*J. Phys. Chem. A* **2000,** *104,* 7099−7105

# Multipole Corrections To Account for Structure and Proximity Effects in Group Contribution Methods: Octanol−Water Partition Coefficients

**Shiang-Tai Lin and Stanley I. Sandler***

*Center for Molecular and Engineering Thermodynamics, Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716*

*Received: January 6, 2000; In Final Form: April 19, 2000*

We propose a new method to correct for structure and proximity effects that are not generally accounted for in group contribution models. Molecular orbital calculations for isolated molecules are performed to obtain the net charge and dipole moment on each group within the molecule. These group multipoles, which vary as a result of the presence of proximate groups, are used as the basis for corrections to simple group contribution methods. We have applied this method to our recently developed GCSKOW model to predict the octanol−water partition coefficients ($K_{OW}$) for complex molecules. We studied this simple, theoretically based method with a total of 450 solutes, including five pharmaceutical compounds. The root-mean-square (RMS) deviation in the $\log_{10} K_{OW}$ calculated from the GCSKOW model with these multipole corrections is found to be 0.18 (which leads to a mean error of 52% in $K_{OW}$), compared with a RMS deviation of 0.44 (177%) when no correction is made. This simpler approach is more accurate than the KOW−UNIFAC model (RMS of 0.28 or an error of 92% in $K_{OW}$ for 344 compounds) and is comparable to the ClogP model (0.18 or 52%), which heretofore had been considered to be the most accurate predictive method for octanol−water partition coefficients.

## 1. Introduction

Group contribution methods have been reasonably successful for estimating many physical and thermodynamic properties of pure substances and of mixtures.[1] These methods dissect a molecule into building blocks or groups, e.g., atoms or groups of atoms. The physical picture then changes from a collection of molecules to a collection of groups that are assumed to be independent of each other. As a result, the overall property of the system becomes additive and is obtained by summing up contributions from all groups. This deconstruction of molecules greatly reduces the number of parameters needed to describe the properties of systems containing these functional groups. Once the group parameters for the property of interest have been determined from available experimental data, they can be used to predict the properties of new, more complex systems. Such methods not only provide a simple and systematic method for approximating the properties of new systems but, in reverse engineering, can also be used to design new compounds with desired properties.

Despite these appealing features, all group contribution methods have three common difficulties that limit their applicability. First, the definition of groups is empirical and arbitrary. Different methods use different groups to represent the same molecule. This raises the question of finding a “best” set of functional groups to describe the system properties. Second, simple group contribution methods do not distinguish between isomers. For example, the predicted properties of 2-butanol and isobutanol are identical from these methods. We refer to this type of deficiency as the structure effect. Finally, group contribution methods fail when a molecule contains two or more strong functional groups in close proximity. The interaction of a strong functional group with the others, which is a result of electrostatic forces, is affected by the presence of other neighboring strong functional groups; thus, the underlying assumption that each group is independent of others is no longer valid. This failure of group contribution methods is referred to as the (intramolecular) proximity effect.

Kehiaian,[2] in the DISQUAC model, accounted for proximity effects by empirically varying the values of the interaction parameters of a group depending on its first and second neighboring groups.[3] Wu and Sandler[4,5] later used quantum mechanical calculations to determine the charges on the atoms of a molecule and suggested that a better definition for groups is that each be electroneutral. However, their approach resolved the proximity effects by forming new, larger groups that contained the interacting functional groups. Gani and coworkers[6−8] proposed the inclusion of second-order groups in addition to the common functional groups, which they referred to as first-order groups. Second-order groups were defined according to the principle of conjugation[9] and comprised several first-order groups. These second-order groups serve as an empirical correction for structure and proximity effects. Many of the second-order groups were found to be identical to those defined by Wu and Sandler.[4]

In this work, we reexamine the definition of groups and propose a simple but theoretically based alternative to correct for both the structure and the proximity effects in a group contribution method recently developed by Lin and Sandler,[10] the GCSKOW model, for predicting the octanol−water partition coefficient, $K_{OW}$. The GCSKOW model was found to provide a more accurate prediction of $K_{OW}$ for monofunctional solutes than other models. However, when applying the same set of group parameters to multifunctional solutes, large deviations were observed as a result of proximity effects. To resolve this problem, rather than redefining a larger group to satisfy the zero-charge requirement, we use quantum mechanics to determine the charge and dipole of the group and on the basis of this make

---

* Corresponding author. E-mail: sandler@udel.edu. Telephone: 302-831-2945. Fax: 302-831-4466.

10.1021/jp000091m CCC: $19.00

corrections for the interaction parameters of the same group in different molecules. Therefore, with the additional information from inexpensive, first-principle calculations for isolated molecules, traditionally defined groups can be retained and $K_{OW}$ can be accurately predicted for complex, multifunctional solutes without constructing new groups.

## 2. Theory

Wu and Sandler[4] suggested the following principles for group identification:

(1) The geometry of a functional group shall be the same independent of the molecules in which the group occurs.

(2) Each atom in a functional group should have approximately the same charge in all molecules in which the group occurs, and the group should have approximately zero charge.

(3) Each functional group should be the smallest entity such that a molecule can be divided into a collection of electroneutral groups.

The requirements of invariant group geometry and atomic charges are reasonable since charge distribution determines the interaction between groups. However, demanding the net charge on a group be zero is unnecessary. As long as the charge distribution on a group remains the same independent of the molecule in which it is contained, it should make the same contribution to the behavior of a mixture.

Instead of constructing new groups when the zero-charge criterion is violated, the way we proceed here is to define a reference charge distribution for each group and explicitly correct for any deviation of this distribution from the reference state. In other words, the deviations of the monopole (charge), dipole, quadrupole, and so on, of a group from their reference values as a result of the presence of proximate strong groups in the same molecule are used to correct for the deviations from the simple group contribution idea. However, there may not be a universal way to make such corrections for all physical and thermodynamic properties. Different parameters have a different dependence on the mulitpole moments. For this reason, here we only apply this new method to predict the octanol−water partition coefficient by using the GCSKOW group contribution model.[10] Later work will consider applications to other properties.

The octanol−water partition coefficient of species *i*, $K_{OW,i}$, can be computed from the difference of solvation free energies of this species in the octanol-saturated water phase (essentially pure water) and in the water-saturated octanol phase. The solvation free energy, which is the work required to transfer a molecule from a fixed position in the ideal gas phase to a fixed position in solution,[11] consists of the free energies of the creation of a cavity in the solution and of the placement of the solute into that cavity. The former contribution is determined by the volume and surface area of the solute, whereas the later is equivalent to establishing the solute electronic configuration in the cavity, usually termed as the charging free energy. On the basis of this, Lin and Sandler derived the following GCSKOW model:[10]

$$\log_{10} K_{OW,i} = -0.126 + 1.031 r_i - 1.208 q_i + \frac{\Delta G^{*chg}_{i/W} - \Delta G^{*chg}_{i/OR}}{1.364} \quad (1)$$

where $r_i$ and $q_i$ are the volume and surface area parameters of species *i* and $\Delta G^{*chg}_{i/W}$ and $\Delta G^{*chg}_{i/OR}$ are the charging free energies in the water and the octanol-rich phases, respectively. Note that the constant coefficients in eq 1 are not fitted parameters; they are dependent on temperature and have been calculated from the water and *n*-octanol volume and surface area parameters and their equilibrium concentrations that are unaffected by the addition of a dilute solute.[10]

The molecular structure parameters, $r_i$ and $q_i$, in eq 1 can be calculated in a group contribution manner:

$$r_i = \sum_{k=1}^{N_i} R_k, \quad q_i = \sum_{k=1}^{N_i} Q_k \quad (2)$$

where the summation is over the $N_i$ functional groups contained in species *i*; $R_k$ and $Q_k$ are the volume and surface area contributions of functional group *k*. The charging free energy $\Delta G^{*chg}_{i/j}$, originating from the attractive interaction between solute *i* and solvent *j*, has three contributions: electrostatic, dispersion, and repulsion:[12]

$$\Delta G^{*chg}_{i/j} = \Delta G^{*es}_{i/j} + \Delta G^{*disp}_{i/j} + \Delta G^{*rep}_{i/j} \quad (3)$$

The dispersion and repulsion contributions are usually assumed to be dependent on the molecular surface area[13−16] and do not vary with different solute electronic configurations. The electrostatic contribution, however, is very sensitive to the charge distribution in a molecule. Kirkwood[17] has derived a general equation for the electrostatic interaction between a distribution of discrete point charges in a spherical cavity and an isotropic dielectric medium. The results were expressed[18] in terms of the multipole moments at the center of the cavity as

$$\Delta G^{*es} = -\frac{\epsilon - 1}{2\epsilon}\frac{e^2}{a} - \frac{\epsilon - 1}{1 + 2\epsilon}\frac{\mu^2}{a^3} + \text{higher order terms} \quad (4)$$

where $\epsilon$ is the dielectric constant, *a* is the radius of the cavity, *e* is the total charge, and $\mu$ is the dipole moment evaluated at the center of the cavity.

The structure and proximity effects that result from the changes of multipole moments of each group can now be considered. Truncating eq 4 at the second term and assuming that the charge distribution of a solute group is independent of the solvent, the electrostatic contribution of group *k* at some reference net charge $e_k^0$ and dipole $\mu_k^0$ in solvent *j* is

$$\Delta G^{*es,0}_{k/j} = -\frac{\epsilon_j - 1}{2\epsilon_j}\frac{(e_k^0)^2}{a_{k/j}} - \frac{\epsilon_j - 1}{1 + 2\epsilon_j}\frac{(\mu_k^0)^2}{{a_{k/j}}^3} \quad (5)$$

where $\epsilon_j$ is the dielectric constant of solvent *j* and $a_{k/j}$ is the effective cavity radius of group *k* in solvent *j*. For a molecule *i*,

$$\begin{aligned}\Delta G^{*chg}_{i/j} &= \sum_{k=1}^{N_i} \Delta G^{*chg}_{k/j} = \sum_{k=1}^{N_i} (\Delta G^{*es}_{k/j} + \Delta G^{*disp}_{k/j} + \Delta G^{*rep}_{k/j}) \\ &= \sum_{k=1}^{N_i} (\Delta G^{*es,0}_{k/j} + \Delta G^{*disp}_{k/j} + \Delta G^{*rep}_{k/j}) + \\ &\qquad \sum_{k=1}^{N_i} (\Delta G^{*es}_{k/j} - \Delta G^{*es,0}_{k/j}) \quad (6) \\ &= \sum_{k=1}^{N_i} \Delta G^{*chg,0}_{i/W} + \sum_{k=1}^{N_i} \left\{ -\frac{(\epsilon_j - 1)}{2\epsilon_j}\frac{[{e_k}^2 - (e_k^0)^2]}{a_{k/j}} - \frac{(\epsilon_j - 1)}{(1 + 2\epsilon_j)}\frac{[{\mu_k}^2 - (\mu_k^0)^2]}{{a_{k/j}}^3} \right\}\end{aligned}$$

Therefore, the charging free energy parameter in eq 1 becomes

$$\Delta G^{*\mathrm{chg}}_{i/\mathrm{W}} - \Delta G^{*\mathrm{chg}}_{i/\mathrm{OR}} = \sum_{k=1}^{N_i} \Delta\Delta G^{*\mathrm{chg},0}_{k/\mathrm{W-OR}} + \sum_{k=1}^{N_i} \{ C^e_k [{e_k}^2 - (e^0_k)^2] + C^\mu_k [{\mu_k}^2 - (\mu^0_k)^2] \} \quad (7)$$

with $\Delta\Delta G^{*\mathrm{chg},0}_{k/\mathrm{W-OR}} = \Delta G^{*\mathrm{chg},0}_{k/\mathrm{W}} - \Delta G^{*\mathrm{chg},0}_{k/\mathrm{OR}}$ and

$$C^e_k = -\frac{\epsilon_\mathrm{W} - 1}{2\epsilon_\mathrm{W} a_{k/\mathrm{W}}} + \frac{\epsilon_\mathrm{OR} - 1}{2\epsilon_\mathrm{OR} a_{k/\mathrm{OR}}} \quad (8)$$

$$C^\mu_k = -\frac{\epsilon_\mathrm{W} - 1}{(1 + 2\epsilon_\mathrm{W}){a_{k/\mathrm{W}}}^3} + \frac{\epsilon_\mathrm{OR} - 1}{(1 + 2\epsilon_\mathrm{OR}){a_{k/\mathrm{OR}}}^3} \quad (9)$$

The reference charge and dipole can be absorbed into the first term in eq 7 to give

$$\Delta G^{*\mathrm{chg}}_{i/\mathrm{W}} - \Delta G^{*\mathrm{chg}}_{i/\mathrm{OR}} = \sum_{k=1}^{N_i} (\Delta\Delta G^{*\mathrm{chg},00}_{k/\mathrm{W-OR}} + C^e_k {e_k}^2 + C^\mu_k {\mu_k}^2) \quad (10)$$

where $\Delta\Delta G^{*\mathrm{chg},00}_{k/\mathrm{W-OR}} = \Delta\Delta G^{*\mathrm{chg},0}_{k/\mathrm{W-OR}} - C^e_k(e^0_k)^2 - C^\mu_k(\mu^0_k)^2$. Equation 10 thus allows us to account for changes in the solvation free energy resulting from variations in the group charge $e_k$ and dipole $\mu_k$ from the reference state as group $k$ appears in different molecules.

## 3. Computational Details

To determine the effect of accounting for only the variation in charge and dipole moment in a group contribution method, the identification of groups remains the same as that in Lin and Sandler.[10] The Gaussian 98 program[19] was used to evaluate the group charges and dipoles of isolated molecules. Geometry optimization in a vacuum was performed by using the Hartree–Fock method with a 6-31G(d,p) basis set, followed by a Mulliken population analysis[20] to obtain the charge on each atom. The net group charge $e_k$ and dipole $\vec{\mu}_k$ were calculated according to

$$e_k = \sum_{m=1}^{N_k} e_m, \quad \vec{\mu}_k = \sum_{m=1}^{N_k} e_m \vec{r}_m \quad (11)$$

where the summation is over the $N_k$ atoms contained in group $k$, $e_m$ is the Mulliken charge of atom $m$, and $\vec{r}_m$ is the position vector of atom $m$ originating at the center of gravity of group $k$.[18]

With the multipole corrections, the GCSKOW model of eq 1 contains three types of parameters: two group structure parameters ($R_k$ and $Q_k$), one free energy parameter ($\Delta\Delta G^{*\mathrm{chg},00}_{k/\mathrm{W-OR}}$), and two multipole coefficients ($C^e_k$ and $C^\mu_k$). The group structure parameters were computed by using the van der Waals radii of each atom.[10] The $\Delta\Delta G^{*\mathrm{chg},00}_{k/\mathrm{W-OR}}$, $C^e_k$, and $C^\mu_k$ were optimized such that the root-mean-square (RMS) deviation of $\log_{10} K_\mathrm{OW}$ was minimized:

$$\mathrm{RMS} = \left[ \frac{\sum_{i=1}^{N} (\log_{10} K^\mathrm{GCSKOW}_{\mathrm{OW},i} - \log_{10} K^\mathrm{expt}_{\mathrm{OW},i})^2}{N} \right]^{1/2} \quad (12)$$

**TABLE 1: Group Parameters for the GCSKOW Model**

| main group | subgroup $k$ | $\Delta\Delta G^{*\mathrm{chg},00}_{k/\mathrm{W-OR}}$ | $C^e_k$ | $C^\mu_k$ | $R_k$ | $Q_k$ |
|---|---|---|---|---|---|---|
| CH$_2$ | —CH$_3$ | 2.201 | −1.968 | −55.315 | 0.887 | 0.840 |
| | —CH$_2$— | 1.077 | −0.342 | −21.809 | 0.665 | 0.523 |
| | >CH— | −0.091 | 1.756 | −6.509 | 0.497 | 0.235 |
| | >C< | −0.729 | 1.720 | | 0.213 | 0.000 |
| cyclic-CH$_2$ | —CH$_2$— | 1.071 | 1.270 | −18.372 | 0.665 | 0.523 |
| | >CH— | −0.051 | −0.093 | 8.892 | 0.497 | 0.235 |
| OH | —OH | −10.405 | 5.581 | 72.945 | 0.532 | 0.572 |
| COOH | —COOH | −1.114 | −37.312 | 10.940 | 1.307 | 1.176 |
| CHO | —CHO | 0.432 | −0.277 | −1.897 | 0.984 | 0.892 |
| C═O | >CO | −3.315 | −10.273 | 5.375 | 0.716 | 0.527 |
| OC═O | —COO— | −0.723 | −5.608 | 2.172 | 1.062 | 0.821 |
| CH$_2$OCH$_2$ | —O— | −2.146 | −0.384 | | 0.401 | 0.327 |
| NO$_2$ | —NO$_2$ | −1.158 | 7.864 | −0.241 | 1.086 | 0.986 |
| CN | —CN | 2.651 | −31.702 | −13.126 | 0.954 | 0.852 |
| NH$_2$ | —NH$_2$ | −9.255 | 0.979 | 71.550 | 0.742 | 0.748 |
| | —NH— | −4.959 | 3.082 | 17.907 | 0.553 | 0.401 |
| | >N(sp3)— | −5.525 | 4.689 | | 0.341 | 0.000 |
| | ═N(sp2)— | −1.301 | 0.003 | | 0.436 | 0.339 |
| CONH$_2$ | —CONH$_2$ | −1.622 | 20.481 | −0.360 | 1.467 | 1.284 |
| C$_6$H$_6$ | ═CH— | 0.468 | 0.135 | −0.955 | 0.537 | 0.431 |
| | ═C< | 0.031 | 5.699 | | 0.316 | 0.114 |
| C═C | —CHCH$_2$ | 1.866 | 1.665 | −16.773 | 1.359 | 1.210 |
| C≡C | —CCH | 1.014 | −0.099 | −10.183 | 1.129 | 1.008 |
| F | —F | 0.946 | −6.997 | | 0.376 | 0.458 |
| Cl | —Cl | 1.129 | −5.583 | | 0.861 | 0.771 |
| Br | —Br | 1.268 | −0.292 | | 1.114 | 0.935 |

Two types of compounds were considered in this work: monofunctional and multifunctional. Monofunctional compounds contain no more than nonalkyl, strong functional group, e.g., species in the homologous series, cyclic alkanes, alkylbenzenes, and branched, secondary, and tertiary species. Species containing more than one strong functional group (that is, multifunctional compounds) include X(CH$_2$)$_n$Y-type species, where X and Y are strong functional groups (OH, COOH, CHO, COCH$_3$, CH$_3$COO, CH$_3$O, NO$_2$, CN, CONH$_2$, C$_6$H$_5$, CHCH$_2$, CCH, F, Cl, and Br); chlorofluorocarbons; multiaromatic rings; multichlorinated benzenes; phenol, aniline, toluene, pyridine, furan, and their derivatives; and complex pharmaceutical compounds such as nicotine, caffeine, theophylline, piracetam, and mexiletine. (An example calculation for nicotine is provided in the Appendix.)

## 4. Results and Discussion

A total of 450 species, comprising 204 monofunctional and 246 multifunctional compounds, were used to obtain the free energy parameters and the multipole coefficients listed in Table 1. The calculated results from the GCSKOW model are compared with experimental measurements in Figure 1. All experimental data are taken from the literature.[21–25] [Experimental values for monofunctional compounds from different sources are generally within 0.2 log unit, except for pentane (2.5,[24] 3.39[21–23,25]), 2,3-dimethylbutane (3.42,[21,25] 3.85[23,24]), cyclohexene (2.18,[24] 2.86[21,23,25]), and hexamethylbenzene (4.31,[23] 4.61,[25] 5.11[22]). For these compounds and all multifunctional compounds, experimental values from ref 25 are used.] The predictions agree very well with experimental data for both types of compounds. Table 2 lists the RMS deviation of $\log_{10} K_\mathrm{OW}$ from the GCSKOW model developed here, together with two other group contribution models: KOW–UNIFAC[26] and ClogP.[25,27] In the KOW–UNIFAC model the octanol–water partition coefficient is calculated on the basis of a statistical mechanical model for infinite dilution activity coefficients, whereas ClogP uses predefined fragments and bond connections. Fewer compounds are considered for the KOW–UNIFAC model because

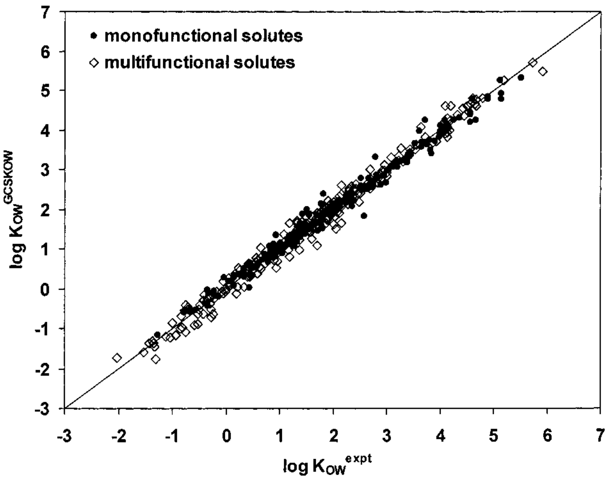

**Figure 1.** Comparison between predictions from the GCSKOW model and experimental data.

**TABLE 2: Root-Mean-Square Deviations of $\log_{10} K_{\mathrm{OW}}$ Predictions of Numerous Compounds from Different Group Contribution Models**

| | GCSKOW | KOW−UNIFAC | ClogP |
|---|---|---|---|
| | | Monofunctional | |
| RMS | 0.18 | 0.20 | 0.17 |
| data points | 204 | 165 | 204 |
| | | Multifunctional Compounds | |
| RMS | 0.19 | 0.34 | 0.19 |
| data points | 246 | 179 | 246 |
| | | Overall | |
| RMS | 0.18 | 0.28 | 0.18 |
| data points | 450 | 344 | 450 |

of the lack of necessary group parameters. Because of the importance of proximity effects in multifunctional solutes, all three predictive models exhibit a greater deviation from experiment for such compounds. Overall, the GCSKOW model, with a RMS deviation of 0.18 or 52% deviation from the experimental $K_{\mathrm{OW}}$, is superior to the KOW−UNIFAC model, 0.28 or 92%, and comparable to ClogP, 0.18 or 52%.

The progressive improvement from inclusion of the group charge and group dipole corrections can be seen by performing two additional optimizations: first the $C_k^e$ and $C_k^\mu$ are set to zero, which corresponds to no correction for variations in group charge or dipole moments (the original GCSKOW model[10]); then only the $C_k^\mu$ are set zero, which corresponds to corrections for group charge only. The RMS errors resulting in $\log_{10} K_{\mathrm{OW}}$ from each of these optimizations for the same set of compounds are found to be 0.44 and 0.28, respectively, compared with 0.18 when both the charge and dipole moments are considered. Therefore, the deviation is approximately reduced by 40% (compared with no multipole corrections) when group charge corrections are included and by 60% when considering both the group charge and group dipole variations. We have not considered higher order corrections since the current predictions with both group charge and dipole corrections are already within the experimental uncertainty.

The structure effects in isomers are also accounted for by using the group contribution method proposed here. Table 3 compares the predictions of some isomers from different models. It can be seen that the GCSKOW model captures essentially all the structure effects for these isomers, whereas ClogP and KOW−UNIFAC do not. The KOW−UNIFAC and ClogP models essentially follow Kehiaian's concept[2] of correcting for structure (and proximity) effects. For instance, KOW−UNIFAC uses different group parameters for primary and secondary hydroxyl groups, and ClogP defines different bond connections between the alkyl and hydroxyl groups. These models consider only the effects from the first nearest neighbors and thus do not distinguish between the isomers listed in Table 3, where second or even third nearest neighboring groups would have to be included. In contrast, the GCSKOW model with mulitpole corrections based on a calculation for the whole molecule distinguishes between all isomers without defining new group parameters or bond connections. This is because the structural variations in isomers result in changes of group charges and dipole moments, and these are conveniently corrected for by using eq 10.

**TABLE 3: Structure Effects on the Octanol−Water Partition Coefficients for Isomers**

| solute | experiment | GCSKOW | KOW−UNIFAC | ClogP |
|---|---|---|---|---|
| 3-methy-1-butanol | 1.42 | 1.35 | 1.48 | 1.22 |
| 2-methyl-1butanol | 1.29 | 1.30 | 1.48 | 1.22 |
| 2-hexanol | 1.76 | 1.68 | 1.70 | 1.66 |
| 3-hexanol | 1.65 | 1.67 | 1.70 | 1.66 |
| 2-pentanone | 0.91 | 1.03 | 0.85 | 0.85 |
| 3-pentanone | 0.99 | 1.08 | 0.85 | 0.85 |
| propyl acetate | 1.23 | 1.18 | 1.32 | 1.24 |
| ethyl propionate | 1.21 | 1.29 | 1.32 | 1.24 |
| methyl butyrate | 1.29 | 1.38 | 1.32 | 1.24 |
| butyl acetate | 1.78 | 1.74 | 1.91 | 1.77 |
| methyl pentanoate | 1.96 | 1.94 | 1.91 | 1.77 |
| 1-naphthol | 2.84 | 2.89 | 2.71 | 2.65 |
| 2-naphthol | 2.70 | 2.62 | 2.71 | 2.65 |

The proximity effects, which are the major cause of the failure of most simple group contribution models, are more significant than the structure effects that result in variations of only about 0.1 log unit for $K_{\mathrm{OW}}$. Here we investigate proximity effects on $\mathrm{X(CH_2)}_n\mathrm{Y}$-type molecules, where X and Y can be the same or different strong end groups. One might anticipate a monotonic increase of $\log_{10} K_{\mathrm{OW}}$ for this kind of series as more methylene groups separate the two end groups. In fact, in some series there exists an unexpected minimum in $\log_{10} K_{\mathrm{OW}}$ as a function of the number of methylene groups. Further discussion will focus on such series. Table 4 shows the predictions for eight families of $\mathrm{X(CH_2)}_n\mathrm{Y}$ compounds by using GCSKOW, KOW−UNIFAC, and ClogP. The GCSKOW model successfully describes the minimum behavior in most $\mathrm{X(CH_2)}_n\mathrm{Y}$ series. The KOW−UNIFAC model is capable of describing only the $\mathrm{X(CH_2)}_n\mathrm{C_6H_5}$ series because new groups were constructed only for $n = 0$. A similar situation is true for the ClogP method, where new bond connections were also defined only for $n = 0$. The success of these models is not surprising since in most cases the minimum occurs when $n = 1$. However, KOW−UNIFAC and ClogP models are both incapable of predicting the minimum at $n = 2$ for the $\mathrm{CN(CH_2)}_n\mathrm{CN}$ series [parameters for the CN group in KOW−UNIFAC are not applicable for $n = 0$ and $n = 1$; the irregular predictions of ClogP are the result of the inclusion of a Y−C−Y proximity correction for $\mathrm{CN(CH_2)_1CN}$ and a Y−CC−Y proximity correction for $\mathrm{CN(CH_2)_2CN}$].

The group charge and dipole correction method proposed here not only correctly predicts the location of minimum in $\log_{10} K_{\mathrm{OW}}$ as a function of the number of methylene groups but also provides an explanation for this phenomenon. Figure 2 shows the variation of the free energy parameter $(\sum_{k=1}^{N_i}\Delta\Delta G^{*\mathrm{chg},00}_{k/\mathrm{W-OR}})$, the charge correction $(\sum_{k=1}^{N_i}C_k^e e_k^2)$, and the dipole correction $(\sum_{k=1}^{N_i}C_k^\mu \mu_k^2)$ in the $\mathrm{CN(CH_2)}_n\mathrm{CN}$ series. The contribution from the free energy term increases with the number of methylene groups contained in the molecule, thereby resulting in a linear increase in $\log_{10} K_{\mathrm{OW}}$. However, the variation of the charge

**TABLE 4: Proximity Effects on the Octanol−Water Partition Coefficients for X(CH$_2$)$_n$Y-Type Compounds**

| X | Y | $n$ | experiment | GCSKOW | KOW−UNIFAC | ClogP |
|---|---|---|---|---|---|---|
| CN | CN | 0 | 0.07 | 0.11 | | 0.07 |
| | | 1 | −0.50 | −0.52 | | −1.20 |
| | | 2 | −0.99 | −0.85 | −1.42 | −0.82 |
| | | 3 | −0.72 | −0.52 | −0.83 | −0.95 |
| | | 4 | −0.32 | −0.19 | −0.23 | −0.42 |
| | | 5 | 0.05 | 0.29 | 0.36 | 0.10 |
| | | 6 | 0.59 | 0.75 | 0.95 | 0.63 |
| CN | Cl | 1 | 0.45 | 0.28 | | 0.22 |
| | | 2 | 0.18 | 0.23 | −0.02 | 0.20 |
| | | 3 | 0.56 | 0.63 | 0.57 | 0.38 |
| OH | C$_6$H$_5$ | 0 | 1.47 | 1.58 | 1.42 | 1.47 |
| | | 1 | 1.10 | 1.03 | 0.83 | 1.10 |
| | | 2 | 1.36 | 1.53 | 1.42 | 1.18 |
| | | 3 | 1.88 | 2.01 | 2.18 | 1.71 |
| | | 4 | 2.35 | 2.54 | 2.94 | 2.24 |
| COOH | C$_6$H$_5$ | 0 | 1.87 | 1.79 | 1.79 | 1.88 |
| | | 1 | 1.41 | 1.70 | 1.41 | 1.41 |
| | | 2 | 1.84 | 1.91 | 2.00 | 1.75 |
| | | 3 | 2.42 | 2.46 | 2.59 | 2.28 |
| | | 4 | 2.70 | 2.98 | 3.18 | 2.81 |
| | | 5 | 3.27 | 3.52 | 3.78 | 3.34 |
| | | 6 | 3.63 | 4.05 | 4.37 | 3.87 |
| | | 7 | 4.09 | 4.59 | 4.96 | 4.40 |
| CH$_3$CO | C$_6$H$_5$ | 0 | 1.58 | 1.67 | 1.32 | 1.58 |
| | | 1 | 1.44 | 1.65 | 1.39 | 1.43 |
| | | 3 | 2.42 | 2.65 | 2.57 | 2.27 |
| CH$_3$O | C$_6$H$_5$ | 0 | 2.11 | 1.74 | 2.06 | 2.06 |
| | | 1 | 1.92 | 1.85 | | 1.92 |
| | | 3 | 2.70 | 2.75 | | 2.29 |
| NO$_2$ | C$_6$H$_5$ | 0 | 1.85 | 1.96 | 1.83 | 1.88 |
| | | 1 | 1.75 | 1.79 | | 1.75 |
| | | 2 | 2.08 | 2.15 | 1.96 | 1.66 |
| CONH$_2$ | C$_6$H$_5$ | 0 | 0.64 | 0.57 | | 0.65 |
| | | 1 | 0.45 | 0.69 | | 0.45 |
| | | 2 | 0.91 | 1.04 | | 0.71 |
| | | 3 | 1.41 | 1.62 | | 1.24 |

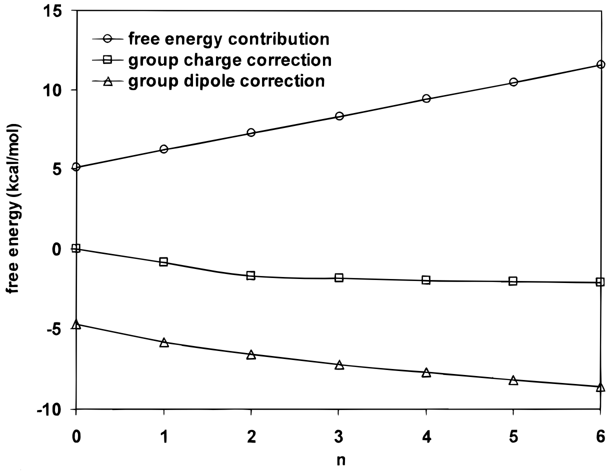

**Figure 2.** The variation of the free energy contribution ($\sum_{k=1}^{N_i}\Delta\Delta G^{*\text{chg},00}_{k/\text{W-OR}}$), the group charge contribution ($\sum_{k=1}^{N_i}C^{e}_{k}e_k^{2}$), and the group dipole contribution ($\sum_{k=1}^{N_i}C^{\mu}_{k}\mu_k^{2}$) to $K_{\text{OW}}$ in the CN(CH$_2$)$_n$CN series.

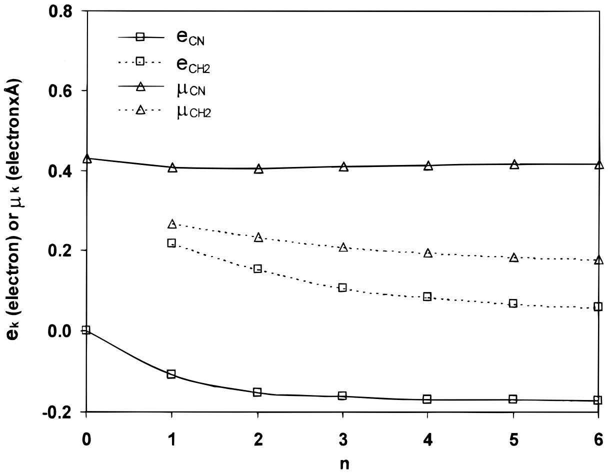

**Figure 3.** The variation of averaged charge and dipole moment on the CN group ($e_{\text{CN}}$ and $\mu_{\text{CN}}$) and CH$_2$ group ($e_{\text{CH}_2}$ and $\mu_{\text{CH}_2}$) in the CN-(CH$_2$)$_n$CN series.

**TABLE 5: Octanol−Water Partition Coefficient Predictions for Pharmaceutical Compounds**

| | experiment | GCSKOW | | KOW−UNIFAC | | ClogP | |
|---|---|---|---|---|---|---|---|
| solute | log$_{10}$ $K_{\text{OW}}$ | log$_{10}$ $K_{\text{OW}}$ | dev | log$_{10}$ $K_{\text{OW}}$ | dev | log$_{10}$ $K_{\text{OW}}$ | dev |
| nicotine | 1.17 | 0.99 | −0.18 | | | 1.32 | 0.15 |
| theophylline | −0.02 | −0.02 | 0.00 | | | −0.06 | −0.04 |
| caffeine | −0.07 | −0.07 | 0.00 | | | −0.06 | 0.01 |
| piracetam | −1.54 | −1.66 | −0.12 | | | −1.36 | 0.18 |
| mexiletine | 2.15 | 2.13 | −0.02 | 2.53 | 0.38 | 2.57 | 0.42 |

correction term with $n$ is not linear. This term decreases significantly between $n = 0$ and 2 and tends to a constant value after $n = 5$. This nonlinear decrease is a direct result from the proximity effect. For cyanogen ($n = 0$) charges on both CN groups ($e_{\text{CN}}$) are zero and the charge correction is the largest (Figure 3). At larger CN separation distances, the two CN groups are less correlated and the charge correction decreases. This term approaches a constant value when $n > 5$ because the addition of zero-charged methylene groups ($e_{\text{CH}_2}$ for large $n$) does not contribute to this term. Similar reasoning applies to the dipole correction, except that the net dipole moment of a methylene group is nonzero, which results in a constant decrease in this term at large $n$. It is the nonlinear decreases of the charge and dipole contributions that cause the log$_{10}$ $K_{\text{OW}}$ to have a minimum as a function of the number of methylene groups. Thus, we see that the proximity effect appears as a nonlinear contribution to $K_{\text{OW}}$ as a result of the electrostatic interactions between groups.

A more severe test of group contribution methods is their application to large pharmaceutical compounds. These compounds are of practical interest and usually difficult to model because of their molecular complexity. Table 5 lists computed values of $K_{\text{OW}}$ for five small/medium-sized pharmaceuticals. The selection of these five compounds was based only on their molecular size, not prejudged by the accuracy of the prediction. (It is not our intention at this point to consider other much larger compounds since our purpose here is to demonstrate the use of group multipoles to improve the current group contribution models.) Predictions from the GCSKOW model agree very well with experimental data. The KOW−UNIFAC model cannot be used for many of these compounds, again as a result of the lack of group parameters. It is also interesting to find that the ClogP model, which is one of the best existing models for $K_{\text{OW}}$ prediction, does not perform better than the GCSKOW model.

While the application of group multipole corrections to the GCSKOW model is very successful and potentially can be used with other group contribution models, we should point out the assumptions and approximations we have made:

(1) Groups with multipole corrections are assumed to be independent of one another. That is, the proximity effects are corrected for only by the changes of multipoles of each group due to the presence of other functional groups.

(2) Solute molecules are not polarized by the solvent; that is, the charge distribution of the solute in the solution phase is

presumed to be the same as in the ideal gas phase. This is not generally true and solute polarization is known to play a role in the solvation free energy. However, as we have shown here, these ideal gas multipole moments, which are much simpler to calculate, can be used as a basis for correcting for the intramolecular proximity effect.

(3) Mulliken charges evaluated at the Hartree−Fock level are presumed to be a good measure for the characterization of structure and proximity effects. We have examined the calculations with a higher level density functional theory[28,29] and found comparable results. A more sophisticated population analysis, such as CHelpG charges,[30] leads to results that are less accurate than those based on the Mulliken charges. The CHelpG charges have been found to be less transferable than Mulliken charges and are often too large in magnitude, thereby resulting in intermolecular interactions that are too strong.[31] Mulliken charges, while highly dependent on the basis set, are consistent among different molecules and may serve as a good measure of multipole variations.

The GCSKOW model with multipole corrections currently contains a total of 70 fitting parameters (the $\Delta\Delta G^{*\mathrm{chg},00}_{k/\mathrm{W-OR}}$, $C^e_k$, and $C^\mu_k$ in Table 1) and 450 compounds were used to obtain these parameters. This set of experimental data includes most combinations of $\mathrm{X(CH_2)}_n\mathrm{Y}$-type species, especially for $n = 0$, 1, and 2, where the proximity effect between groups X and Y should be the strongest. The success of the multipole correction method for the types of compounds considered here suggests that the method should be applicable to more complex compounds, although as with any new method predictions for other compounds still should be treated with caution. Although a broader database should be used to explore the potential and limitations of the multipole correction method, the results we have obtained so far demonstrate the utility of this new approach to account for the structure and proximity effects in group contribution methods.

## 5. Conclusions

Most group contribution models ignore the structure and proximity effects considered here, or correct for them on the basis of proximate neighbors by constructing new and larger groups to include the interacting groups, or use a combination of these two methods. The first approach handles the problem at the cost of introducing a large number of empirical parameters (potentially $[N(N+1)/2]m$ parameters for a set of $N$ groups when up to the $m$th nearest neighbors are considered). The second approach, although possessing a theoretical basis,[4] requires the continual creation of new groups whenever a new proximity effect is encountered. The knowledge from known proximity effects may not be useful when a new type of proximity effect is encountered.

The multipole correction method we propose here obviates both of the above difficulties. The structure and proximity effects are effectively accounted for through the calculation of group charges and dipole moments by using quantum mechanics for the isolated molecules. This theoretically based approach requires significantly fewer parameters than others because the changes in the electronic configuration of each group and, hence, the changes in the electrostatic interaction parameters are rigorously taken into account. Application of this method to the prediction of the octanol−water partition coefficients proves its superiority over the empirical correction methods. This method should also be applicable to other group contribution models for physical and thermodynamic property predictions, provided that the group parameters can be expressed in terms of multipole expansion.

**Acknowledgment.** The authors thank the National Science Foundation (CTS-9521406) and the Department of Energy (DE-FG02-85ER13436) for financial support of this research. We also thank the BoiByte Corp. and Professor J. Gmehling of Universität Oldenburg for providing us with their computer programs of their models.

## Appendix: Example Calculation for Nicotine

Nicotine contains four aromatic CH groups, one N(sp2) group, one aromatic C group, one cyclic CH group, three cyclic $\mathrm{CH_2}$ groups, one N(sp3) group, and one $\mathrm{CH_3}$ group. The coordinates (in angstroms) of each atom from the optimized geometry and Mulliken charges (in electrons) are calculated from Gaussian 98 and shown in Table 6.

**TABLE 6**

| group | atom | mass | *x* | *y* | *z* | charge |
|---|---|---|---|---|---|---|
| ACH | C | 12 | 0.29089 | 0.05381 | −2.05295 | 0.11754 |
| | H | 1 | −0.66464 | 0.09515 | −2.54989 | 0.14549 |
| ACH | C | 12 | 2.53684 | −0.07856 | −2.27749 | 0.12459 |
| | H | 1 | 3.38208 | −0.13714 | −2.94204 | 0.14819 |
| ACH | C | 12 | 2.72296 | −0.07615 | −0.90339 | −0.20708 |
| | H | 1 | 3.71340 | −0.13589 | −0.49007 | 0.15567 |
| ACH | C | 12 | 1.61095 | 0.00225 | −0.08947 | −0.07283 |
| | H | 1 | 1.70587 | −0.00278 | 0.98145 | 0.17762 |
| N (sp2) | N | 14 | 1.34884 | −0.01613 | −2.84410 | −0.54078 |
| AC | C | 12 | 0.34914 | 0.07355 | −0.66847 | −0.08360 |
| cyclic-CH | C | 12 | −0.90677 | 0.20274 | 0.16584 | 0.08621 |
| | H | 1 | −1.76283 | −0.02907 | −0.47761 | 0.10432 |
| cyclic-$\mathrm{CH_2}$ | C | 12 | −1.11401 | 1.60928 | 0.77020 | −0.22568 |
| | H | 1 | −1.70806 | 2.23014 | 0.10934 | 0.12573 |
| | H | 1 | −0.15972 | 2.10416 | 0.90238 | 0.13150 |
| cyclic-$\mathrm{CH_2}$ | C | 12 | −1.79559 | 1.35975 | 2.13397 | −0.24524 |
| | H | 1 | −2.74666 | 1.87138 | 2.22659 | 0.12180 |
| | H | 1 | −1.16118 | 1.70414 | 2.94237 | 0.12888 |
| cyclic-$\mathrm{CH_2}$ | C | 12 | −1.96120 | −0.16078 | 2.19580 | −0.00505 |
| | H | 1 | −2.95018 | −0.45161 | 1.82474 | 0.08911 |
| | H | 1 | −1.85915 | −0.56202 | 3.19846 | 0.12285 |
| N (sp3) | N | 14 | −0.90892 | −0.66573 | 1.33389 | −0.60598 |
| $\mathrm{CH_3}$ | C | 12 | −1.02175 | −2.07665 | 1.04564 | −0.12810 |
| | H | 1 | −0.19313 | −2.40250 | 0.42944 | 0.12908 |
| | H | 1 | −0.99142 | −2.64033 | 1.97173 | 0.11884 |
| | H | 1 | −1.95259 | −2.32606 | 0.52651 | 0.08693 |

The center of mass, charge, and dipole moment (in electron × angstrom) for each group can then be determined, as shown in Table 7. (A computer program for this part of the calculation is available upon request.)

**TABLE 7**

| group | mass | *x* | *y* | *z* | charge | dipole |
|---|---|---|---|---|---|---|
| ACH | 13 | 0.21691 | 0.05701 | −2.09142 | 0.26304 | 0.13487 |
| ACH | 13 | 2.60228 | −0.08309 | −2.32894 | 0.27277 | 0.13683 |
| ACH | 13 | 2.79964 | −0.08077 | −0.87139 | −0.05142 | 0.17160 |
| ACH | 13 | 1.61830 | 0.00186 | −0.00657 | 0.10479 | 0.18225 |
| N (sp2) | 14 | 1.34884 | −0.01613 | −2.84410 | −0.54078 | 0.00000 |
| AC | 12 | 0.34914 | 0.07355 | −0.66847 | −0.08360 | 0.00000 |
| cyclic-CH | 13 | −0.97305 | 0.18479 | 0.11602 | 0.19052 | 0.09814 |
| cyclic-$\mathrm{CH_2}$ | 14 | −1.08813 | 1.68945 | 0.73221 | 0.03156 | 0.16258 |
| cyclic-$\mathrm{CH_2}$ | 14 | −1.81834 | 1.42126 | 2.19871 | 0.00544 | 0.16037 |
| cyclic-$\mathrm{CH_2}$ | 14 | −2.02493 | −0.21051 | 2.24119 | 0.20691 | 0.12093 |
| N (sp3) | 14 | −0.90892 | −0.66573 | 1.33389 | −0.60598 | 0.00000 |
| $\mathrm{CH_3}$ | 15 | −1.02657 | −2.15300 | 1.03161 | 0.20676 | 0.11954 |
| nicotine | 162 | 0.04500 | 0.00421 | −0.05296 | 0.00000 | 0.58109 |

We are now ready to calculate all group parameters. First, the molecular structure parameters are computed according to eq 2:

$$\begin{aligned} r_{\text{nicotine}} &= 4R_{\text{ACH}} + R_{\text{N(sp2)}} + R_{\text{AC}} + R_{\text{c-CH}} + 3R_{\text{c-CH}_2} + R_{\text{N(sp3)}} + R_{\text{CH}_3} \\ &= (4 \times 0.537) + 0.436 + 0.316 + 0.497 + (3 \times 0.665) + 0.341 + 0.887 = 6.620 \end{aligned}$$

$$\begin{aligned} q_{\text{nicotine}} &= 4Q_{\text{ACH}} + Q_{\text{N(sp2)}} + Q_{\text{AC}} + Q_{\text{c-CH}} + 3Q_{\text{c-CH}_2} + Q_{\text{N(sp3)}} + Q_{\text{CH}_3} \\ &= (4 \times 0.431) + 0.339 + 0.114 + 0.235 + (3 \times 0.523) + 0.000 + 0.840 = 4.821 \end{aligned}$$

Next, the three contributions in eq 10 are determined individually:

$$\begin{aligned} \sum_k \Delta G^{*\text{chg},00}_{k/\text{W-OR}} &= 4\Delta G^{*\text{chg},00}_{\text{ACH/W-OR}} + \Delta G^{*\text{chg},00}_{\text{N(sp2)/W-OR}} + \Delta G^{*\text{chg},00}_{\text{AC/W-OR}} + \Delta G^{*\text{chg},00}_{\text{c-CH/W-OR}} + 3\Delta G^{*\text{chg},00}_{\text{c-CH}_2\text{/W-OR}} + \Delta G^{*\text{chg},00}_{\text{N(sp3)/W-OR}} + \Delta G^{*\text{chg},00}_{\text{CH}_3\text{/W-OR}} \\ &= (4 \times 0.468) - 1.301 + 0.031 - 0.051 + (3 \times 1.071) - 5.525 + 2.201 = 0.440 \end{aligned}$$

$$\begin{aligned} \sum_k C^e_k {e_k}^2 &= C^e_{\text{ACH}}({e_{\text{ACH}}}^2 + {e_{\text{ACH}}}^2 + {e_{\text{ACH}}}^2 + {e_{\text{ACH}}}^2) + C^e_{\text{N(sp2)}}{e_{\text{N(sp2)}}}^2 + C^e_{\text{AC}}{e_{\text{AC}}}^2 + C^e_{\text{c-CH}}{e_{\text{c-CH}}}^2 + C^e_{\text{c-CH}_2}({e_{\text{c-CH2}}}^2 + {e_{\text{c-CH2}}}^2 + {e_{\text{C-CH2}}}^2) + C^e_{\text{N(sp3)}}{e_{\text{N(sp3)}}}^2 + C^e_{\text{CH}_3}{e_{\text{CH}_3}}^2 \\ &= 0.135[0.263^2 + 0.273^2 + (-0.051)^2 + 0.105^2] + 0.003(-0.541)^2 + 5.699(-0.084)^2 - (0.093 \times 0.191^2) + 1.270(0.032^2 + 0.005^2 + 0.207^2) + 4.689(-0.606)^2 - (1.968 \times 0.207^2) = 1.752 \end{aligned}$$

$$\begin{aligned} \sum_k C^\mu_k {\mu_k}^2 &= C^\mu_{\text{ACH}}({\mu_{\text{ACH}}}^2 + {\mu_{\text{ACH}}}^2 + {\mu_{\text{ACH}}}^2 + {\mu_{\text{ACH}}}^2) + C^\mu_{\text{N(sp2)}}{\mu_{\text{N(sp2)}}}^2 + C^\mu_{\text{AC}}{\mu_{\text{AC}}}^2 + C^\mu_{\text{c-CH}}{\mu_{\text{c-CH}}}^2 + C^\mu_{\text{c-CH}_2}({\mu_{\text{c-CH}_2}}^2 + {\mu_{\text{c-CH}_2}}^2 + {\mu_{\text{c-CH}_2}}^2) + C^\mu_{\text{N(sp3)}}{\mu_{\text{N(sp3)}}}^2 + C^\mu_{\text{CH}_3}{\mu_{\text{CH}_3}}^2 \\ &= -0.955(0.135^2 + 0.137^2 + 0.171^2 + 0.182^2) + 0 \times 0 + 0 \times 0 + (8.892 \times 0.098^2) - 18.372(0.163^2 + 0.160^2 + 0.121^2) + 0 \times 0 - (55.315 \times 0.120^2) = -2.033 \end{aligned}$$

Then, we use eq 10 to obtain

$$\Delta G^{*\text{chg}}_{\text{nicotine/W}} - \Delta G^{*\text{chg}}_{\text{nicotine/OR}} = 0.440 + 1.752 - 2.033 = 0.159$$

Finally, substituting these parameters into eq 1 yields the octanol−water partition coefficient of nicotine:

$$\log_{10} K^{\text{GCSKOW}}_{\text{OW,nicotine}} = -0.126 + (1.031 \times 6.620) - (1.208 \times 4.821) + \frac{0.159}{1.364} = 0.99$$

The experimental value is 1.17.

## References and Notes

(1) Reid, R. C.; Prausnitz J. M.; Poling B. E. *The Properties of Gases & Liquids*; McGraw-Hill: New York, 1987.
(2) Kehiaian, H. V. *Fluid Phase Equilib*. **1983**, *13*, 243.
(3) Kehiaian, H. V.; Marongiu, B. *Fluid Phase Equilib*. **1985**, *21*, 197.
(4) Wu, H. S.; Sandler, S. I. *Ind. Eng. Chem. Res*. **1991**, *30*, 881.
(5) Wu, H. S.; Sandler, S. I. *Ind. Eng. Chem. Res*. **1991**, *30*, 889.
(6) Constantinou, L.; Gani, R. *AIChE J*. **1994**, *40*, 1697.
(7) Abildskov, J.; Constantinou, L.; Gani, R. *Fluid Phase Equilib*. **1996**, *118*, 1.
(8) Abildskov, J.; Gani, R.; Rasmussen, P.; O'Connell, J. P. *Fluid Phase Equilib*. **1999**, *160*, 349.
(9) Constantinou, L.; Prickett, S. E.; Mavrovouniotis, M. L. *Ind. Eng. Chem. Res*. **1993**, *32*, 1734.
(10) Lin, S. T.; Sandler, S. I. *Ind. Eng. Chem. Res*. **1999**, *38*, 4081.
(11) Ben-Naim, A. *Solvation Thermodynamics*; Plenum Press: New York, 1987.
(12) Tomasi, J.; Rersico, M. *Chem. Rev*. **1994**, *94*, 2027.
(13) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A., III; Honig, B. *J. Am. Chem. Soc*. **1994**, *116*, 11875.
(14) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005.
(15) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. *J. Phys. Chem. A* **1998**, *102*, 5074.
(16) Cramer, C. J.; Truhlar, D. G. *Chem. Rev*. **1999**, *99*, 2161.
(17) Kirkwood, J. G. *J. Chem. Phys*. **1934**, *2*, 351.
(18) Rivail, J. L.; Rinaldi, D. *Chem. Phys*. **1976**, *18*, 233.
(19) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, Jr., R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98, Revisions A.3 A.6*; Gaussian, Inc.: Pittsburgh, PA, 1998.
(20) Mulliken, R. S. *J. Chem. Phys*. **1955**, *12*, 1833.
(21) Sangster, J. *Octanol−Water Partition Coefficients: Fundamentals and Physical Chemistry*; Wiley and Sons: New York, 1997.
(22) Kamlet, M. J.; Doherty, R. M.; Abraham, M. H.; Marcus, Y.; Taft, R. W. *J. Phys. Chem*. **1988**, *92*, 5244.
(23) Suzuki, T.; Kudo, Y. *J. Comput.-Aided Mol. Des*. **1990**, *4*, 155.
(24) Isnard, P.; Lambert, S. *Chemosphere*. **1989**, *18*, 1837.
(25) *ClogP for Windows V2.0.0b*; Biobyte Corp.: Claremont, CA, 1997.
(26) Wienke, G.; Gmehling, J. *Toxicol. Environ. Chem*. **1998**, *65*, 57.
(27) Leo, A. *Chem. Rev*. **1993**, *93*, 1281.
(28) Hohenberg, P.; Kohn, W. *Phys. Rev*. **1964**, *136*, B864−B871.
(29) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1989.
(30) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem*. **1990**, *11*, 361.
(31) Sigfridsson, E.; Ryde, U. *J. Comput. Chem*. **1998**, *19*, 377.