Theor Chem Acc (1998) 99:34–43

Theoretical
Chemistry Accounts

Regular article

Variation of surface partition in GEPOL: effects on solvation free energy and free-energy profiles

Christian S. Pomelli[1], **Jacopo Tomasi**[2]

[1] Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy
[2] Dipartimento di Chimica e Chimica Industriale, Via Risorgimento 35, I-56124 Pisa, Italy,
e-mail: tomasi@dcci.unipi.it

Received: 6 August 1997 / Accepted: 23 October 1997

Abstract. The problem of defining efficient strategies for partitioning the cavity surface in QM solvation procedures based on boundary elements methods is considered here. The GEPOL procedure to get the cavity surface, and its partition into tesserae is adopted as a starting point: a version with variable tesselation is presented. The procedure to build the new sphere tesselations is described and several different options to select the surface partition have been implemented. The effects of the variation of the surface partition on the free energy of solvation of several solutes are also presented. Two free energy of solvation profiles evaluated with several different cavity partitions are analysed. We find that a radius-driven tesselation for every sphere reduces the number and extension of the cavity artefacts.

Key words: Solvent Effect – Boundary elements method

1 Introduction

The concept of molecular surface, through lacking a precise definition, is so rich in heuristic significance that it has been used in a large variety of models in almost all subfields of molecular science.

Among these models, we focus our attention here to models in which this surface is used to define a portion of space in which other molecules are not allowed, and to describe properties related to the outer components of a complex material system. This is the realm of solvation models, in which a separation between a "solute" and a "solvent" is introduced. Among these models, which can be based on discrete or continuous descriptions of the solvent, we limit our attention to continuous models, and in particular to the polarizable continuum model (PCM) we have developed. For a general review on continuum solvation methods (including PCM) see Ref [1].

Correspondence to: J. Tomasi

PCM uses a boundary element method (BEM) to describe the interaction between the solute and the averaged distribution of the solvent.

To implement BEM computational procedures one has to define the surface limiting the solute, and partition it into small elements. The surface can also be considered a "molecular surface" (the molecule is the solute, actually a single molecule or a small cluster of molecules) and the partition of the surface, called tesselation, must satisfy some basic requisites. It must cover the whole surface without juxtapositions or unaccounted portions of the surface, give a precise value of the area of tessera and the location of representative points within the tessera. They must be one point at least, the centre of the tessera, and in some applications more, as the vertices of the polyhedron defining that tessera.

Other requirements are less imperative: the surface can be locally concave or convex, and then the tesserae are defined in terms of polyhedra on a curved surface, or can be replaced by a discrete set of planar surfaces, with planar tesserae.

The number of tesserae is another parameter which can be fixed as a balance between accuracy of the integration given via a BEM procedure, and the computational demands (computation time and occupation of the computer memory).

In PCM methods (there are a number of variants, see [2] for a recent account) use is generally made of the GEPOL procedure to define the surface and its tesselation.

GEPOL was conceived and firstly elaborated in our group as a component of PCM [3] and later revised and updated, and distributed as an independent code [4]. GEPOL has proven to be successful and is now employed in many solvation procedures [5, 6] and in other models not directly related to solvation. We currently use a derived version, called GEPOL-GB giving the more detailed information we need to compute the derivatives of tessera with respect to nuclear coordinates but otherwise equivalent to the most recent available GEPOL versions.

In this paper we document some other partitions of the GEPOL surface, and so it is necessary here to review a few points concerning GEPOL.

The cavity is defined in terms of spheres centred at the nuclei of the solute. When the radii, R_k, of such spheres are equivalent to the van der Waals radii $R_{vdw}(k)$, the surface is the van der Waals surface of the molecule Σ_{vdw}. In PCM calculations of the solvent reaction potential and the electrostatic portion of the free energy, the solvent-excluding surface Σ_E is adopted, in which the radii R_k are scaled by a factor α_k (typically 1.2 for all atoms) and additional spheres are used. Positions and radii of these new spheres are determined by the position of the original spheres and by an additional parameter: the radius R_s of the solvent. It is not necessary to report the details of this procedure here; suffice to say it is an iterative procedure giving origin to several "generations of spheres". Some cut-off parameters avoid an excessive proliferation of spheres. These additional spheres will fill crevices on the molecular surface too small to accommodate solvent molecules and to describe the changes in the solvent forbidden space during processes of molecular dissociation/association or conformational changes of a special type which we shall examine later. In both cases the number of additional spheres can be quite large. For other applications use is made of the solvent accessible surface Σ_A, in which all radii R_k are supplemented by a constant term equal to the mean radius of the solvent R_s: $R'_k = R_k + R_s$.

Tesselation in GEPOL is always given in terms of an inscribed pentakisdodecahedron for each sphere present in the model. Some tesserae are eliminated because they are completely within the inner volume, others are cut, and replaced by spherical polygons of irregular shape, whose geometric characteristics are determined with the aid of the Gauss-Bonnet procedure used in GEPOL-GB.

This procedure is quite efficient, but somewhat rigid. In some cases the tesserae of the additional spheres have a small area when compared to others. The number of the tesserae is an important parameter in assessing the computational effectiveness of the solvation method, which scales as N_t^3 (N_t = total number of tesserae) for the inversion of the matrix in the BEM procedure and scales as N_t^2 in the evaluation of the solute-solvent interaction integrals. Here we present a method giving more freedom in the selection of tesserae, supplemented with a study to find rules to give the best computational results with the lowest number of tesserae. This paper is organized in the following way:

1. A description of the procedure used to generate the surface partition.
2. The description of the implementation within our PCM-BEM method.
3. The exposition and the discussion of the results obtained.

2 Polyhedra and geodesic notation

In this work we consider only tesselations with triangular tesserae derived from pentakisdodecahedrons, tetrahedrons and icosahedrons as particular cases of sphere tesselation. It is well known the sphere tesselation is a fundamental problem of algebraic topology [7].

Any tesselation of the sphere is related to a plane polyhedron inscribed in the sphere by a bijective relation:

1. The vertices of polyhedron and the vertices of tesselation are the same.
2. Any spherical polygon of tesselation corresponds to a polyhedron face. The two corresponding objects share the same vertices.
3. Any great circle arc that connects two vertices of any spherical polygon of tesselation corresponds to an edge of the polyhedron with the same two vertices.

The polyhedron properties we shall describe in the following also apply to tesselations.

A generic polyhedron is topologically determined by the number of faces (N_t), edges (N_e) and vertices (N_v). These three quantities are related by the Euler-Poincaré relation [7]:

$$\chi = N_v + N_t - N_e \ . \tag{1}$$

The Euler-Poincaré characteristic χ for a polyhedron inscribed within a sphere is 2 [7].

In a triangular polyhedron

$$N_e = \tfrac{3}{2} N_t \ , \tag{2}$$

thus

$$N_v = 2 + \tfrac{1}{2} N_t \ . \tag{3}$$

A useful tool to describe a polyhedron is the so-called geodesic notation [8]: $\{p, q\}$ where p is the number of faces that share a vertex and q the number of vertices for each face. If a "+" is placed after the q value then all faces must be divided into triangles by adding an edge between all vertices of the faces to the medium point of the corresponding spherical polygon.

The numbers of vertices, triangles, edges and the geodesic notations for platonic polyhedra and pentakisdodecahedron are reported in Table 1.

A large ensemble of polyhedra can be obtained by equilateral division of any triangular polyhedron [8].

In equilateral division of the order M any primitive triangle is replaced by M^2 triangles:

$$N_t = N_{t_0} M^2 \ , \tag{4}$$

$$N_v = 2 + \tfrac{1}{2} N_{t_0} M^2 \ , \tag{5}$$

$$N_e = \tfrac{3}{2} N_{t_0} M^2 \ , \tag{6}$$

Table 1. Faces, vertices, edges and geodesic notations (Geodesic) for the five platonic polyhedra and for pentakisdodecahedron

Polyhedra	Faces	Vertices	Edges	Geodesic
Tetrahedron	4	4	6	{3,3}
Cube	6	8	12	{3,4}
Octahedron	8	6	12	{4,3}
Icosahedron	20	12	30	{5,3}
Dodecahedron	12	20	30	{3,5}
Pentakisdodecahedron	60	32	90	{3,5+}

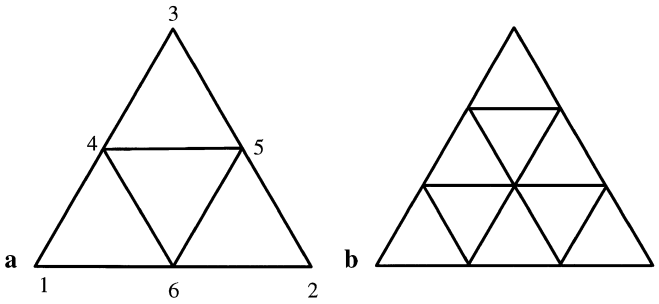


Fig. 1. **a** Equilateral division of the second order: the triangle (1; 2; 3) is replaced by the triangles (1; 4; 6), (4; 5; 6), (2; 5; 6) and (3; 4; 5). **b** Equilateral division of the third order

where N_t is the final and N_{t_0} the primitive number of triangles.

The geodesic notation can be extended to these polyhedra: $\{p,q\}_{M,M}$ the index M is double because a non-equilateral division also exists. This second procedure leads to polyhedra constituted of triangles greater in breadth than height as faces, not appropriate for BEM methods.

The procedure of equilateral division of a triangle is shown in Fig. 1.

All edges are divided into M equal segments by $M-1$ new vertices. Between any pair of new vertices we draw a new segment (parallel to a primitive edge). These segments intersect and thus create new vertices but the vertices do not lie on primitive edges.

Now we derive the expressions for the new vertices coordinates. The new vertices divide the primitive arc into M equal arcs. If $\vec{v}_1$ and $\vec{v}_2$ are the primitive vertices of the primitive arc, thus

$$\theta_0 = \arccos\left(\frac{\vec{v}_1 \cdot \vec{v}_2}{|\vec{v}_1||\vec{v}_2|}\right) \tag{7}$$

is the angle, with respect to the sphere centre, between the two primitive vertices. A new vertex $\vec{v}(k)$ subtends an angle

$$\theta_1 = \frac{k}{N}\theta_0 \ , \tag{8}$$

to $\vec{v}_1$, and at an angle of

$$\theta_2 = \frac{N-k}{N}\theta_0 \tag{9}$$

to $\vec{v}_2$.

If we write

$$\vec{v}(k) = a\vec{v}_1 + b\vec{v}_2 \tag{10}$$

then, according to Eqs. (7–9):

$$a = \frac{\cos\left(\frac{k}{N}\theta_0\right) - \cos(\theta)\cos\left(\frac{N-k}{N}\theta_0\right)}{\sin^2(\theta)} \ , \tag{11}$$

$$b = \frac{\cos\left(\frac{N-k}{N}\theta_0\right) - \cos(\theta)\cos\left(\frac{k}{N}\theta_0\right)}{\sin^2(\theta)} \ . \tag{12}$$

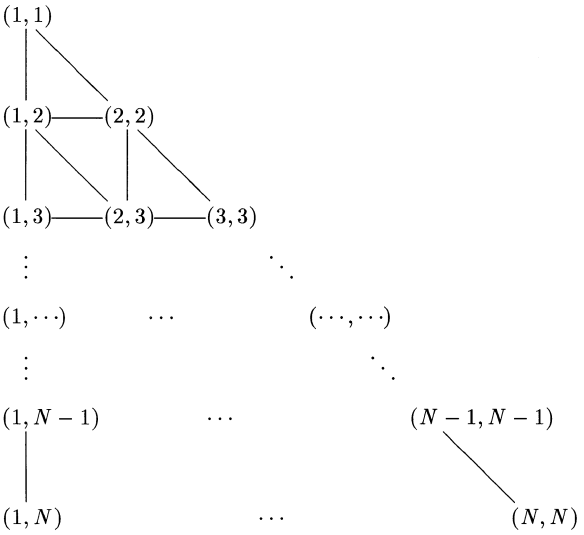


Fig. 2. Graphical representation of the equilateral division procedure. The vertices are ordered in the low-triangular part of a square matrix. The lines are the curvilinear sides (arcs) of the new triangles

The new vertices not lying on primitive arcs are determined by a similar procedure. Between any pair of new vertices, equidistant from a primitive vertex, $\vec{v}(l)$ and $\vec{v}'(l)$, there exist $l-1$ vertices. It is sufficient to substitute $\vec{v}_1$ and $\vec{v}_2$ with $\vec{v}(k)$ and $\vec{v}'(k)$, and N with l in Eqs. (10–12).

The vertices establish the metric of the polyhedron. To complete the polyhedron description we have to define the triangles. They are represented by a triad of positive integers each associated to a vertex.

Now consider Fig. 2. This is a matrix representation of vertices of the equilateral division of a triangle. The symbols $(1,1)$, $(1,N)$, (N,N) represent the primitive vertices, the other elements of the matrix represent new vertices. The lines shown in the figure connect vertices of the new triangles: these lines represent the arcs. The primitive triangle is represented by

$$(1,1);\ (1,N);\ (N,N)\ . \tag{13}$$

Simple algebraic observations lead to the following expressions for new triangles:

$$\begin{aligned} &(i,j);\ (i+1,j);\ (i+1,j+1) \quad i=1,N \qquad j=1,i \\ &(i,j);\ (i,j+1);\ (i+1,j+1) \quad i=2,N \quad j=1,i-1 \end{aligned} \tag{14}$$

the first row leads to $N(N+1)/2$ triangles, the second to $N(N-1)/2$.

This procedure builds a tesselation of the sphere as a set of triangles. These triangles have, in general, no equal area except if we use the basic polyhedra. This is due to the fact that the projection on the sphere magnifies the new triangles by different factors: these are larger for new triangles that are placed at the centre than for those placed on the border of the triangle subjected to division. The new triangles are similar in area if we start from a polyhedron with a large number of tesserae like pentakisdodecahedron.

3 Implementation

This method has been implemented in our version of PCM included in the Gaussian 94 set of programs [9].

There are two options to build the tesselation:

1. *TsN* option. Each atomic or additional sphere is partitioned into the same number of tesserae. We indicate this choice with the notation $TsN = N$ where N is the number of tesserae in the chosen sphere tesselation. By setting $TsN = 60$ we have the traditional GEPOL cavity.
2. *TsA* option. For each atomic or additional sphere, the partition is chosen according to the sphere size: the program selects, for each sphere, the tesselation that leads to a tesselar mean area most similar to that of a given input value. We indicate this choice with the notation $TsA = A$ where A is the input area in Å^2.

We shall limit ourselves to consider the influence of the cavity partitioning on the electrostatic component of the free energy of solvation. There are some non-electrostatic components of the free energy of solvation, generally they are expressed in the PCM computational scheme as cavitation, dispersion and repulsion terms each separately computed. The cavitation term G_{cav}, computed with the aid of the Pierotti-Claverie approximation [1], formally depends on the exposed area of each sphere, and not on this tesselation. The Floris-Tomasi [10] formulas for the dispersion and repulsion contributions, $G_{\text{dis}} + G_{\text{rep}}$, actually use a partition of the surface in tesserae. There is however no important difference between the use of this partitioning for the electrostatic and for dispersion-repulsion contributions. In the electrostatic term the local elements are inserted into an iterative BEM procedure, quite delicate to handle and time consuming. In the repulsion and dispersion terms the local elements are directly obtained and summed, and easy to compute. A change in the tesselation strategy changes $G_{\text{dis}} + G_{\text{rep}}$ by an amount less than 0.5% for the whole set of partition examined. The change in the computation time is completely irrelevant, given the very short time necessary for these calculations (less than 1% of the total CPU time).

We do not report a full exposition of the electrostatic part of the PCM method but we focus our attention on the more CPU-consuming step: evaluation of the apparent surface charges (ASC). There are several methods to evaluate ASC [1], we adopt here the direct-BEM approach:

$$\mathbf{q} = \mathbf{D}^{-1}\mathbf{f} \tag{15}$$

where $\mathbf{f}$, $\mathbf{q}$ are column vectors of size N (where N is the number of tesserae) which respectively contain the flux of the electric field through any tessera of the cavity and the corresponding ASC.

$\mathbf{D}$ is a square $N \times N$ matrix with elements

$$D_{kl} = \begin{cases} \frac{4\pi\epsilon}{\epsilon-1}(1 - A_k) & k = l \\ \frac{4\pi\epsilon}{\epsilon-1} B_{kl} & k \neq l \end{cases} , \tag{16}$$

where A_k and B_{kl} have the following expressions:

$$A_k = \frac{\epsilon - 1}{2\epsilon}\left(1 - \sqrt{\left(\frac{\Delta S_k}{4\pi R_k^2}\right)}\right) , \tag{17}$$

$$B_{kl} = \frac{\epsilon - 1}{4\pi\epsilon}\Delta S_k \frac{(\vec{r}_k - \vec{r}_l)\cdot \hat{n}_k}{|\vec{r}_k - \vec{r}_l|} . \tag{18}$$

Here ΔS_k is the area of the kth tessera, R_k is the curvature radius, r_k is the medium point and $\hat{n}_k$ is the unit extern normal vector. For a more detailed presentation of the PCM-BEM approach see [1] and references quoted therein.

4 Results and discussion

4.1 Fix geometry solutes

We report on the results of systematic ab initio calculations for five molecules with an extended set of cavities built with different choices of sphere tesselation.

The basis set 6-31G* and the standard Pauling radii [11] for atoms have been used. Only the electrostatic contribution is considered because the other terms do not sensibly depend on the type and granularity of the cavity tesselation.

The values of the electrostatic free energy of solvation for water, hydrogen fluoride, acetate anion, glycine (zwitterionic form) and dioxane in water at standard pressure/temperature conditions are reported in Tables 2–6. These tables are organized in the following way:

1. The first row of each table reports the various options for the sphere partition: the *TsN* options indicated with $4M^2$, $20M^2$ and $60M^2$ refer to partitions of the sphere starting from a tetrahedron, a icosahedron and a pentakisdodecahedron, respectively.
2. For each tesselation option given in the first row we report the parameter (A or M) on which the calculation depends, the resulting number of tesserae N and the unsigned difference δ with respect to the convergence value of ΔG_{el} in k cal/mol.
3. The last row contains the convergence value of ΔG_{el}. This value is obtained as the mean of the elements of each set with the largest number of tesserae.

Some values for coarse partitions of the cavity are not reported because the relative calculation does not converge. For glycine we also report a graphical representation of these data in Fig. 3.

As a reference we now consider the standard GEPOL tesselation: the pentakisdodecahedron ($TsN = 60$). The elements of the other sets that have a comparable number of tesserae are $TsN = 4 \cdot 4^2$ for the tetrahedron series and $TsN = 20 \cdot 2^2$ for the icosahedron series. For the *TsA* series the answer depends upon the kind of atoms that constitute the molecule. By examining the reported cases we obtain a value of $TsA = 0.4{-}0.5\text{Å}^2$. In all these cases the relative error is limited to 1–2%. This error is lower than that affecting the experimental

Table 2. Electrostatic free energy of solvation at the HF/6-31G* level with several cavity options for H_2O. $|\delta|$ in kcal/mol

$TsA = A$			$TsN = 4M^2$			$TsN = 20M^2$			$TsN = 60M^2$		
A	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$
5.0	12	2.04	1	12	2.04	1	48	0.04	1	126	0.05
4.5	22	0.37	2	36	1.02	2	158	0.07	2	427	0.01
4.0	22	0.37	3	76	0.40	3	334	0.01	3	922	0.00
3.5	22	0.37	4	130	0.04	4	558	0.02			
3.0	36	1.02	5	198	0.01	5	853	0.00			
2.5	36	1.02	6	266	0.04						
2.0	40	0.34	7	356	0.03						
1.5	40	0.34	8	450	0.01						
1.0	60	0.13	9	570	0.00						
0.9	76	0.40	10	688	0.00						
0.8	96	0.14	11	823	0.01						
0.7	96	0.14	12	976	0.00						
0.6	100	0.09	13	1124	0.00						
0.5	136	0.05	14	1300	0.00						
0.4	160	0.00	15	1480	0.00						
0.3	210	0.12									
0.2	300	0.04									
0.1	550	0.00									
0.05	1082	0.00									

Convergence value: −6.41 kcal/mol

Table 3. Electrostatic free energy of solvation at the HF/6-31G* level with several cavity options for HF. $|\delta|$ in kcal/mol

$TsA = A$			$TsN = 4M^2$			$TsN = 20M^2$			$TsN = 60M^2$		
A	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$
5.0	8	–	1	8	–	1	35	0.49	1	85	0.11
4.5	8	–	2	24	0.06	2	115	0.00	2	330	0.02
4.0	20	0.82	3	54	0.16	3	235	0.00	3	720	0.00
3.5	20	0.82	4	96	0.03	4	415	0.00	4	1245	0.01
3.0	20	0.82	5	138	0.10	5	680	0.01			
2.5	20	0.82	6	200	0.05	6	940	0.00			
2.0	28	0.43	7	268	0.02	7	1250	0.00			
1.5	28	0.43	8	348	0.03						
1.0	51	0.12	9	436	0.03						
0.9	51	0.12	10	514	0.02						
0.8	51	0.12	11	632	0.02						
0.7	70	0.08	12	750	0.01						
0.6	73	0.19	13	862	0.02						
0.5	82	0.14	14	1010	0.00						
0.4	93	0.06	15	1146	0.00						
0.3	128	0.11	16	1298	0.00						
0.2	200	0.01	17	1478	0.00						
0.1	381	0.03									
0.05	786	0.01									

Convergence value: −6.80 kcal/mol

data. The standard tesselation for the spheres can be thus considered a good choice.

Several sets of calculations show an oscillatory behaviour before reaching convergence. This is due to "roughness" introduced into the system by the PCM-BEM approach: every partition of the cavity leads to a different taxonomy of tesserae cuttings, and thus leads to a different spectrum of tesserae areas and positions of medium points. This can give rise to "local catastrophes" that alter values of the apparent surface charges introducing a change in the energy. This phenomenon disappears when we use a number of tesserae 2–4 times greater than the standard one.

In Table 7 we report some synthetic data about the rotational invariance of PCM. The tesselation is built, for each sphere, taking the principal axis of rotation of the polyhedra parallel to the principal axis of inertia of the molecule. If we choose a different relative orientation between molecule and polyhedra there are different situations of cutting between tesserae and spheres, and thus the partition of the cavity into tesserae is different. This procedure does not affect the measure of the total surface area nor the volume of the cavity. We have taken six different orientations of the water molecule, corresponding to the six permutations of the cartesian axes. Table 7 has an organization similar to that used for Tables 2–6: we report, instead of the absolute unsigned value, the standard deviation of the six values obtained at fixed-sphere partition with the six different polyhedra orientations. The standard deviation corresponds to

Table 4. Electrostatic free energy of solvation at the HF/6-31G* level with several cavity options for CH_3COO^-. $|\delta|$ in kcal/mol

$TsA = A$			$TsN = 4M^2$			$TsN = 20M^2$			$TsN = 60M^2$		
A	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$
5.0	25	–	1	25	0.77	1	102	0.10	1	269	0.15
4.5	58	2.73	2	98	0.07	2	331	0.06	2	908	0.02
4.0	58	2.73	3	191	0.23	3	683	0.06			
3.5	58	2.73	4	301	0.02	4	1174	0.03			
3.0	98	1.27	5	437	0.04						
2.5	98	1.27	6	591	0.01						
2.0	93	0.08	7	759	0.05						
1.5	93	0.08	8	976	0.04						
1.0	148	0.07	9	1200	0.01						
0.9	191	0.07	10	1459	0.01						
0.8	217	0.06									
0.7	217	0.06									
0.6	231	0.10									
0.5	282	0.12									
0.4	362	0.09									
0.3	447	0.02									
0.2	641	0.02									
0.1	1149	0.01									

Convergence value: −79.80 kcal/mol

Table 5. Electrostatic free energy of solvation at the HF/6-31G* level with several cavity options for $^+H_3NCH_2COO^-$. $|\delta|$ in kcal/mol

$TsA = A$			$TsN = 4M^2$			$TsN = 20M^2$			$TsN = 60M^2$		
A	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$
5.0	48	–	1	37	–	1	143	0.14	1	336	0.06
4.5	63	0.88	2	130	2.98	2	436	0.02	2	1103	0.02
4.0	85	1.65	3	231	0.62	3	874	0.04			
3.5	93	1.50	4	372	0.31	4	1434	0.03			
3.0	130	2.98	5	525	0.45						
2.5	130	2.98	6	719	0.03						
2.0	136	0.02	7	940	0.01						
1.5	140	0.23	8	1205	0.01						
1.0	193	0.83	9	1471	0.02						
0.9	237	0.47									
0.8	252	0.21									
0.7	281	0.18									
0.6	305	0.27									
0.5	383	0.07									
0.4	427	0.05									
0.3	532	0.29									
0.2	809	0.03									
0.1	1435	0.03									

Convergence value: −30.96 kcal/mol

about 0.5% of ΔG_{el} for TsA in the range 4.5−0.3 and TsN in the range 16−200, and practically zero for the finest sphere partitions. The $\sigma_{\Delta G_{\mathrm{el}}}$ obtained for $TsA = 5.0$ as well as for $TsN = 4$ (leading to two equal cavity partitions) are zero because the group of rotations applied to the molecule is as that applied to the cavity. This accidental rotational invariance leads, however, to a value of ΔG_{el} very far from the convergence value. To have a ΔG_{el} free from rotational artefacts we have to use a number of tesserae 2−4 times larger than the "standard" value to avoid "local catastrophes".

4.2 Free-energy profiles

We now pass on to examine a couple of free-energy profiles, the first corresponding to a chemical reaction, the second to a conformational transformation. Both examples show typical variations of the cavity shape occurring when a molecular system changes its molecular geometry.

The S_N2 reaction

$$Cl_a^- + CH_3Cl_b \rightarrow CH_3Cl_b + Cl_a^- \ , \tag{19}$$

selected as first example (Fig. 4), is representative of all bimolecular reactions $A + B \rightarrow C + D$. During the reaction there are two solvent separated species which merge into a complex structure and then separate again.

The symmetry of our S_N2 reaction allows us to examine half of the energy profile only. The unique cavity present at the transition state is split into two when the reaction coordinate

$$r_c = r_{C-Cl_a} - r_{C-Cl_b} \tag{20}$$

Table 6. Electrostatic free energy of solvation at the HF/6-31G* level with several cavity options for dioxane. $|\delta|$ in kcal/mol

$TsA = A$			$TsN = 4M^2$			$TsN = 20M^2$			$TsN = 60M^2$		
A	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$	M	N_t	$\|\delta\|$
5.0	46	–	1	46	–	1	211	0.10	1	476	0.04
4.5	83	3.82	2	186	0.04	2	648	0.12			
4.0	97	2.60	3	327	0.38	3	1181	0.03			
3.5	97	2.60	4	518	0.19						
3.0	171	0.03	5	720	0.09						
2.5	171	0.03	6	1012	0.03						
2.0	183	0.18	7	1309	0.03						
1.5	198	0.18									
1.0	238	0.30									
0.9	314	0.36									
0.8	329	0.38									
0.7	358	0.13									
0.6	376	0.12									
0.5	513	0.21									
0.4	548	0.07									
0.3	675	0.08									
0.2	1052	0.01									
Convergence value: −7.40 kcal/mol											

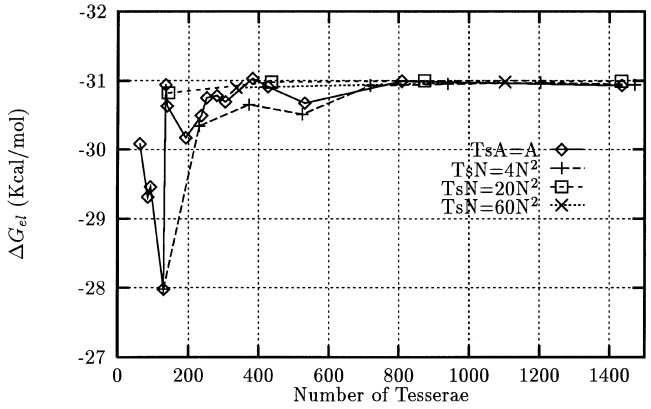


Fig. 3. Free energy of solvation for glycine (zwitterionic form) with different tesselations

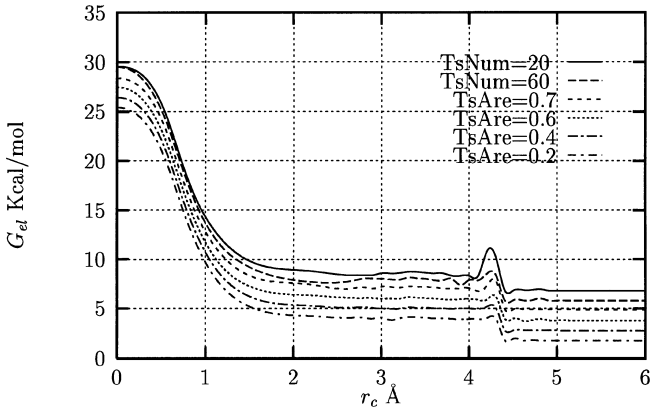


Fig. 4. Free energy profiles for the S_N2 reaction. G_{el} is set equal to zero for $TsA = 0.2$ at $r_c = \infty$. The other curves are shifted by 1 kcal/mol each to avoid juxtapositions

Table 7. Standard deviation of ΔG_{el} with respect to axes permutation for H_2O. $\sigma_{\Delta G_{el}}$ in kcal/mol

$TsA = A$		$TsN = 4M^2$		$TsN = 20M^2$		$TsN = 60M^2$	
A	$\sigma_{\Delta G_{el}}$	M	$\sigma_{\Delta G_{el}}$	M	$\sigma_{\Delta G_{el}}$	M	$\sigma_{\Delta G_{el}}$
5.0	0.00	1	0.00	1	0.06	1	0.02
4.5	0.03	2	0.02	2	0.01	2	0.00
4.0	0.03	3	0.03	3	0.01	3	0.00
3.5	0.03	4	0.01	4	0.01		
3.0	0.02	5	0.02	5	0.00		
2.5	0.02	6	0.00				
2.0	0.03	7	0.00				
1.5	0.03	8	0.00				
1.0	0.05	9	0.00				
0.9	0.03	10	0.00				
0.8	0.03	11	0.00				
0.7	0.03	12	0.00				
0.6	0.02	13	0.00				
0.5	0.03	14	0.00				
0.4	0.02	15	0.00				
0.3	0.03						
0.2	0.00						
0.1	0.00						
0.05	0.00						

has a value larger than 4.5 A. To gain a good description of the effect of the cavity variation along the reaction coordinate we have performed 37 ab initio calculations, at the HF/6-31G* level. All the geometries are fully optimized. The recent versions of PCM are all provided with geometry optimization subroutines [2]. In the case of the S_N2 reaction considered here, the changes in geometry in passing from vacuum to aqueous solution are extremely modest: at the TS there is a change in the R_{C-Cl} distance of 0.017 A (for more details see [12]). The energy profile of this reaction is not affected by the level at which geometry optimization has been performed.

The analysis of Fig. 4 leads to the following considerations:

1. All the profiles show some small "roughness" especially with r_c in the 4−4.5 A range. We attribute these artefacts to the local cavity situation similar to those described to explain the oscillation in the convergence of the ΔG_{el} in the first part of this section. The number and intensity of these phenomena is lower if we use a larger number of tesserae and the *TsA* option.
2. All profiles exhibit a step corresponding with cavity merging. In Fig. 5a–c we report some cavity-related quantities versus the reaction coordinate: the surface area, the number of spheres and the number of tessera. All these quantities show a peak in correspondence with the cavity disjunction.

This phenomenon can be considered to be an intrinsic characteristic of the solvent-excluding surface: when the two reagents cavities begin to merge together the solvent-excluding surface becomes larger because the solvent cannot pass through the two reagents. This creates a "bridge" between the two former separated cavities and thus the area becomes larger. A pictorial example is reported in Fig. 6. In our system the "bridge" was established when $r_c = 4.534$ A corresponding to $r_{C-Cl'} = 6.330$ A. At values of $r_c < 4.534$ A the "bridge" importance decreases until $r_{C-Cl'}$ becomes equal to the sum of the *C* and *Cl'* spheres radii. At this point the "bridge" disappears ($r_{C-Cl} = 4.140$ Å or $r_c = 2.344$). From a physical point of view when $r_c = 4.534$ a desolvation process begins because an increasing amount of the two fragments surfaces becomes inaccessible to the solvent. This explains why the step is negative when r_c increases.

The variation of the number of tesserae with respect to the reaction coordinate (reported in Fig. 5c) shows a significant difference between the *TsN* and the *TsA* options. In the first case the number of tesserae becomes four times larger in the cavity merging zone with respect to the transition state ($r_c = 0$) and reagents ($r_c = \infty$) zones. With the second option the ratio is only 1.2. Because the increase of the area in the cavity merging zone does not depend on the cavity partitioning, if we use the *TsN* option there is a large number of small tesserae because the radii of the added spheres are generally smaller than the atomic ones. With the *TsA* option we create tesserae of comparable area. This leads to a better formulated BEM problem: the **D** matrix off-diagonal elements are proportional to the tesserae area and when

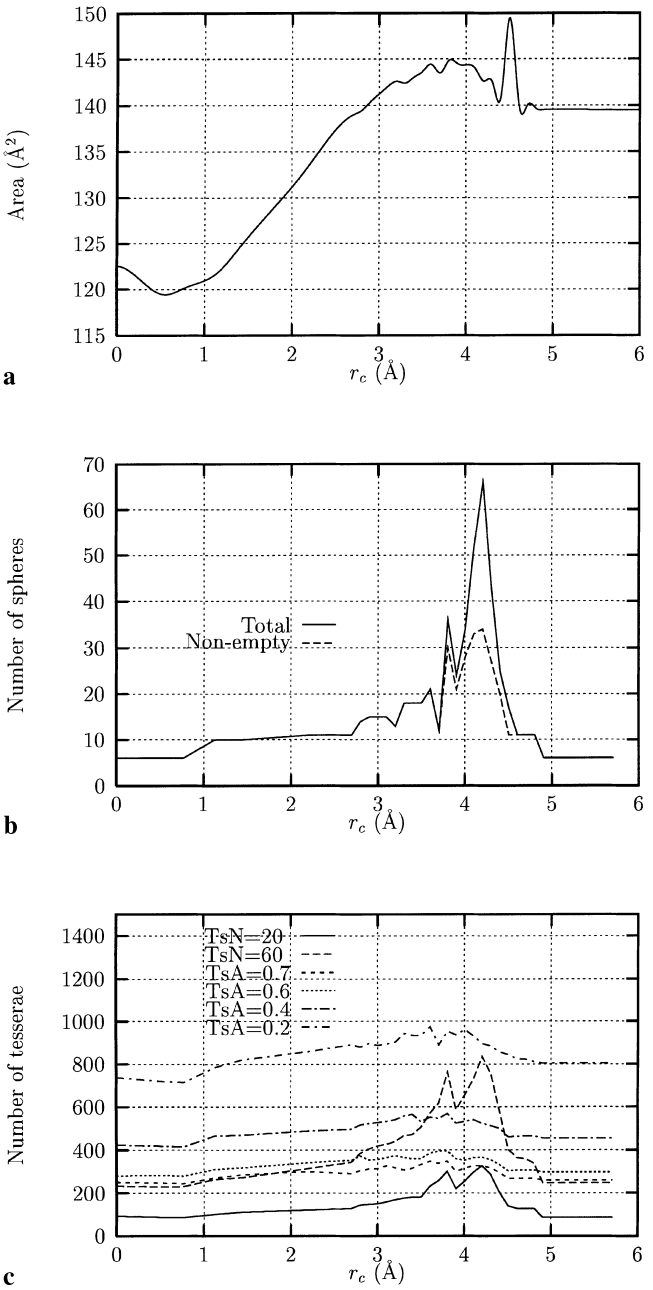


Fig. 5. a–c Variations of area **a** number of spheres **b** and number of tesserae **c** along the reaction coordinate for the S_N2 reaction

these elements are similar the BEM problem is better determined.

These considerations lead to the conclusion that the *TsA* option give tesselations of better quality than those obtained with *TsN*. A value of TsA of 0.4−0.5 A appears to be a good compromise between computational cost and numerical precision.

The second example shows similar phenomena, albeit of different origin. In conformational changes of molecules of medium-large size it is quite common to have rotations around a simple sigma bond connecting conformations where two groups R_1 and R_2 of the molecule are in close contact. They both contribute to define a

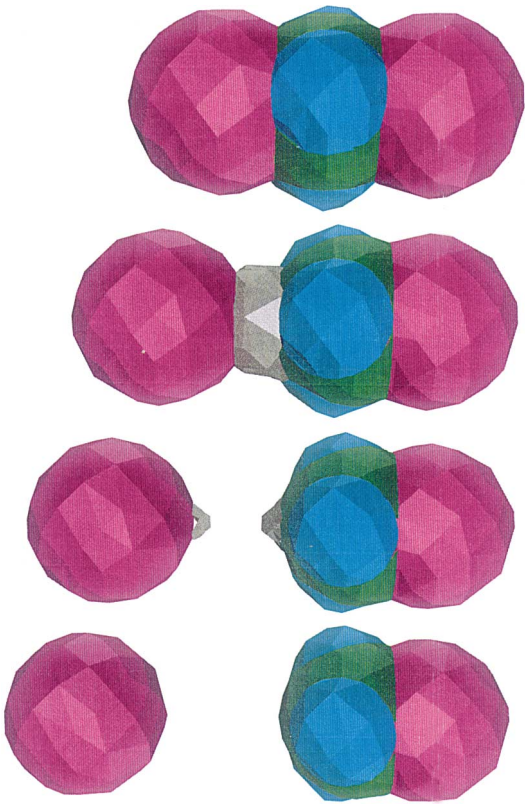


Fig. 6. Pictorial representation of the cavity evolution along the r_c coordinate with $TsA = 0.4$. From top to bottom: 1. The transition state ($r_c = 0$ A) there are no added spheres; 2. The added spheres connect the cavity ($r_c = 4.205$ A); 3. There are two disconnected cavities but there are solvent-excluding zones ($r_c = 4.706$ A); 4. The solvent can move freely between the two cavities ($r_c = 5.206$ A). The tessarae displayed here refer to the $TsA = 0.4$ option. Note that for numerical calculations it is in general sufficient to have a sphere partitioning lower than that necessary for a good visual representation

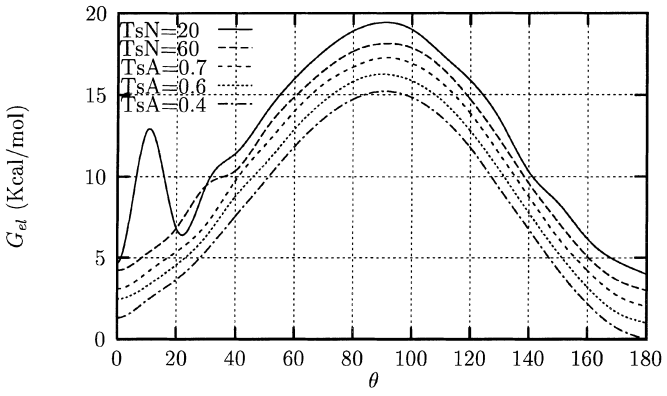


Fig. 7. Free-energy profiles for the conformational rotation of 1,2,3,5,6,7-octahexene with different cavity options. G_{el} is set equal to zero for $TsA = 0.4$ for $\theta = 180°$. The other curves are shifted by 1 kcal/mol each to avoid juxtapositions

portion of the cavity surface, with other conformations where the groups R_1 and R_2 are distant contribute to describe two separate portions of the cavity surface.

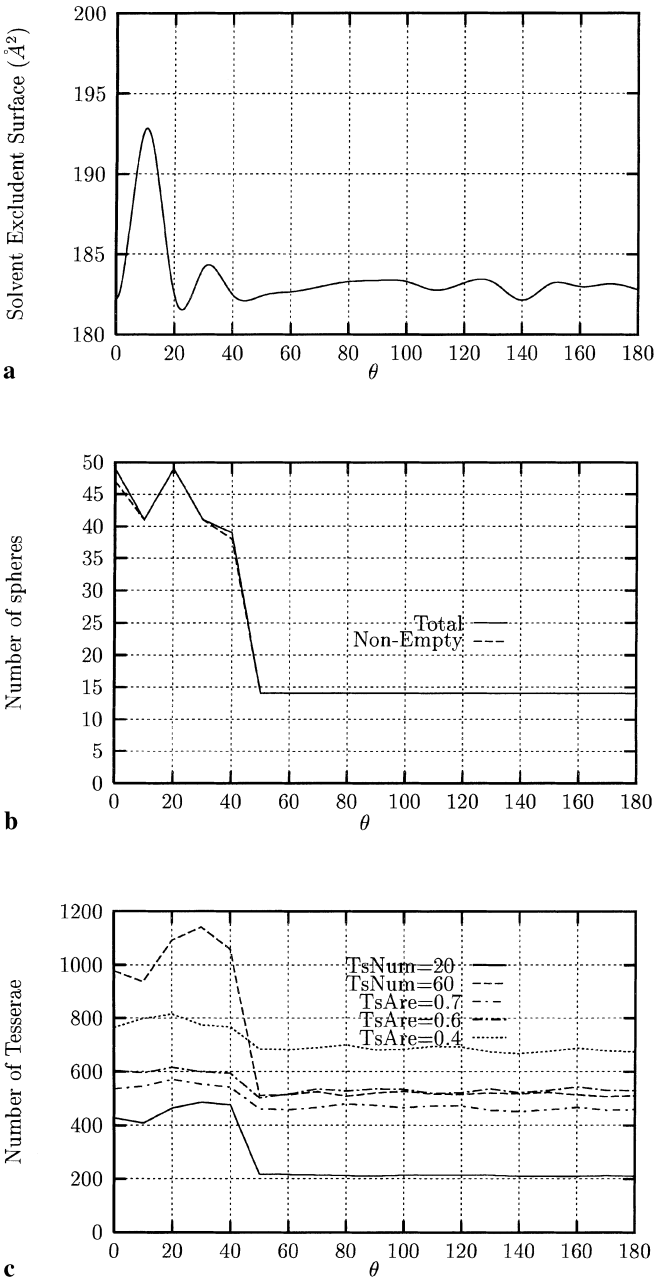
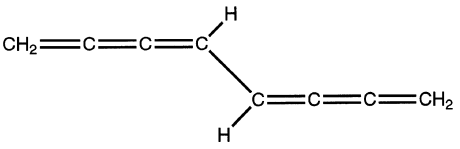


Fig. 8. Variation of area **a** number of spheres **b** and number of tesserae along the torsion angle θ for the conformational rotation in



Scheme 1

To describe this process we have selected a molecule, reported in Scheme 1: the 1,2,3,5,6,7-octahexene.

This is a system of limited chemical interest, but exhibits this conformationl behaviour with the minimal number of internal rotation degrees of freedom. The free energy profiles of this process are reported in Fig. 7.

When the torsion angle θ is in the range $0-40^\circ$ there are some added spheres that disappear when θ is greater than 40°. For this reaction we have performed ab initio calculations with 10° steps at the HF/6-31G* level. The *TsA* profiles show less "roughness" than the those obtained with *TsN*. In this case we do not observe a desolvation associated step but the choises only a slope variation around $\theta = 27^\circ$. For $\theta > 40^\circ$ $TsN = 60$ or $TsA = 0.6$ give a similar number of tesserae ($\simeq 500$), for $\theta = 20-40$ the number of tesserae becomes $\simeq 1300$ with $TsN = 60$ and $\simeq 600$ with $TsA = 0.6$, as Fig. 8 shows. Here again the convenience of using the *TsA* procedure is clear.

5 Conclusions

We have tested the convergence properties of the PCM-BEM electrostatic problem. The standard pentakisdodecahedral tesselation is a good choice for the fixed geometry solutes, equivalent to the $TsA = 0.4-0.5$ option. The *TsA* procedure exhibits better behaviour when we study the free energy variations with respect to some molecular coordinates and it reduces the number and the entity of numerical artefacts intrinsic to the method. These artefacts can be a serious obstacle to the application of automatic geometry optimization procedures to these systems.

We have found it very beneficial to use the *TsA* partition in the geometry optimization calculations. A remarkable number of computer-driven geometry optimizations of solutes fail if the *TsA* option is not used (unpublished calculations with V. Barone and M. Cossi). In a separate paper we shall document our progress in establishing efficient computational codes for geometry optimizations starting from the analytical PCM gradient formulas [13] and exploiting other procedures established in more recent times [2], to which that presented here has been added.

Acknowledgements. The financial support of CNR is acknowledged.

References

1. Tomasi J, Persico M (1994) Chem Rev 94:2027
2. Amovilli C, Barone V, Cammi R, Cancès E, Cossi M, Mennucci B, Pomelli CS, Tomasi J (in press) Adv Quantum Chem
3. Pascual-Ahuir JL, Silla E, Tomasi J, Bonaccorsi R (1987) J Comp Chem 8:778
4. Pascual-Ahuir JL, Silla E, Tuñon I (1994) J Comp Chem 15:1127
5. Rivail JL, Rinaldi D (1995) In: Leszczynski J (ed) Computational chemistry, review of current trends. World Scientific Publishing, New York, vol 7, pp 175
6. Cramer CJ, Truhlar DG (1996) In: Tapia O, Bertrán J (eds) Solvent effects and chemical reactivity. Kluver, Dordrecht, pp 1
7. Massey WS (1967) Algebraic topology: an introduction. Harcourt, Brace and World, New York, chapter 1
8. Wenninger MJ (1979) Spherical models. Cambridge University Press, Cambridge, chapter 4
9. Gaussian 94, Revision B.3. Frisch MJ, Trucks GW, Schlegel HB, Gill PMW, Johnson BG, Robb MA, Cheeseman JR, Keith T, Petersson GA, Montgomery JA, Raghavachari K, Al-Laham MA, Zakrzewski VG, Ortiz JV, Foresman JB, Peng CY, Ayala PY, Chen W, Wong MW, Andres JL, Replogle ES, Gomperts R, Martin RL, Fox DJ, Binkley JS, Defrees DJ, Baker J, Stewart JP, Head-Gordon M, Gonzalez C, Pople JA (1995) Gaussian, Inc., Pittsburgh, Pa
10. Floris FM, Tomasi J (1989) J Comp Chem 10:616
11. Barone V, Cossi M, Tomasi J (1996) Chem Phys Lett 255:327
12. Pomelli CS, Tomasi J (1997) J Phys Chem 101:3561
13. Cammi R, Tomasi J (1994) J Chem Phys 101:3888