Infinite Dilution Activity Coefficients from *Ab Initio* Solvation Calculations

Shiang-Tai Lin and Stanley I. Sandler
Center for Molecular and Engineering Thermodynamics, Dept. of Chemical Engineering, University of Delaware,
Newark, DE 19716

A Group Contribution Solvation (GCS) model was developed to calculate infinite dilution activity coefficients (γ^{∞}) based on modern computational chemistry. The GCS model results in an average error of 7% in γ^{∞} for the limited number of data points among water, n-hexane, acetonitrile and n-octanol, whereas the errors are 47% and 52% with the UNIFAC model and the modified UNIFAC model, respectively. GCS was also used to calculate infinite dilution partition coefficients, which can be used to determine how a dilute solute partitions between two solvents. Solutes were examined in three different liquid-liquid systems: water/n-hexane, water/acetonitrile, and water/n-octanol. With GCS, the average errors are 22% (for 18 solutes), 18% (for 14 solutes) and 14% (for 15 solutes) for these solvent systems, while comparable errors are 237%, 286% and 226% with UNIFAC; and 342%, 414% and 306% with modified UNIFAC. The GCS model is a powerful new tool to predict the octanol-water partition coefficients.

Introduction

Infinite dilution activity coefficients (γ^{∞}) are of great value in both chemical and environmental engineering. If both infinite dilution activity coefficients are known for a binary system, parameters in a two-parameter activity coefficient model can be determined and then predictions of vapor-liquid equilibria over the entire composition range can be made. Also, infinite dilution activity coefficients provide important information for the design of high purity distillation and extraction processes. Moreover, knowledge of infinite dilution activity coefficients of chemical pollutants enables us to estimate how they partition in the environment (Sandler, 1996).

A large amount of experimental data on infinite dilution activity coefficients have been collected and can be found in the literature (Tiegs et al., 1986; Hait et al., 1993; Kojima et al., 1997). Although direct measurement is the most reliable way to obtain infinite dilution activity coefficients, it is costly, time-consuming and, sometimes, dangerous (for example, toxic species, and high temperatures or pressures may be involved). It is therefore desirable to have a means to predict activity coefficients at infinite dilution.

A variety of thermodynamic models have been proposed to describe the dependence of the activity coefficient on the mixture composition, molecular structure, properties of pure substances or intermolecular interactions. Examples include the MOSCED model (Thomas and Eckert, 1984), the SPACE model (Hait et al., 1993), the ASOG model (Derr and Deal, 1969), and the UNIFAC model (Fredenslund et al., 1975). MOSCED and SPACE, derived from the regular solution theory (Hildebrand et al., 1970), have the advantage of using only pure substance properties such as polarizability, polarity, acidity, and basicity. Their applicability is currently limited to nonaqueous systems; therefore, they are not considered in this work. Both ASOG and UNIFAC are group contribution models, however the latter has received more attention as it has been constantly examined and refined with new experimental data. The UNIFAC model was developed from the statistical-mechanically derived UNIQUAC model (Abrams and Prausnitz, 1975; Maurer and Praunitz, 1978). In this model molecules are considered to be a collection of functional groups, and the values of adjustable group-group interaction energy parameters have been obtained by correlating data for simple mixtures containing these groups.

Correspondence concerning this article should be addressed to S. I. Sandler.

Efforts have been made to obtain values for the molecular interaction parameters in the UNIQUAC model using various computational methods. Jónsdóttir et al. (1994, 1996, 1997) used molecular mechanics methods to find the minimum energy configurations of all possible combinations of conformers for a pair of molecules and then used Boltzmann averaging of the energies to determine the interaction parameters. Sum and Sandler (1999) applied quantum mechanical methods to calculate the pair interaction energies in a small cluster of molecules that were then used as the interaction parameters. These methods were for the calculation of the interaction energies between two molecules, and therefore can be classified as a direct method.

Here we propose an indirect approach. Using the infinite dilution activity coefficient as a bridge, we relate the two energy parameters in the UNIQUAC model to the attractive part of the solvation free energy. The electrostatic part of this free energy can be obtained from various continuum solvation models. In the method developed by Honig and coworkers (Tannor et al., 1994; Marten et al., 1996) the electrostatic interactions are determined by solving the Poisson-Boltzmann equation with atom-centered charges obtained from fitting to the quantum-mechanically calculated solute charge distribution. In the Polarizable Continuum Model (PCM) of Tomasi and his colleagues (Miertuš et al., 1981; Cossi et al., 1996) the Poisson equation is solved in terms of virtual charges appearing at the solute/solvent boundary, and the solute charge distribution is obtained from quantum mechanical calculations. Rather than solving the Poisson equation exactly, Cramer and Truhlar use the generalized Born equation to approximate the electrostatic interactions and develop the Solvation Model (SM) (Giesen et al., 1996; Chambers et al., 1996; Hawkins et al., 1998). Klamt treats the solvent as a conductor, which results in a simpler boundary condition for the Poisson equation, and establishes the Conductor-like-Screening Model (COSMO) (Klamt, 1995; Klamt et al., 1998). Despite the different accuracies and efficiencies, the quantum mechanical calculations in these models is limited to the computation of the electrostatic interactions. Recent work of Amovilli and Mennucci (1997) includes nonelectrostatic interactions, that is, dispersion and repulsion, in PCM and no empirical or fitting parameters are needed. We, therefore, adopt this last approach for the solvation calculation, and this leads to the development of our simple-to-use Group Contribution Solvation (GCS) model. In this way we take advantage of recent advances in quantum mechanical solvation calculational methods to determine the UNIQUAC parameters, and then make accurate predictions of the infinite dilution activity and partition coefficients.

Theory

Activity coefficient from solvation energy

The activity coefficient is used to describe the change in value of the chemical potential of a species in a mixture of mole fractions x from that as a pure component as follows

$$\mu_1(T,P,x) = \mu_1^0(T,P) + kT \ln x_1\gamma_1(T,P,x) \tag{1}$$

Ben-Naim (1978) related the chemical potential of a species to the work required to add one molecule into the system, and derived the following expression for chemical potential

$$\mu_1(T,P,x) = \mu_1^*(T,P,x) + kT \ln \rho_1\Lambda_1^3 \tag{2}$$

with

$$\rho_1 = \frac{N_1}{V} \quad \text{and} \quad \Lambda_1^3 = \frac{h^3}{(2\pi m_1 kT)^{3/2}}$$

where μ_1^*, called pseudochemical potential, is the chemical potential of species 1 at a fixed position in the fluid; N_1 is the number of molecules of species 1; V is the volume of the mixture; and m_1 is the mass of molecule 1. Equation 2 states that the free energy required to insert one molecule of species 1 into a system (μ_1) is the free energy of adding that molecule at a fixed point (μ_1^*) plus the energy of molecular motion and entropy that the molecule acquires ($kT \ln \rho_1\Lambda_1^3$) after it is released.

The solvation process, as defined by Ben-Naim (1987), is the process of transferring one molecule from a fixed position in an ideal gas phase to a fixed position in the fluid phase at constant temperature and pressure. The Gibbs free energy of solvation of species 1 in a fluid then is

$$\Delta g_1^{*\text{sol}}(T,P,x) = \mu_1^*(T,P,x) - \mu_1^{*,IG}(T,P) \tag{3}$$

Combining Eqs. 1 to 3, the activity coefficient of species 1 can be expressed in terms of the solvation energies of species 1 in mixture and in its pure liquid, and the ratio of the number densities as follows

$$\begin{aligned} kT \ln x_1\gamma_1(T,P,x) &= \mu_1(T,P,x) - \mu_1^0(T,P) \\ &= \mu_1^*(T,P,x) + kT\ln\rho_1\Lambda_1^3 - \mu_1^*(T,P) - kT\ln\rho_1^0\Lambda_1^3 \\ &= \mu_1^*(T,P,x) - \mu_1^*(T,P) + kT\ln\frac{\rho_1}{\rho_1^0} \\ &= \left[\mu_1^*(T,P,x) - \mu_1^{*,IG}(T,P)\right] - \left[\mu_1^*(T,P) - \mu_1^{*,IG}(T,P)\right] \\ &\quad + kT\ln\frac{\rho_1}{\rho_1^0} = \Delta g_1^{*,\text{sol}}(T,P,x) - \Delta g_1^{*\text{sol},0}(T,P) \\ &\quad + kT\ln\frac{\rho_1}{\rho_1^0} \end{aligned} \tag{4}$$

Considering the limiting case that $x_1 \to 0$ in a binary system, Eq. 4 becomes

$$\begin{aligned} kT\ln\gamma_1^\infty &= \Delta g_1^{*\text{sol},\infty} - \Delta g_1^{*\text{sol},0} + kT\ln\frac{\rho_2^0}{\rho_1^0} \\ &= \Delta g_{1/2}^{*\text{sol}} - \Delta g_{1/1}^{*\text{sol}} + kT\ln\frac{\rho_2^0}{\rho_1^0} \end{aligned} \tag{5}$$

where we have used the relation $\rho_1(x_1 \to 0) = x_1\rho_2^0$ and subscript $1/j$ to represent the property of solute 1 in pure solvent

j. Equation 5 is based on a single molecule; for convenience, we change to a molar basis by multiplying by Avogadro's number.

$$RT \ln \gamma_1^{\infty} = \Delta G_{1/2}^{*\,\mathrm{sol}} - \Delta G_{1/1}^{*\,\mathrm{sol}} + RT \ln \frac{\rho_2^0}{\rho_1^0} \tag{6}$$

This is the fundamental equation that allows the calculation of γ^{∞} from the solvation energies.

Solvation energy

The solvation process can be decomposed into two steps: first, the charges on the solute are turned off and the remaining hard particle is inserted into the solvent. This is equivalent to creating a cavity of the necessary volume in the solvent. Secondly, after the solute is embedded in the solvent, charges are turned on. In the first step, the free energy change is a result of the repulsive interactions between the solute and the solvent and is called the cavity formation free energy $\Delta G^{*\,\mathrm{cav}}$. In the second step, the free energy change is a result of attractive interactions and is referred to as the charging free energy $\Delta G^{*\,\mathrm{chg}}$. The rotational and vibrational motions of the solute may also change when it is transferred from an ideal gas to a solution. However, the corresponding free energy change usually has a negligible effect on the total solvation energy, except for large molecules. Therefore, the change of molecular motions upon solvation has been neglected in our study and we have

$$\Delta G^{*\,\mathrm{sol}} = \Delta G^{*\,\mathrm{cav}} + \Delta G^{*\,\mathrm{chg}} \tag{7}$$

In this work the charging free energy is evaluated by the EHCD method (effective Hamiltonian methods that use a continuum distribution of the solvent) (Tomasi and Persico, 1994). In such continuum models, the solvent is considered to be a dielectric medium without any internal structure. The total Hamiltonian of the system $\hat{H}$ is the sum of the Hamiltonian of the solute in vacuum $\hat{H}_1^{(0)}$, and a perturbation term $\hat{V}_{\mathrm{int}}$ accounting for the interaction between the solute and the continuum solvent

$$\hat{H} = \hat{H}_1^{(0)} + \hat{V}_{\mathrm{int}} \tag{8}$$

Solving the following Schrödinger equations

$$\hat{H}_1^{(0)} \Psi_1^{(0)} = E_1^{(0)} \Psi_1^{(0)} \tag{9}$$

$$\hat{H} \Psi = E \Psi \tag{10}$$

we obtain the energies and wave functions for both the solute in vacuum ($E_1^{(0)}$ and $\Psi_1^{(0)}$) and in solution (E and Ψ). The interaction energy ΔE_{int} between the solute and the solvent is thus

$$\Delta E_{\mathrm{int}} = E - E_1^{(0)} = \langle \Psi | \hat{H} | \Psi \rangle - \langle \Psi_1^{(0)} | \hat{H}_1^{(0)} | \Psi_1^{(0)} \rangle \tag{11}$$

where the definition of the notation bra ($\langle|$) and ket ($|\rangle$) can be found in quantum chemistry textbooks (Szabo and Ostlund, 1989).

Using the concept of the charging process (Langlet et al., 1988; Kirkwood, 1968), we can convert ΔE_{int} to a free energy change as follows. Let λ be a charging parameter with $\lambda = 0$ corresponding to the charges on the solute turned off, and $\lambda = 1$ corresponding to the charges fully turned on. By integrating the interaction energy from $\lambda = 0$ to $\lambda = 1$, the charging free energy is obtained

$$\Delta G_{1/2}^{*\,\mathrm{chg}} = \int_{\lambda=0}^{\lambda=1} \Delta E_{\mathrm{int}}(\lambda)\, d\lambda \tag{12}$$

To carry out the integration in Eq. 12 we need to know the interaction potential $\hat{V}_{\mathrm{int}}$. Formulations for the four components of $\hat{V}_{\mathrm{int}}$ (electrostatic, induction, dispersion and repulsion) are given by Tomasi and Persico (1994), Amovilli (1994), and Amovilli and Mennucci (1997); and will not be reproduced here.

For the cavity formation free energy, we have developed a new model based on UNIQUAC to take into account the molecular shape effect and solvent-solvent interactions. The derivation is tedious but straightforward and is presented in the Appendix. Briefly, we combined Eq. 6 with the expression from UNIQUAC, and by varying the charges on the solute and the solvent, we are able to separate the charging free energy from the other two free energy contributions in the UNIQUAC model. The results we obtained are

$$\frac{\Delta G_{1/2}^{*\,\mathrm{chg}}}{RT} = q_1(\tau_2 - \tau_{12} - \ln \tau_{21} + \ln \tau_1) \tag{13}$$

$$\frac{\Delta G_{1/2}^{*\,\mathrm{cav}}}{RT} = \ln \gamma_{1/2}^{\infty}(\mathrm{comb}) + q_1(1 - \tau_2 - \ln \tau_1) - \ln \frac{\rho_2^0 kT}{P_1^{\mathrm{vap}}} \tag{14}$$

where

$$\tau_{ij} = \exp\left(-\frac{u_{ij} - u_{jj}}{RT}\right); \qquad \tau_i = \exp\left(\frac{u_{ii}}{RT}\right);$$

$$RT \ln \gamma_{1/2}^{\infty}(\mathrm{comb}) = \ln \frac{r_1}{r_2} + \frac{z}{2} q_1 \ln \frac{q_1 r_2}{q_2 r_1} + \frac{z}{2}\left(\frac{r_1}{r_2} q_2 - q_1\right) + \left(1 - \frac{r_1}{r_2}\right)$$

is the combinatorial part of $RT \ln \gamma^{\infty}$ in the UNIQUAC model; r_i and q_i are the UNIQUAC volume and surface area parameters for species i; z is the coordination number, usually taken to be 10; u_{ij} is the average interaction energy between species i and species j; P_i^{vap} is the vapor pressure of species i at temperature T. Further, τ_i can be calculated from the following equation

$$\frac{\Delta G_{1/1}^{*\,\mathrm{chg}}}{RT} = q_1(\tau_1 - 1 + \ln \tau_1) \tag{15}$$

The two UNIQUAC energy parameters τ_{12} and τ_{21} are implicit variables in Eq. 13. Together with a similar equation for $\Delta G_{2/1}^{*\,\mathrm{chg}}$ (see Eq. A27 in the Appendix), we can solve these two equations for τ_{12} and τ_{21}.

Combining Eqs. 6, 7, 13, and 14, we have the new Group Contribution Solvation (GCS) model for the calculation of γ^∞

$$RT \ln \gamma_{1/2}^{\infty} = RT \ln \gamma_{1/2}^{\infty}(\text{comb}) + RTq_1(\tau_1 - \tau_2) + \left(\Delta G_{1/2}^{*\text{chg}} - \Delta G_{1/1}^{*\text{chg}}\right) \quad (16)$$

The infinite dilution partition coefficient, $\gamma_{1/2}^{\infty}/\gamma_{1/3}^{\infty}$, can be calculated from the following equation

$$RT \ln \frac{\gamma_{1/2}^{\infty}}{\gamma_{1/3}^{\infty}} = RT \ln \frac{\gamma_{1/2}^{\infty}(\text{comb})}{\gamma_{1/3}^{\infty}(\text{comb})} + RTq_1(\tau_3 - \tau_2) + \left(\Delta G_{1/2}^{*\text{chg}} - \Delta G_{1/3}^{*\text{chg}}\right) \quad (17)$$

Computational procedure

The quantum chemistry package General Atomic and Molecular Electronic Structure System (GAMESS) (Schmidt et al., 1993) was used to calculate the charging free energy. The procedure of obtaining necessary parameters for calculating γ^∞ is described below.

(1) The equilibrium structure of the solute in the vacuum is obtained by performing a geometry optimization using the Hartree Fock (HF) method with DZPsp(df) basis set (Double Zeta basis set with polarization functions (DZP) (Hunzinaga, 1965; Dunning, 1970) plus one set of diffuse and polarization functions with exponents of one-third of the most diffuse DZP set). The same geometry is used in subsequent solvation calculations without further optimization. The DZPsp(df) basis set was chosen, because it is conventionally used in this type of calculations (Amovilli and McWeeny, 1990; Amovilli, 1994; Amovilli and Mennucci, 1997; Mennucci et al., 1998) and our preliminary studies showed that its performance on solvation energy is comparable to a much larger basis set, such as 6-31 + + G(3d,3p) or 6-311 + + G(3d,3p).

(2) The solvation calculation is then carried out using the HF method at the same DZPsp(df) level. PCM (Miertuš et al., 1981; Cossi et al., 1996) was used for the electrostatic and induction contributions, and the method of Amovilli and Mennucci (1997) was used for the dispersion and repulsion contributions. The shape of the solute molecule embedded in the homogeneous dielectric solvent was described by fused spheres centered on the solute nuclei. The radii of the spheres used were the van der Waals radii of the atoms multiplied by a scale factor α, listed in Table 1 and discussed in detail later. The physical properties of the solvent (dielectric constant, ionization potential, refractive index, and density) were obtained from the CRC Handbook of Chemistry and Physics (Lide, 1995).

(3) The molecular structure parameters r and q were obtained using the method that Abrams and Prausnitz (1975) proposed for the UNIQUAC model; however, we used their values for water ($r_w = 0.92$, $q_w = 1.40$), which were determined from data regression. The volume and surface area of a molecule were calculated using the optimized geometry from step 1 and the van der Waals radius for each atom. These values were then normalized using the van der Waals volume and surface area of a standard segment, 15.17 cm^3/mol and 2.5×10^9 cm^2/mol, respectively, to yield

$$r = \frac{\text{van der Waals volume}}{15.17\ \text{cm}^3/\text{mol}}$$

$$q = \frac{\text{van der Waals surface area}}{2.5\times10^9\ \text{cm}^2/\text{mol}}$$

In continuum models, such as PCM, a scale factor is used to account for the fact that the dielectric of the first solvation layer around a molecule or a group is different from that in the bulk solution (Tomasi and Persico, 1994). Hence, instead of directly using van der Waals radii, the radii of the fused spheres are multiplied by the scale factors in EHCD calculations. Previously, it was found that various contributions to the solvation energy require different scale factors. For the electrostatic and induction contributions, $\alpha = 1.2$ was recommended (Miertuš et al., 1981; Tomasi and Persico, 1994; Cossi et al., 1996); and $\alpha = 1.6$ was suggested for the dispersion and repulsion contributions (Amovilli, 1994; Amovilli and Mennucci, 1997). Amovilli and Mennucci (1997) adapted their method so that the same scale factor $\alpha = 1.2$ can be used for all energy contributions.

Conventionally, the same scale factor was assigned to all types of atoms regardless of their chemical environment, for example, carbonyl oxygen, hydroxyl oxygen, carbonyl carbon and hydroxyl hydrogen all had the same value of α. However, this led to acceptable results for only a very small number of substances. Since the scale factor is used to account for the dielectric behavior of the solvent in the vicinity of the solute, its value depends on how the solvent molecules distribute around the solute, which should be dependent on both the solvent and functional group of the solute. For example, we do not expect the distribution of water molecules around the methylene group of ethanol to be the same as the distribution around the hydroxyl group in that molecule. Therefore, instead of the assumption that the scale factors are the same for all atoms, we adopt the idea of group contributions to the solvation calculation and find the optimum value of α for each functional group interaction that results in the best prediction of infinite dilution activity coefficients. The same value of α was used for all atoms belonging to that functional group. Thus, as in UNIFAC, there is a group contribution part of the model.

The optimization procedure of the group scale factors was as follows:

(1) First, we considered the ratio of the infinite dilution activity coefficients of alkanes in water and in *n*-hexane. The scale factors for the methyl and methylene groups in the two solvents were optimized to produce minimum errors in the ratios of γ^∞ for the solutes C_3H_8, C_4H_{10}, C_5H_{12} and C_6H_{14}.

(2) The scale factors of all functional groups in *n*-hexane were then assumed to be the same as found in step (1) based on the rationale that *n*-hexane does not have strong local interactions, such as hydrogen bonding, with most functional groups and its distribution is not much different around different functional groups.

(3) To find the other scale factors, such as for the OH group in water, we take C_2H_5OH as the solute and calculate the ratio of its infinite dilution activity coefficients in water and *n*-hexane. We then optimized the scale factor for the

Table 1. Optimized Values of the Group Scale Factor α for Functional Groups in Water, Acetonitrile, n-Octanol, and n-Hexane

Functional Group	α Value in Water	α Value in Acetonitrile	α Value in n-Octanol	α Value in n-Hexane
CH_3, CH_2, CH, C	1.42	1.24	1.14	1.10
Cyclic CH_2	1.35	1.21	1.14	1.10
ACH*	1.33	1.24	1.22	1.10
H_2O	1.11	1.19		1.10
OH	0.88	1.15	1.15	1.10
CN	1.10	1.23	1.38	1.10
NO_2	1.12		1.33	1.10
$Cl(CCl_4)$	1.26	1.18	1.16	1.10
$Cl(C_nH_{2n+1}CCl_3)$	1.23	1.18	1.20	1.10
$Cl(C_nH_{2n+1}CHCl_2)$	1.20	1.18	1.24	1.10
$Cl(C_nH_{2n+1}CH_2Cl)$	1.17	1.18	1.28	1.10

*ACH represents the CH group in an aromatic ring.

Examples: C_2H_5CN in water: $\alpha = 1.42$ for each atom in group CH_3 and CH_2; $\alpha = 1.10$ for each atom in the CN group. CH_3CHCl_2 in water: $\alpha = 1.42$ for all C and H atoms; $\alpha = 1.20$ for the two Cl atoms.

OH group in water until the error in the predicted activity coefficient ratio was minimum.

The optimized values of α for various functional groups in n-hexane, water, acetonitrile and in n-octanol, the only solvents we have so far considered, are listed in Table 1. It is noteworthy that the solvation energy is very sensitive to α, especially in low dielectric constant solvents. For example, varying 0.01 units of α for a CH_2 group would result in a free energy change of approximately 0.02 kcal/mol in water (dielectric constant = 78.39) whereas the change is about 0.5 kcal/mol in hexane (dielectric constant = 1.89).

Results

A total of eleven activity coefficients at infinite dilution are calculated using the GCS model and compared with predictions from UNIFAC and modified UNIFAC (Gmehling et al., 1993) in Table 2. No experimental data (Tiegs et al., 1986; Hait et al., 1993; Kojima et al., 1997) for the infinite dilution activity coefficients of n-octanol in acetonitrile and n-octanol in n-hexane could be found in the literature. The absolute average percentage deviation (AA%D) for 9 data points from GCS is 7%, which is much smaller than that from UNIFAC 47%, or modified UNIFAC 52%. The infinite dilution activity coefficient of n-octanol in water ($\gamma^{\infty}_{O/W}$), n-octanol in acetonitrile ($\gamma^{\infty}_{O/A}$) and n-cotanol in hexane ($\gamma^{\infty}_{O/H}$) were not obtained from EHCD calculations, but rather from a linear extrapolation of the charging free energies. We found that the charging free energy is proportional to the surface area parameter q for a homologous series. For example, following relations were found for alcohols in water, n-octanol, hexane and in acetonitrile

$$\Delta G^{chg}_{Alcohol/W}\ (kcal/mol) = -2.636 \cdot q_{Alcohol} - 8.190 \quad (18)$$

$$\Delta G^{chg}_{Alcohol/O}\ (kcal/mol) = -4.047 \cdot q_{Alcohol} - 5.361 \quad (19)$$

$$\Delta G^{chg}_{Alcohol/H}\ (kcal/mol) = -4.249 \cdot q_{Alcohol} - 2.750 \quad (20)$$

$$\Delta G^{chg}_{Alcohol/A}\ (kcal/mol) = -3.869 \cdot q_{Alcohol} - 5.292 \quad (21)$$

all with a correlation coefficient of essentially 1.00. Equations 18 and 19 were determined from the results of charging free energy calculations for CH_3OH, C_2H_5OH, n-C_3H_7OH and n-C_4H_9OH; Eq. 20 for C_2H_5OH, n-C_3H_7OH and n-C_4H_9OH; and Eq. 21 for CH_3OH and C_2H_5OH.

The advantage of using these simple linear equations between the charging free energy and the surface area parameter of the solute for $\gamma^{\infty}_{O/W}$ is to extend our calculational method to other, larger systems without computational resource problems. The computational time of solvation calculations is mainly dependent on the number of basis functions used for the solute. In Figure 1 we plot the CPU time vs. the number of basis functions used. The data points correspond to the calculations of CH_4, C_2H_6, C_3H_8, C_4H_{10}, C_5H_{12} and

Table 2. Comparison of Predicted Infinite Dilution Activity Coefficients from GCS, UNIFAC, and Modified UNIFAC at 298.15 K

	GCS		UNIFAC		Modified UNIFAC		Experiment
Solute/Solvent*	$\gamma^{\infty}_{1/2}$	%D**	$\gamma^{\infty}_{1/2}$	%D	$\gamma^{\infty}_{1/2}$	%D	$\gamma^{\infty}_{1/2}$
O/A	$1.93\times10^{0\ddagger}$		5.88×10^{0}		8.40×10^{0}		N/A
H/O	3.01×10^{0}	7	1.95×10^{0}	−31	2.54×10^{0}	−10	2.81×10^{0}
A/O	8.38×10^{0}	12	2.79×10^{0}	−63	1.99×10^{0}	−73	7.48×10^{0}
A/W	1.17×10^{1}	6	1.32×10^{1}	19	1.19×10^{1}	7	1.11×10^{1}
O/H	$1.37\times10^{1\ddagger}$		9.40×10^{0}		2.65×10^{1}		N/A
W/A	1.23×10^{1}	9	8.40×10^{0}	−26	8.46×10^{0}	−25	1.13×10^{1}
H/A	2.50×10^{1}	−2	2.41×10^{1}	−6	2.60×10^{1}	2	2.55×10^{1}
A/H	2.99×10^{1}	8	4.76×10^{-1}	−98	3.45×10^{-1}	−99	2.76×10^{1}
W/H	1.63×10^{3}	−1	1.40×10^{3}	−15	1.36×10^{2}	−92	1.65×10^{3}
O/W	$1.01\times10^{4\ddagger}$	−13	3.16×10^{3}	−73	4.08×10^{3}	−65	1.16×10^{4}
H/W	3.97×10^{5}	−1	1.06×10^{4}	−97	6.62×10^{3}	−98	4.00×10^{5}
AA%D***		7		47		52	

*A: acetonitrile, H: n-hexane, O: n-octanol, W: water.

**%D = percentage deviation, that is, $(\gamma^{\infty}_{predicted} - \gamma^{\infty}_{experiment}/\gamma^{\infty}_{experiment})\times 100$.

***AA%D = Absolute average percentage deviation.

‡Charging free energies calculated from Eqs. 18 to 21.

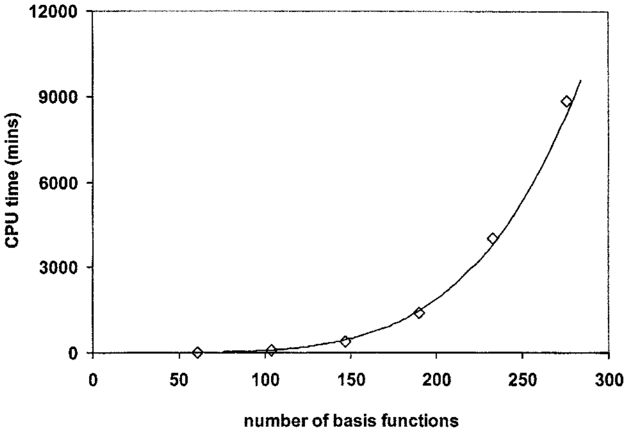

Figure 1. CPU time vs. number of basis functions used in solvation calculation.

Open diamonds are calculations of CH_4, C_2H_6, C_3H_8, C_4H_{10}, C_5H_{12}, and C_6H_{14} in n-octanol on a 200 MHz R10000 processor of Silicon Graphics Power Challenge. The solid curve is calculated from Eq. 22.

C_6H_{14} in n-octanol using one 200 MHz R10000 processor of Silicon Graphics Power Challenge. The dependence of computational time on the number of basis functions is approximately

$$\text{CPU time (mins)} = 3.74 \cdot 10^{-8} \times (\text{number of basis functions})^{4.65} \quad (22)$$

Although the exact CPU time is also affected by the scale factor α, solvent properties (dielectric constant, ionization potential, and so on) and the type of solute, Eq. 22 gives a rough estimate of the time needed. Accordingly, it would take approximately 28 days to complete the solvation energy calculation for n-octanol (387 basis functions using DZPspd basis set). This intensive computational demand is easily avoided by using Eqs. 18 to 21, which were determined from high level calculations for smaller molcules in each homologous series. As will be seen later in the discussion of the octanol-water partition coefficient, these simple equations provide a very efficient and accurate means of predicting the infinite dilution behavior of large molecules.

At present, the GCS model cannot provide predictions of the infinite dilution activity coefficient for systems not considered here because of the limited α-database. To calculate γ^∞ of species 1 in solvent 2, we need the information of α for all the functional groups in both liquid 1 and liquid 2. For example, to predict the infinite dilution activity coefficient of 1-butanol in water, we need these scale factors for CH_3, CH_2 and OH in both water and 1-butanol. As we have not yet computed the α values for the functional groups in 1-butanol, we are not able to calculate γ^∞ of 1-butanol in water. We are now investigating methods of estimating values for the scale factors α of functional groups in various solvents without doing extensive quantum chemistry calculations.

Despite the limited ability of predicting infinite dilution activity coefficients, GCS can accurately predict the ratio of γ^∞ in two solvents, that is, the infinite dilution partition coefficient. These partition coefficients are of great value in industry, particularly in distillation and extraction processes, because they can be used to determine how a dilute solute partitions between solvents. We have considered three dual solvent systems: water/n-hexane, water/acetonitrile and water/n-octanol. The results for the ratio of infinite dilution activity coefficients in each solvent pair are presented in Tables 3 to 5, Figures 2 to 4, and compared with the predictions from the UNIFAC and modified UNIFAC models. It can be

Table 3. Infinite Dilution Partition Coefficient in Water/*n*-Hexane at 298.15 K

	GCS		UNIFAC		Modified UNIFAC		Experiment
Solute	$\gamma^\infty_{1/W}/\gamma^\infty_{1/H}$	dev*	$\gamma^\infty_{1/W}/\gamma^\infty_{1/H}$	dev	$\gamma^\infty_{1/W}/\gamma^\infty_{1/H}$	dev	$\gamma^\infty_{1/W}/\gamma^\infty_{1/H}$
C_3H_8	2.59×10^{3}	−0.29	3.04×10^{2}	−1.56	1.65×10^{2}	−1.92	4.23×10^{3}
C_4H_{10}	1.51×10^{4}	−0.22	9.94×10^{2}	−1.83	5.62×10^{2}	−2.17	2.19×10^{4}
C_5H_{12}	8.76×10^{4}	−0.11	3.25×10^{3}	−2.07	1.92×10^{3}	−2.38	1.06×10^{5}
C_6H_{14}	3.97×10^{5}	0.00	1.06×10^{4}	−2.15	6.62×10^{4}	−2.43	4.00×10^{5}
C_6H_{12}	8.16×10^{4}	0.05	5.98×10^{3}	−1.49	2.48×10^{3}	−2.02	7.46×10^{4}
H_2O	6.13×10^{-4}	0.01	7.13×10^{-4}	0.10	7.36×10^{-3}	1.48	6.06×10^{-4}
C_2H_5OH	5.89×10^{-2}	−0.11	2.75×10^{-1}	0.80	9.00×10^{-2}	0.14	7.08×10^{-2}
C_3H_7OH	3.43×10^{-1}	0.04	9.00×10^{-1}	0.61	3.08×10^{-1}	−0.02	3.21×10^{-1}
C_4H_9OH	1.38×10^{0}	−0.02	2.97×10^{0}	0.43	1.06×10^{0}	−0.18	1.43×10^{0}
CH_3NO_2	7.20×10^{-1}	−0.36	1.03×10^{0}	−0.15	1.36×10^{0}	0.02	1.32×10^{0}
$C_2H_5NO_2$	4.37×10^{0}	−0.01	7.48×10^{0}	0.31	7.89×10^{0}	0.34	4.45×10^{0}
$C_3H_7NO_2$	2.90×10^{1}	0.18	2.45×10^{1}	0.08	2.71×10^{1}	0.15	2.12×10^{1}
CH_3CN	3.92×10^{-1}	−0.01	4.76×10^{-1}	0.10	3.45×10^{-1}	−0.09	4.02×10^{-1}
C_2H_5CN	2.55×10^{0}	0.17	4.10×10^{0}	0.46	3.11×10^{0}	0.29	1.90×10^{0}
CCl_4	7.86×10^{3}	−0.06	9.32×10^{3}	0.04	1.71×10^{3}	−0.96	8.66×10^{3}
$CHCl_3$	6.09×10^{2}	0.07	5.80×10^{2}	0.04	4.13×10^{2}	−0.16	5.42×10^{2}
C_4H_9Cl	4.82×10^{3}	−0.02	1.70×10^{3}	−0.63	1.05×10^{3}	−0.92	4.96×10^{3}
CH_2ClCH_2Cl	4.22×10^{2}	0.42	2.73×10^{2}	0.16	1.67×10^{2}	−0.13	2.07×10^{2}
AAD**		0.12		0.72		0.88	

*dev = $RT \ln(\gamma^\infty_{1/W}/\gamma^\infty_{1/H})^{\text{calc}} - RT \ln(\gamma^\infty_{1/W}/\gamma^\infty_{1/H})^{\text{expt}}$ (kcal/mol).
**AAD = absolute average deviation.

Table 4. Infinite Dilution Partition Coefficient in Water/Acetonitrile at 298.15 K

	GCS		UNIFAC		Modified UNIFAC		Experiment
Solute	$\gamma^{\infty}_{1/W}/\gamma^{\infty}_{1/A}$	dev	$\gamma^{\infty}_{1/W}/\gamma^{\infty}_{1/A}$	dev	$\gamma^{\infty}_{1/W}/\gamma^{\infty}_{1/A}$	dev	$\gamma^{\infty}_{1/W}/\gamma^{\infty}_{1/A}$
C_4H_{10}	1.17×10^{3}	-0.14	7.88×10^{1}	-1.74	3.44×10^{1}	-2.23	1.48×10^{3}
C_5H_{12}	4.64×10^{3}	-0.04	1.87×10^{2}	-1.95	9.36×10^{1}	-2.35	4.98×10^{3}
C_6H_{14}	1.59×10^{4}	0.01	4.42×10^{2}	-2.11	2.54×10^{2}	-2.44	1.57×10^{4}
C_5H_{10}	1.79×10^{3}	0.07	1.51×10^{2}	-1.39	1.42×10^{2}	-1.43	1.59×10^{3}
C_6H_{12}	6.59×10^{3}	0.26	3.58×10^{2}	-1.47	3.85×10^{2}	-1.43	4.26×10^{3}
C_6H_6	8.46×10^{2}	0.05	7.79×10^{2}	0.00	7.94×10^{2}	0.01	7.79×10^{2}
H_2O	8.11×10^{-2}	-0.05	1.19×10^{-1}	0.18	1.18×10^{-1}	0.17	8.84×10^{-2}
C_2H_5OH	9.84×10^{-1}	-0.01	3.05×10^{0}	0.66	1.21×10^{1}	0.11	1.01×10^{1}
CH_3CN	1.17×10^{1}	0.03	1.32×10^{1}	0.10	1.19×10^{1}	0.04	1.11×10^{1}
CCl_4	1.98×10^{3}	0.12	2.02×10^{3}	0.13	3.78×10^{2}	-0.86	1.61×10^{3}
$CHCl_3$	7.36×10^{2}	0.02	6.06×10^{2}	-0.09	5.42×10^{2}	-0.16	7.08×10^{2}
CH_2Cl_2	2.53×10^{2}	-0.04	2.19×10^{2}	-0.12	1.63×10^{2}	-0.30	2.70×10^{2}
C_3H_7Cl	8.35×10^{2}	0.29	2.21×10^{2}	-0.50	1.12×10^{2}	-0.90	5.12×10^{2}
C_4H_9Cl	3.33×10^{3}	0.33	5.22×10^{2}	-0.77	3.05×10^{2}	-1.09	$1.92\times10^{3*}$
AAD		0.10		0.89		1.06	

*Experimental data for $\gamma^{\infty}_{C_4H_9Cl/A}$ at 348.15 K was used.

easily seen from the figures that predictions from the GCS model agree very well with experimental measurements, while the other two models show considerable scatter. To quantify the error of prediction, we compare the value of $RT\ln\gamma^{\infty}_{1/2}/\gamma^{\infty}_{1/3}$ with experimental data. Since $RT\ln\gamma^{\infty}_{1}$ is the excess Gibbs free energy of species 1 at infinite dilution,

Table 5. Infinite Dilution Partition Coefficient in Water/n-Octanol at 298.15 K and Experimental Data for K_{OW}

	GCS		UNIFAC		Modified UNIFAC		Experiment	
Solute	$\gamma^{\infty}_{1/W}/\gamma^{\infty}_{1/O}$	dev	$\gamma^{\infty}_{1/W}/\gamma^{\infty}_{1/O}$	dev	$\gamma^{\infty}_{1/W}/\gamma^{\infty}_{1/O}$	dev	$\gamma^{\infty}_{1/W}/\gamma^{\infty}_{1/O}$	$\log_{10}K_{OW}$
CH_4	8.29×10^{1}		N/A		N/A		N/A	1.09*
C_2H_6	5.86×10^{2}		8.89×10^{1}		2.81×10^{1}		N/A	1.18**
C_3H_8	2.39×10^{3}		2.49×10^{2}		8.57×10^{1}		N/A	2.36**
C_4H_{10}	9.88×10^{3}		6.96×10^{2}		2.65×10^{2}		N/A	2.89**
C_5H_{12}	3.97×10^{4}	0.02	1.95×10^{3}	-1.76	8.27×10^{2}	-2.27	3.82×10^{4}	3.39**
C_6H_{14}	1.32×10^{5}	-0.04	5.45×10^{3}	-1.93	2.60×10^{3}	-2.37	1.42×10^{5}	4.11***
C_7H_{16}	5.14×10^{5}	-0.11	1.52×10^{4}	-2.20	8.24×10^{3}	-2.56	6.21×10^{5}	4.66***
C_5H_{10}	1.61×10^{4}		1.23×10^{3}		6.25×10^{2}		N/A	3.00**
C_6H_{12}	3.73×10^{4}	0.02	3.45×10^{3}	-1.39	1.97×10^{3}	-1.39	3.63×10^{4}	3.44**
C_6H_6	1.13×10^{3}	-0.02	1.51×10^{3}	0.16	1.18×10^{3}	0.16	1.16×10^{3}	2.15**
CH_3OH	7.28×10^{-1}		1.62×10^{0}		1.43×10^{0}		N/A	-0.77**
C_2H_5OH	3.66×10^{0}	0.04	6.59×10^{0}	0.39	4.04×10^{0}	0.10	3.44×10^{0}	-0.31**
C_3H_7OH	1.49×10^{1}		1.85×10^{1}		1.26×10^{1}		N/A	0.25**
C_4H_9OH	3.98×10^{1}		5.22×10^{1}		3.96×10^{1}		N/A	0.88**
$C_5H_{11}OH$	$1.80\times10^{2\dagger}$		1.44×10^{2}		1.25×10^{2}		N/A	1.56**
$C_6H_{13}OH$	$6.88\times10^{2\dagger}$		4.04×10^{2}		3.99×10^{2}		N/A	2.03**
$C_8H_{17}OH$	$1.00\times10^{4\dagger}$		3.16×10^{3}		4.08×10^{3}		N/A	3.15**
$C_{12}H_{25}OH$	$2.14\times10^{6\dagger}$		1.93×10^{5}		4.42×10^{5}		N/A	5.13**
CH_3NO_2	4.01×10^{0}	0.05	4.76×10^{0}	0.15	4.06×10^{0}	0.06	3.68×10^{0}	-0.34***
$C_2H_5NO_2$	1.28×10^{1}		2.27×10^{1}		1.51×10^{1}		N/A	0.18***
$C_3H_7NO_2$	3.61×10^{1}		6.36×10^{1}		4.75×10^{1}		N/A	0.87***
CH_3CN	1.40×10^{0}	-0.04	4.71×10^{0}	0.68	1.99×10^{0}	0.17	1.48×10^{0}	-0.34*
C_2H_5CN	7.10×10^{0}	0.13	2.28×10^{1}	0.82	1.17×10^{1}	0.42	5.74×10^{0}	0.16**
CCl_4	7.11×10^{3}	0.10	6.95×10^{3}	0.09	1.24×10^{3}	-0.93	5.99×10^{3}	2.64***
$CHCl_3$	8.32×10^{2}	-0.01	7.86×10^{2}	-0.04	5.31×10^{2}	-0.27	8.40×10^{2}	1.90***
CH_3CCl_3	1.91×10^{3}	-0.26	1.95×10^{3}	-0.25	N/A	N/A	2.98×10^{3}	2.49**
CH_2Cl_2	1.84×10^{2}	0.07	2.57×10^{2}	0.26	1.10×10^{2}	-0.24	1.65×10^{2}	1.25***
CH_3CHCl_2	5.17×10^{2}	-0.02	7.33×10^{2}	0.18	3.17×10^{2}	-0.31	5.37×10^{2}	1.79**
C_3H_7Cl	1.05×10^{3}		4.07×10^{2}		2.18×10^{2}		N/A	2.04**
C_4H_9Cl	4.13×10^{3}		1.59×10^{3}		6.68×10^{2}		N/A	2.64**
CH_2ClCH_2Cl	1.53×10^{2}	-0.28	3.32×10^{2}	0.18	1.80×10^{2}	-0.18	2.44×10^{2}	1.48**
AAD		0.08		0.70		0.83		

*Data from Sangster (1997).
**Data from Suzuki and Kudo (1990).
***Data from Isnard and Lambert (1989).
†Charging free energies calculated from Eqs. 18 and 19.

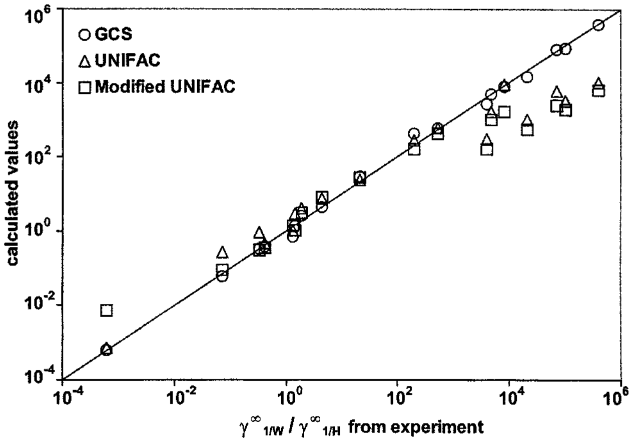

Figure 2. Prediction of the infinite dilution partition coefficient in water/n-hexane ($\gamma_{1/W}^\infty/\gamma_{1/H}^\infty$) at 298.15 K.

$RT \ln \gamma_{1/2}^\infty/\gamma_{1/3}^\infty$ directly reflects the accuracy of the solvation energy calculation. Absolute average deviations (AAD) for the three models are shown in the last lines of the tables. For the water/n-hexane system, the absolute average deviations are 0.12, 0.72 and 0.89 kcal/mol for the GCS, UNIFAC and modified UNIFAC models, respectively, corresponding to errors of 22%, 237% and 349% in the infinite dilution partition coefficient. For the water/acetonitrile system, the AADs are 0.10, 0.80 and 0.97 kcal/mol from the three models, corresponding to errors of 18%, 286% and 414% in γ^∞. For the water/n-octanol system, the three models give AADs of 0.08, 0.70 and 0.83 kcal/mol, respectively, corresponding to 14%, 226%, and 306% error in γ^∞. Generally, GCS gives reliable infinite dilution partition coefficients over a wide range (10^{-4} to 10^6), whereas UNIFAC and modified UNIFAC fail for large values. Overall, the average error of the excess Gibbs free energy from the GCS model developed here is smaller than errors from the other two models by at least 0.5 kcal/mol, and the average error of the ratio is smaller by an order of magnitude.

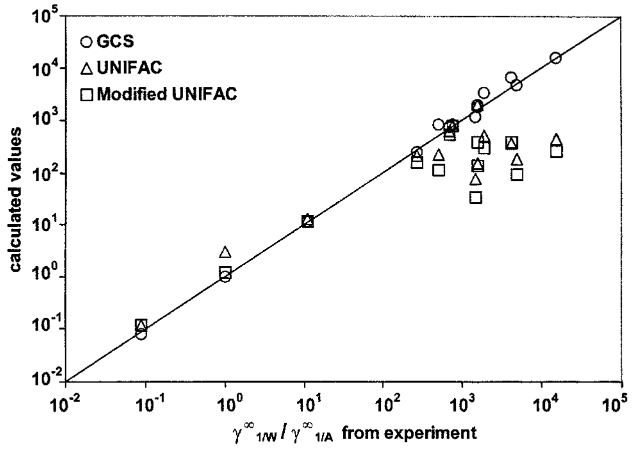

Figure 3. Prediction of the infinite dilution partition coefficient in water/acetonitrile ($\gamma_{1/W}^\infty/\gamma_{1/A}^\infty$) at 298.15 K.

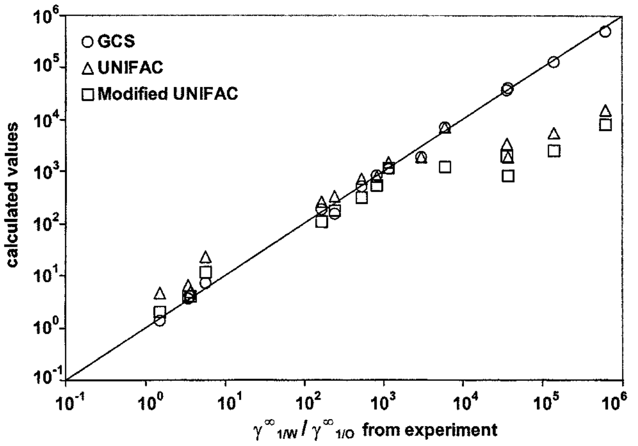

Figure 4. Prediction of the infinite dilution partition coefficient in water/n-octanol ($\gamma_{1/W}^\infty/\gamma_{1/O}^\infty$) at 298.15 K.

Another important application of the infinite dilution partition coefficient is to determine the distribution of chemical pollutants in the environment, which is largely determined by the octanol-water partition coefficient K_{OW}. In particular, K_{OW} is known to give reasonable estimates of how a chemical partitions between water and the lipids of biota, and also is used to estimate the soil-water and sediment-water partition coefficients (Sandler, 1996). Tse and Sandler (1994) showed that K_{OW} is strongly correlated with the ratio of infinite dilution activity coefficients in water and n-octanol. We used GCS to predict the infinite dilution partition coefficient in the water/n-octanol system for 31 solutes in Table 5 and then found the correlation to the octanol-water partition coefficient to be (shown in Figure 5)

$$\log_{10} K_{OW} = -0.71 + 0.92 \log_{10} \frac{\gamma_{1/W}^\infty}{\gamma_{1/O}^\infty} \tag{23}$$

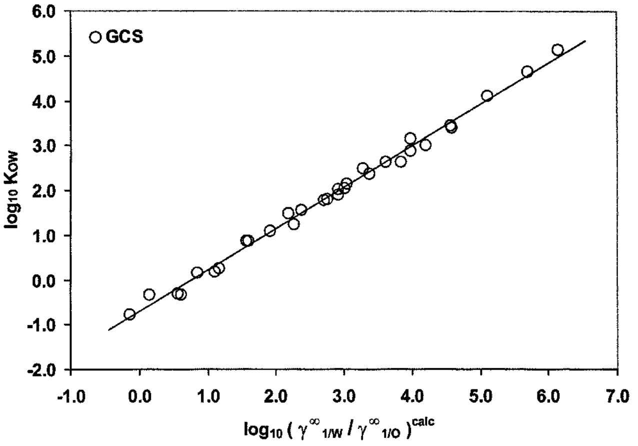

Figure 5. Correlation of the logarithm of the predicted activity coefficient ratio in water and in n-octanol to the literature values of the octanol-water partition coefficient.

The line is the correlation $\log_{10} K_{OW} = -0.71 + 0.92 \log_{10}(\gamma_{1/W}^\infty/\gamma_{1/O}^\infty)$ with a correlation factor (R^2) of 0.99.

compared to the similar correlation $\log_{10} K_{OW} = -0.65 + 0.91 \log_{10} \gamma^{\infty}_{1/W}/\gamma^{\infty}_{1/O}$ of Tse and Sandler (1994) based upon 12 measured activity coefficient ratios. Note that calculations for n-$C_5H_{11}OH$, n-$C_6H_{13}OH$, n-$C_8H_{17}OH$ and n-$C_{12}H_{25}OH$ are done with Eqs. 18 and 19. The prediction for $\log_{10} K_{OW}$ of n-dodecanol, the largest value having been considered so far, is 5.08 which is in very good agreement with the experimental value 5.13.

Discussion

The GCS method proposed here has been shown to yield accurate values for infinite dilution activity coefficients and partition coefficients for a number of systems. The success of GCS can be attributed to two factors. First, a difference of charging free energies rather than single charging free energy is used to calculate γ^{∞}. As a result, computational errors from the use of the HF method, which neglects electron correlation contributions, and the use of an incomplete basis set, which are important in individual calculated charging free energies, are of much less importance when a difference is considered. Some researchers have tried to account for these factors by adjusting the atomic radii of the solute so that correct solvation energies are obtained (Sitkoff et al., 1994; Marten et al., 1996). However, in that approach errors of different origins are artificially made to cancel each other, that is, errors from neglecting electron correlation and an incomplete basis set cancel errors from an artificial choice of solute atomic radii. While some success has been achieved with this approach in predicting solvation energies of various solutes in water, a different set of atomic radii were then required for each solvent (Sitkoff et al., 1996). In our model, by using a difference between two charging free energies, there is a separate error cancellation between the two calculations of neglecting electron correlation, and of using an incomplete basis set. That is, errors of the same origin cancel. Moreover, the use of physically meaningful group scale factors which account for different solvent distributions around each functional groups avoids the inconsistency of using solvent-dependent atomic radii.

The second reason for the success of the GCS model results from the replacement of the Scaled Particle Theory (SPT) (Pierotti, 1976) used by others (Pfund et al., 1990; Tomasi and Persico, 1994; van Duijnen et al., 1996; Barone et al., 1997) with the more realistic cavity formation free energy derived from the UNIQUAC model. SPT was developed to estimate the free energy required to create a spherical cavity in a hard sphere solvent. Claverie (1978) later extended this model to account for nonspherical solutes. Despite all these efforts, the accuracy of the SPT is questionable. The study of Floris et al. (1997) shows that the free energy difference between SPT and Monte Carlo simulation for creating a pentane-shaped cavity in water is about 2.8 kcal/mol, corresponding to an error of 11,134% in γ^{∞}. Researchers using SPT calculations have redefined the shape of the cavity in order to obtain a more accurate solvation energy (Barone et al., 1997). However, as shown from the study of Cossi et al. (1996), the charging free energy is very sensitive to the choice of atomic radii. Refinement of the cavity shape not only changes the cavity formation contribution, but also changes the charging free energy contribution. Therefore, their limited success probably stemmed from the cancellation of errors of different origins. Rather than modifying or extending SPT, we have separated the charging free energy contribution from the UNIQUAC model and thereby developed a new model for the cavity formation free energy. The capability of UNIQUAC in correlating the phase behavior for a wide variety of compounds suggests that this may be a more accurate cavity formation free energy. This postulate is supported by the success of the Group Contribution Solvation model in describing activity coefficients for both large, for example, C_7H_{16}, and small, for example, CH_3CN, solutes in different solvents (water, n-hexane, acetonitrile and n-octanol).

Conclusions

There has been considerable effort devoted to predicting the phase behavior of mixtures. Unlike other work which attempt to directly calculate the unknown energy parameters in thermodynamic models, we use an indirect approach to achieve this goal. By equating the γ^{∞} expressions from both the UNIQUAC and solvation models, we were able to relate τ_{12} and τ_{21} in UNIQUAC to the charging free energies, which were then determined from *ab initio* solvation calculations. This approach has several advantages. First, it avoids the ambiguous meaning of the parameters in this thermodynamic model. In the derivation of the UNIQUAC model, u_{ij} is the average interaction between species i and j. However, it is not clear how to orient two molecules to get correct average interaction energy, how to average the energies of various orientations, or the effect of neighboring molecules. Secondly, the indirect approach we have used here is not model specific. We have chosen to use the UNIQUAC model because of its success with various types of mixtures. Any other liquid solution model could have been used, and relations obtained between its energy parameters and the charging free energy. Finally, the indirect approach used here can be more efficient than the direct calculation of interaction energies. In our work, the energy of only one solute molecule in a continuum is computed in each calculation, whereas at least two molecules, both the solute and the solvent, need to be included in the direct approach. Because the computational time increases with the number of molecules, or, to be exact, the total number of basis functions used, our method allows one to consider larger solute molecules.

We have shown that the Group Contribution Solvation (GCS) model is capable of accurately predicting infinite dilution activity coefficients because of the use of differences in charging free energies, and the use of a better model for the cavity formation free energy. The introduction of group scale factor α allows us to study solvation in various solvents using a coherent molecular description, and to predict the behavior of large molecules from the behavior of small molecules. These scale factors are currently obtained from fitting the computational chemistry calculations to experimental data for infinite dilution activity coefficients. While we know the scale factors are closely related to the distribution of the solvent molecules around a functional group, their dependence on the molecular properties and intermolecular interactions is not yet known, and will require more research. The present database of the group scale factors has been used to successfully predict infinite dilution partition coefficients, and we are in the process of extending this database to include additional solvents. A strong correlation between the predicted infinite dilution partition coefficients from the GCS and the octanol-water partition coefficients has also been found. This suggests the utility of the GCS model in environmental studies, as well as the more traditional applications in chemical engineering.

Acknowledgment

The authors would like to thank the National Science Foundation (CTS-9521406) and the Dept. of Energy (DE-FG02-85ER13436) for financial support of this research. Also, the assistance of Dr. C. Amovilli of Università di Pisa in quantum mechanical solvation calculations is gratefully acknowledged.

Notation

g = Gibbs free energy per molecule
G = Gibbs free energy per mole of molecules
h = Planck's constant
k = Boltzmann's constant
R = gas constant
α = group scale factor, the ratio of the atomic radii used in solvation calculation and van der Waals radii
Λ = thermal De Broglie wavelength
ρ = number density

Superscripts

0 = pure substance
(0) = in vacuum
∞ = infinite dilution
$*$ = the molecule is fixed at a certain position
cav = cavity formation
chg = charging
IG = ideal gas
sol = solvation

Subscripts

1, 2 = species 1, 2
1/1 = species 1 in pure fluid 1
1/2 = species 1 in pure fluid 2 (infinite dilution condition)
A = acetonitrile
H = n-hexane
O = n-octanol
W = water

Literature Cited

Abrams, D. S., and J. M. Prausnitz, "Statistical Thermodynamics of Liquid Mixtures: A New Expression for the Excess Gibbs Energy of Partly or Completely Miscible Systems," *AIChE J.*, **21**, 116 (1975).

Amovilli, C., and R. McWeeny, "A Matrix Partitioning Approach to the Calculation of Intermolecular Potentials. General Theory and Some Examples," *Chem. Phys.*, **140**, 343 (1990).

Amovilli, C., "Calculation of the Dispersion Energy Contribution to the Solvation Free Energy," *Chem. Phys. Lett.*, **229**, 244 (1994).

Amovilli, C., and B. Mennucci, "Self-Consistent-Field Calculation of Pauli Repulsion and Dispersion Contributions to the Solvation Free Energy in the Polarizable Continuum Model," *J. Phys. Chem. B*, **101**, 1051 (1997).

Barone, V., M. Cossi, and J. Tomasi, "Geometry Optimization of Molecular Structures in Solution by the Polarizable Continuum Model," *J. Chem. Phys.*, **107**, 3210 (1997).

Ben-Naim, A., "Standard Thermodynamics of Transfer. Uses and Misuses," *J. Phys. Chem.*, **82**, 792 (1978).

Ben-Naim, A., *Solvation Thermodynamics*, Plenum Press, New York (1987).

Chambers, C. C., G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, "Model for Aqueous Solvation Based on Class IV Atomic Charges and First Solvation Shell Effects," *J. Phys. Chem.*, **100**, 16385 (1996)

Claverie, O., *Intermolecular Interactions: from Diatomics to Biomolecules*, Wiley, Chichester, U.K. (1978).

Cossi, M., V. Barone, R. Cammi, and J. Tomasi, "*Ab Initio* Study of Solvated Molecules: A New Implementation of the Polarizable Continuum Model," *Chem. Phys. Lett.*, **255**, 327 (1996).

Derr, E., and C. H. Deal, "Analytical Solutions of Groups: Correlation of Activity Coefficients Through Structural Group Parameters: I," *Inst. Chem. Eng. Symp. Ser.*, London, **3**, 40 (1969).

Dunning, T. H., "Gaussian Basis Functions for Use in Molecular Calculations: I. Contraction of (9s5p) Atomic Basis Sets for the First-Row Atoms," *J. Chem. Phys.*, **53**, 2823 (1970).

Floris, F. M., M. Selmi, A. Tani, and J. Tomasi, "Free Energy and Entropy for Inserting Cavities in Water: Comparison of Monte Carlo Simulation and Scaled Particle Theory Results," *J. Chem. Phys.*, **107**, 6353 (1997).

Fredenslund, A., R. L. Jones, and J. M. Prausnitz, "Group-Contribution Estimation of Activity Coefficients in Nonideal Liquid Mixtures," *AIChE J.*, **21**, 1086 (1975).

Giesen, D. J., M. Z. Gu, C. J. Cramer, and D. G. Truhlar, "A Universal Organic Solvation Model," *J. Org. Chem.*, **61**, 8720 (1996).

Gmehling, J., J. Li, and M. Schiller, "A Modified UNIFAC Model. 2. Present Parameter Matrix and Results for Different Thermodynamic Properties," *Ind. Eng. Chem. Res.*, **32**, 178 (1993).

Hait, M. J., C. L. Liotta, C. A. Eckert, D. L. Bergmann, A. M. Karachewski, A. J. Dallas, D. I. Eikens, J. J. Li, P. W. Carr, R. B. Poe, and S. C. Rutan, "Space Predictor for Infinite Dilution Activity Coefficients," *Ind. Eng. Chem. Res.*, **32**, 2905 (1993).

Hawkins, G. D., C. J. Cramer, and D. G. Truhlar, "Universal Quantum Mechanical Model for Solvation Free Energies Based on Gas-Phase Geometries," *J. Phys. Chem. B*, **102**, 3257 (1998).

Hildebrand, J. H., J. M. Prausnitz, and R. L. Scott, *Regular and Related Solutions*, Van Nostrand Reinhold, New York (1970).

Hunzinaga, S., "Gaussian-Type Functions for Polyatomic Systems. I," *J. Chem. Phys.*, **42**, 1293 (1965).

Isnard, P., and S. Lambert, "Aqueous Solubility and n-Octanol/ Water Partition Coefficient Correlations," *Chemosphere*, **18**, 1837 (1989).

Jónsdóttir, S. O., K. Rasmussen, and A. Fredenslund, "UNIQUAC Parameters Determined by Molecular Mechanics," *Fluid Phase Equil.*, **100**, 121 (1994).

Jónsdóttir, S. O., R. A. Klein, and K. Rasmussen, "UNIQUAC Interaction Parameters for Alkane/Amine Systems Determined by Molecular Mechanics," *Fluid Phase Equil.*, **115**, 59 (1996).

Jónsdóttir, S. O., and R. A. Klein, "UNIQIAC Interaction Parameters for Molecules with -OH Groups on Adjacent Carbon Atoms in Aqueous Solution Determined by Molecular Mechanics—Glycols, Glycerol and Glucose," *Fluid Phase Equil.*, **132**, 117 (1997).

Kirkwood, J. G., *Theory of Liquids*, Gordon and Breach, New York (1968).

Klamt, A., "Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena," *J. Phys. Chem.*, **99**, 2224 (1995).

Klamt, A., V. Jonas, T. Bürger, and J. C. W. Lohrenz, "Refinement and Parametrization of COSMO-RS," *J. Phys. Chem. A*, **102**, 5074 (1998).

Kojima, K., S. Zhang, and T. Hiaki, "Measuring Methods of Infinite Dilution Activity Coefficients and a Database for Systems Including Water," *Fluid Phase Equil.*, **131**, 145 (1997).

Langlet, J., P. Claverie, and A. Pullman, "Improvements of the Continuum Model. 1. Application to the Calculation of the Vaporization Thermodynamic Quantities of Nonassociated Liquids," *J. Phys. Chem.*, **92**, 1617 (1988).

Lide, D. R., *CRC Handbook of Chemistry and Physics*, 76th ed., CRC Press, New York (1995, 1996).

Marten, B., K. Kim, C. Cortis, R. A. Friesner, R. B. Murphy, M. N. Ringnalda, D. Sitkoff, and B. Honig, "New Model for Calculation of Solvation Free Energies: Correction of Self-Consistent Reaction Field Continuum Dielectric Theory for Short-Range Hydrogen-Bonding Effects," *J. Phys. Chem.*, **100**, 11775 (1996).

Maurer, G., and J. M. Prausnitz, "On the Derivation and Extension of the UNIQUAC," *Fluid Phase Equil.*, **2**, 91 (1978).

Mennucci, B., C. Amovilli, and J. Tomasi, "On the Effect of Pauli Repulsion and Dispersion on Static Molecular Polarizabilities and

Hyperpolarizabilities in Solution," *Chem. Phys. Lett.*, **286**, 221 (1998).

Miertuš, S., E. Scrocco, and J. Tomasi, "Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of *Ab Initio* Molecular Potentials for the Prevision of Solvent Effects," *Chem. Phys.*, **55**, 117 (1981).

Pfund, D. M., L. L. Lee, and H. D. Cochran, "Chemical Potential Prediction in Realistic Fluid Models with Scaled Particle Theory," *Int. J. Thermophys.*, **11**, 73 (1990).

Pierotti, R. A., "A Scaled Particle Theory of Aqueous and Non Aqueous Solutions," *Chem. Rev.*, **76**, 717 (1976).

Sandler, S. I., "Infinite Dilution Activity Coefficients in Chemical, Environmental and Biochemical Engineering," *Fluid Phase Equil.*, **116**, 343 (1996).

Sangster, J., *Octanol-Water Partition Coefficients: Fundamentals and Physical Chemistry*, Wiley, New York (1997).

Schmidt, M. W., K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, "General Atomic and Molecular Electronic Structure System," *J. Comput. Chem.*, **14**, 1347 (1993).

Sitkoff, D., K. A. Sharp, and B. Honig, "Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models," *J. Phys. Chem.*, **98**, 1978 (1994).

Sitkoff, D., N. Ben-Tal, and B. Honig, "Calculation of Alkane to Water Solvation Free Energies Using Continuum Solvent Models," *J. Phys. Chem.*, **100**, 2744 (1996).

Sum, A. K., and S. I. Sandler, "A Novel Approach to Phase Equilibria Predictions Using Ab Initio Methods," *Ind. Eng. Chem. Res.*, **38**, 2849 (1999).

Suzuki, T., and Y. Kudo, "Automatic LogP Estimation Based on Combined Additive Modeling Methods," *J. Comput.-Aided Mol. Des.*, **4**, 155 (1990).

Szabo, A., and N. S. Ostlund, *Modern Quantum Chemistry—Introduction to Advanced Electronic Structure Theory*, McGraw-Hill, New York (1989).

Tannor, D. J., B. Marten, R. Murphy, R. A. Friesner, D. Sitkoff, A. Nicholls, M. Ringnalda, W. A. Goddard, III, and B. Honig, "Accurate First Principles Calculation of Molecular Charge Distributions and Solvation Energies from *Ab Initio* Quantum Mechanics and Continuum Dielectric Theory," *J. Amer. Chem. Soc.*, **116**, 11875 (1994).

Tiegs, D., J. Gmehling, A. Medina, M. Soares, J. Bastos, P. Alessi, and I. Kicic, *Activity Coefficients at Infinite Dilution*, Chemistry Data Series Vol. IX, Parts 1 and 2, DECHEMA, Frankfurt/Main, Germany (1986).

Thomas, E. R., and C. A. Eckert, "Prediction of Limiting Activity Coefficients by a Modified Separation of Cohesive Energy Density Model and UNIFAC," *Ind. Eng. Chem. Process*, **23**, 194 (1984).

Tomasi, J., and M. Persico, "Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvents," *Chem. Rev.*, **94**, 2027 (1994).

Tse, G., and S. I. Sandler, "Determination of Infinite Dilution Activity Coefficients and 1-Octanol/Water Partition Coefficients of Volatile Organic Pollutants," *J. Chem. Eng. Data*, **39**, 354 (1994).

Van Duijnen, P. Th., and A. H. de Vries, "Direct Reaction Field Force Field: A Consistent Way to Connect and Combine Quantum-Chemical and Classical Descriptions of Molecules," *Int. J. Quantum Chem.*, **60**, 1111 (1996).

Appendix: Derivation of the Cavity Formation Free Energy from UNIQUAC

In the UNIQUAC model, the excess Gibbs free energy is calculated from a combinatorial term, which accounts for size and shape differences, and a residual term, which arises from different intermolecular interactions between unlike molecules

$$\frac{G^{\text{ex}}}{RT} = \frac{G^{\text{ex}}(\text{comb})}{RT} + \frac{G^{\text{ex}}(\text{res})}{RT} \tag{A1}$$

$$\frac{G^{\text{ex}}(\text{comb})}{RT} = \sum_i x_i \ln\frac{\phi_i}{x_i} + \frac{z}{2}\sum_i x_i q_i \ln\frac{\theta_i}{\phi_i} \tag{A2}$$

$$\frac{G^{\text{ex}}(\text{res})}{RT} = -\sum_i x_i q_i \ln\left(\sum_j \theta_j \tau_{ji}\right) \tag{A3}$$

where

$$\theta_i = \text{area fraction of species } i = \frac{x_i q_i}{\sum_j x_j q_j}$$

and

$$\phi_i = \text{volume fraction of species } i = \frac{x_i r_i}{\sum_j x_j r_j}$$

The activity coefficient is obtained by taking derivative of the excess Gibbs energy with respect to the mole number of species of interest which gives

$$\ln\gamma_i = \ln\gamma_i(\text{comb}) + \ln\gamma_i(\text{res}) \tag{A4}$$

$$\ln\gamma_i(\text{comb}) = \ln\frac{\phi_i}{x_i} + \frac{z}{2} q_i \ln\frac{\theta_i}{\phi_i} + l_i - \frac{\phi_i}{x_i}\sum_j x_j l_j \tag{A5}$$

and

$$\ln\gamma_i(\text{res}) = q_i\left[1 - \ln\left(\sum_j \theta_j \tau_{ji}\right) - \sum_j \frac{\theta_j \tau_{ij}}{\sum_k \theta_k \tau_{kj}}\right] \tag{A6}$$

where $l_i = z/2(r_i - q_i) - (r_i - 1)$. For binary mixtures at infinite dilution, the activity coefficient is

$$\begin{aligned}\ln\gamma_1^\infty &= \ln\frac{r_1}{r_2} + \frac{z}{2} q_1 \ln\frac{q_1 r_2}{q_2 r_1} + \frac{z}{2}\left(\frac{r_1}{r_2} q_2 - q_1\right) + \left(1 - \frac{r_1}{r_2}\right)\\ &\quad + q_1(1 - \tau_{12} - \ln\tau_{21}) = \ln\gamma_1^\infty(\text{comb}) + q_1(1 - \tau_{12} - \ln\tau_{21})\end{aligned} \tag{A7}$$

The only unknowns in Eq. A7 are the energy parameters τ_{12} and τ_{21}. If we have the infinite dilution activity coefficients of each species in the other, we can determine these two parameters and make predictions of vapor-liquid equilibria over the whole composition range. It is possible to obtain the infinite dilution activity coefficients from the solvation calculation. From Eq. 6

$$\ln\gamma_1^\infty = \frac{\Delta G_{1/2}^{*\,\text{sol}}}{RT} - \frac{\Delta G_{1/1}^{*\,\text{sol}}}{RT} + \ln\frac{\rho_2^0}{\rho_1^0} \tag{A8}$$

We can find the relationship between τ_{12}, τ_{21} and the charging free energy of the solvation energy by equating Eqs. A7 and A8

$$\left(\frac{\Delta G^{*\,\mathrm{chg}}_{1/2}}{RT} - \frac{\Delta G^{*\,\mathrm{chg}}_{1/1}}{RT}\right) + \left(\frac{\Delta G^{*\,\mathrm{cav}}_{1/2}}{RT} - \frac{\Delta G^{*\,\mathrm{cav}}_{1/1}}{RT}\right) + \ln\frac{\rho_2^0}{\rho_1^0} = \ln\gamma_1^\infty(\mathrm{comb}) + q_1(1-\tau_{12}-\ln\tau_{21}) \tag{A9}$$

where we have used the relation $\Delta G^{*\,\mathrm{sol}} = \Delta G^{*\,\mathrm{cav}} + \Delta G^{*\,\mathrm{chg}}$.

We can separate the different contributions to the solvation energy using the following four four steps:

(1) We first turn off charges on component 1.

In the solvation model: $\Delta G^{*\,\mathrm{chg}}_{1/2} = \Delta \hat{G}^{*\,\mathrm{chg}}_{1/1} = 0$

$$\ln\gamma_1^\infty = \frac{\Delta G^{*\,\mathrm{cav}}_{1/2}}{RT} - \frac{\Delta \hat{G}^{*\,\mathrm{cav}}_{1/1}}{RT} + \ln\frac{\rho_2^0}{\rho_1^0} \tag{A10}$$

The superscript ˆ indicates that the charges on the solvent molecules are turned off, that is, a hard-particle solvent is used. (Note that $\Delta G^{*\,\mathrm{chg}}_{1/1} + \Delta G^{*\,\mathrm{cav}}_{1/1} = \Delta\hat{G}^{*\,\mathrm{cav}}_{1/1} + \Delta\hat{G}^{*\,\mathrm{cav}}_{1/1} = \Delta G^{*\,\mathrm{sol}}_{1/1}$, where $\Delta G^{*\,\mathrm{chg}}_{1/1} + \Delta G^{*\,\mathrm{cav}}_{1/1}$ corresponds to the process that only the charge on one molecule 1, which is considered as the solute, is turned off; and $\Delta\hat{G}^{*\,\mathrm{chg}}_{1/1} + \Delta\hat{G}^{*\,\mathrm{cav}}_{1/1}$ corresponds to the process in which the charges on both solute and solvent molecules are turned off.)

In UNIQUAC for this case: $u_{11} = u_{12} = u_{21} = 0$, $\tau_{21} = 1$, $\tau_{12} = \exp(u_{22}/RT) = \tau_2$

$$\ln\gamma_1^\infty = RT\ln\gamma_1^\infty(\mathrm{comb}) + q_1(1-\tau_2) \tag{A11}$$

Equating Eqs. A10 and A11, we obtain an expression for the cavity formation free energies

$$\frac{\Delta G^{*\,\mathrm{cav}}_{1/2} - \Delta\hat{G}^{*\,\mathrm{cav}}_{1/1}}{RT} + \ln\frac{\rho_2^0}{\rho_1^0} = RT\ln\gamma_1^\infty(\mathrm{comb}) + q_1(1-\tau_2) \tag{A12}$$

Subtracting Eq. A12 from Eq. A9, we obtain an equation for the charging free energies

$$\frac{\Delta G^{*\,\mathrm{chg}}_{1/2} - \Delta\hat{G}^{*\,\mathrm{chg}}_{1/1}}{RT} = q_1(\tau_2 - \tau_{12} - \ln\tau_{21}) \tag{A13}$$

(2) Now we change component 2 to component 1 in the above mixture so that

$$\tau_2 = \exp\left(\frac{u_{11}}{RT}\right) = \tau_1$$

and Eq. A12 becomes

$$\frac{\Delta G^{*\,\mathrm{cav}}_{1/1} - \Delta\hat{G}^{*\,\mathrm{cav}}_{1/1}}{RT} = q_1(1-\tau_1) \tag{A14}$$

Since $\Delta G^{*\,\mathrm{chg}}_{1/1} + \Delta G^{*\,\mathrm{cav}}_{1/1} = \Delta\hat{G}^{*\,\mathrm{chg}}_{1/1} + \Delta\hat{G}^{*\,\mathrm{cav}}_{1/1}$, with Eq. A14, we obtain the following

$$\frac{\Delta G^{*\,\mathrm{chg}}_{1/1} - \Delta\hat{G}^{*\,\mathrm{chg}}_{1/1}}{RT} = q_1(\tau_1 - 1) \tag{A15}$$

We can now eliminate the carats in Eqs. A12 and A14 by taking their difference

$$\frac{\Delta G^{*\,\mathrm{cav}}_{1/2} - \Delta G^{*\,\mathrm{cav}}_{1/1}}{RT} + \ln\frac{\rho_2^0}{\rho_1^0} = \ln\gamma_1^\infty(\mathrm{comb}) + q_1(\tau_1 - \tau_2) \tag{A16}$$

Similarly, for Eqs. A13 and A15, we have

$$\frac{\Delta G^{*\,\mathrm{chg}}_{1/2} - \Delta G^{*\,\mathrm{chg}}_{1/1}}{RT} = q_1(1 + \tau_2 - \tau_1 - \tau_{12} - \ln\tau_{21}) \tag{A17}$$

The Group Contribution Solvation model can be readily obtained from subtracting Eq. A16 from Eq. A8.

(3) We now turn off the charges on component 2. In the solvation model: $\Delta\hat{G}^{*\,\mathrm{chg}}_{1/2} = 0$

$$\ln\gamma_1^\infty = \frac{\Delta\hat{G}^{*\,\mathrm{cav}}_{1/2}}{RT} - \frac{\Delta G^{*\,\mathrm{chg}}_{1/1} - \Delta G^{*\,\mathrm{cav}}_{1/1}}{RT} + \ln\frac{\rho_2^0}{\rho_1^0} \tag{A18}$$

In the UNIQUAC model: $u_{12} = u_{21} = u_{22} = 0$, $\tau_{12} = 1$, $\tau_{21} = \exp(u_{11}/RT) = \tau_1$

$$\ln\gamma_1^\infty = \ln\gamma_1^\infty(\mathrm{comb}) + q_1\ln\tau_1 \tag{A19}$$

Equating the above two equations, we obtain

$$-\frac{\Delta G^{*\,\mathrm{chg}}_{1/1}}{RT} + \frac{\Delta\hat{G}^{*\,\mathrm{cav}}_{1/2} - \Delta G^{*\,\mathrm{cav}}_{1/1}}{RT} + \ln\frac{\rho_2^0}{\rho_1^0} = \ln\gamma_1^\infty(\mathrm{comb}) - q_1\ln\tau_1 \tag{A20}$$

Then, subtracting Eq. A20 from Eq. A9, we obtain

$$\frac{\Delta\hat{G}^{*\,\mathrm{chg}}_{1/2}}{RT} = q_1(1 - \tau_{12} + \ln\tau_1 - \ln\tau_{21}) \tag{A21}$$

(4) Again, we change component 2 in the mixture in step 3 to component 1 and Eq. A20 becomes

$$-\frac{\Delta G^{*\,\mathrm{chg}}_{1/1}}{RT} + \frac{\Delta\hat{G}^{*\,\mathrm{cav}}_{1/1} - \Delta G^{*\,\mathrm{cav}}_{1/1}}{RT} = -\,q_1\ln\tau_1 \tag{A22}$$

and Eq. A21 becomes

$$\frac{\Delta\hat{G}^{*\,\mathrm{chg}}_{1/1}}{RT} = q_1\ln\tau_1 \tag{A23}$$

With above equations, we can find the relation between the energy parameters and the charging free energies. From the addition of Eqs. A15 and A23, we get

$$\frac{\Delta G_{1/1}^{*\,\mathrm{chg}}}{RT} = q_1(\ln\tau_1 + \tau_1 - 1) \tag{A24}$$

While from the addition of Eqs. A13 and A23, we have

$$\frac{\Delta G_{1/2}^{*\,\mathrm{chg}}}{RT} = q_1(\tau_2 - \tau_{12} - \ln\tau_{21} + \ln\tau_1) \tag{A25}$$

By interchanging the indices, we obtain the two additional equations

$$\frac{\Delta G_{2/2}^{*\,\mathrm{chg}}}{RT} = q_2(\ln\tau_2 + \tau_2 - 1) \tag{A26}$$

and

$$\frac{\Delta G_{2/1}^{*\,\mathrm{chg}}}{RT} = q_2(\tau_1 - \tau_{21} - \ln\tau_{12} + \ln\tau_2) \tag{A27}$$

We can now find expressions for the cavity formation free energy from the UNIQUAC model. First, the solvation energy of pure species 1 at low pressure is

$$\frac{\Delta G_{1/1}^{*\,\mathrm{sol}}}{RT} = \frac{\Delta G_{1/1}^{*\,\mathrm{chg}}}{RT} + \frac{\Delta G_{1/1}^{*\,\mathrm{cav}}}{RT} = \ln\frac{\rho_1^{IG}}{\rho_1^0} = \ln\frac{P_1^{\mathrm{vap}}}{\rho_1^0\,kT} \tag{A28}$$

where we have used the ideal gas law $\rho_1^{\mathrm{IG}} = P_1^{\mathrm{vap}}/kT$.

Combining Eqs. A9 and A28, we have

$$\frac{\Delta G_{1/2}^{*\,\mathrm{chg}} + \Delta G_{1/2}^{*\,\mathrm{cav}}}{RT} - \ln\frac{P_1^{\mathrm{vap}}}{\rho_2^0 kT}$$

$$= \ln\gamma_1^{\infty}(\mathrm{comb}) + q_1(1 - \tau_{12} - \ln\tau_{21}) \tag{A29}$$

Finally, we combine Eqs. A25 and A29 to obtain the expression for the cavity formation free energy

$$\frac{\Delta G_{1/2}^{*\,\mathrm{cav}}}{RT} = \ln\gamma_1^{\infty}(\mathrm{comb}) + q_1(1 - \tau_2 - \ln\tau_1) - \ln\frac{\rho_2^0 kT}{P_1^{\mathrm{vap}}}$$

$$= \ln\frac{r_1}{r_2} + \frac{z}{2}q_1\ln\frac{q_1 r_2}{q_2 r_1} + \frac{z}{2}\left(\frac{r_1}{r_2}q_2 - q_1\right) + \left(1 - \frac{r_1}{r_2}\right)$$

$$+ q_1(1 - \tau_2 - \ln\tau_1) - \ln\frac{\rho_2^0 kT}{P_1^{\mathrm{vap}}} \tag{A30}$$

By changing the indices 2 to 1 in the above equation, we also obtain

$$\frac{\Delta G_{1/1}^{*\,\mathrm{cav}}}{RT} = q_1(1 - \tau_1 - \ln\tau_1) - \ln\frac{\rho_1^0 kT}{P_1^{\mathrm{vap}}} \tag{A31}$$

Manuscript received Mar. 12, 1999, and revision received July 23, 1999.