Dispersion and Repulsion Contributions to the Solvation Energy: Refinements to a Simple Computational Model in the Continuum Approximation

F.M. Floris and J. Tomasi*

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

J.L. Pascual Ahuir

Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York 10029 New York

A computational method for the evaluation of dispersion and repulsion contributions to the solvation energy is here presented in a formulation which makes use of a continuous distribution of the solvent, without introducing additional assumptions (e.g., local isotropy in the solvent distribution). The analysis is addressed to compare the relative importance of the various components of the dispersion energy ($n = 6, 8, 10$) and of the repulsion term, to compare several molecular indicators (molecular surface and volume, number of electrons) which may be put in relation to the dispersion-repulsion energy, and to define simplified computational strategies. The numerical examples refer to saturated hydrocarbons in water, treated with the *homogeneous approximation* of the distribution function which for this type of solution appears to be acceptable.

INTRODUCTION

A preceding article[1] hereafter called article I, presented and analyzed a computational method for the evaluation of the dispersion energy contributions to the solvation energy. The main features of that method are represented by (1) the use of a continuous distribution function of the solvent around the solute, (2) the evaluation of the energy under the form of a surface integral, (3) the use of atom-atom potentials.

There are several reasons which support the introduction of these characterizing features.

First, a continuous distribution of the solvent, which is not necessarily a constant distribution, is quite a flexible model, corresponding to the form under which experimental and computer simulation data are generally expressed. A continuous model for dispersion-repulsion terms complements analogous continuous models for electrostatic solute-solvent interactions. Also the continuum models for electrostatic interactions[2,3] are not limited to a homogenous dielectric medium: in future communications we shall present extensions of our computational method for electrostatic interactions to cases in which local anisotropies, corresponding to the solvent organization around the solute, are considered. Other types of anisotropies, due to the presence of a second liquid phase have already been considered in other articles, for the electrostatic as well as for the dispersion terms.[4-6]

Second, the reduction of the interaction energy to a surface integral allows us to better exploit a portion of the computations necessary to evaluate solute-solvent interactions. The computational code for the simultaneous evaluation of the quantum-mechanical solute wave function and of the semiclassical electrostatic problem of the insertion of the solute charge distribution in the dielectric cavity[2] requires in fact the numerical evaluation of integrals defined over the molecular surface; a large portion of the necessary calculations may be employed so as to get the dispersion-repulsion contribution to the energy. In addition, the expression of the various contributions to the energy of the system (electrostatic, dispersion, etc.) may be exploited to reduce the pertinent information at the level of the molecular surface. A graphical representation of physicochemical information encoded at the level of the molecular surface may be exploited for molecular modelling and molecular engineering studies.[7,8]

Third, the atom-atom expressions of the dispersion-repulsion terms are widely employed in studies on interacting systems performed at the discrete molecular level. There are no theoretical or practical objections to the extension of this approach to systems in which a part of the matter is described by continuous functions.

---

*To whom all correspondence should be addressed.

A different approach makes explicit use of the solvent fluctuating reaction potential and of its effect on the solute polarizability. This approach, pioneered by Linder[9,10], has been exploited by Rivail and co-workers[11,12] and, more recently, recast by Aguilar and Olivares del Valle[13] into a computational setup which preserves the other features of the present computational method, and in particular the other two points discussed above. The use of the reaction field for the evaluation of dispersion contributions to the solvation energy makes the whole method more compact, but at present more information about the dependence of the results on other computational parameters (basis set, intramolecular correlation, repulsion terms, etc.) is needed. To this end a comparison with the results obtained via the atom-atom potentials may be of considerable help. The use of dispersion terms involving the explicit consideration of submolecular units makes a noticeable flexibility in the calculations in fact possible. There is in literature a wide choice of parameters (atom-atom, bond-bond, group-group) not only limited to the first term in the r^{-n} expansion. The dispersion parameters are often accompanied by repulsion terms; the repulsion contribution is not yet present in the reaction potential formulation. The parameters have two distinct sources: experimental data or *ab initio* calculations (see article I for a selected bibliography). A judicious use of both sources may give some information about the basis set dependence of the results.

For these reasons we consider of some interest to extend the analysis initiated in article I. The present article is addressed at the examination of the effects due to the additional terms in the r^{-n} expansion of the dispersion potential and to the consideration of a repulsive term.

OUTLINE OF THE METHOD

As said in the introduction, the method has been exposed in article I, where we limit to its application to the first term ($n = 6$) alone of the solute-solvent dispersion energy. We resume here the most important points, following a general formulation that includes the other terms of the dispersion energy and the repulsion energy. The solute-solvent interaction potential (dispersion and repulsion contributions only) is described by two-body functions of the type

$$V_{ml}(r_{ml}) = \sum_{n=6,8,10} -\frac{d_{ml}^{(n)}}{r_{ml}^{n}} + c_{ml} \exp(-\gamma_{ml} r_{ml}) \quad (1)$$

the indexes m and l stand for a subunit of the solute and of the solvent (in the present article they are atoms, but they may also be bonds or chemical groups). The dispersion coefficients $d_{ml}^{(n)}$ ($n = 6, 8, 10$) and the repulsion coefficients c_{ml} and γ_{ml} are

taken from the literature. The sum on all m and all l subunits of the solute and solvent molecules gives the expression for the dispersion-repulsion interaction in the dimer system, for a particular space configuration of the two molecules.

To obtain the corresponding solute-solvent interaction energy we need to know the structure of the solvent, namely the space configuration of all molecules of the solvent around the solute molecule. This information, which generally is obtained from fluid simulations or from experimental measurements, is expressed as continuous distribution functions that may be written in the following form

$$\rho_{ml}(\mathbf{r}_{ml}) = N_l \rho_L g_{ml}(\mathbf{r}_{ml}). \quad (2)$$

Here N_l is the number of subunits of type l in each molecule of the solvent, ρ_l is the macroscopic density of the solvent and $g_{ml}(\mathbf{r}_{ml})$ is a correlation function, depending on the position of l with respect to m.

The interaction energy has the nature of a free energy, Gibbs or Helmoltz, according to the definition of the distribution function which appears in the expression of the interaction energy. We shall make use of the Gibbs free energy, in agreement with article I. The dispersion-repulsion contribution to the solvation free-energy may be written as a sum of volume integrals:

$$\mathbf{G}_{dis-rep}(M \text{ in } L)$$

$$= \sum_{l \in L} \sum_{m \in M} \int \rho_{ml}(\mathbf{r}_{ml}) V_{ml}(\mathbf{r}_{ml}) d\mathbf{r}_{ml}$$

$$= \rho_L \sum_{l \in L} N_l \sum_{m \in M} \left[\sum_{n=6,8,10} -d_{ml}^{(n)} \int \frac{g_{ml}(\mathbf{r}_{ml})}{r_{ml}^{n}} d\mathbf{r}_{ml} \right.$$

$$\left. + c_{ml} \int g_{ml}(\mathbf{r}_{ml}) \exp(-\gamma_{ml} r_{ml}) d\mathbf{r}_{ml} \right] \quad (3)$$

We have here introduced a change of variables: $\mathbf{r}_{ml} = \mathbf{r}_l - \mathbf{r}_m$. The integrals in eq. (3) are defined over the entire space, but the presence of short range repulsions allows us to define a portion of the space, containing the solute molecule M, in which there are no atoms l of the solvent. In other words, we may introduce for each l a closed volume C_l so that

$$g_{ml}(\mathbf{r}_{ml}) = 0, \text{ if } \mathbf{r}_{ml} \in C_l \quad (4)$$

On the basis of this consideration we may develop approximate models, which assume the existence of C_l cavities and use analytical expressions of the correlation functions of variable complexity.

For example, in article I we have presented a model of cavities C_l related to the van der Waals spheres centered in the nuclei of M, and enlarged to take into account also the region of space forbidden to the nuclei of the solvent atoms, using at the same time the simplest expression of the correlation function $g_{ml}(\mathbf{r}_{ml}) = 1$, if $\mathbf{r}_{ml} \notin C_l$, corresponding to a homogeneous density of the solvent. We call this *homogeneous approximation*. When the integrals of eq. (3) are limited to the portion of space outside C_l,

it is convenient to transform them into surface integrals, computed on the Σ_l surfaces, which delimit the cavities C_l. The mathematical derivation, accompanied by some comments, is resumed in the Appendix. The resulting expression may be summarised in the following way

$$G_{\text{dis}-\text{rep}}(M \text{ in } L)$$

$$= \rho_L \sum_{l \in L} N_l \sum_{m \in M} \left[\sum_{n=6,8,10} \int_{\Sigma l} \mathbf{A}_{ml}^{(n)}(\mathbf{r}_{ml}) \right.$$

$$\left. \cdot \mathbf{n}_{\Sigma_l} \, ds + \int_{\Sigma_l} \mathbf{A}_{ml}^{\text{rep}}(\mathbf{r}_{ml}) \cdot \mathbf{n}_{\Sigma_l} \, ds \right] \quad (5)$$

The integrands contain the scalar products of the vector functions $\mathbf{A}_{ml}^{(n)}$ and $\mathbf{A}_{ml}^{\text{rep}}$, computed at the cavity surface, with the unit inward normal vector $\mathbf{n}_{\Sigma_l}$ to surface Σ_l. The vector functions, which permit the replacement of the volume integrals with surface integrals preserve in their definition (see the Appendix) the generality of the description of $g_{ml}(\mathbf{r}_{ml})$ functions introduced in eq. (2). The integrals of eq. (5) are computed numerically. The spheres defining the C_l volumes are replaced by convex regular polyedra, the details are given in reference 14. It is sufficient here to recall that the surface Σ_l is replaced by a set of tesseras with areas $\Delta s_{lk(j)}$; the area of surface is given with a good approximation by

$$S_{\Sigma_l} = \sum_{j,k(j)} \Delta s_{lk(j)}. \quad (6)$$

where j runs on the spheres defining C_l. The eq. (5) is thus replaced by

$$G_{\text{dis}-\text{rep}}(M \text{ in } L) = \rho_L \sum_{l \in L} N_l \sum_{m \in M} \sum_{j,k(j)}$$

$$\times \left[\sum_{n=6,8,10} \mathbf{A}_{ml}^{(n)}(\mathbf{r}_{mk}) + \mathbf{A}_{ml}^{\text{rep}}(\mathbf{r}_{mk}) \right] \cdot \mathbf{n}_{\Sigma_l} \Delta s_{lk(j)} \quad (7)$$

The vector functions are computed at the center k of each tessera, and are obtained via a one dimensional integration on r_{ml} (see Appendix).

RESULTS

We shall examine here a set of hydrocarbons as solutes in water. The analysis of the results aim at answering the following questions

1. what is the relative importance of the dispersion term with $n = 6, 8, 10$ and of the repulsion term;
2. how good is the correlation between the dispersion-repulsion contribution to the free solvation energy and some molecular parameters, namely the area of the molecular surface S_M, the molecular volume V_M and the number of electrons N_M;
3. how large is the error in the calculation of $G_{\text{dis}-\text{rep}}$ introduced by the truncation of the surface integral, neglecting in eq. (7) the contributions given

by the k tesseras belonging to the j sphere with $r_{mj} \geq R_{\text{cut}}$, where R_{cut} is a parameter of cutoff.

The cavity model adopted here uses the *homogeneous approximation* of the correlation function.

The calculations refer to molecules in their standard geometry.[15] The linear hydrocarbons are in the all-staggered conformations. The molecular cavity is composed by spheres centred on the carbon C atoms of the solute. The radius of these spheres correspond to the van der Waals radii for C, CH, CH_2, CH_3, and CH_4, supplemented by an increment δ_l (see article I). The van der Waals radii used are 1.6 Å for the carbon atom and 2.0 Å for the groups CH_n with $n = 1, 2, 3, 4$. The δ_l increments used for the water solvent, $\delta_O = \delta_H = 1.40$ Å, give values of $G_{\text{dis}-\text{rep}}$ in reasonable agreement with calculations obtained by use in eq. (3) of the realistic correlation functions g_{ml} drawn from computer simulation and from the application of the integral equation approach.[16] The comparison of the two sets of calculations has been performed on a smaller set of alkanes in water. Note that in the present work the set of C_l cavities is reduced to a cavity alone, the increment δ_l being equal for the two solvent atoms: H and O. In a few cases, the specific conformation of the solute induces the creation of additional spheres, to take into account solvent excluded volume effects. These additional spheres are treated as the original ones (i.e., with the addition of δ_l increment). Surfaces (S_M) and volumes (V_M) employed in the following refer to the original set of spheres, i.e., without enlargement of the corresponding radii by δ_l. The values of S_M and V_M are obtained with the same program which computes $G_{\text{dis}-\text{rep}}$. The portion of the program computing S_M and V_M has been distributed as a separate program through the QCPE.[17] This subroutine gives values of better precision than other subroutines of current use.[18]

The dispersion-repulsion parameters used in this work are empirical parameters, optimized by comparison with vaporisation and sublimation free energies, taken from the compilation done by Claverie et al.[19] Note that they are of atomic nature, and consequently the position of each atom m of the solute molecule within the cavity is computed with the necessary precision.

Table I reports the total value of G_{dis}, and the percentage of G_{dis} due to the contributions with $n = 6$, 8, 10. G_{rep} has the opposite sign, and is recorded in this table as the percentage of the total value of G_{dis}.

In the relatively large variety of dimensions and shapes present in the set of solutes of Table I there is a remarkable invariance of the relative weight of the different components of G_{dis}. The mean values of $G_{dis}^{(n)}$ and the corresponding standard deviations are reported in Table II. The relative weight of $G_{dis}^{(6)}$ increases a little in passing to large solutes. The repulsion contribution has a relatively larger deviation

TABLE I. Values of G_{dis}^{tot} and its components $G^{(6)}$, $G^{(8)}$, and $G^{(10)}$ as percentage of the total dispersion contribution. The repulsion contributions are reported as percentage of G_{dis}^{tot}.

Solute	$G_{dis}^{tot(a)}$	$G_{dis}^{(6)}$ (%)	$G_{dis}^{(8)}$ (%)	$G_{dis}^{(10)}$ (%)	G_{rep} (%)
Methane	−4.368	88.04	9.75	2.20	−19.01
Ethane	−6.388	88.38	9.52	2.13	−18.39
Propane	−8.131	88.60	9.33	2.08	−17.97
Isobutane	−9.637	88.78	9.18	2.03	−17.64
n-butane	−9.825	88.70	9.22	2.04	−17.88
Cyclopentane	−9.947	89.23	8.88	1.90	−15.72
Neopentane	−10.934	88.97	9.05	1.98	−17.32
Cyclohexane	−11.882	88.97	9.05	1.98	−17.23
Isopentane	−11.063	88.97	9.05	1.98	−16.97
n-pentane	−11.494	88.79	9.29	2.05	−17.70
Me-cyclopentane	−11.347	89.37	8.67	1.86	−15.33
2,2-diMe-butane	−12.171	89.22	8.87	1.91	−16.27
(e)Me-cyclohexane	−13.317	89.04	8.99	1.97	−17.14
3-Me-pentane	−12.496	89.11	8.95	1.94	−16.44
2-Me-pentane	−12.707	88.99	9.03	1.98	−17.05
n-hexane	−13.290	88.81	9.16	2.02	−17.71
2,4-diMe-pentane	−13.837	89.19	8.89	1.93	−16.54
1-cis-2-diMe-cyclohexane	−14.219	89.32	8.79	1.89	−15.99
2,2-diMe-pentane	−13.776	89.18	8.89	1.93	−16.48
2,2,4-triMe-penthane	−14.556	89.49	8.67	1.85	−15.44
n-heptane	−14.941	88.85	9.13	2.01	−17.66
2,2,5-triMe-hexane	−16.426	89.28	8.82	1.91	−16.30
decaline(T)	−16.669	89.21	8.68	1.93	−16.77
3-Et-3-isoPr-pentane	−16.549	89.86	8.40	1.74	−13.89
n-decane	−19.812	88.95	9.06	2.00	−17.53
n-pentadecane	−28.086	89.01	9.00	1.98	−17.46

^{a}In kcal/mol.

with respect to the mean value (see Table II again). The positive contribution to free energy by G_{rep} is to a good extent cancelled by the $G_{dis}^{(8)}$ and $G_{dis}^{(10)}$ contributions: the use of the $G_{dis}^{(6)}$ alone, instead of $G_{dis-rep}$ introduces a mean error of -0.777 kcal/mol, with a standard error of 0.301 kcal/mol. The relations between $G_{dis-rep}$ and S_M, V_M, and N_M are illustrated in the Figures 1, 2, and 3. In every case there is a good linear relationship which may be expressed by the following equations

$$G_{dis-rep} = -0.03208 - 0.0767\ S_M, \quad R = 0.9959,$$
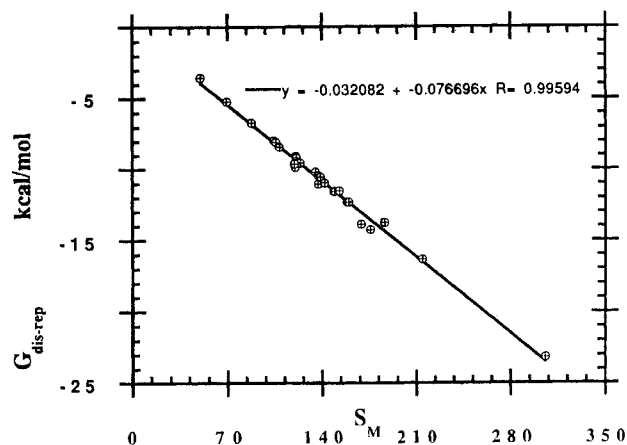$$\sigma = 0.345 \text{ kcal/mol}$$

$$G_{dis-rep} = -1.26850 - 0.07469\ V_M, \quad R = 0.9969,$$
$$\sigma = 0.302 \text{ kcal/mol}$$

$$G_{dis-rep} = -2.366 - 0.16109\ N_M, \quad R = 0.9854,$$
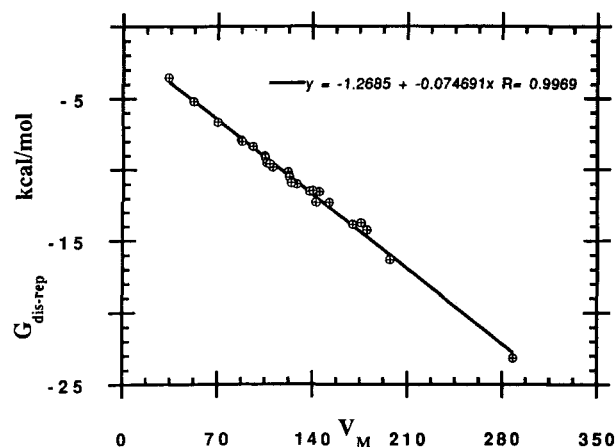$$\sigma = 0.653 \text{ kcal/mol}$$

Surface and volume seem to be predictors of $G_{dis-rep}$ for hydrocarbons in water of completely equivalent quality.

The neglect of distant contributions to the surface integrals has been performed by introducing one

additional parameter, R_{cut}. When $r_{mj} \geq R_{cut}$ the corresponding contribution is eliminated; r_{mj} is the distance between one atom m of the solute and the center j of the sphere to which the contribution to the surface integrals belongs. We have considered convenient for computational purposes, and for clarity of interpretation, to use the r_{mj} distance instead of the distance between m and the center of the tessera k on the cavity surface. With the present definition, contributions from distant atoms are accepted, or rejected in total, for each subunit defined by a van der Waals sphere. The analysis is here limited to the solutes of largest size considered in Table I, these being the most interesting cases. Table III

Table II. Statistical analysis of the separate $G_{dis}^{(n)}$ and G_{rep} components.

	$G_{dis}^{(6)}$ (%)	$G_{dis}^{(8)}$ (%)	$G_{dis}^{(10)}$ (%)	G_{rep} (%)
Mean	89.01	9.01	1.97	−16.92
Range	1.82	1.35	0.46	−5.12
Std. deviation	0.36	0.28	0.10	−1.09

Figure 1. Plot of $G_{dis-rep}$ (M in L) against S_M, area of the van der Waals molecular surface.

Figure 2. Plot of $G_{dis-rep}$ (M in L) against V_M, van der Waals molecular volume.
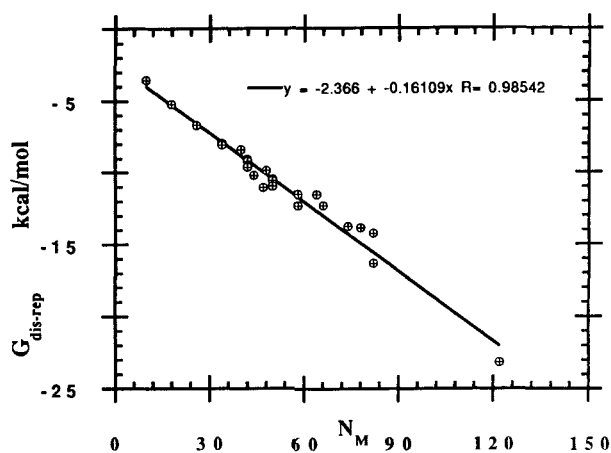
Figure 3. Plot of $G_{dis-rep}$ (M in L) against N_M, number of electrons of the solute molecule.

reports the percentage of $G_{dis-rep}$ recovered with different values of R_{cut}. Table IV displays analogous values, but referred to the $G_{dis}^{(6)}$ contribution alone. To make more evident the effect of the truncation we report the same data in a graphic form in Figures 4 and 5, adding some values not reported in the tables. It results that with $R_{cut} = 3.5$ Å one obtains reliable estimates of $G_{dis}^{(6)}$ as well as $G_{dis-rep}$.

CONCLUSIONS

The article is addressed to document the formal elaboration of a simple method to compute dispersion-repulsion contributions to the solvation energy of molecular solutes. The formal exposition has been reported in detail because we consider this approach applicable also to other more complex models of solution, including e.g., local dishomogeneities (cybotactic regions), dishomogeneities on a large scale (phase boundaries) and the presence of other bodies in solution. Some of these extensions have already been considered in unpublished studies. The mathematical method may also be applied to other aspects of the solute-solvent interaction, e.g., electrostatic contributions with nonhomogeneous or nonisotropic dielectrics, spin-spin couplings, etc. A generalized reduction of the solvation terms to surfaces integrals may also be exploited, as suggested in the introduction, in expanding the applications of the reduction of the chemical information at the molecular surface level.[7,8,20] The use of atom-atom parameters in computing $G_{dis-rep}$, adopted in the present study, is supported by several studies on dimeric systems in vacuo: the reader is referred to the analysis done by Szczęśniak and Scheiner[21] as a

Table III. Percentage of $G_{dis-rep}$, for some fixed R, radius of cutoff.

Solute	$R = 1.50$ Å	$R = 3.00$ Å	$R = 4.50$ Å	$R = 6.00$ Å	$R = 7.50$ Å
n-heptane	59.23	96.16	98.66	99.63	99.91
Decaline(T)	49.96	95.11	99.15	99.97	100.00
n-decane	55.46	95.21	98.15	99.43	99.77
3-Et-3-isoPr-pentane	52.04	93.53	99.49	100.00	100.00
n-pentadecane	52.56	94.43	97.73	99.25	99.67

Table IV. Percentage of G_{dis}, $n = 6$ only, for some fixed R, radius of cutoff.

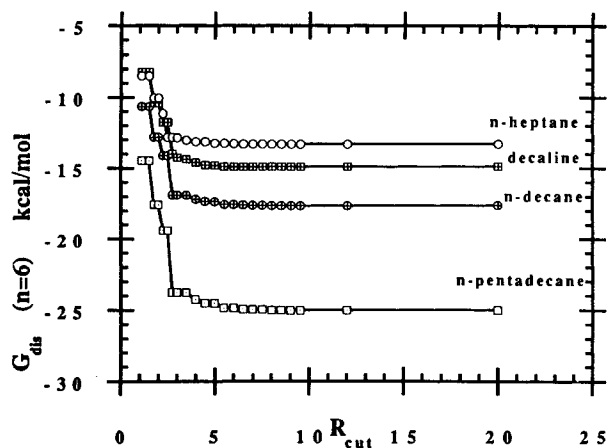| Solute	$R = 1.50$ Å	$R = 3.00$ Å	$R = 4.50$ Å	$R = 6.00$ Å	$R = 7.50$ Å
n-heptane	63.80	96.56	98.79	99.67	99.92
Decaline(T)	55.22	95.58	97.96	99.97	100.00
n-decane	60.43	95.72	98.34	99.49	99.80
3-Et-3-isoPr-pentane	55.89	94.01	99.53	100.00	100.00
n-pentadecane	57.85	95.01	97.95	99.31	99.70

Figure 4. Dependence of G_{dis}, the term $n = 6$ only, on the cutoff parameter.
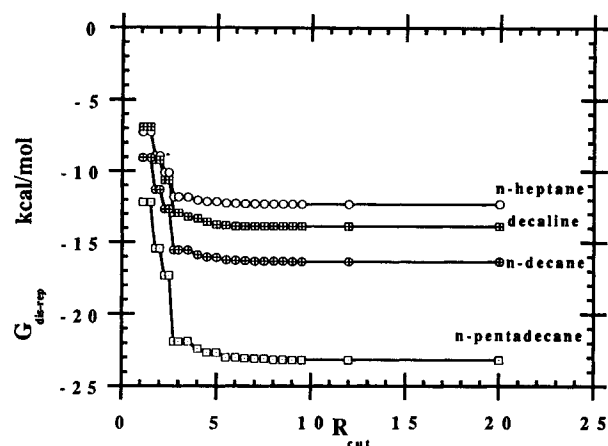

Figure 5. Dependence of $G_{dis-rep}$ (M in L) on the cutoff parameter.

typical example of comparison of the performance of empirical atom-atom potential with respect to *ab initio* calculation. The application of these potentials to solvation problems seems to give sensible results: article I and an article by Langlet et al.[22] give abundant information on this point. The atom-atom potentials we have employed may also be replaced, at least in the systems we have studied, by simpler group-atom potentials (in the present case CH-l, CH$_2$-l, CH$_3$-l, with l = H,O) without appreciable loss of information and remarkable gain in computational efficiency. This topic will be treated in detail in another article.

Empirical potentials may be replaced by the computational procedure elaborated by Aguilar and Olivares del Valle[13] which still reduces the calculation of $G_{dis}^{(6)}$ to the evaluation of an integral over the molecular surface, but starting from a different approach. A comparative analysis of the approach employed here with that given by reference 13 should give more insight to the important problem

of the dispersion contributions to the solvation phenomena.

The cases considered here, in spite of the simplicity of the model we have employed (homogeneous solvent distribution, solutes limited to the hydrocarbons class) give new information about some features of $G_{dis-rep}$ contribution to the solvation free energy. The repulsion term is to a good extent compensated by the $G_{dis}^{(8)}$ and $G_{dis}^{(10)}$ contributions. $G_{dis-rep}$, as well as the leading term $G_{dis}^{(6)}$ may be approximated, with a good precision, by contributions lying within a radius $R_{cut} \simeq 3.5$ Å. We recall that in liquid systems containing large bodies the dispersion contributions govern the long range behavior of the solute-solvent interaction, in spite of their local short-range characteristics when compared with electrostatic interactions. The reasons for this apparent contradiction are well known.[23] We may add that the short-range character of $G_{dis-rep}$ makes itself manifest as the *source* of $G_{dis-rep}$ is considered: the *source* has a local character, as shown by the small value of R_{cut} sufficient to give a reliable description of $G_{dis-rep}$. Another test performed in this article regards the correlation between $G_{dis-rep}$ and the area of the surface or the volume of the cavity. The use of S_M or V_M as molecular predictors has been the object of numerous papers (see e.g., Meyer[24]) among which a certain number of them express contrasting views on the relative superiority of S_M^{25-30} or of V_M^{31-34} as predictors. Our results, limited to the $G_{dis-rep}$ contribution for hydrocarbons in water, show that the two predictors are equivalent. A similar conclusion reached from a different starting point, is implicitly presented in a recent paper by Hall and Smith.[35] This simple and limited example should indicate the potential utilizations of the present method, which provides reliable information on S_M, V_M and $G_{dis-rep}$ computed at a comparable level of accuracy. In addition, these quantities are obtained within the framework of a coordinated computational setup which provides estimates of the electrostatic component of free solvation energy as well as the separate evaluation of the enthalpic and entropic contributions of all the components of the solvation energy, all at a comparable level of accuracy. We are confident that these computational methods will be able to shed more light on the subtle and complex physico-chemical features of molecular solutions.

APPENDIX

Transformation of the Volume Integrals of eq. (3) into Surface Integrals

The transformation is just a simple application of the Green–Ostragradsky theorem. We report the derivation, already done in article I and in references 13 and 36 to document the expression given in the

test and to show other extensions of the basic formulation. The simplest case of integration of eq. (3) corresponds to the case of a convex solute in an infinitely dilute solution. The volume integration regards all the space, with the exception of the molecular space C_l. We designate with Σ_l the internal boundary and we introduce a closed surface, completely containing the cavity C_l, defining the external boundary of the solvent, say a sphere $S(R)$; R will tend to infinity. Making use of the G-O theorem one obtains the following expression:

$$\int_{C_l} \nabla \cdot \mathbf{A}(\mathbf{r}_{ml}) dr_{ml} = \int_{\Sigma_l} \mathbf{A}(\mathbf{r}_{ml}) \cdot \mathbf{n}_{\Sigma_l} ds$$

$$+ \lim_{R \to \infty} \int_{S(R)} \mathbf{A}(\mathbf{r}_{ml}) \cdot \mathbf{n}_S ds \quad (A.1)$$

where $\mathbf{n}_{\Sigma_l}$ is a unit vector directed from Σ_l to the interior of the cavity containing the solute, and $\mathbf{n}_S$ is a unit vector directed towards larger R. We therefore need an expression of a vector function $\mathbf{A}(\mathbf{r}_{ml})$ in such a way that

$$\nabla \cdot \mathbf{A}(\mathbf{r}_{ml}) = F(\mathbf{r}_{ml}) \quad (A.2)$$

where $F(\mathbf{r}_{ml})$ is a generic integrand of eq. (3).

The simplest form of $\mathbf{A}$ may be obtained by imposing

$$\mathbf{A}(\mathbf{r}_{ml}) = f(r_{ml})\mathbf{r}_{ml} \quad (A.3)$$

the function $f(r_{ml})$ no longer depends on the vector $\mathbf{r}_{ml}$ but on its modulus. This simplification is reasonable in most physical models. The current expressions for the dispersion-repulsion potentials give a dependence of the distance alone, and ρ function mainly depends on the distance between m and l. Using the approximation (A.3) one has:

$$\nabla \cdot \mathbf{A}(\mathbf{r}_{ml}) = r_{ml}f'(r_{ml}) + 3f(r_{ml}) \quad (A.4)$$

and

$$d[r_{ml}^3 f(r_{ml})] = r_{ml}^2 F(r_{ml}) dr_{ml} \quad (A.5)$$

from which:

$$f(r_{ml}) = \frac{1}{r_{ml}^3} \int_{r^0_{ml}}^{r_{ml}} x^2 F(x) dx \quad (A.6)$$

$$\mathbf{A}(\mathbf{r}_{ml}) = \left[\frac{1}{r_{ml}^3} \int_{r^0_{ml}}^{r_{ml}} x^2 F(x) dx \right] \mathbf{r}_{ml} \quad (A.7)$$

The lower integration limit in (A.7) can be fixed in different ways, according to the boundary conditions of the specific case under investigation. For dilute isotropic solutions it is convenient to define it in such way that the second integral of eq. (A.1) vanishes. To do so it is sufficient to introduce a restriction on g_{ml} $(\mathbf{r}_{ml})$ (see eq. (4)) consistent with the physical properties of dilute isotropic solutions, namely that $g_{ml} \to 1$ as $R \to \infty$. This restriction corresponds to put $r^0_{ml} = \infty$. We may thus write explicitly

the expression for the several vector functions we need to solve eq. (5)

$$\mathbf{A}^{(n)}(\mathbf{r}_{ml}) = d_{ml}^{(n)} \left[\frac{1}{r_{ml}^3} \int_{r_{ml}}^{\infty} \frac{g_{ml}}{x^{(n-2)}} dx \right] \mathbf{r}_{ml} \quad (A.8)$$

with $n = 6, 8, 10$ and:

$$\mathbf{A}^{(\text{rep})}(\mathbf{r}_{ml}) = -c_{ml}$$
$$\times \left[\frac{1}{r_{ml}^3} \int_{r_{ml}}^{\infty} g_{ml} \exp(-\gamma_{ml} r_{ml}) dx \right] \mathbf{r}_{ml} \quad (A.9)$$

In the present article we have presented results related to a simpler expression of $g_{ml}(\mathbf{r}_{ml})$, replacing the one given in eq. (4)

$$g_{ml}(\mathbf{r}_{ml}) = \begin{cases} 0, & \text{if } \mathbf{r}_{ml} \in C_l; \\ 1, & \text{if } \mathbf{r}_{ml} \notin C_l. \end{cases} \quad (A.10)$$

For this homogeneous approximation the integration of eq. (A.8) and eq. (A.9) give the following vector functions

$$\mathbf{A}^{(n)}(\mathbf{r}_{ml}) = \frac{d_{ml}^{(n)}}{(n-3)r_{ml}^n} \mathbf{r}_{ml} \quad (A.11)$$

with $n = 6, 8, 10$ and

$$\mathbf{A}^{(\text{rep})}(\mathbf{r}_{ml}) = -c_{ml}$$
$$\times \left[\frac{1}{\gamma_{ml} r_{ml}} + \frac{2}{\gamma_{ml} r_{ml}^2} + \frac{3}{\gamma_{ml} r_{ml}^3} \right]$$
$$\exp(-\gamma_{ml} r_{ml})\mathbf{r}_{ml} \quad (A.12)$$

A vector function for a repulsion contribution given under the form $c_{ml} r_{ml}^{-12}$ (not used in the present article) is

$$\mathbf{A}^{\text{rep}}(\mathbf{r}_{ml}) = -\frac{c_{ml}}{9r_{ml}^{12}} \mathbf{r}_{ml}. \quad (A.13)$$

The relationship between the results obtained with the homogeneous approximation and those obtained with more realistic expressions of $g_{ml}(\mathbf{r}_{ml})$ will be discussed in another article. So far we have used the approximation introduced with eq. (A.3), namely of the dependence of F on the modulus of $\mathbf{r}_{ml}$ alone. As remarked before, there is dependence on $\mathbf{r}_{ml}$ in the m-l potential as well as in the $\rho(\mathbf{r}_{ml})$ function. The isotropy of the interaction potential may be considered acceptable, especially when the potential is of the atom-atom type. There are, on the contrary, many physical systems which require anisotropic expressions for $g_{ml}(\mathbf{r}_{ml})$. To get the appropriate expressions for the transformation of the volume integrals (see eq. 3) into surface integrals it is necessary to examine the various possible cases or at least some of the most important ones. An important type of anisotropy is related to the presence of the boundary in the isotropic solution. Such a boundary may represent the occurrence of a gas-solution separation surface, or the occurrence of a separation between two immiscible liquids. In both cases the

separation may be represented by a flat surface, with an abrupt discontinuity. In these cases the close surface $S(R)$ introduced in connection with eq. (1) may be considered as corresponding to a cube. It is sufficient to recast our formulation in Cartesian coordinates and to introduce separate conditions for the face of the cube corresponding to the solution boundary with respect to the other five, going to infinity. This has been done in references 6 and 37 using in the first reference *homogeneous approximation* and in the second also more realistic descriptions. The same transformation may be used when the solution is limited by a flat solid. Different, but straightforward, analytical definitions of the boundary are necessary for solids having regular shapes, like spheres or cylinders (see e.g., for a semiquantitative discussion Israelachvili[23]). For irregular solids (as e.g., a large biomolecule) a different approach is necessary, which requires a numerical integration, similar to the one introduced here. A second type of anisotropy in solution regards the cybotactic region, i.e., the region of the solvent in which the distribution function is modified by the solute. For small solutes the cybotactic effects are fairly well represented by changes in the radial g_{ml} function. This description may no longer be sufficient for larger solutes with irregular shape especially when groups with a large cybotactic effect, such as cation groups are present. One could then pass to more detailed, and more complex models, in which the relevant cybotactic regions are singled out and described by an appropriate distribution function. For the calculations one may resort to the discrete expression of $G_{dis-rep}$ and to perform the numerical integrations outlined in eq. (A.8), (A.9) for each tessera. This relatively lengthy operation may be applied in general, and we have tested the performances for the case of the solute near a flat surface, for which, as said above, there is a speedier formulation. The solvent interactions in presence of irregular solids, mentioned above, may be treated in such a way. We have not considered here all the cases of interest. The description of the reduction of the dispersion-repulsion integral to a surface integral for systems containing an appreciable number of molecules, as well as for a solution covered by a thin surface film, will be discussed in a future article.

References

1. F. Floris and J. Tomasi, *J. Comp. Chem.*, **10**, 616 (1989).
2. S. Miertuš, E. Scrocco, and J. Tomasi, *Chem. Phys.*, **55**, 117 (1981).
3. R. Bonaccorsi, R. Cimiraglia, and J. Tomasi, *J. Comp. Chem.*, **4**, 567 (1983).
4. R. Bonaccorsi, M. Hodošček, and J. Tomasi, *J. Mol. Struct. (Theochem)*, **164**, 105 (1988).
5. R. Bonaccorsi, E. Ojalvo, and J. Tomasi, *Coll. Czech. Chem. Comm.*, **53**, 2320 (1988).
6. R. Bonaccorsi, F. Floris, P. Palla, and J. Tomasi, *Thermochimica Acta*, **162**, 213 (1990).
7. G. Alagona, R. Bonaccorsi, C. Ghio, and J. Tomasi, 8th Annual Conference of Molecular Graphics Society, St. Andrews (Scozia) 29-31 March 1989.
8. G. Alagona, R. Bonaccorsi, C. Ghio, and J. Tomasi, TOC at Mun '89: Theory of Environmental effects on Molecular Structure and Chemical Reactivity, St. John's Newfoundland (Canada) 25 June–1 July 1989.
9. B. Linder and D.A. Rabhenold, *Adv. Quant. Chem.*, **6**, 203 (1972).
10. T.B. Mac Rury and B. Linder, *J. Chem. Phys.*, **58**, (1973) 5388, 5398.
11. B.J. Costa Cabral, D. Rinaldi, and J.L. Rivail, *C.R. Accad. Sci. Paris.*, **208II**, 675 (1984).
12. D. Rinaldi, B.J. Costa Cabral, and J.L. Rivail, *Chem. Phys. Lett.*, **125**, 495 (1986).
13. M.A. Aguilar, and F.J. Olivares del Valle, *Chem. Phys.*, **138**, 327 (1990).
14. J.L. Pascual Ahuir, E. Silla, J. Tomasi, and R. Bonaccorsi, *J. Comp. Chem.*, **8**, 778 (1987).
15. S. Mariott, and R.D. Topsom, *J. Mol. Struct.*, **110**, 337 (1984).
16. F. Floris, A. Tani, and J. Tomasi, to be published.
17. J.L. Pascual Ahuir, E. Silla, J. Tomasi, and R. Bonaccorsi, GEPOL/87. (QCPE progr. 554) *QCPE Bull.*, **8**, 133 (1988).
18. J.L. Pascual Ahuir, and E. Silla, in *Proc. Workshop Quant. Chem.*, Girona (Spain) 1988.
19. F. Vigné-Maeder and P. Claverie, *J. Am. Chem. Soc.*, **109**, 24 (1987).
20. G.A. Arteca, and P.J. Mezey, *J. Comp. Chem.*, **9**, 554 (1988).
21. M.M. Szczęśniak and S. Scheiner, *J. Chem. Phys.*, **80**, 1535 (1984).
22. J. Langlet, P. Claverie, J. Caillet, and A. Pullmann *J. Phys. Chem.*, **92**, 1617 (1988).
23. J.N. Israelachvili, *Intermolecular and Surface Forces. With Application to Colloidal and Biological Systems*, Academic Press, London, 1985.
24. J.A.Y. Meyer, *J. Mol. Struct. (Theochem)*, **124**, 93 (1985).
25. H.H. Huhlig, *J. Phys. Chem.*, **41**, 1215 (1937).
26. R.B. Hermann, *J. Phys. Chem.*, **76**, 2754 (1972).
27. G.L. Amidon, S.H. Yalkowsky, and S.C. Valvani, *J. Phys. Chem.*, **79**, 2239 (1975).
28. S.C. Valvani, S.H. Yalkowsky, and J.L. Amidon, *J. Phys. Chem.*, **80**, 829 (1976).
29. W.L. Jorgensen, J. Gao, and C. Ravimohan, *J. Phys. Chem.*, **89**, 3470 (1985).
30. S.J. Gill, G. Olafsson, and I. Wadso, *J. Phys. Chem.*, **89**, 3759 (1985).
31. McAuliffe, *J. Phys. Chem.*, **70**, 1267 (1966).
32. N. Funasaki, S. Hada, and S. Neya, *J. Phys. Chem.*, **89**, 3046 (1985).
33. A.Y. Meyer, D. Farin, and D. Aunir, *J. Amer. Chem. Soc.*, **108**, 7897 (1986).
34. Y.K. Kang, G. Namethy, and H.A. Sheraga, *J. Phys. Chem.*, **91**, 4105 (1987).
35. G.C. Hall and C.M. Smith, *J. Mol. Struct. (Theochem)*, **179**, 293 (1988).
36. M.J. Huron and P. Claverie, *J. Phys. Chem.*, **76**, 2123 (1972).
37. F.M. Floris, Thesis of laurea, University of Pisa, (1989).