

# Self-Consistent-Field Calculation of Pauli Repulsion and Dispersion Contributions to the Solvation Free Energy in the Polarizable Continuum Model

Claudio Amovilli\* and Benedetta Mennucci

Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via Risorgimento 35, 56126 Pisa, Italy

Received: July 22, 1996; In Final Form: September 16, 1996<sup>®</sup>

By using the theory of intermolecular forces, two new expressions for Pauli repulsion and dispersion contributions to the solvation free energy are derived. These expressions contain explicitly the solute electron density and, therefore, can be used directly in the SCF calculation of the solute wave function within the polarizable continuum model (PCM). The final expressions are very simple and include also some intrinsic solvent properties which are, for repulsion, the density, the molecular weight, the number of valence electrons, and for dispersion, the refractive index and the ionization potential. This new approach does not depend on any given intermolecular potential and it can be adapted to any choice of basis set. For small-size basis sets, even minimal, the dispersion contribution is obtained in two steps and includes the effect of adding diffuse and polarization functions, not used in the wave function itself. This method has been implemented in our HONDO package, in a version which includes the cavitation contribution, determined by the Pierotti–Claverie method, and the polarization contribution determined by the Miertus–Scrocco–Tomasi method. Some preliminary results on solutes containing C, H, O, and N are presented for solvation in water, *n*-hexane, and 1-octanol. The quality of these results, given the simplicity of the PCM, is acceptable and of great interest for future developments.

## 1. Introduction

The use of effective Hamiltonian methods, which assume a continuum distribution of the solvent (EHCD),<sup>1</sup> forms one of the most powerful theoretical tools for the investigation of solvation processes. The alternative methods are mainly based on computer simulation techniques<sup>2</sup> and are strongly limited by the number of solvent molecules involved in the calculation. In the EHCD models the solvation free energy is decomposed into different contributions according to the usual decomposition of intermolecular potentials. The calculation is then performed in two steps: (i) the computation, by direct methods, of cavitation,<sup>3,4</sup> repulsion, and dispersion<sup>5–7</sup> contributions and (ii) the quantum-mechanical treatment, by standard *ab initio*<sup>8</sup> or semiempirical<sup>7</sup> techniques, of the solute embedded in a polarizable continuum. The repulsion and dispersion terms are strongly dependent on the solute wave function and it would thus be more appropriate to consider their evaluation in the second step. Several interesting attempts in this direction have been made<sup>9–17</sup>:

here two new and very simple expressions, containing explicitly the solute electron density, for the calculation of repulsion and dispersion are presented, within the EHCD method known as polarizable continuum model (PCM<sup>8</sup>).

## 2. Theory

The polarizable continuum model,<sup>8</sup> in its original form, is based on a description of the solvent as a macroscopic continuum medium having suitable properties (dielectric constant, thermal expansion coefficient, etc.). In this procedure (Figure 1), the solute molecule is embedded in a cavity in the dielectric medium which is defined in terms of interlocking spheres centered on the solute nuclei and is studied *ab initio*; solute–solvent interactions are described in terms of the reaction field due to the presence of the dielectric medium. The latter acts as a perturbation on the Hamiltonian of the solute through

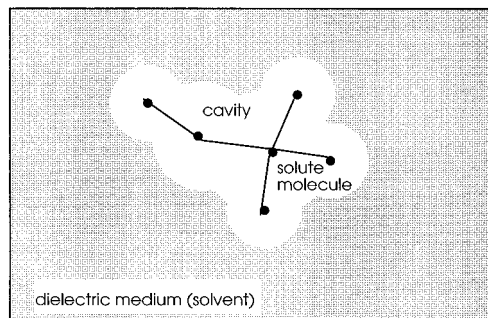


Figure 1. Scheme of solvation in the polarizable continuum model.

its reaction potential  $\hat{V}_R$ :

$$\begin{aligned}\hat{H}^o\Psi^o &= E^o\Psi^o && \text{in vacuo} \\ (\hat{H}^o + \hat{V}_R)\Psi &= E\Psi && \text{in solution}\end{aligned}\quad (1)$$

where  $\hat{H}^o$  is the Hamiltonian of the solute in vacuo (including nuclear repulsion terms),  $\Psi^o$  and  $\Psi$  are the solute wave functions in vacuo and in solution, respectively. The solvent reaction potential can be partitioned into many contributions of different physical origin, related to dispersion, repulsive, and electrostatic forces between solute and solvent molecules. In the standard original PCM only the electrostatic term, here referred to as polarization contribution, is explicitly considered as an interaction potential  $\hat{V}_{\text{pol}}$  to be added to  $\hat{H}^o$  in order to get the effective Hamiltonian. The other two terms are computed by resorting to separate calculations based on empirical parameters or other simplifying assumptions.

The potential  $\hat{V}_{\text{pol}}$  is described in terms of a set of induced point charges  $\{q_i\}$  placed at the center of small surface elements (tesserae) covering the cavity in the dielectric medium, where the solute is embedded. The electrostatic formulas giving the  $\{q_i\}$  charges may be expressed as a set of linear equations and can be written in a matrix formulation as

\* Author to whom correspondence should be addressed.

<sup>®</sup> Abstract published in *Advance ACS Abstracts*, January 1, 1997.

$$\mathbf{q} = -\mathbf{A}\mathbf{D}^{-1}\mathbf{E} \quad (2)$$

$\mathbf{D}$  is a nonsymmetric square matrix with dimension equal to the number of tesserae, and elements depending on geometrical cavity parameters and on the dielectric constant.  $\mathbf{A}$  is a diagonal matrix whose elements are the areas of surface tesserae, while  $\mathbf{E}$  is the column vector containing the normal components of the electric field due to the solute.

The procedures for the calculation of solute energy and wave function with the PCM method have been reported in several papers<sup>1,8,22</sup> and we do not repeat here all the details.

Suffice to say that, in order to get solvation quantities, one has to resort to a direct minimization of the functional of the free energy,  $G$ , of the whole system solute–solvent. In the former PCM this functional included only the polarization contribution; but with the new implementation procedure, to be treated here, the functional to be minimized will be

$$G = G_{\text{solute}} + G_{\text{solvent}} + G_{\text{pol}} + G_{\text{rep}} + G_{\text{dis}} \quad (3)$$

The remaining term, i.e., the cavitation term, is not shown in eq 3, as it is not computed by a quantum-mechanical treatment; it is obtained from the so-called Pierotti–Claverie' formula,<sup>3,4</sup> in which the scaled particle theory (SPT) expression for a single sphere  $G_{\text{cav}}(R_i)$  is weighted with a factor proportional to the area  $A_i$  exposed to the solvent:

$$G_{\text{cav}} = \sum_i^{\text{sph}} \frac{A_i}{4\pi R_i^2} G_{\text{cav}}(R_i) \quad (4)$$

In sections 2.1 and 2.2 the new expressions for  $G_{\text{rep}}$  and  $G_{\text{dis}}$  are presented and finally in section 2.3 their contribution to the reaction potential  $\hat{V}_R$  is derived.

**2.1. Repulsion.** For Pauli repulsion we start from the work of Floris et al.<sup>6</sup> in which the expression for the contribution to the solvation free energy was given as

$$G_{\text{rep}} = \rho_B \int d\vec{R} U_{\text{rep}}^{\text{AB}}(\vec{R}) g_{\text{AB}}(\vec{R}) \quad (5)$$

Here the label A refers to the solute, B to the solvent,  $\vec{R}$  is an appropriate set of coordinates defining the internal geometry of the complex AB,  $\rho_B$  is the number density, and  $g_{\text{AB}}$  is a correlation function which is 0 inside the solute cavity ( $R \in C$ ) and 1 outside ( $R \notin C$ ). This equation is actually used with calculated or estimated potentials  $U_{\text{rep}}^{\text{AB}}$  available from the literature. We now suggest the substitution of  $U_{\text{rep}}^{\text{AB}}$  by a suitable expression taken from the theory of intermolecular forces. The repulsion forces between two interacting molecules originate mainly from the Pauli exclusion principle, increasing with the overlap of the two distributions and being related to the density of electrons with the same spin. Describing the supersystem in terms of antisymmetrized products of group functions, the latter effect corresponds to the violation of strong orthogonality<sup>18</sup> and an expression in terms of the densities of the interacting fragments can be derived.<sup>19</sup> This expression is normally decomposed<sup>20</sup> into two terms: the “exchange” contribution depends on one-electron density matrices and is given by

$$E_{\text{ex}} = -\frac{1}{2} \int \frac{d\vec{r}_1 d\vec{r}_2}{r_{12}} P_{\text{A}}(\vec{r}_1; \vec{r}_2) P_{\text{B}}(\vec{r}_2; \vec{r}_1) \quad (6)$$

which is always negative; the “penetration” contribution, always positive, is more complicated, and depends on the one- two- and three-particle densities of the two fragments. [Here

“exchange” and “penetration” are not the “exchange repulsion” and the “penetration effect in Coulomb interaction” commonly referred to in textbooks; we refer to the matrix partitioning approach to the calculation of intermolecular potentials given in ref 20.] In the region of van der Waals interactions, the penetration term is normally about twice the exchange term but of opposite sign; we therefore suggest an approximate relation

$$U_{\text{rep}}^{\text{AB}}(\vec{R}) = \frac{1}{2} \int \frac{d\vec{r}_1 d\vec{r}_2}{r_{12}} P_{\text{A}}(\vec{r}_1; \vec{r}_2) P_{\text{B}}(\vec{R}|\vec{r}_2; \vec{r}_1) \quad (7)$$

At this point, because in the PCM the electron density of the solvent is not given, it is useful to make the following two assumptions: (i) each valence electron pair of the solvent molecules can be localized in bond and lone pair regions and (ii) each pair, owing to the thermal motion of the solvent molecules, will have the same probability to be found at any point of the solution not occupied by the solute. [In the Hartree–Fock approximation this localization corresponds to a unitary transformation of the occupied orbitals which leaves  $P_{\text{B}}(\vec{R}|\vec{r}_2; \vec{r}_1)$  unchanged.] With these assumptions eq 5 can be rewritten as follows

$$G_{\text{rep}} = \rho_B n_{\text{pair}}^{\text{B}} \int_{\vec{R} \notin C} d\vec{R} \frac{1}{2} \int \frac{d\vec{r}_1 d\vec{r}_2}{r_{12}} P_{\text{A}}(\vec{r}_1; \vec{r}_2) P_{\text{pair}}(\vec{R}|\vec{r}_2; \vec{r}_1) \quad (8)$$

where  $n_{\text{pair}}^{\text{B}}$  is the number of valence electron pairs of the solvent and  $P_{\text{pair}}(\vec{R}|\vec{r}_2; \vec{r}_1)$  a suitable pair density matrix. Here  $\vec{R}$  is the coordinate of the centroid of the localized orbital containing the pair (bond or lone pair) in a reference frame fixed on the solute molecule. The simplest way of defining such a density, each pair being localizable, is through a Gaussian representation of localized orbitals. Assuming

$$P_{\text{pair}}(\vec{R}|\vec{r}_2; \vec{r}_1) = 2N_{\xi}^2 e^{-\xi(\vec{r}_1 - \vec{R})^2} e^{-\xi(\vec{r}_2 - \vec{R})^2} \quad (9)$$

where  $N_{\xi}$  is a normalization factor, eq 8 becomes

$$G_{\text{rep}} = \rho_B n_{\text{pair}}^{\text{B}} \int_{\vec{r}_1, \vec{r}_2 \notin C} d\vec{r}_1 d\vec{r}_2 P_{\text{A}}(\vec{r}_1; \vec{r}_2) e^{-(1/2)\xi r_{12}^2} / r_{12} \quad (10)$$

Formally, this is not an exact result, some border effect in the integration over  $\vec{R}$  should be expected, but these effects can be recovered subsequently with the definition of the cavity. A further very useful approximation is given by the following substitution

$$\frac{e^{-(1/2)\xi r_{12}^2}}{r_{12}} \leftrightarrow \frac{4\pi}{\xi} \delta(\vec{r}_1 - \vec{r}_2) \quad (11)$$

which brings eq 10 to the simple expression

$$G_{\text{rep}} = \alpha \int_{\vec{r} \notin C} d\vec{r} P_{\text{A}}(\vec{r}) \quad (12)$$

with

$$\alpha = (4\pi/\xi) \rho_B n_{\text{pair}}^{\text{B}} \quad (13)$$

A suitable value for  $\xi$  is 0.7, about the best value for helium, the simplest neutral two-electron system. Equation 12 shows a proportionality, in this regime, between the repulsion and the fraction of solute electrons outside the cavity, therefore the minimization of the total free energy, which contains this contribution, provides a confinement of the electronic cloud of

the solute. This effect makes the compensation procedures required in standard PCM calculations<sup>8</sup> less important.

**2.2. Dispersion.** As for the dispersion contribution, we start from the recent method developed by one of us (Amovilli<sup>15</sup>). In this method the suggested equation for  $G_{\text{dis}}$  is the following

$$G_{\text{dis}} = \frac{1}{\pi} \int_0^\infty d\omega \sum_{K \neq 0} \frac{\omega_{0K}^A}{\omega_{0K}^A{}^2 + \omega^2} \times \int d\vec{r}_1 \oint_{S(C)} \frac{d\vec{r}_2}{r_{12}} P_A(0K|\vec{r}_1) \sigma_B[\epsilon_B(i\omega), P_A(0K|\vec{r})](\vec{r}_2) \quad (14)$$

where  $P_A(0K|\vec{r})$  and  $\omega_{0K}^A$  are respectively transition densities and energies for the solute A (for transition to state K),  $\sigma_B$  the surface charge density, induced by the electric field of the charge distribution  $P_A(0K|\vec{r})$ , depending on a dielectric constant calculated at imaginary frequencies. This equation can be manipulated in order to obtain a simplified expression containing explicitly the ground state electron density of the solute. First of all it is convenient to simplify the expression for  $\sigma_B$ .

In the present context  $\epsilon_B$  is always smaller than the square of the estimated zero frequency refractive index,  $\eta_B^2$  (normally this value is very close to the refractive index measured in the visible spectrum far from electron transitions (see ref 15)); consequently we suggest writing directly

$$\sigma_B[\epsilon_B(i\omega), P_A(0K)](\vec{r}) = - \frac{\Omega_B^2}{(\Omega_B^2 + \omega^2)} \frac{(\eta_B^2 - 1)}{4\pi\eta_B^2} E_{0K}^A(\vec{r}) \quad (15)$$

where  $\Omega_B = \eta_B I_B$ ,  $I_B$  being the ionization potential of the solvent.  $E_{0K}^A(\vec{r})$  is the outward component, orthogonal to the cavity surface of the solute at the point  $\vec{r}$ , of the electric field associated with the transition density  $P_A(0K|\vec{r})$ . Equation 15 arises from the direct use of the dispersion relation for the dielectric constant (see for example ref 21) neglecting the autopolarization effects and the contribution at plasma frequencies. We remark that both the terms in this calculation are small and of opposite sign. Equation 15 allows the integration over the frequency giving

$$G_{\text{dis}} = - \frac{1}{8\pi} \frac{(\eta_B^2 - 1)}{\eta_B^2} \sum_{K \neq 0} \frac{\Omega_B}{(\Omega_B + \omega_{0K})} \times \oint_{S(C)} d\vec{r} V_{0K}^A(\vec{r}) E_{0K}^A(\vec{r}) \quad (16)$$

where  $V_{0K}^A$  is the potential associated with the field  $\vec{E}_{0K}^A$ . The surface integral in (16) is positive definite and therefore it is useful to introduce an average transition energy  $\omega_A$  for the solute in order to have

$$G_{\text{dis}} = - \frac{(\eta_B^2 - 1)}{8\pi\eta_B \left( \eta_B + \frac{\omega_A}{I_B} \right)} \sum_{K \neq 0} \oint_{S(C)} d\vec{r} V_{0K}^A(\vec{r}) E_{0K}^A(\vec{r}) \quad (17)$$

To evaluate the sum over the excited states of the solute, we now adopt an uncoupled Hartree–Fock (UCHF) scheme. Following such a scheme, considering only single excitations between occupied and virtual canonical orbitals, one can write

$$P_A(0K|\vec{r}) = \sqrt{2} \phi_i^A(\vec{r}) \phi_m^A(\vec{r}) \quad (i \rightarrow m) \quad (18)$$

and using the following definitions, in terms of orbitals of the starting basis set,

$$P_{rs} = 2 \sum_i^{\text{occ}} C_{ri} C_{si} \quad (S^{-1})_{rs} = \sum_i^{\text{all}} C_{ri} C_{si} \quad (19)$$

eq 17 takes the simple form

$$G_{\text{dis}} = - \frac{\beta}{2} \sum_{rstu} [rs|tu] P_{tu} \left( (S^{-1})_{st} - \frac{1}{2} P_{st} \right) \quad (20)$$

in which

$$\beta = \frac{(\eta_B^2 - 1)}{4\pi\eta_B \left( \eta_B + \frac{\omega_A}{I_B} \right)} \quad (21)$$

and

$$[rs|tu] = \frac{1}{2} \oint_{S(C)} d\vec{r} [V_{rs}(\vec{r}) E_{tu}(\vec{r}) + V_{tu}(\vec{r}) E_{rs}(\vec{r})] \quad (22)$$

Here  $\mathbf{P}$  is the solute electron density matrix,  $\mathbf{S}$  the overlap matrix, and  $\mathbf{C}$  the coefficient matrix of the Hartree–Fock canonical orbitals, and  $V_{rs}$  and  $E_{rs}$  are respectively the potential and the outward component of the field of the distribution  $\chi_r \chi_s$ ,  $\{\chi\}$  being the orbital basis set. Finally, in order to use eq 20 in practical applications, it is important to define some appropriate and simple method for the evaluation of the solute average transition energy  $\omega_A$ . Since the most important transitions in polarization of molecules occur between “frontier” orbitals—the highest occupied and the lowest virtual orbitals—the average transition energy can be approximated as follows

$$\omega_A = \frac{1}{\nu'_{\text{virt}}} \sum_{\epsilon_m^A < a} \epsilon_m^A - \frac{1}{\nu'_{\text{occ}}} \sum_{\epsilon_i^A > -a} \epsilon_i^A \quad (23)$$

where  $\epsilon_k^A$  are Fock eigenvalues and  $\nu'_{\text{occ}}$  and  $\nu'_{\text{virt}}$ , respectively, the number of occupied and virtual orbitals falling in the energy interval  $(-a, +a)$ , and  $a$  is a fitting parameter. With the same basis sets as used in ref 15, we have fitted the results and the best value for  $a$  has been found to be 1.1 hartree.

Use of eq 20 for large basis sets becomes rapidly prohibitive, the number of two-electron integrals in (22) being of order  $n^4/8$ . In such cases it is convenient to split the basis set into two parts, one needed to construct  $\mathbf{P}$  and the other, orthogonalized by a Schmidt procedure to the first, describing a complementary space. Within this space a further orthogonalization is then applied. Calling these two sets  $W_1$  and  $W_2$ , respectively, we have

$$\begin{aligned} \langle r|s \rangle &= S_{rs} & r, s \in W_1 \\ \langle r|s \rangle &= \delta_{rs} & P_{rs} = 0 & r, s \in W_2 \\ \langle r|s \rangle &= 0 & P_{rs} = 0 & r \in W_1, s \in W_2 \end{aligned} \quad (24)$$

$W_1$  brings a contribution to the dispersion free energy of the same type as that given by eq 20 while  $W_2$  gives only the following contribution

$$G_{\text{dis}}(W_1, W_2) = -(\beta/2) \sum_{r,u \in W_1} \sum_{t \in W_2} [rt|tu] P_{ru} \quad (25)$$

The number of integrals  $[rs|tu]$ , to be computed in this case, is greatly reduced.

**2.3. Reaction Field.** The conditions for the variational minimum of  $G$  leads to the formulation of a pseudo Hartree–Fock equation, whenever the trial function is taken as a single determinant. By using an expansion over a finite basis set  $\{\chi\}$  one arrives at a matrix formulation equivalent to that of the corresponding HF problem in vacuo:

$$\mathbf{FC} = \epsilon \mathbf{SC} \quad (26)$$

where

$$\mathbf{F} = \mathbf{h} + \mathbf{G} \quad (27)$$

with

$$\mathbf{h} = \mathbf{h}^\circ + \frac{1}{2}(\mathbf{j} + \mathbf{y}) + \mathbf{h}_{\text{rep}} + \mathbf{h}_{\text{dis}} \quad (28)$$

$$\mathbf{G} = \mathbf{G}^\circ(\mathbf{P}) + \mathbf{X}_{\text{pol}}(\mathbf{P}) + \mathbf{X}_{\text{dis}}(\mathbf{P}) \quad (29)$$

$\mathbf{h}^\circ$  and  $\mathbf{G}^\circ$  being the one-electron and electron interaction matrices of the corresponding HF problem in vacuo.

Here the matrix dimensions are related to the number of basis set functions: matrices  $\mathbf{y}$ ,  $\mathbf{j}$ , and  $\mathbf{X}_{\text{pol}}(\mathbf{P})$  collect one-electron integrals corresponding to the components of  $\hat{V}_{\text{pol}}$  and describe, in that order, interactions of the solute nuclei with the apparent charges generated by the electrons, of the solute electrons with the charges generated by the nuclei, and of the electrons with the charges generated by themselves. Remark that the last term depends on the density matrix  $\mathbf{P}$ , as the two-electron terms collected in  $\mathbf{G}^\circ$ . A fourth solute–solvent interaction term, solute nuclei with the charges generated by the nuclei themselves, does not depend on the electronic wave function, and it may be absorbed into the nuclear repulsion term to give a contribution,  $V_{\text{nn}}$ , constant for a given solute nuclear geometry. More details can be found elsewhere, e.g. in ref 22. The terms indicated with the suffixes *dis* and *rep* can be obtained easily from eqs 12 and 20.

In particular, for the repulsion term we have to evaluate the solute charge distribution outside the cavity  $q_{\text{out}}^A$ . By applying the Gauss theorem we obtain

$$q_{\text{out}}^A = \int_{\vec{r} \notin C} P_A(\vec{r}) d\vec{r} = n_{\text{el}} + \frac{1}{4\pi} \oint_{S(C)} \vec{E}^A(\vec{r}) \cdot \hat{n}_r d\vec{r} \quad (30)$$

where  $\hat{n}_r$  is the outward unit vector perpendicular to the cavity surface  $S(C)$  at point  $\vec{r}$  and  $\vec{E}_A$  the solute electric field due to the electrons only.

Equation 12 can then be rewritten in matrix form as follows:

$$G_{\text{rep}} = \alpha \text{Tr} \mathbf{P} [\mathbf{S} - \mathbf{S}^{(\text{in})}] = \text{Tr} \mathbf{P} \mathbf{h}_{\text{rep}} \quad (31)$$

where

$$S_{rs}^{(\text{in})} = -(1/4\pi) \oint_{S(C)} E_{rs}(\vec{r}) d\vec{r} \quad (32)$$

Equation 31 shows that repulsion interactions contribute to the Fock matrix only through the one-electron matrix

$$\mathbf{h}_{\text{rep}} = \alpha [\mathbf{S} - \mathbf{S}^{(\text{in})}] \quad (33)$$

The dispersive contributions are more complex and lead to both one-electron and electron interaction matrices. From eq 20 we easily obtain

$$G_{\text{dis}} = \text{Tr} \mathbf{P} [\mathbf{h}_{\text{dis}} + \frac{1}{2} \mathbf{X}_{\text{dis}}(\mathbf{P})] \quad (34)$$

where:

$$\begin{aligned} h_{\text{dis},ru} &= -(\beta/2) \sum_{st} [rs|tu] (S^{-1})_{st} \\ X_{\text{dis},ru}(\mathbf{P}) &= (\beta/2) \sum_{st} [rs|tu] P_{st} \end{aligned} \quad (35)$$

The integrals  $[rs|tu]$  and  $S_{rs}^{(\text{in})}$  are computed by exploiting the partitioning of the cavity as described above and discretizing the surface in terms of the tesserae. Finally some consideration of the solute cavity radii must be made. For the polarization contribution the recommended radii are 1.2 times<sup>8</sup> the van der Waals radii<sup>23,24</sup> while for dispersion a factor 1.6 has been suggested.<sup>15</sup> Here we stress that even for repulsion, which depends strongly on diffuse functions, a factor 1.6 seems to be more appropriate.

In the calculations, described in this work, the same radii have been used for all contributions. In order to use the factor 1.2 for dispersion and repulsion we have scaled the expressions 12 and 20 by two different factors smaller than one. Supposing for the dispersion and repulsion contributions the behavior<sup>6</sup>

$$G_\kappa = \sum_i^{\text{sph}} \frac{C_i^{(\kappa)}}{R_i^{n_\kappa}} \quad (\kappa = \text{rep, dis}) \quad (36)$$

with  $i$  running over the interlocking spheres forming the cavity and scaling the radii by the same ratio  $\lambda = (R_{oi}/R_i)$  one has

$$G_\kappa = \lambda^{n_\kappa} \sum_i^{\text{sph}} \frac{C_i^{(\kappa)}}{R_{oi}^{n_\kappa}} \quad (37)$$

To compute  $\lambda^{n_\kappa}$ , instead of using  $(1.2/1.6)^{n_\kappa}$  with  $n_\kappa$  derived from Lennard-Jones type potentials, we have preferred to resort to comparisons with experimental solvation free energy data. The expressions we suggest to be used with the same radii of the polarization contribution, for  $\alpha$  and  $\beta$  in atomic units, are then the following

$$\alpha = 0.063 \rho_B \frac{n_{\text{val}}^B}{M_B} \quad (38)$$

$$\beta = 0.036 \frac{(\eta_B^2 - 1)}{\eta_B \left( \eta_B + \frac{\omega_A}{I_B} \right)} \quad (39)$$

where  $\rho_B$  is now the density of the solvent relative to density of water at 298 K, and  $n_{\text{val}}^B$  and  $M_B$  are the number of valence electrons and the molecular weight of the solvent while for  $\beta$  everything remains as specified above.

The scaling factors  $\lambda^{n_\kappa}$  included in the numerical constants of eqs 38 and 39 differ from the corresponding Lennard-Jones values by less than 5%.

### 3. Calculations

The new approach presented in this paper has been tested in calculations on a selected group of solutes containing C, H, O, and N for solvation in water, *n*-hexane, and 1-octanol. The method has been implemented in a recent version of the HONDO package.<sup>25</sup> In order to test the reliability of the method, we have compared the experimental solvation free energy, for a transfer process, at 298 K, from gas to solution in which the density of the solute in the gas phase is the same as

**TABLE 1: Cavity Radii (Å) for Atoms of Solutes Considered in This Work**

bond	single	double	triple
C	1.7	1.7	1.8
N	1.6	1.6	1.7
O	1.5	1.5	1.6
X <sup>a</sup>	C	N	O
H	1.2	1.0	1.0 <sup>b</sup>

<sup>a</sup> X: atom bonded to hydrogen. <sup>b</sup> In H<sub>2</sub>O: 1.05.

the concentration in the liquid phase, and the computed value

$$\Delta G_{\text{sol}} = G_{\text{iec}} + G_{\text{pol}} + G_{\text{rep}} + G_{\text{dis}} + G_{\text{cav}} \quad (40)$$

where

$$G_{\text{iec}} = \langle \Psi | \hat{H}^{\circ} | \Psi \rangle - \langle \Psi^{\circ} | \hat{H}^{\circ} | \Psi^{\circ} \rangle \quad (41)$$

is the internal energy change of the solute,  $\hat{H}^{\circ}$  being the solute Hamiltonian in vacuo,  $\Psi^{\circ}$  the wave function in vacuo and  $\Psi$  the wave function in the presence of dielectric.  $G_{\text{rep}}$  and  $G_{\text{dis}}$  have been calculated as specified in sections 2.1 and 2.2, and  $G_{\text{pol}}$  has been calculated by means of Miertus–Scrocco–Tomasi method and  $G_{\text{cav}}$  by Pierotti–Claverie formula. The basis sets used for solutes have been of DZP quality<sup>26,27</sup> plus one more set of diffuse and polarization functions with exponents one-third of the most diffuse DZP set. For the larger systems a MINI basis set<sup>28</sup> has been employed for the description of the electron density but in all cases diffuse and polarization functions have been included for the evaluation of the dispersion contribution. The decomposition of the basis set described at the end of section 2.2 has been used when required by our program limitations. The solute geometries have been optimized in vacuo except for *n*-octane for which the linear *anti* conformation has been used. To our knowledge these geometries do not change significantly in solution. The cavity radii for solutes are given in Table 1, these values have been multiplied by 1.2 for the quantum mechanical part of the calculation and they have been used as given in Table 1 for the cavitation contribution. For cavitation we have also used a radius 1.35 Å for water, 2.97 Å for *n*-hexane, and 3.25 Å for 1-octanol. As for the

solvent ionization potential  $I_{\text{B}}$ , we have used the following values in hartrees: 0.451 for water, 0.366 for *n*-hexane, and 0.4 for 1-octanol. Tables 2–4 contain calculated and experimental free energies. For CH<sub>4</sub> in water we have used all the different basis sets considered in this work; in the three cases, the resulting errors in solvation free energies have been less than 0.3 kcal/mol, showing an acceptable accuracy even with less quality basis sets, provided polarization and diffuse functions are used in  $G_{\text{dis}}$  calculation.

For benzene in water, a simple test calculation employing only the MINI basis set was made; the resulting  $G_{\text{dis}}$ , −5.6 kcal/mol, 35% of the  $G_{\text{dis}}$  value reported in Table 2, confirms the importance of diffuse and polarization functions in this calculation and shows, at least for large molecules, that the dispersion reaction potential is mainly determined by the one-electron operator  $\hat{h}_{\text{dis}}$  whose matrix elements are given in eq 35.

Turning to the results of Tables 2–4, it is evident that for most of the systems considered a good precision has been achieved but there are solutes which show a large error, in some cases really unexpected. For carbonyl compounds a possible source of error is the neglect of intramolecular electron correlation: it has been shown that the polarization contribution can differ by 1–2 kcal/mol if computed by density functional theory<sup>29</sup> instead of standard SCF calculations. To explain the error in other cases we have considered the cavitation contribution, this contribution being less well connected with our development of the solvation free energy. In order to study in some detail this problem we have fitted the following quantity

$$G' = \Delta G_{\text{exp}} - G_{\text{iec}} - G_{\text{pol}} - G_{\text{rep}} - G_{\text{dis}} \quad (42)$$

with the function

$$A + BS \quad (43)$$

where  $A$  and  $B$  are fitting parameters depending on the solvent and  $S$  the surface of the solute cavity calculated with radii of Table 1. This approach to cavitation has already been described in other methods of solvation (see for example ref 7). Figure 2 shows the result of this fitting for solvation in water. The agreement, except for aldehydes, is improved especially for solutes with large cavities. We remark that in such cases the

**TABLE 2: Computed Solvation Free Energies, and Their Components, of Different Solutes in Water at 298 K**

solute	$W_1^a$	$W_2^a$	$\Delta G_{\text{sol}}$ components (kcal/mol)						exp <sup>b</sup>	err
			iec	pol	rep	disp	cav	tot		
H <sub>2</sub> O	DZP+s,p,d		1.0	−8.2	1.7	−4.6	4.2	−6.0	−6.3	0.3
NH <sub>3</sub>	DZP+s,p,d		0.8	−6.1	2.3	−5.7	4.6	−4.2	−4.3	0.1
H <sub>2</sub> O <sub>2</sub>	DZP+s,p,d		1.2	−10.2	1.9	−6.7	5.6	−8.3	−8.7	0.4
N <sub>2</sub>	DZP+s,p,d		0.0	−0.2	0.7	−3.9	5.4	2.0	2.3	−0.3
CH <sub>4</sub>	DZP+s,p,d		0.0	−0.2	2.0	−5.5	6.0	2.3	2.0	0.3
CH <sub>4</sub>	DZP	s,p,d	0.0	−0.2	2.0	−6.0	6.0	1.8	2.0	−0.2
CH <sub>4</sub>	MINI	2s,2p,2d	0.0	−0.4	2.0	−6.6	6.0	2.0	2.0	0.0
CO	DZP+s,p,d		0.1	−0.9	0.7	−4.0	5.5	1.4	2.2	−0.8
HCHO	DZP+s,p,d		1.3	−7.2	1.4	−5.6	6.2	−3.9	−1.7	−2.2
HCN	DZP+s,p,d		1.2	−6.9	1.6	−4.7	6.0	−2.8	−3.2	0.4
N <sub>2</sub> H <sub>4</sub>	DZP+s,p,d		1.7	−10.0	2.2	−8.2	6.5	−7.8	−9.3	1.5
CH <sub>3</sub> OH	DZP+s,p,d		0.9	−6.8	2.2	−7.5	7.2	−4.0	−5.1	1.1
C <sub>2</sub> H <sub>2</sub>	DZP+s,p,d		0.2	−3.2	1.5	−5.1	6.6	0.0	0.0	0.0
C <sub>2</sub> H <sub>4</sub>	DZP+s,p,d		0.1	−1.3	2.8	−7.1	7.5	1.9	1.3	0.6
CH <sub>3</sub> NH <sub>2</sub>	DZP	s,p,d	0.6	−5.3	2.2	−8.4	7.7	−3.2	−4.6	1.4
C <sub>2</sub> H <sub>6</sub>	DZP	s,p,d	0.0	−0.2	2.5	−8.8	8.8	2.2	1.8	0.4
CH <sub>3</sub> CHO	DZP	s,p,d	1.5	−7.9	2.0	−8.6	8.8	−4.2	−3.5	−0.7
C <sub>2</sub> H <sub>5</sub> OH	DZP	s,p,d	0.8	−6.8	2.6	−10.0	9.9	−3.5	−5.0	1.5
C <sub>3</sub> H <sub>8</sub>	MINI	2s,2p,2d	0.0	−0.5	4.3	−12.9	11.5	2.4	2.0	0.4
CH <sub>3</sub> COCH <sub>3</sub>	MINI	2s,2p,2d	0.8	−4.8	3.0	−13.1	11.6	−2.5	−3.8	1.3
C <sub>6</sub> H <sub>6</sub>	MINI	2s,2p,2d	0.2	−2.7	4.7	−15.8	12.7	−0.9	−0.9	0.0
<i>n</i> -C <sub>8</sub> H <sub>18</sub>	MINI	2s,2p,2d	0.0	−1.3	8.6	−28.3	27.0	5.9	3.0	2.9

<sup>a</sup> See text for details on the basis set decomposition. <sup>b</sup> Data from refs 30 and 31.

**TABLE 3: Computed Solvation Free Energies, and Their Components, of Different Solutes in *n*-Hexane at 298 K**

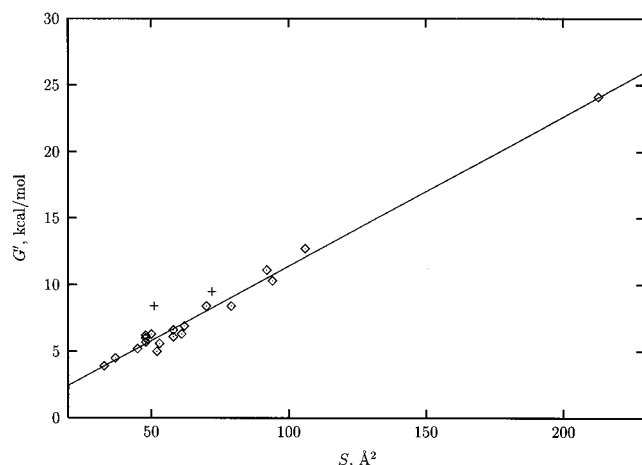
solute	$W_1^a$	$W_2^a$	$\Delta G_{\text{sol}}$ components (kcal/mol)						$\text{exp}^b$	err
			iec	pol	rep	disp	cav	tot		
CH <sub>3</sub> OH	DZP+s,p,d		0.1	-2.2	1.5	-7.3	6.0	-1.9	-1.4	-0.5
C <sub>2</sub> H <sub>5</sub> OH	DZP	s,p,d	0.1	-2.1	1.7	-9.8	7.7	-2.4	-2.0	-0.4
C <sub>2</sub> H <sub>5</sub> NH <sub>2</sub>	DZP	s,p,d	0.1	-1.5	1.8	-10.7	8.0	-2.4	-2.1	-0.3
CH <sub>3</sub> COCH <sub>3</sub>	MINI	2s,2p,2d	0.1	-1.5	2.0	-12.6	8.8	-3.2	-2.5	-0.7
C <sub>6</sub> H <sub>6</sub>	MINI	2s,2p,2d	0.0	-0.9	3.1	-15.4	9.6	-3.6	-4.0	0.4
aniline	MINI	2s,2p,2d	0.1	-2.9	3.6	-17.5	11.2	-5.5	-5.4	-0.1
<i>p</i> -cresol	MINI	2s,2p,2d	0.1	-2.9	3.8	-19.3	12.9	-5.4	-5.9	0.5
<i>o</i> -cresol	MINI	2s,2p,2d	0.1	-2.6	3.8	-19.3	12.7	-5.3	-6.3	1.0

<sup>a</sup> See text for details on the basis set decomposition. <sup>b</sup> Data from ref 7.

**TABLE 4: Computed Solvation Free Energies, and Their Components, of Different Solutes in 1-Octanol at 298 K**

solute	$W_1^a$	$W_2^a$	$\Delta G_{\text{sol}}$ components (kcal/mol)						$\text{exp}^b$	err
			iec	pol	rep	disp	cav	tot		
CH <sub>4</sub>	DZP+s,p,d		0.0	-0.1	1.6	-6.4	5.4	0.4	0.5	-0.1
CH <sub>3</sub> OH	DZP+s,p,d		0.2	-2.4	1.8	-8.8	6.2	-3.0	-4.0	1.0
C <sub>2</sub> H <sub>2</sub>	DZP+s,p,d		0.0	-1.2	1.2	-5.9	5.5	-0.4	-0.5	0.1
C <sub>2</sub> H <sub>4</sub>	DZP+s,p,d		0.0	-0.5	2.3	-8.1	6.5	0.1	-0.2	0.3
CH <sub>3</sub> NH <sub>2</sub>	DZP	s,p,d	0.1	-1.9	1.8	-9.8	6.7	-3.1	-3.8	0.7
C <sub>2</sub> H <sub>6</sub>	DZP	s,p,d	0.0	-0.1	2.0	-10.2	7.7	-0.6	-0.7	0.1
C <sub>2</sub> H <sub>5</sub> OH	DZP	s,p,d	0.1	-2.3	2.1	-11.7	8.6	-3.2	-4.6	1.4
C <sub>3</sub> H <sub>8</sub>	MINI	2s,2p,2d	0.0	-0.2	3.7	-14.7	10.0	-1.1	-1.2	0.1
CH <sub>3</sub> COCH <sub>3</sub>	MINI	2s,2p,2d	0.2	-3.0	1.3	-12.4	10.0	-3.9	-3.5	-0.4
C <sub>6</sub> H <sub>6</sub>	MINI	2s,2p,2d	0.0	-1.0	3.9	-18.3	10.8	-4.6	-3.8	-0.8

<sup>a</sup> See text for details on the basis set decomposition. <sup>b</sup> Data from ref 32.



**Figure 2.**  $G'$  (eq 42) values with respect to the function  $A + BS$  where  $S$  is the cavity surface for the solutes listed in Table 1,  $A = 0.198$  and  $B = 0.112$ . The two values indicated with crosses refer to aldehydes (HCHO and CH<sub>3</sub>CHO) and have not been considered in the fitting.

Pierotti–Claverie formula for cavitation does not work properly; it is well-known that the contribution to the entropy coming from cavitation, calculated by the Pierotti–Claverie method, increases too rapidly with the dimensions of the solute molecule if compared with experimental data, the other terms being essentially enthalpy contributions. A final interesting remark concerns the transfer process from water to 1-octanol. The logarithm of the partition coefficient

$$\log P = \log(C_{\text{oct}}/C_{\text{w}}) \quad (44)$$

where  $C_{\text{oct}}$  and  $C_{\text{w}}$  are equilibrium concentrations for a given solute, respectively, in 1-octanol and water, is largely used in structure activity relationships studies for molecules of biological interest. For the solutes considered in this work we have reported in Table 5 the  $\log P$  values calculated by means of the relation

**TABLE 5: Calculated  $\log P$  (Eq 45) at 298 K. Comparison with Experimental Data**

solute	$\log P$	$\text{exp}^a$	err
CH <sub>4</sub>	1.17	1.09	0.08
CH <sub>3</sub> OH	-0.73	-0.77	0.04
C <sub>2</sub> H <sub>2</sub>	0.29	0.37	-0.08
C <sub>2</sub> H <sub>4</sub>	1.32	1.13	0.19
CH <sub>3</sub> NH <sub>2</sub>	-0.07	-0.57	0.50
C <sub>2</sub> H <sub>6</sub>	2.05	1.81	0.24
C <sub>2</sub> H <sub>5</sub> OH	-0.22	-0.31	0.09
C <sub>3</sub> H <sub>8</sub>	2.57	2.36	0.21
CH <sub>3</sub> COCH <sub>3</sub>	1.03	-0.24	1.27
C <sub>6</sub> H <sub>6</sub>	2.71	2.15	0.56

<sup>a</sup> Data from ref 32.

$$\log P = - \frac{(\Delta G_{\text{sol,oct}} - \Delta G_{\text{sol,w}})}{2.302RT} \quad (45)$$

where the solvation free energies are taken from Tables 2 and 4. The results show, for this transfer process, a cancellation of errors in many cases; only for acetone is the error large. Acetone, being a carbonyl compound, could be influenced by the neglect of electron correlation discussed above.

#### 4. Conclusion

In this paper an extension of the polarizable continuum model, for the calculation of solvation free energy, to the repulsion and dispersion contributions has been presented. Starting from the theory of intermolecular forces two new expressions, depending explicitly on the ground state electron density of the solute, have been derived for these contributions.

The Pauli repulsion term is shown to be proportional to the fraction of solute electrons outside the cavity and to the average number of solvent valence electrons per unit volume. The dispersion term depends on solvent refractive index and ionization potential and on solute response functions approximated in terms of the ground state electron density and of an averaged transition energy. In both expressions simple surface integrals must be evaluated. This calculation is

performed by partitioning the cavity in terms of spheres centered on each atom of the solute molecule and discretizing the surface in terms of small surface elements.

The new approach has been tested on several solutes containing C, H, O, and N for solvation in water, *n*-hexane, and 1-octanol. For most of the molecules considered in this work the agreement with experimental results is quite good. In the other cases the results could be affected by errors related to the neglect of intramolecular electron correlation and to the use of the Pierotti–Claverie formula for the evaluation of the cavitation free energy.

The electron correlation in some systems has a strong effect on the polarization contribution and these effects are of the same order of magnitude as the errors we have encountered in our calculations. The inclusion of the present approach in density functional programs for solvation energy could improve the level of accuracy. About cavitation we remark that the known methods of calculation are not yet completely independent of the repulsion contribution; work in this area is at present in progress in our laboratory.<sup>33</sup>

It is possible to refine the method used in this work in order to define systematically cavity radii, basis sets, average transition energies for solutes, and ionization potentials for solvents. Such refinements would allow calculations for molecules in biological systems where experimental data are not available.

Another field in which the impact of the present approach could be important is the study of excited states and of polarizabilities and hyperpolarizabilities of solutes, because in such cases the effect of the repulsion contribution could be dominant.

**Acknowledgment.** We thank Prof. J. Tomasi for discussions on this general area and for his continuing interest in this work. We are also indebted to Prof. R. McWeeny for helpful hints during the preparation of the manuscript. This work takes part of collaborative projects with C.N.R.; the financial support from I.C.Q.E.M. of C.N.R. is gratefully acknowledged.

## References and Notes

(1) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027, and many references therein.

- (2) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638.
- (3) Pierotti, R. A. *Chem. Rev.* **1976**, *76*, 717.
- (4) Langlet, J.; Claverie, P.; Caillet, J.; Pullman, A. *J. Phys. Chem.* **1988**, *92*, 1617.
- (5) Floris, F. M.; Tomasi, J. *J. Comput. Chem.* **1989**, *10*, 616.
- (6) Floris, F. M.; Tomasi, J.; Pascual Ahuir, J. L. *J. Comput. Chem.* **1991**, *12*, 784.
- (7) Giensen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1995**, *99*, 7137.
- (8) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.
- (9) Linder, B. J. *Chem. Phys.* **1960**, *33*, 668.
- (10) Thole, B. T.; Van Duijnen, P. T. *Chem. Phys.* **1982**, *71*, 211.
- (11) Rinaldi, D.; Ruiz-Lopez, M. F.; Martins Costa, M. T. C.; Rivail, J. L. *Chem. Phys. Lett.* **1986**, *128*, 177.
- (12) Aguilar, M. A.; Olivares del Valle, F. J. *Chem. Phys.* **1989**, *138*, 327.
- (13) Sharp, K.; Jean-Charles, A.; Honig, B. J. *J. Phys. Chem.* **1992**, *96*, 3822.
- (14) Olivares del Valle, F. J.; Aguilar, M. A. *J. Mol. Struct. (THEOCHEM)* **1993**, *99*, 25.
- (15) Amovilli, C. *Chem. Phys. Lett.* **1994**, *229*, 244.
- (16) McCreery, J. H.; Christoffersen, R. E.; Hall, G. G. *J. Am. Chem. Soc.* **1976**, *98*, 7191.
- (17) Bernhardsson, A.; Lindh, R.; Korlström, G.; Roos, B. O. *Chem. Phys. Lett.* **1996**, *251*, 141.
- (18) McWeeny, R. *Proc. R. Soc.* **1959**, A253, 242.
- (19) (a) McWeeny, R.; Stuchliffe, B. T. *Proc. R. Soc.* **1963**, A273, 103.  
(b) See also: McWeeny, R. *Methods of Molecular Quantum Mechanics*, 2nd ed.; Academic Press: London, 1992; Chapter 13, for a simpler presentation.
- (20) Amovilli, C.; McWeeny, R. *Chem. Phys.* **1990**, *140*, 343.
- (21) Mahanty, J.; Ninham, S. W. *Dispersion Forces*; Academic Press: London, 1976.
- (22) Cammi, R.; Tomasi, J. *J. Comput. Chem.* **1995**, *16*, 1449.
- (23) Bondi, A. J. *Chem. Phys.* **1964**, *68*, 441.
- (24) Fanci, M. M.; Hout, R. F.; Hehre, W. J. *J. Am. Chem. Soc.* **1984**, *106*, 563.
- (25) Dupuis, M.; Farazdel, A.; Karna, S. P.; Maluendes, S. A. *HONDO-8*; IBM Corp.: Kingston, NY.
- (26) Huzinaga, S. *J. Chem. Phys.* **1965**, *42*, 1293.
- (27) Dunning, T. H. *J. Chem. Phys.* **1970**, *53*, 2823.
- (28) Tatewaki, H.; Huzinaga, S. *J. Comput. Chem.* **1980**, *1*, 205.
- (29) Fortunelli, A.; Tomasi, J. *Chem. Phys. Lett.* **1994**, *231*, 34.
- (30) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016.
- (31) Pearson, R. G. *J. Am. Chem. Soc.* **1986**, *108*, 6109.
- (32) Rekker, R. F. *The Hydrophobic Fragmental Constant*; Elsevier: Amsterdam, 1977.
- (33) Floris, F. M.; Tani, A.; Selmi, M.; Tomasi, J., manuscript in preparation.