Parametrizing PCM to Obtain Solvation Free Energies from Group Contributions. 1

Carmen Lucia de O. Mendes,† Clarissa O. da Silva,‡ and Edilson C. da Silva*,†

Departamento de Físico-Química, Instituto de Química, Universidade Federal do Rio de Janeiro,
Cidade Universitária, Centro de Tecnologia, Rio de Janeiro, RJ CEP-21949-900,Brazil, and
Departamento de Química-ICE, Universidade Federal RURAL do Rio de Janeiro,
Rodovia BR 465 km 7, Seropédica, RJ-CEP 23890-000, Brazil

Received: June 28, 2005; In Final Form: January 4, 2006

A parametrization methodology for evaluating the solvation free energy, using the polarizable continuum model implemented in Gamess software, is presented in a formulation which makes use of a group contribution conception to construct the cavities. The systems studied include alkanes, alcohols, aldehydes and ketones embebed in a continuous medium simulating the water as the solvent. For each family, the CH_2, OH, and C═O moieties of atoms are put together in single spheres forming a group. The cavities are constructed in two different ways, one for the electrostatic component and the other for nonelectrostatic contributions, i.e., the cavitation, dispersion, and repulsion components of free energy of solvation. A multivariate analysis is performed to obtain an assembly of variables, for each homologous series, able to give the results which are close to experiment. The analysis is addressed in order to (i) compare the theoretical free energy of solvation with the experimental trends of the solutes in aqueous media, when the chain is increased, (ii) compare the behavior of each component of free energy with the increasing CH_2 number, (iii) investigate the influence of the oxygen atom on the components, and (iv) quantify the relative contribution of each component to the final free energy of solvation for some homologous series.

1. Introduction

In recent years, there has been an increasing interest in group contribution methods for estimating some thermodynamic properties of pure substances and solutions when no experimental data is available.[1−7] In these methods, the molecule is formed by groups of atoms which 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 such groups.

All group contribution methods have common difficulties that limit their applicability, for instance: (a) different methods use different groups to represent the same molecules because the definition of groups is empirical and arbitrary; (b) simple group contribution methods do not distinguish between isomers; (c) the inclusion of electrostatic interactions can largely change the definition of the part of the molecule assumed to be a group.

Over the last years theoretical methods have been developed for the calculation of fragment contributions to the solvation free energy, particularly in the framework of quantum mechanical continuum solvation methods. Cramer and Thrular[8−11] developed fractional methods based on GB/SA model in order to obtain solvation free energy values for systems in aqueous solution and in organic solvents, and more recently[12] they have used this model—the series SM5.*x*—to perform solubility calculations of liquid and solid solutes in aqueous solution. Luque et al.[13,14] have developed and parametrized similar algorithms by using the Miertus−Scrocco−Tomasi (MST) model for the solvent. They have recently examined the group contributions to the solvation free energy in water and in octanol[15] for different solutes, all of them relating to tacrine. In their work they tested the influence of the normalization charges process and the transferability of data to calculations based on group contributions, within a series of structurally similar molecules.

Sandler et al.[16,17] have combined the PCM method to another group contribution method in order to obtain partition coefficients for the octanol/water heterogeneous system for many different solutes. They have observed that their model could be applied to solutes with very polarizable functional groups that suffer very much the influence of the neighboring groups. In another work,[3] Sandler et al. have demonstrated, from the determination of Henry's constant through calculations of the solvation free energy, that the group contributions are influenced by the type of neighboring group (i.e., the type of functional group when it is located in different molecules or in different positions on the same molecule). By having use of a multipole correction method in order to correct for intramolecular proximity effects, they have obtained results that are better than those obtained by simple group contribution methods in which little or no effect of neighboring groups are taken into consideration.

Among many different models dedicated to describing the solvent effects, the polarizable continuum model (PCM)[18−21] has been widely used since its appearance in 1981. Since then, several other extensions have been published and a breakthrough occurred in 1994,[22,23] when the first and second derivatives of the energy were obtained and allowed analytical calculations for geometry optimizations, frequency and reaction path calculations in solutions.[24−26] The model is currently used to study the absorption and emission spectra[27] of some amino acids in solution,[28,29] solvent effects on nuclear magnetic shielding parameters,[30] optical rotations of chiral molecules,[31] and so

---

* Corresponding author. Fax: +55-21-25627265. E-mail: ecds@chaer.iq.ufrj.br.

† Universidade Federal do Rio de Janeiro.

‡ Universidade Federal RURAL do Rio de Janeiro.

10.1021/jp053520v CCC: $33.50

on.[32,33] There are several applications of PCM, in its different versions, involving different types of solvents: in nonelectrolyte solutions,[5,14,34−42] in liquid crystals[43] and in ionic solutions.[44−46] A. A. C. C. Pais et al.[47] have described the behavior of the different components of the solvation free energy for alkanes and alcohols in aqueous solution by using PCM, considering many different basis sets.

Hence, considering all the relevant results that can be found in the literature while adopting this versatile and robust model and the possibility of using the solvation free energies to obtain thermodynamic information relating to the system's macroscopic behavior such as activity coefficient, solubility, or pK_a, the PCM model has been chosen to describe the solvent effects and the idea of group contributions to defining the cavity and the solvation energy values of each moiety of the system identified as a group. The final purpose of this method of parametrization is to predict, through group contributions, properties of new and more complex systems, which are not possible to be treated, at least as yet, at a quantum mechanical level.

2. Computational Details

The geometries of solute molecules have been optimized in a vacuum using Gaussian 98 software.[48] All calculations, either in a vacuum or in water, have been performed using the Hartree−Fock method and 6-31G(d,p) basis set.

In the PCM framework, the solvation free energy (G_{solv}) of a solute immersed in a polarizable continuum can be written, neglecting the molecular movement, as a contribution of four components:

$$G_{\mathrm{solv}} = G_{\mathrm{el}} + G_{\mathrm{cav}} + G_{\mathrm{disp}} + G_{\mathrm{rep}} \tag{1}$$

Since its first formulation, PCM has exploited cavities built with interlocking spheres centered on atoms, or group of atoms, according to the well-known GEPOL93[49] procedure. The terms in eq 1 are calculated with slightly different definitions of the solute cavity. For the electrostatic component, G_{el}, the best surface—"excluding surface—SE"—is formed starting from the van der Waals atomic radii scaled by an "α" factor, taking into account the first solvation layer.

The electrostatic component, G_{el}, is obtained self-consistently from

$$G_{\mathrm{el}} = \langle \Psi_{\mathrm{el}} | H^0 | \Psi_{\mathrm{el}} \rangle + \frac{1}{2} U_{\mathrm{el}} \tag{2}$$

where U_{el} refers to the solute−solvent interaction electrostatic potential. The solvent is represented by a set of charges located on the surface of the cavity (SE) that hosts the solute molecule. Thus, the cavity shape directly affects the solute−solvent interaction potential.

The cavitation energy is computed by using a van der Waals (vdW) surface formed by spheres centered on atoms or groups of atoms with proper radii, and is not computed by a quantum-mechanical treatment; it is obtained from Pierrotti−Claverie's formula to which the scaled particle theory (STP) is applied.[50] In this framework, G_{cav} is expanded in powers of R_{MS}, i.e., the radius of the sphere which excludes the centers of the solvent molecules ($R_{\mathrm{MS}} = R_{\mathrm{M}} + R_{\mathrm{S}}$, where R_{M} and R_{S} are the radii of the solute and solvent, respectively):

$$G_{\mathrm{cav}} = K_0 + K_1 R_{\mathrm{MS}} + K_2 R_{\mathrm{MS}}^2 + K_3 R_{\mathrm{MS}}^3 \tag{3}$$

Claverie et al.[51] have proposed the following expression to consider the fact that the solute molecules are described as overlapping spheres:

$$G_{\mathrm{cav}} = \sum_{i=1}^{N} ((A_i)/(4\pi R^2)) G_{\mathrm{cav}}(R_i) \tag{4}$$

where each sphere with radius R_i contributes with a weight depending on the portion of the surface which is exposed to the solvent (A_i). The K_i coefficients are defined in Persico and Tomasi's review.[18]

The dispersion and repulsion terms are calculated with adopting the "solvent accessible surface—SAS " i.e., a vdW surface with the radii augmented by the solvent radius. According to Floris,[52,53] the following expression is used for calculating the disp-rep component of free energy.

$$\begin{aligned} G_{\mathrm{disp-\ rep}}(\mathrm{M\ in\ S}) = {} & \\ \rho_S \sum_{\mathrm{s}\in\mathrm{S}} N_{\mathrm{S}} \sum_{\mathrm{m}\in\mathrm{M}} \Bigg[\sum_{6} - {d_{\mathrm{ms}}}^6 \int \frac{g_{\mathrm{ms}}\,(r_{\mathrm{ms}})}{{r_{\mathrm{ms}}}^6}\, \mathrm{d}r_{\mathrm{ms}} + {} & \\ c_{\mathrm{ms}} \int g_{\mathrm{ms}}(r_{\mathrm{ms}}) \exp(-\,\gamma_{\mathrm{ms}}\, r_{\mathrm{ms}})\, \mathrm{d}r_{\mathrm{ms}} \Bigg] & \end{aligned} \tag{5}$$

where ${d_{\mathrm{ms}}}^6$ is the atom−atom dispersion coefficient of order 6, N_{s} is the number of atoms of type s for each molecule of the solvent S, ρ_{S} is the macroscopic numeral density of solvent, the r_{ms} distance is $r_{\mathrm{ms}} = r_{\mathrm{s}} - r_{\mathrm{m}}$, where m stands for the solute subunits and s for the solvent subunits, γ_{ms} is a repulsion coefficient taken from the literature, and g_{ms} is the correlation function, which depends on the m−s distance only.

2.1. Parametrization Details. The calculations were performed by using the PC GAMESS version[54] of the GAMESS (US) QC package[55] in order to obtain the free energy of solvation and its components through a single point calculation, in aqueous medium, by applying the D-PCM at 298.15 K.

Initially, the first step of parametrization was performed with all atoms of the solute having individual spheres, including the hydrogens. The solutes considered in this first part belong to the class of alkanes (methane, ethane, *n*-propane, *n*-butane and *n*-pentane) and alcohols (methanol, ethanol, *n*-propanol, *n*-butanol, and *n*-pentanol). The parametrization protocol involved simultaneous changes in the following parameters: radius and α factors for all the atoms of solutes and of "r_{diff}" parameters (Gamess parameter = distances between the first atoms of each type and the cavity) and of "r_{w}" parameter (Gamess parameter = array of atomic radii for the solute to compute dis-rep) related to the SAS. The radius values considered for them were based on the van de Waals values (C = 1.70, 1.72, and 1.73 Å; O = 1.40, 1.50, and 1.60 Å and finally H = 1.0, 1.2, and 1.3 Å). The initial values considered for "rdiff" parameter were for O = 1.4, 1.5, and 1.6 Å and for H = 1.0, 1.2, and 1.3 Å, with all possible combinations being taken into account. Finally, the α factors used were for C = 1.2 and 1.4, and for both O and H = 1.0 and 1.2. By combination of all of these possibilities, 324 input files were prepared for each alkane and 972 for each alcohol molecule considered in this work. It is important to stress that the solute radius alters all the solvation free energy components; the α factors only alter the electrostatic component and the "r_{diff}" parameters alter only the dispersion and repulsion components and are related to the solvent properties.

In the multivariate analysis, a matrix was constructed where each line was a run for a particular solute and each column was a variable. To this matrix a column was added in order to report the errors between the theoretical ΔG_{solv} and the experiments. After ordering each line of this matrix by the error column values, it was possible to select the lines (or sets of variables) that have the ΔG_{solv} values closer to experiment.

Unfortunately, this methodology was not able to provide a unique set of optimal variables comprising all the members of the same family (alkanes or alcohols).

Besides that, this protocol has proved to be inefficient since the set of values selected, for different groups of solutes in the same family, provided ΔG_{solv} values that are aligned in a homologous series by a very different slope from that obtained when the experimental values are plotted vs the number of C atoms.

At a second step, the protocol was modified in order to account for such experimental profile of the solvation energy in a homologous series. The solute molecules of alkanes (*n*-propane, *n*-butane, *n*-pentane, *n*-hexane, *n*-heptane and *n*-octane) were used in this stage. The cavities were constructed using the group contribution idea as follows: each one of CH_2 (and OH for alcohols) groups is included in a single sphere. By transferring the best collection of variables adjusted in the first step for alkanes, we see that the only variable now adjusted is related to a new group, the CH_2. It is important to notice that the set of solutes are different in both steps, and so the predictive ability of this parametrization was here indirectly tested for *n*-hexane, *n*-heptane, and *n*-octane.

As alcohol molecules are different from alkane molecules only by replacement of one hydrogen atom with one OH group, the values of all the variables already obtained for alkanes were maintained and only the OH group variables parametrized. This procedure was performed for the remaining families, i.e, the variables relating to the carbon and oxygen atoms belonging to C═O were adjusted and so were those for the hydrogen atom bonded to C═O.

It is important to stress that the parametrization protocol was performed for a set of solute molecules of alkanes, used to only obtain the best values for CH_2 group. These values were fully transferred for alcohols, aldehydes and ketones and kept frozen, with only those corresponding to the cavities for OH and C═O, respectively, being adjusted. However, it is important to mention that every time that such groups underwent a strong neighboring effect, such CH_2 group was modified, as better explained in the following paragraphs.

TABLE 1: Best Conditions[a]

family	atom or group	on	bonded to	*R* (Å)	α
n-alkanes	C	CH_3	CH_2	1.700	1.390
	H	CH_3	CH_2	1.050	
	CH_2		CH_3 or CH_2	1.900	1.390
n-alcohols	C	CH_2	OH	1.700	1.390
	H	CH_2	OH	1.050	
	OH		CH_2	1.700	1.060
snd-alcohols	C	CH	OH	1.700	1.200
	H	CH	OH	1.080	
	CH_2		CH_3 or CH_2	1.900	1.380
	OH		CH	1.700	1.060
	CH_2		CHOH	1.900	1.380
aldehydes (type I)	C	HC═O	CH_2	1.700	1.390
	H	HC═O	CH_2	1.050	
	O	HC═O	CH_2	1.700	1.016
aldehydes (type II)	C═O	HC═O	CH_2	1.700	1.124
	H	HC═O	CH_2	1.050	1.010
ketones (type III)	C	C═O	CH_3 and CH_2	1.680	1.600
	O	C═O	CH_3 and CH_2	1.700	0.974
	CH_2		C═O	1.900	1.140
ketones (type IV)	C═O		CH_3 and CH_2	1.900	1.390
	CH_2		C═O	1.900	1.142

[a] Type I, for aldehydes, and type III, for ketones, respectively correspond to cavities, where carbon and oxygen atoms are surrounded by two separated spheres; type II carbon and oxygen atoms are surrounded by one sphere centered on the oxygen atom; type IV carbon and oxygen atoms are surrounded by one sphere centered on the carbon atom.

3. Results

3.1. The "Best" Conditions and the Second Protocol. The multiple parametrization results obtained with the second protocol are presented in Table 1. They were obtained under the following conditions as set forth in the first protocol by using the Gamess software:

(i) $r_{diff} = 1.00$ Å for H and 1.50 Å for O,

(ii) $r_w = 1.00$ Å for the H atoms.

(iii) The values of all the other variables not specified were set to their default values in the program.

This set of optimum parameters is associated with a different model for constructing the cavity. In this model, only individual spheres centered on the heavy atoms form the electrostatic cavity and the cavities associated with the other components of the solvation energy are constructed using the aforementioned second protocol. Figure 1 depicts what such new cavities are like.

3.2. Discussion of the Parametrization. *3.2.1. For the Alkane Family.* As the α factor for carbon atoms is equal to 1.39, the spheres involving the carbon atoms of the CH_3 groups are large; so the hydrogen atoms of the CH_3 groups are incorporated into the carbon sphere. Thus, in the calculations of the electrostatic component, such hydrogens are not individually exposed to the solvent.

The α parameter is a correction factor applied only to the electrostatic components of ΔG_{solv}. Thus, for the cavitation, dispersion and repulsion components, the H atoms of CH_3 began to hold individual spheres. This happens because the spheres containing carbon atoms have radius equal to 1.70 Å, and as they are not α-corrected, they are not large enough to encompass the H atoms. The hydrogen atoms of the CH_2 groups, on the other hand, are included in one individual sphere centered on the C atom, since they are treated as a group, in this protocol.

3.2.2. For the Alcohol Family. The same cavity definition as already set forth for the alkanes is adopted for alcohols. The only difference is related to the OH, here considered as an individual group with its own sphere (Figure 1, parts b and d). All the parameters were fully transferred, but an exception is made for the CH_2 bonded to OH group; in this case the OH group replaces the H of a CH_3 of alkanes, and thus the hydrogens of CH_2OH must have separate spheres.

In secondary alcohols, the CH_2 group has an unusual behavior: the α factor for the CH_2 group is equal to 1.38. Although this value does not seem to be very different than 1.39, if one uses the latter values for C atoms of CH_2 in secondary alcohols, the final value of ΔG_{solv} is far from the experimental one. This result indicates that there are differences in the CH_2 group behavior depending on the position of the OH group. Another important difference present in secondary alcohols refers to the radius value for the H bonded to C atom on CHOH. It is slightly greater (1.08 Å) than the values used in all other cases (1.05 Å).

3.2.3. For Aldehyde and Ketone Families. Again, all the previous parameters were transferred from alkanes and alcohols to aldehydes and ketones. However, two different types of cavity are constructed in each case (types I and II for aldehydes and types III and IV for ketones), as shown in Table 1. For

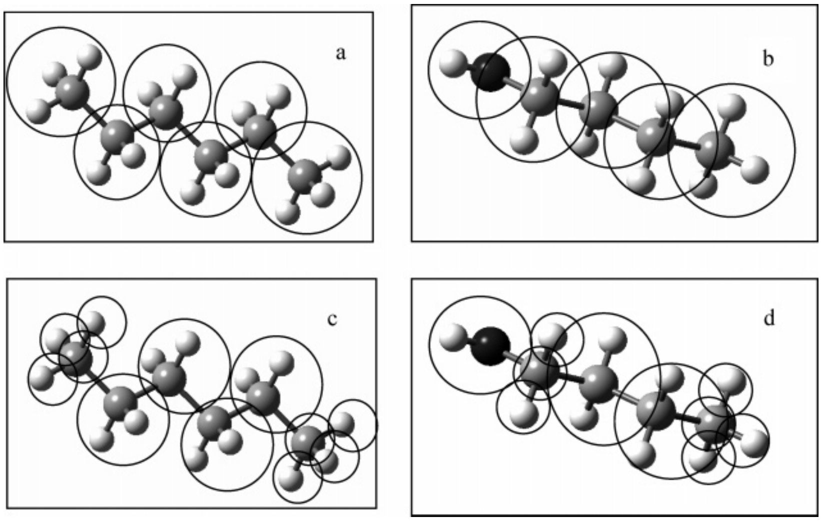

Figure 1. Electrostatic cavities for alkanes (a) and alcohols (b). Cavitation, dispersion and repulsion cavities for alkanes (c) and alcohols(d).

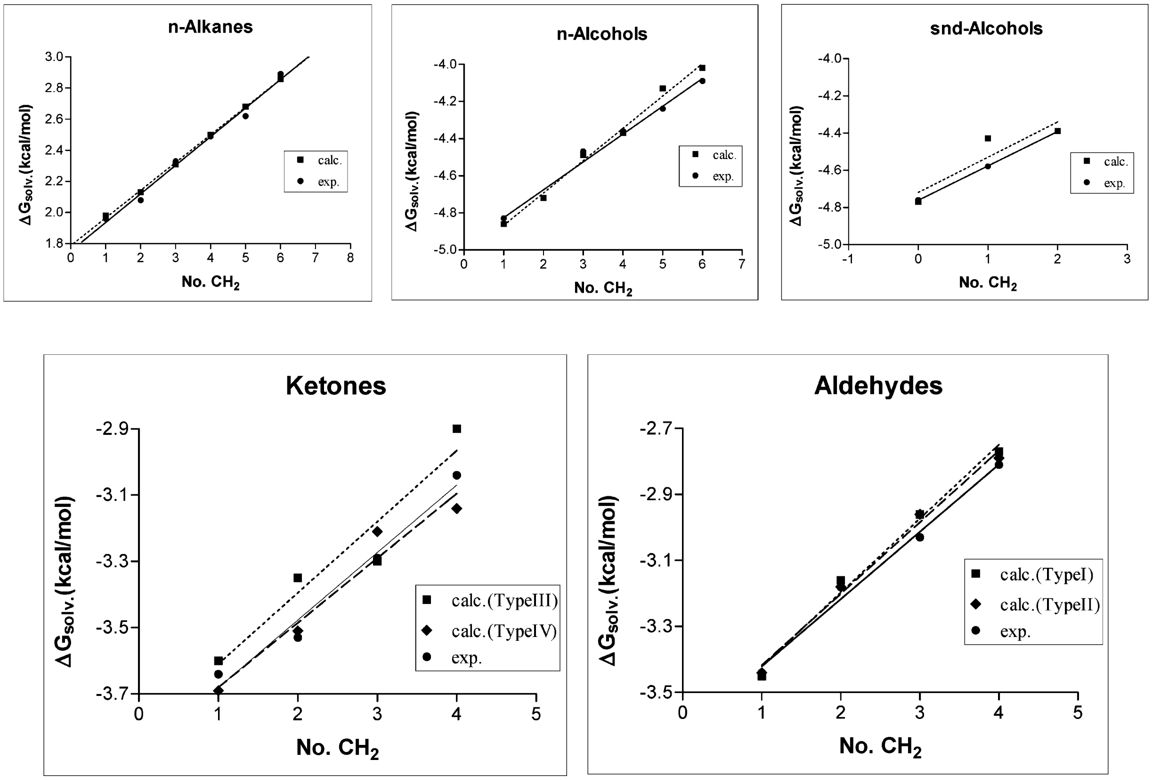

Figure 2. Relationship between the number of CH_2 groups in the chains and the solvation free energy in aqueous solution, at 298.15 K. (●−●) experimental and (■- - - -■) theoretical, respectively.

aldehydes, only the α factor for oxygen atoms (1.016 for type I and 1.124 for type II) is different than that of *n*-alcohols (1.06). In one case, the smaller value of α indicates the contraction of a double bond. In the other case, it is reasonable that the α is larger due the inclusion of the carbon atom into the oxygen sphere. In type II aldehydes the hydrogens bonded to $C=O$ have own spheres, different than those of type I, as the spheres of the oxygens are not big enough to include the hydrogens. These two types of cavities give very similar results, as will be seen later.

In the case of type III ketones, the radius of the carbon atom of $C=O$ is equal to 1.68 Å, very close to the other carbon atoms. However, a large α factor (1.60) was obtained for this carbon atom. Because the oxygen atom and the CH_3 and CH_2 groups surround this carbon atom, if α is lower it will be wrapped by the other atoms and its influence will not be seen. On the other hand, the oxygen atom has an α factor equal to 0.974, which is very different than other α factors for oxygen. Here again, if this α is greater than that obtained, the oxygen sphere will largely wrap the carbon atom bonded thereto.

The presence of the oxygen in the middle of the chain alters the chemistry of the nearest group. One can notice that values of the parameters found for the CH_2 groups vicinal to the carbonyl in ketones, are different than those used for the same group vicinal to carbonyl of aldehydes. In the latter case, the parameters were the same as those adopted for alkanes and alcohols. Thus, in the aldehydes this does not occur and in secondary alcohols it does, but on a minor scale.

In the second case (type IV), the sphere surrounding the $C=O$ group is centered on the C atom. The values found are reported in Table 1 and clearly very different than type II, as could be expected, since the proposed cavities are very different.

TABLE 2: Results of Curves $\Delta G_{solv} = a$ (Number of CH_2) $+ b$

family	calculated *a*	calculated *b*	experimental *a*	experimental *b*
n-alkanes	0.178	1.786	0.184	1.752
n-alcohols	0.174	−5.040	0.150	−4.977
snd-alcohols	0.190	−4.72	0.185	−4.762
aldehydes (I)	0.224	−3.645	0.204	−3.625
aldehydes (II)	0.217	−3.635	0.204	−3.625
ketones (III)	0.215	−3.825	0.204	−3.885
ketones(IV)	0.195	−3.875	0.204	−3.885

An important consequence is that the CH_2 group vicinal to the carbonyl is also modified as compared to the other hydrocarbons.

3.3. Consequences of Parametrization. *3.3.1. Influence of CH_2 Number on Solvation Free Energies and on Its Components.* By using the optimized input variables for all alkanes (C3 to C8) it was possible to obtain the results tabulated in Table 2 and illustrated in Figure 2. All the experimental data are found in refs 56 and 57.

The regression coefficient of the curves indicates good estimates of ΔG_{solv} for the parametrized series. A possible deviation of some results in a homologous series can be explained: the multiple parametrization is a complex process due to an enormous way of combinations which can give a result close to the experimental one. Additionally, the best condition for a solute/solvent system is not necessarily the best condition for another member of the homologous series. Thus, the final result is a combination of various input variables applicable to all the solute molecules belonging to the same family.

The best conditions for PCM inputs, optimized by a multivariate analysis, are obtained by using the experimental data reported in the literature as the reference values. More relevant than obtaining ΔG_{solv} values matching the experimental values, was obtaining results which follow the experimental trend, as can be seen in Table 2 and Figure 2.

The range of errors between calculated and experimental values are as follows: for alkanes, −1% to 3%; for alcohols, −3% to 1%; for secondary alcohols, −3% to 0%; for aldehydes, −2% to 0% (types I and II); for ketones, −5% to 0% (type III) and −2% to 3% (type IV). These results indicate a good agreement between the two curves—calculated and experimental.

As a means of checking the consistency of the results, the experimental and calculated curves were extrapolated. In the case of secondary alcohols, extrapolation for number of CH_2 = 3 (2-hexanol) the values obtained from each curve (−4.21 kcal/mol for the experimental and −4.15 kcal/mol for the theoretical) are in good agreement with the percent error of −1,4%, a value belonging to the error range drawn for this family. In the case of aldehydes by extrapolating the theoretical curve to number of $CH_2 = 6$ (*n*-octanal) a value of −2.30 kcal/mol is obtained when using type I cavity and of −2.33 kcal/mol for type II. Both of these values are close to the experimental value: −2.29 kcal/mol.[56]

Again, in the ketone series, by extrapolating the theoretical curves to number of $CH_2 = 5$ (octanone), the ΔG_{solv} values obtained are −2.75 kcal/mol for type III and −2.90 for type IV. Both values are close to the experimental value: −2.88 kcal/mol.[56]

3.3.2. Growth of Chain and Its Influence on the Free Energy Components. Plotting the values of each solvation free energy component vs the number of CH_2 groups shows that, except for the electrostatic component, all the rest are linear functions of the number of CH_2, as seen in Figure 3 (already discussed in ref 47).

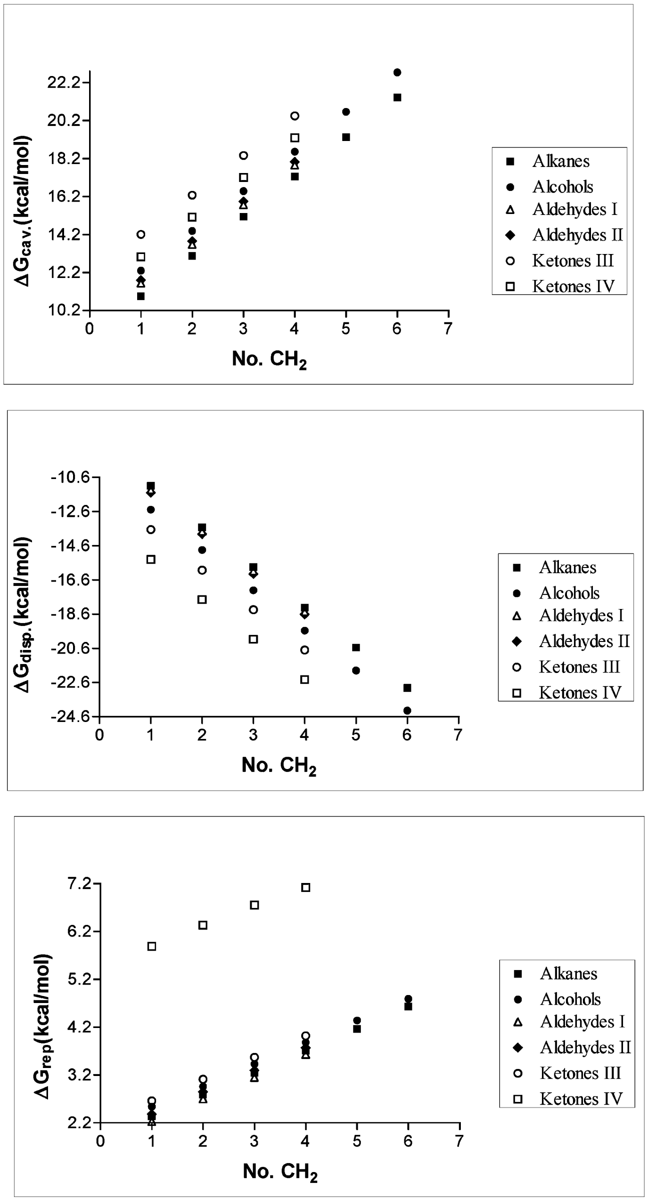

Figure 3. Plot of ΔG_{cav}, ΔG_{disp}, and ΔG_{rep} vs number of CH_2 groups. All linear regression coefficients are $r^2 = 1.000$.

It is clear from the aforementioned figures that the influence of oxygen atom on the chains is more pronounced in ketones for all but the electrostatic free energy components. For type IV ketones, where carbon and oxygen are involved in the same sphere, the ΔG_{rep} plot shows that the difference to be more pronounced. This can be explained as follows.

For type IV ketones, the oxygen atom is included in the carbon sphere and therefore, in the dispersion and repulsion calculations, its area is equal to zero. In the case of type III ketones, the oxygen has a large area (~40 to 50 Å^2) which encompasses much of the carbon bonded thereto (~3 Å^2 area). As shown in eq 5 and in its related expressions,[52,53] repulsion is in proportion to

$$-c_{ms}\left[\frac{1}{\gamma_{ms}r_{ms}}+\frac{1}{\gamma_{ms}r_{ms}^{2}}+\frac{1}{\gamma_{ms}r_{ms}^{3}}\right]\exp(-\gamma_{ms}r_{ms})\mathrm{r}_{ms}$$

should the solute radius (r_m) increase, G_{rep} will decrease, which explains what happens to the results obtained for the repulsion components in cases III and IV for ketones. As regards aldehydes this would not occur because in those of type II,

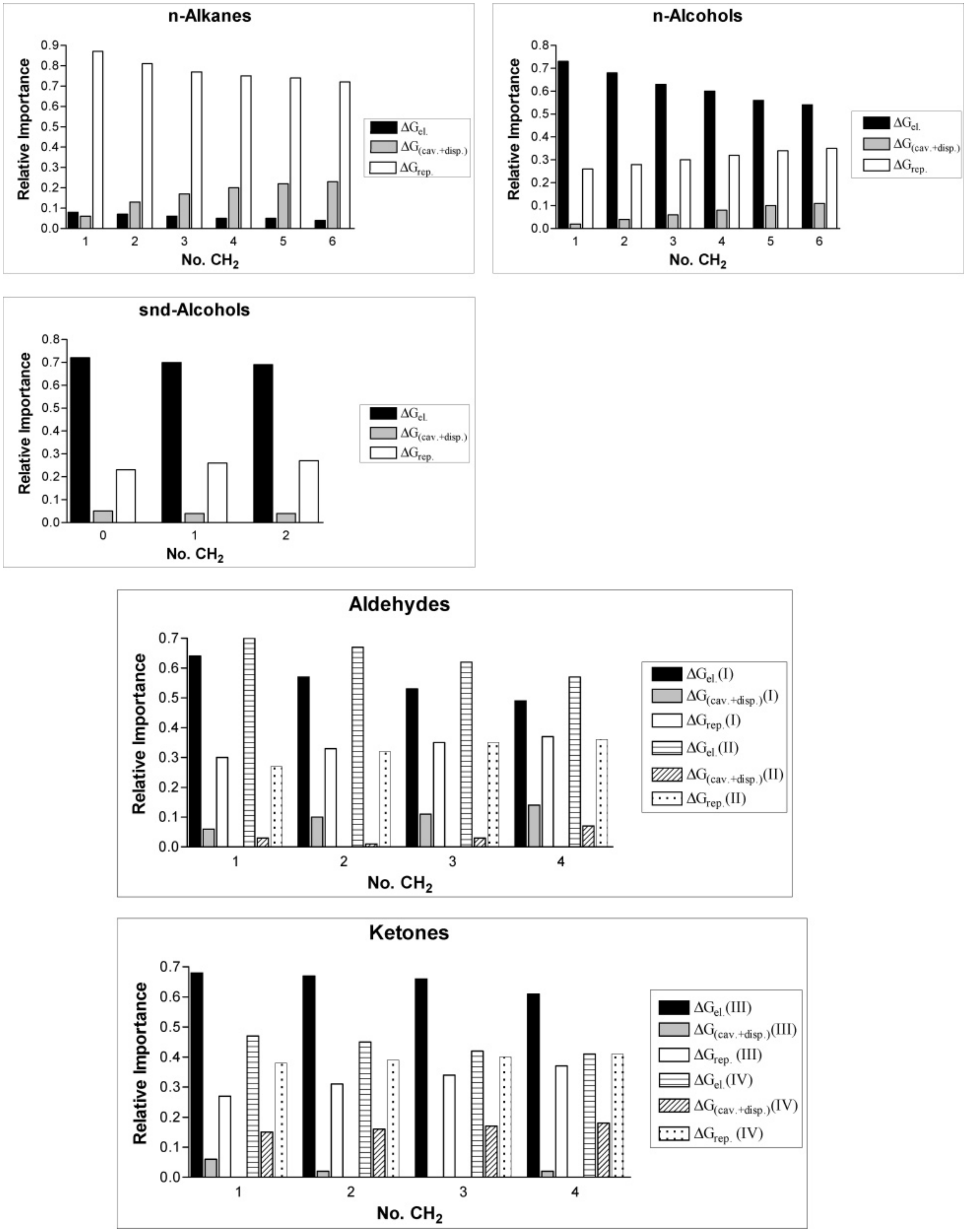

Figure 4. Relative importance of the components of ΔG_{solv}.

where a single sphere centered on the oxygen atom surrounds the C═O group, the areas of the spheres are allocated to the oxygen. In type I aldehydes, the oxygen is surrounded by a sphere and the carbon bonded thereto by another one. The comparison of the areas used for calculating ΔG_{disp} and ΔG_{rep} shows that in both cases the cavities of the oxygen atoms have similar areas, and for this reason the two types of cavities generate similar results for the aldehydes.

Taking into account the previous results, it seems quite interesting to adopt type II and type IV cavities as those to be used for aldehydes and ketones, respectively. In both parametrizations, oxygen atom is part of the C═O group, the difference being that in the former the sphere is centered on the oxygen (type I) and in the latter, on the carbon (type IV).

3.3.2.1. Oxygen Influence on Electrostatic Component. Considering that the relationship between of the ΔG_{el} and the

TABLE 3: Oxygen Influence on the Electrostatic Components of ΔG_{solv}[a]

family	mean of $\Delta G_{el.diff.}$	std dev
alcohols	−6.99	0.02
snd-alcohols	−7.57	0.20
aldehydes (type I)	−5.62	0.11
aldehydes (type II)	−5.78	0.08
ketones (type III)	−6.52	0.10
ketones (type IV)	−6.92	0.06

[a] Mean of $\Delta G_{el.diff.}$ is the arithmetic mean of the $\Delta G_{el.dif.}$ of all members of each family and std dev is the standard deviation in relation to the mean value.

number of CH_2 groups in a homologous series is not a linear function of the number of CH_2, it was necessary to express the oxygen influence as a difference ($\Delta G_{el\ diff}$) between ΔG_{el} of each solute of the alcohol family and each corresponding alkane. This procedure was employed for each family, and the results are reported in Table 3, where the mean values of $\Delta G_{el\ diff}$ for each family is calculated; the standard deviations are also reported. The data show that the influence of oxygen is the same for all solutes in the same family.

One can see from Table 3 that $\Delta G_{el\ diff}$ decreases in that order: aldehydes > ketones > *n*-alcohols > snd-alcohols. This behavior can be explained by analyzing the values of the charges of oxygen atom obtained in the ab initio calculations for aldehydes, ≈−0.55 au; for ketones, ≈−0.59 au; for *n*-alcohols, ≈−0.70 au, and for snd-alcohols, ≈−0.71 au.

3.3.3. Relative Contributions of the Free Energy Components (ΔG_i). To evaluate the relative importance of the free energy components the following methodology was applied. At first, the values of ΔG_{cav} and ΔG_{disp} were added into one term, $\Delta G_{(cav+disp)}$. Next, the absolute values of ΔG_{el}, $\Delta G_{(cav+disp)}$ and ΔG_{rep} were added, and the relative importance of these components was calculated. The first step can be justified by the similar absolute values of the components of ΔG_{solv}. Figure 4 depicts these results as relative contributions to all the systems under study.

For the alkane family, the electrostatic component remains virtually unchanged as the chain increases and its relative importance is approximately 0.08, while the contribution of the repulsive component lies between 0.8 and 0.9. As the solute is an apolar system and the solvent, a medium which simulates a polar compound, the high relative contribution of this term is explained. The term ($\Delta G_{cav+disp}$) has a slightly increasing importance in the homologous series, from 0.06 to 0.22, showing that as the carbon chain increases, the cavitation component becomes more important than the dispersion one. The results support the findings by Cheong and co-workers[58] when they say that, "*the only possible attraction force between an alkane and a solvent are dispersive interactions; so* $\Delta G_{solv} = \Delta G_{cav} + \Delta G_{int}$."

For the primary alcohols, the electrostatic component plays a key role due to the presence of the oxygen atom, and its relative importance decreases as the size of the aliphatic chain increases. The relative importance of ($\Delta G_{cav+disp}$) component is small when the number of CH_2 groups is small (approximately 0.02) and increases up to 0.1 for *n*-heptanol and *n*-octanol. The term ΔG_{rep} increases from 0.26 to 0.35 in the homologous series.

For the secondary alcohols, all the relative contributions do not seem to be affected by the size of the system. Calculations performed with 2-hexanol give support to these observations but are not included in Figure 4 (for secondary alcohols) because there was no experimental value to compare with.

The relative contribution of ΔG_{solv} components was computed for both type I and II cavities for the aldehyde family. For both types of cavities the electrostatic component is the most important one and its relative contribution does not change very much in the homologous series. The ($\Delta G_{cav+disp}$) contribution seems to vary randomly for type II aldehydes and are small for both types and, finally, the term ΔG_{rep} remains virtually unchanged in both situations.

For type III ketones, the electrostatic term has a high relative importance around 0.70, while this value for type IV is about 0.45. The ($\Delta G_{cav+disp}$) term is virtually negligible for type III ketones. For those of type IV, its relative importance is constant around 0.17. The repulsive term grows slowly as the size of the chain increases for both types III and IV.

4. Conclusion

This paper describes the methodology and the results obtained from a PCM parametrization based on the group contribution conception—in the sense of constructing some cavities, assuming that CH_2, OH, and C═O are groups of atoms, for calculating solvation free energy for the homologous series of alkanes, alcohols and aldehydes. The protocols and results are reported herein and discussed in details. With these features, the experimental results and their trends were well reproduced for all the families under study. This achievement is believed to be due to a new model of cavity construction; one type for electrostatic components, where the hydrogens do not have individual spheres and, consequently, have no areas to account for ΔG_{el}, and the other for the cavitation, dispersion, and repulsion components, where the hydrogens have their own spheres and are considered in the free energy components.

The variation of the free energy components with the number of CH_2 groups and the influence of the oxygen on the ΔG_{el} for the oxygenated solutes compared to alkanes, are also reported and a global picture is provided about the behavior of the cavity model in a homologous series.

Another contribution of this work is concerned to the relative importance of each component to the total free energy. As reported, the electrostatic component has the major contribution to polar solutes, except for the alkane family and that the repulsion component is accentuated for type IV ketones. Generally speaking, the $\Delta G_{(cav+disp)}$ terms have little influence.

The authors are confident that, once extended to other solute families, this parametrization may be applied to complex systems as oligipeptides, with a good estimation, precluding the necessity of performing any computational calculation.

It is important to recall that the predictive character of the parametrization here proposed was also observed, and the results found have motivated the authors to extend it to other families in other solvents.

References and Notes

(1) Lin, S T.; Sandler, S. I. *J. Phys. Chem. A* **2000**, *104*, 7099.
(2) Klamt, A.; Eckert, F. *Fluid Phase Equilib.* **2000**, *172*, 43.
(3) Lin, S. T.; Sandler, S. I. *Chem. Eng. Sci.* **2002**, *57*, 2727.
(4) Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 835.
(5) Cappelli, C.; Mennucci, B.; da Silva, C. O.; Tomasi, J. *J. Chem. Phys.* **2000**, *112*, 5382.
(6) Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 71.
(7) Nanu, D. E.; Mennucci, B.; de Loos, T. W. *Fluid Phase Equilib.* **2004**, *221*, 127.
(8) Giesen, D. J.; Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1997**, *98*, 85.
(9) Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *288*, 293.
(10) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1998**, *102*, 3257.

(11) Winget, P.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2000**, *104*, 4726.

(12) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **2003**, *119*, 1661.

(13) Luque, F. J.; Zhang, Y.; Alemán, C.; Bachs, M.; Gao, J.; Orozco, M. *J. Phys. Chem.* **1996**, *100*, 4269.

(14) Curutcher, C.; Orozco, M.; Luque, F. J. *J. Comput. Chem.* **2001**, *22*, 1180.

(15) Soteras, I.; Morreale, A.; López, J. M.; Orozco, M.; Luque, F. J. *Braz. J. Phys.* **2004**, *34*, 48.

(16) Sandler, S. I.; Lin, S.-T.; Sum, A. K. *Fluid Phase Equilib.* **2002**, *194−197*, 61.

(17) Sandler, S. I. *Fluid Phase Equilb.* **2003**, *210*, 147.

(18) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.

(19) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.

(20) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.

(21) Tomasi, J. *Theor. Chem. Acc.* **2004**, *112*, 184.

(22) Cammi, R.; Tomasi, J. *J. Chem. Phys.* **1994**, *101*, 3888.

(23) Cammi, R.; Tomasi, J. *J. Chem. Phys.* **1994**, *100*, 7495.

(24) Cancès, E.; Mennucci, B. *J. Chem. Phys.* **1998**, *109*, 249.

(25) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1998**, *109*, 260.

(26) Cammi, R.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1996**, *104*, 4611.

(27) Mennuci, B. *J. Am. Chem. Soc.* **2002**, *124*, 1506.

(28) Mennucci, B.; Toniolo, A.; Tomasi, J. *J. Chem. Phys. A* **2001**, *105*, 4749.

(29) Mennucci, B.; Toniolo, A.; Tomasi, J. *J. Chem. Phys. A* **2001**, *105*, 7126.

(30) Mennucci, B.; Martínez, J. M.; Tomasi, J. *J. Phys. Chem. A* **2001**, *105*, 7287.

(31) Mennuci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Gabriel, S.; Stephens, P. J. *J. Phys. Chem. A* **2002**, *106*, 6102.

(32) Martínez, J. M.; Papparlado, R. R.; Marcos, E. S.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. B* **2002**, *106*, 1118.

(33) Tomasi, J.; Cammi, R.; Mennucci, B.; Cappelli, C.; Corni, S. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5697.

(34) Curutcher, C.; Bidon-Chanal, A.; Orozco, M.; Luque, F. J.*Chem. Phys. Lett.* **2004**, *384*, 299.

(35) Klant, A.; Eckert, F. *Fluid Phase Equilib.* **2000**, *172*, 43.

(36) Klamt, A. *Fluid Phase Equilib.* **2003**, *206*, 223.

(37) Klant, A.; Eckert, F. *Fluid Phase Equilib.* **2004**, *217*, 53.

(38) Diedenhofen, M.; Eckert, F.; Klamt, A. *J. Chem. Eng. Data* **2003**, *48*, 475.

(39) Mennucci, B.; Cossi, M.; Tomasi, J. *J. Phys. Chem.* **1995**, *102*, 6837.

(40) Cappelli, C.; da Silva, C. O.; Tomasi, J. *J. Mol. Struct.* **2001**, *544*, 191.

(41) Chuman, H.; Mori, A.; Tanaka, H. *Anal. Sci.* **2002**, *18*, 1015.

(42) Namazian, M.; Heidany, H. *J. Mol. Struct.* **2003**, *620*, 257.

(43) Mennucci, B.; Cossi, M.; Tomasi, J. *J. Phys. Chem.* **1996**, *100*, 1807.

(44) Mennuci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101*, 10506.

(45) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 253.

(46) Cossi, M.; Mennucci, B.; Tomasi, J.*Chem. Phys. Lett.* **1994**, *228*, 165.

(47) Pais, A. A. C. C.; Sousa, A.; Eusébio, M. E.; Redinha, J. S. *Phys. Chem. Chem. Phys.* **2001**, *3*, 4001.

(48) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, 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. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.7. Gaussian Inc: Pittsburgh, PA, 1998.

(49) Pascual-Ahuir, J. L.; Silla, E.; Tuñon, I. *J. Comput. Chem.* **1994**, *15*, 1127.

(50) Pierotti, R. A. *J. Phys. Chem.* **1963**, *67*, 1840.

(51) Langlet, J.; Claverie, P.; Caillet, J.; Pullman, A. *J. Phys. Chem.* **1988**, *92*, 1617.

(52) Floris, F.; Tomasi, J. *J. Comput. Chem.* **1989**, *10*, 784.

(53) Floris, F.; Tomasi, J. *J. Comput. Chem.* **1991**, *12*, 616.

(54) Granovsky, A. A. http://classic.chem.msu.su/gran/gamess/index.html.

(55) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.

(56) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1999**, *103*, 9.

(57) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comput. Chem.* **2002**, *23*, 517.

(58) Cheong, W. J. *Bull. Kor. Chem. Soc.* **2003**, *24*, 1207.