF-PCM model, and description of the Poisson−Boltzmann continuum-solvation model. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.
(2) Orozco, M.; Javier Luque, F. *Chem. Rev.* **2000**, *100*, 4187.
(3) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.
(4) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem A* **2001**, *105*, 7241.
(5) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem A* **2006**, *110*, 2493.
(6) Collins, K. D.; Neilson, G. W.; Enderby, J. E. *Biophys. Chem.* **2007**, *128*, 95.
(7) Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A. *J. Phys. Chem. B.* in press.
(8) Nadig, G.; Van Zant, L. C.; Dixon, S. L.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1998**, *120*, 5593.
(9) van der Vaart, A.; Merz, K. M., Jr. *J. Chem. Phys.* **2002**, *116*, 7380.
(10) Tunon, I.; Rinaldi, D.; Ruiz-Lopez, M. F.; Rivail, J. L. *J. Phys. Chem.* **1995**, *99*, 3798.
(11) Karelson, M.; Tamm, T.; Zerner, M. *J. Phys. Chem.* **1993**, *97*, 11901.
(12) Asthagiri, D.; Pratt, L. R.; Kress, J. D. *Phys. Rev. E* **2003**, *68*, 041505.
(13) Asthagiri, D.; Pratt, L. R.; Ashbaugh, H. S. *J. Chem. Phys.* **2003**, *119*, 2702.
(14) Asthagiri, D.; Pratt, L. R.; Paulaitis, M. E.; Rempe, S. B. *J. Am. Chem. Soc.* **2004**, *126*, 1285.
(15) Pliego, J. R., Jr.; Riveros, J. M. *Chem.−Eur. J.* **2002**, *8*, 1945.
(16) Westphal, E.; Pliego, J. R., Jr. *J. Chem. Phys.* **2005**, *123*, 074508.
(17) Zhan, C.-G.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 11534.
(18) Zhan, C.-G.; Dixon, D. A. *J. Phys. Chem. A* **2002**, *106*, 9737.
(19) Zhan, C.-G.; Dixon, D. A. *J. Phys. Chem. A* **2004**, *108*, 2020.
(20) Wang, B.; Raha, K.; Liao, N.; Peters, M. B.; Kim, H.; Westerhoff, L. M.; Wollacott, A. M.; van der Vaart, A.; Gogonea, V.; Suarez, D.; Dixon, S. L.; Vincent, J. J.; Brothers, E. N.; Merz K. M., Jr. 2007, DivCon.
(21) Cances, M. T.; Mennucci, V.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032.
(22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. *Gaussian 03*, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(23) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *AMBER 9*, University of California, San Francisco, 2006.
(24) Kelly, C.; Cramer, C.; Truhlar, D. *J. Phys. Chem. B* **2006**, *110*, 16066.
(25) Kelly, C.; Cramer, C.; Truhlar, D. *J. Chem. Theory Comput.* **2005**, *1*, 1133.
(26) Pliego, J. R., Jr.; Riveros, J. M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622.
(27) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(28) Kelly, C.; Cramer, C.; Truhlar, D. *J. Phys. Chem. A* **2006**, *110*, 2493.