Journal of Computer-Aided Molecular Design, **14:** 329–339, 2000.
KLUWER/ESCOM

Mixed QM/MM molecular electrostatic potentials

Begoña Hernández[a], F. Javier Luque[b,*] & Modesto Orozco[a,*]
[a]*Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain;* [b]*Departament de Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avgda Diagonal s/n, E-08028 Barcelona, Spain*

Received 25 June 1999; Accepted 10 December 1999

Key words: MEP, mixed methods, molecular electrostatic potential, molecular modelling, QM/MM

Summary

A new method is presented for the calculation of the Molecular Electrostatic Potential (MEP) in large systems. Based on the mixed Quantum Mechanics/Molecular Mechanics (QM/MM) approach, the method assumes both a quantum and classical description for the molecule, and the calculation of the MEP in the space surrounding the molecule is made using this dual treatment. The MEP at points close to the molecule is computed using a full QM formalism, while a pure classical evaluation of the MEP is used for points located at large distances from the molecule. The algorithm allows the user to select the desired level of accuracy in the MEP, so that the definition of the regions where the MEP is computed at the classical or QM levels is adjusted automatically. The potential use of this QM/MM MEP in molecular modeling studies is discussed.

Introduction

The molecular electrostatic potential (MEP) is defined as the electrostatic interaction energy between a particle and the unit positive point charge. If the charge distribution of the molecule is defined as a set of point charges, the MEP can be computed following Coulomb's law (Equation 1):

$$V(\mathbf{r_1}) = \sum_i \frac{Q_i}{|\mathbf{r_1} - \mathbf{R}_i|}, \tag{1}$$

where $\{Q\}$ denote the set of point charges that describe the molecular charge distribution, $\mathbf{R}_i$ is the vector defining the position of Q_i, and $\mathbf{r}_1$ is the vector defining the position of the unit positive point charge.

Similarly, for a continuous distribution of the molecular charge distribution the MEP can be determined using Equation 2, where ρ_{TOT} contains both the electron and nuclear charge distributions.

$$V(\mathbf{r_1}) = \int \frac{\rho_{TOT}(\mathbf{r})}{|\mathbf{r_1} - \mathbf{r}|} d\mathbf{r} \tag{2}$$

Treating the nuclear charge by means of point charges located on nuclei, Equation 2 adopts the more usual expression given by Equation 3 (see [1–6]).

$$V(\mathbf{r_1}) = \sum_k \frac{Z_k}{|\mathbf{r_1} - \mathbf{R_k}|} - \int \frac{\rho(\mathbf{r})}{|\mathbf{r_1} - \mathbf{r}|} d\mathbf{r}, \tag{3}$$

where Z_k and R_k are the atomic number and position of atom k, and ρ is the molecular electron charge density.

The calculation of the integral in the right hand term of Equation 3 is simple within the MO – LCAO approach, since it is expressed in terms of integrals involving the product of atomic orbitals (Equation 4, where $P_{\mu\nu}$ is the element $\mu\nu$ of the first-order density matrix and $\{\phi\}$ denotes the basis set of atomic orbitals).

$$V(\mathbf{r_1}) = \sum_k \frac{Z_k}{|\mathbf{r_1} - \mathbf{R_k}|} - \sum_\mu \sum_\nu P_{\mu\nu} \int \frac{\phi_\mu(\mathbf{r})\phi^*_\mu(\mathbf{r})}{|\mathbf{r_1} - \mathbf{r}|} d\mathbf{r}. \tag{4}$$

Since the MEP contains all the information regarding the electrostatic characteristics of a molecule, it is

*To whom correspondence should be addressed. E-mail: javier@far1.far.ub.es (F.J.L.); modesto@luz.bq.ub.es (M.O.).

widely used as an index in many research areas such as molecular structure [7–9], solvation [10–14], crystalline state [15–17], chemical reactivity in organic, inorganic and bioorganic systems [18–30], force-field parametrization [31–36], structure-activity relationships [20, 23–26, 37–42], and molecular similarity studies [43–48]. It has large impact in drug design as a tool for 'lead' optimization and pharmacophore search [38, 39, 45, 47]. Indeed, the MEP has been recently generalized to include van der Waals [49–52] and even polarization [53–56] interactions between a QM particle and any classical probe molecule. These formalisms [49–56] have extended the field of application to areas such as the analysis of very large non-covalent interacting systems in Monte Carlo QM/MM simulation studies [54, 57].

The MEP can be computed within the ab initio, density functional or semiempirical frameworks [58–64]. A large series of systematic studies have shown that in general a HF/6-31G(d) calculation provides a detailed description of the MEPs accurate enough for most purposes [62–65]. When higher accuracy is needed, extension of the basis set and treatment of electron correlation effects or non-local hybrid functionals are necessary [63–65]. The use of MEPs computed from semiempirical methods [58–62] is also acceptable for very large systems where HF/6-31G(d) calculations are not feasible.

For certain kinds of applications, such as molecular similarity or pharmacophore search, the use of QM MEPs is, nevertheless, extremely demanding in CPU time and becomes prohibitive, since the MEP has to be computed in very large grids of points, especially when drugs of different sizes are considered. Current computational facilities enable this type of calculations, but at a speed which is not acceptable when large sets of compounds have to be analyzed in a very short time. This makes it necessary to resort to MEPs computed classically, which gives a reliable description of long-range electrostatic effects, but fails in regions near the molecule. This precludes the use of classical MEPs for an accurate description of molecular structure and reactivity. Different strategies have been suggested to improve the quality of the classical MEP, such as the addition of higher multipole moments to the monopole term [66, 67], the introduction of Gaussian functions [68], or the use of multicentric charges [69–71]. However, even in these cases classical and QM MEPs present notable differences in chemically relevant regions close to the molecule.

In this paper we present a new method to compute efficiently and accurately the MEP using a mixed QM/MM approach. The method computes QM MEPs in regions close to the molecule, and in all other regions where the expected difference between classical and QM MEPs is greater than a given threshold, whereas for the remaining points (typically the largest number of grid points) the MEP is determined at the classical level. The QM/MM MEPs obtained in this way are fully comparable in accuracy to the pure QM ones, but are obtained at a fraction of their computational cost.

QMM/MM approach for the calculation of MEPs

The suitability of the mixed QM/MM approach for computing the MEP is examined by using the error function, $\zeta(\mathbf{r}_1)$, given in Equation 5, which computes the difference between QM (MEP_{QM}) and classical (MEP_{MM}) MEPs at a given point ($\mathbf{r}_1$) in the space surrounding the molecule as a function of the distance from the nuclei (Equation 6). In turn, we examined how such an error function depends on the distance of the point $\mathbf{r}_1$ from the nuclei of the molecule, whose position is designated by $\mathbf{R}_k$ (Equation 6).

$$\zeta(\mathbf{r}_1) = \left|\mathrm{MEP}_{\mathrm{QM}}(\mathbf{r}_1) - \mathrm{MEP}_{\mathrm{MM}}(\mathbf{r}_1)\right|, \tag{5}$$

$$\zeta(\mathbf{r}_1) = f\,|\mathbf{r}_1 - \mathbf{R}_\mathrm{k}|, \tag{6}$$

A thorough variety of analytical functions f were examined to describe the dependence of $\zeta(\mathbf{r}_1)$ on the distance $|\mathbf{r}_1 - \mathbf{R}_k|$. The best results were obtained by a simple exponential function (Equation 7), where α and β are parameters to fit.

$$f = \alpha\, \mathrm{e}^{-\beta(|\mathbf{r}_1 - \mathbf{R}_\mathrm{k}|)} \tag{7}$$

For a polyatomic molecule the error at a point $\mathbf{r}_1$ is expressed as a sum of exponential functions centered at the individual atoms (Equation 8).

$$\zeta(\mathbf{r_1}) = \sum_k \alpha_k\, \mathrm{e}^{-\beta_\mathrm{k}(|\mathbf{r}_1 - \mathbf{R}_\mathrm{k}|)} \tag{8}$$

Since the preceding function facilitates the a priori estimation of the error in the classical MEP at a given point $\mathbf{r}_1$, all the grid points can be classified into three regions depending on the magnitude of the estimated error: (i) QM, (ii) QM/MM, and (iii) MM. The QM region contains all the points where the estimated error is greater than a previously selected energy value (CUT1), and the MEP is directly computed at the QM

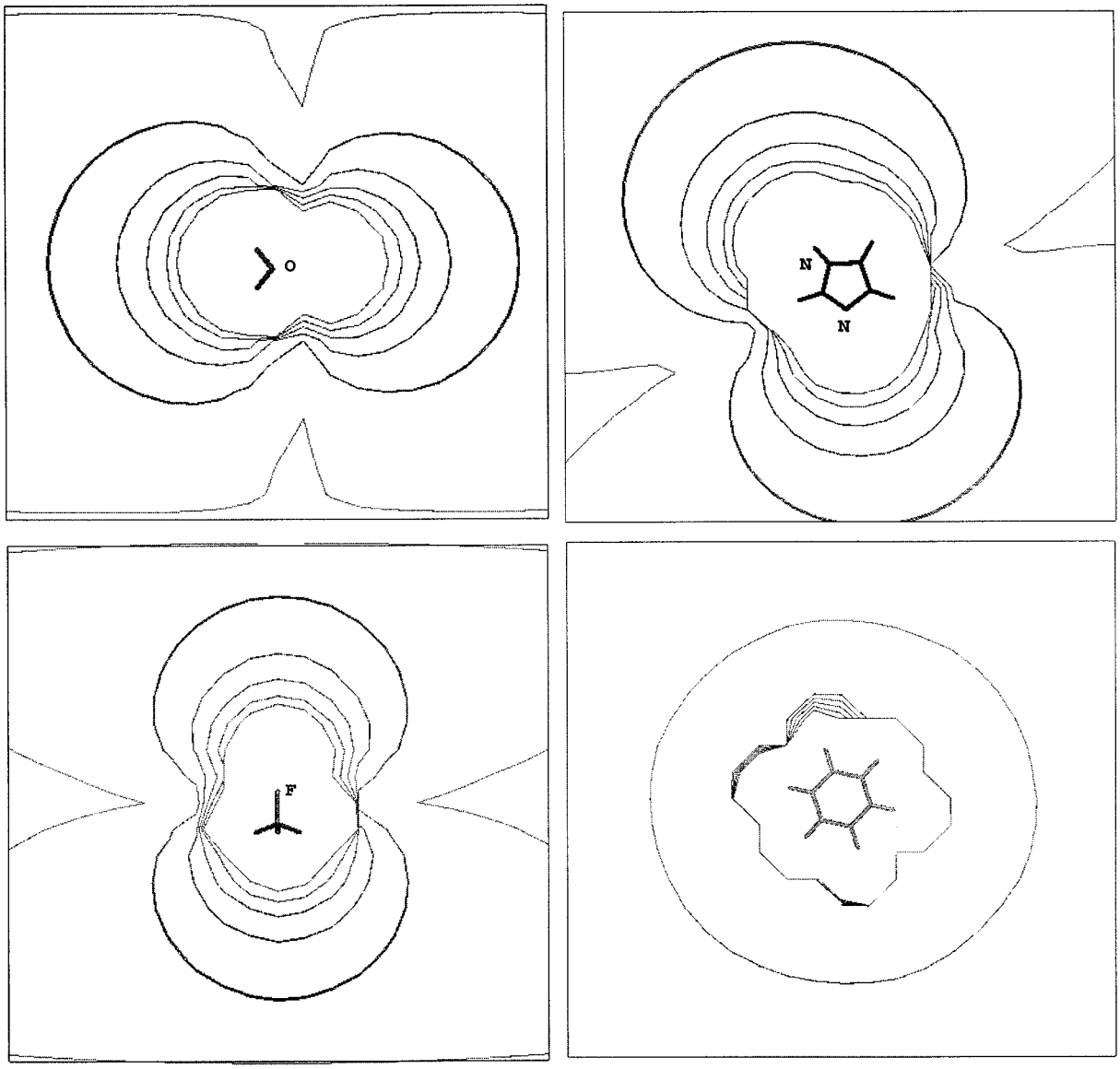

Figure 1. Differences between QM and MM MEPs for water, benzene, fluoromethane and imidazole. The isocontour lines correspond to 0.1, 1.0, 5.0, 10.0, 13.0 and 15.0 kcal mol^{-1}.

level. The QM/MM part is a buffer region containing all the points in a layer that encloses the QM region. The MEP in QM/MM points is computed both at the QM and MM levels. Finally, the rest of the points define the MM part, where the MEP is computed classically. Clearly, this initial classification of the points cannot lead to a very accurate representation of the MEP owing to the large anisotropy of this property. Thus, a searching algorithm is necessary to redefine the boundary between QM and MM parts according to the shape of the MEP distribution around the molecule. To this end, the differences between MM and QM MEPs in the QM/MM points are computed. If for a given point (i, j, k) the error is greater than a previously selected cutoff (CUT2), such a point is included in the QM region, all its neighbors ($i\pm1$, $j\pm1$, $z\pm1$) in the MM region are included in the QM/MM region, and the MEP at these latter points is also computed at the QM level. The searching algorithm continues until no point in the QM/MM part has an error between QM and MM MEPs larger than CUT2. The results show that few cycles are needed to achieve convergence if

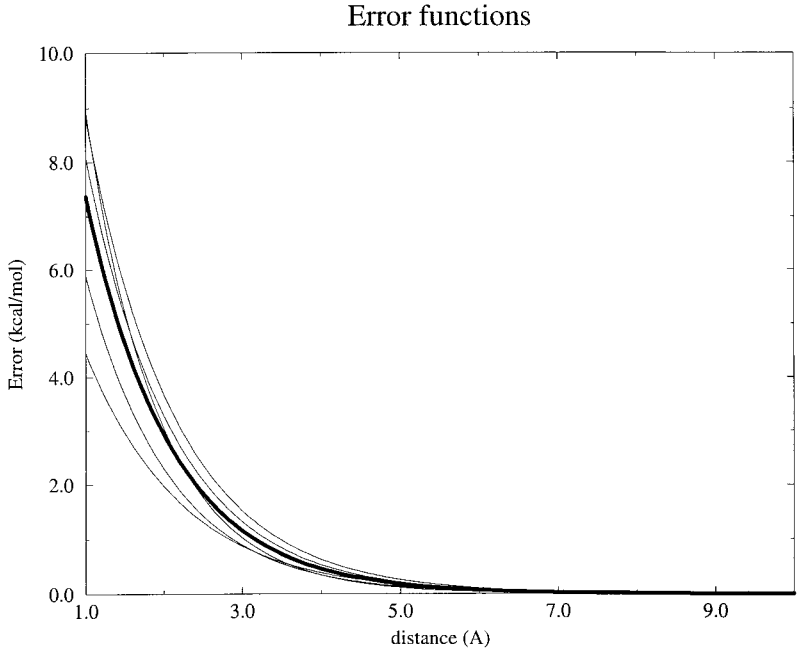

Figure 2. Plot of the fitted error functions. The thicker line corresponds to average values.

a suitable error function, $\zeta(\mathbf{r}_1)$, is used. This process guarantees that the classical MEP is computed only at those points where the difference between MM and QM MEPs is very small.

Computational details

The QM MEP was computed from ab initio HF/6-31G(d) wavefunctions using geometries optimized at the same level of theory. A cubic grid, extending 10 Å beyond the molecule in each direction of the space with a grid step of 0.5 Å was used in all cases. ESP-fitted charges [72, 73] were used to determine classical MEPs. Geometries and wavefunctions were determined using Gaussian-94 [74], and MEPs were determined with MOPETE-99 [75]. All calculations were performed on the ORIGIN-2000 of the Centre de Supercomputació de Catalunya (CESCA), as well as on workstations in our laboratory.

Results and discussion

Inspection of the difference maps between QM and MM MEPs shows that the largest deviations appear generally in inner regions (see Figure 1). However, also in regions relevant for chemical reactivity and intermolecular interactions, i.e., outside the van der Waals radii, the errors due to the use of an MM MEP are notably large, as was demonstrated by MIP [49] calculations in selected molecules (data not shown, but available upon request from the authors).

Inspection of Figure 1 reveals the large anisotropy of the MEP difference maps, which stems from the intrinsic anisotropy of the QM MEP and the difficulty to represent accurately the features of the QM MEP from the set of atom-centered charges used to compute the MM MEP. This effect can be clearly noted in the large errors that appear in regions associated with the presence of lone pairs. Therefore, the errors between QM and MM MEPs can be expected to be described only approximately by means of a function dependent on the distance of a given point from the nuclei. In turn, this makes it necessary to implement a searching algorithm to further reduce the deviation between QM and MM MEPs in specific regions around the molecule, and to guarantee a desired level of accuracy (given by CUT2) in the mixed QM/MM MEP.

Fitting the errors for a set of molecules (H_2O, NH_3, CH_3Cl, CH_3F, CH_2O, HCOOH, $HCONH_2$, CH_3OH, CH_3NH_2, CH_3OCH_3, CH_2FCH_2F and CH_2ClCH_2Cl) to exponential functions allowed us to obtain optimized coefficients and exponents (see Table 1), which can be used to obtain a reasonable, but approximate estimate of the error in MM MEPs. For our purposes,

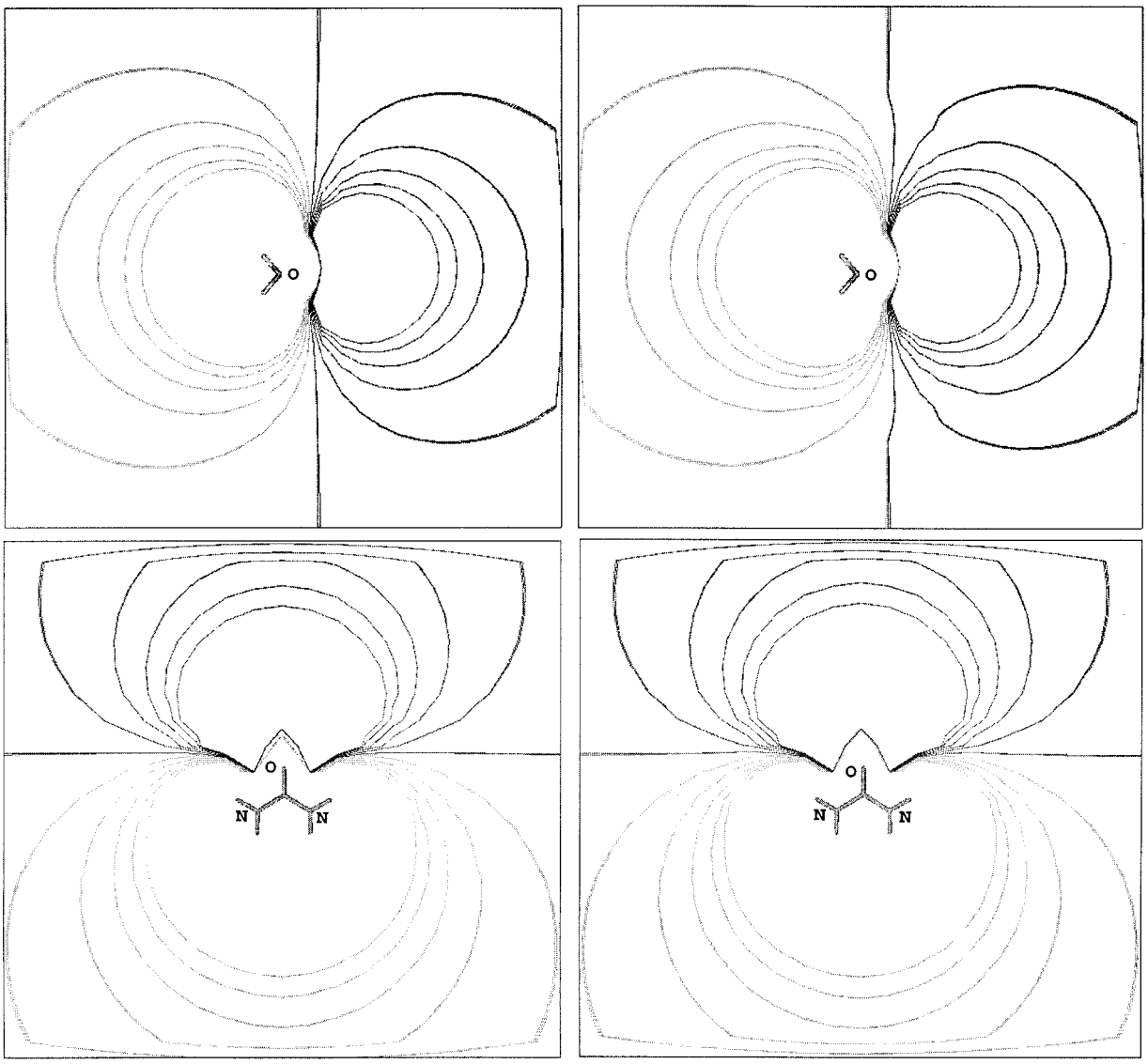

Figure 3. Representation of QM (right) and QM/MM (left) MEPs for selected molecules (H_2O and $CO(NH_2)_2$).

Table 1. Fitted parameters for the exponential error function (see Equation 7)

Atom	α (kcal mol^{-1})	β (Å)
C	26.4	1.08
N	20.0	0.90
O	15.1	0.94
F	10.3	0.81
Cl	21.2	0.88
Average	18.5	0.92

these error functions provide us with an analytical formula to define the grid points that constitute the QM/MM boundary (see above) in the beginning of our iteration process (see above). From a practical point of view, let us note that even though these parameters have been optimized from MEPs determined at the HF/6-31G(d) level, they can be used in the framework of the QM/MM method to compute MEPs at other levels of theory. For instance (see Table 2), we have used HF/6-31G(d)-optimized α and β values to obtain the QM/MM MEP of $HCONH_2$ and CH_3OCH_3, at the STO-3G level. Results in Table 2 demonstrate the

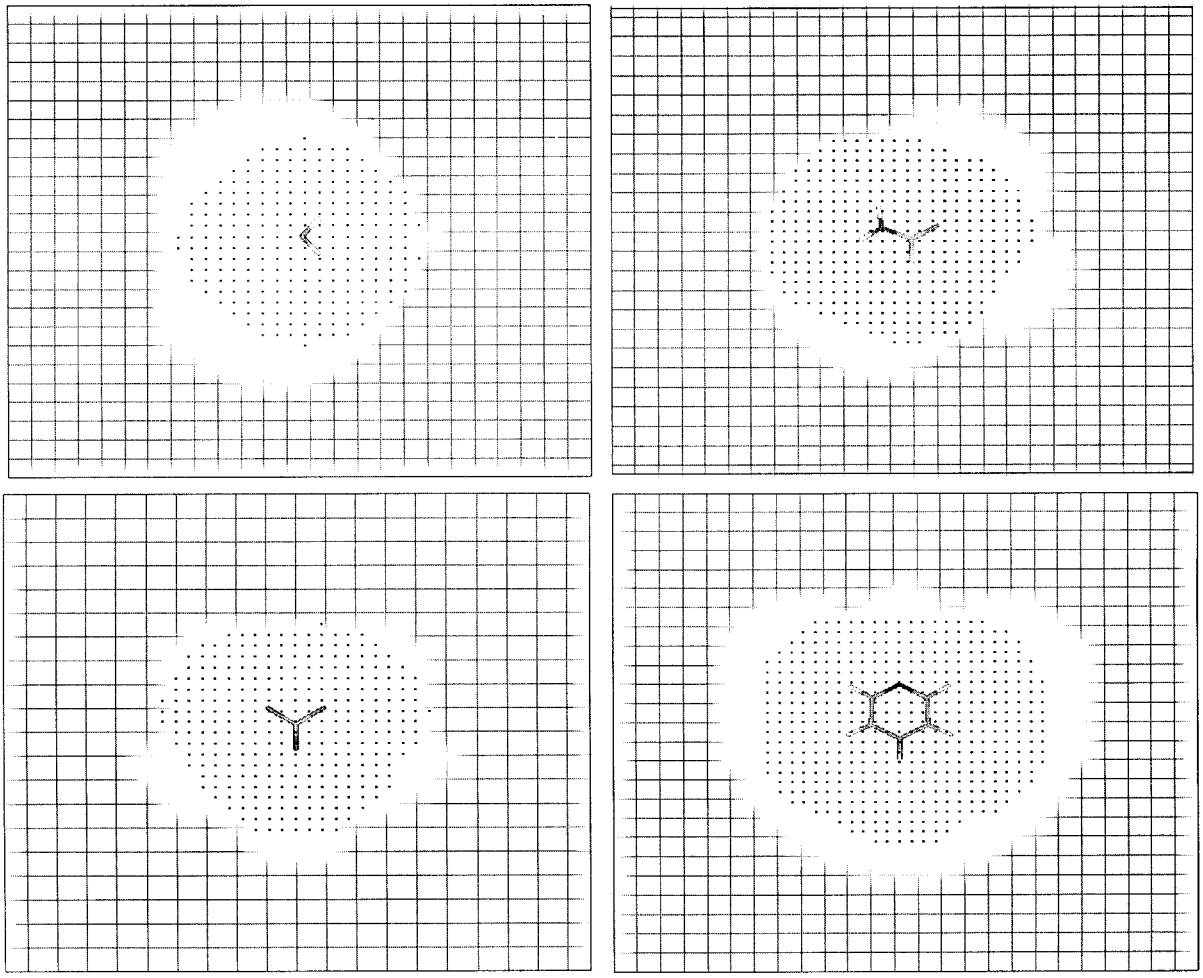

Figure 4. Representation of the partition of QM and MM regions for selected molecules. Dotted regions are QM regions defined from the use of the error functions (CUT). The regions in white correspond to points which were introduced in the QM part by the searching algorithm controlled by CUT2. The regions marked with a grid correspond to portions of the MM parts.

transferability of the α and β parameters in the context of our calculations. These optimized values of α and β are used in the rest of the paper, even though the similarity of the fitted error functions (see Figure 2) suggests that average values for α and β can be used when CUT1 values equal or lower than 0.5 kcal mol^{-1} are considered to define the first QM/MM boundary in our iteration process. We should emphasize here that α and β are just approximate values, which are not used to define the final QM/MM boundary, but only the first guess of a boundary which is then refined in our 3D iteration process.

To verify the usefulness of the algorithm, the MEPs were computed in cubic grids for a series of small neutral molecules (see Table 2) following a pure QM or MM treatment, as well as the mixed QM/MM approach mentioned above. In this latter case, values of CUT1 and CUT2 equal to 0.5 and 0.2 kcal mol^{-1} were chosen to define the first QM/MM boundary region and the final accuracy of QM/MM MEPs, respectively.

The use of the QM/MM approach greatly reduces the cost of calculation, which stems from a dramatic reduction in the number of grid points where the QM MEP has to be computed, even for small grids like those considered here (see Table 2). Thus, for a quite restrictive value of CUT2 (0.2 kcal mol^{-1}), the MEP is computed at the QM level only in around 15% of the total number of grid points. In general, the QM/MM method involves around 10% of the cost of the pure QM calculation, and the results have an accuracy fully comparable to the QM MEP (see Table 2). The reduction of computational effort is more remarkable when

Table 2. Comparison of pure QM and mixed QM/MM MEPs (HF/6-31G(d)) for selected molecules. In all cases a cutoff error (CUT2) of 0.2 kcal mol^{-1} was used. Root mean square deviation and maximum errors are in kcal mol^{-1}

Molecule	RMSD	Max error	Ratio of points (MM/QM)	CPU ratio QM-MM/QM
H_2O	0.04	0.17	13.6	0.12
NH_3	0.06	0.18	11.6	0.12
C_6H_6	0.05	0.17	6.6	0.14
CH_3Cl	0.03	0.17	13.4	0.08
Urea	0.03	0.17	15.8	0.06
CF_2O	0.02	0.16	17.9	0.06
CH_3F	0.02	0.15	21.2	0.05
CH_2O	0.02	0.15	22.8	0.05
$HCONH_2$	0.03	0.17	14.0	0.09
HCOOH	0.03	0.17	16.8	0.06
HF	0.03	0.17	16.8	0.09
Imidazole	0.04	0.18	9.3	0.10
CH_3NH_2	0.04	0.18	11.2	0.09
CH_3OH	0.03	0.17	16.0	0.07
CH_3OCH_3	0.03	0.17	15.4	0.06
Pyridine	0.04	0.17	7.8	0.11
CH_2FCH_2F	0.02	0.15	19.2	0.05
CH_2ClCH_2Cl	0.04	0.17	8.5	0.11
Pyrimidine	0.04	0.17	10.2	0.09
$(CH_3)_3COH$	0.02	0.16	17.9	0.07
CF_4	0.01	0.13	21.0	0.04
CH_3OCH_3[a]	0.03	0.16	15.0	0.06
$HCONH_2$[a]	0.03	0.18	15.5	0.08

[a]Values obtained at the STO-3G level.

Table 3. Root mean square deviation (RMSD in roman), and maximum error (in italic) between QM/MM and QM MEPs when the final CUT2 was 0.1, 0.2 and 0.5 kcal mol^{-1}. The CPU times (relative to the pure QM calculations) are shown in bold. Error values are in kcal mol^{-1}

Molecule	CUT2=0.1	CUT2=0.2	CUT2=0.5
H_2O	0.03	0.04	0.06
	0.09	*0.17*	*0.38*
	0.24	**0.12**	**0.05**
NH_3	0.04	0.06	0.07
	0.09	*0.18*	*0.41*
	0.36	**0.12**	**0.05**
CH_3F	0.01	0.02	0.02
	0.08	*0.15*	*0.23*
	0.10	**0.05**	**0.04**
$HCONH_2$	0.02	0.03	0.05
	0.09	*0.17*	*0.40*
	0.13	**0.09**	**0.05**
HCOOH	0.02	0.03	0.04
	0.08	*0.17*	*0.40*
	0.10	**0.06**	**0.04**
Imidazole	0.03	0.04	0.06
	0.09	*0.18*	*0.41*
	0.19	**0.10**	**0.06**
Pyridine	0.03	0.04	0.07
	0.09	*0.17*	*0.41*
	0.22	**0.11**	**0.07**
Pyrimidine	0.02	0.04	0.06
	0.09	*0.17*	*0.43*
	0.16	**0.09**	**0.05**

the level of the QM calculation and the number of grid points increase. Futhermore, the largest reduction of computational time is found for polar molecules, while the reductions are less important for aromatic compounds, likely due to the problems of the monopole expansion to represent the π-charge distribution. The accuracy of the QM/MM strategy is very good, as stated from the MEP maps shown in Figure 3 and from the statistical results in Table 2. Thus, the RMS deviation between QM and QM/MM MEPs is extremely small (typically 0.02–0.04 kcal mol^{-1}). Accordingly, we believe that the method is very suitable for drug design studies, where the MEP of large molecules has to be computed in very large cubic grids.

Choice of different values for CUT2 allows us to select the accuracy of the QM/MM MEP. Table 3 contains results for selected molecules obtained for CUT2 values of 0.5, 0.2 and 0.1 kcal mol^{-1} (CUT1 was always fixed at 0.5 kcal mol^{-1}). The results show that even for the more accurate value of CUT2 there is a very remarkable reduction in CPU time for the QM/MM strategy. Inspection of maximum errors, RMS deviations and CPU times suggests that a value of CUT2 equal to 0.2 kcal mol^{-1} is likely the best compromise between accuracy and computational cost. However, for most applications a less restrictive value of 0.5 kcal mol^{-1} can be accurate enough, since no RMS deviation larger than 0.07 kcal mol^{-1} is found, whereas the calculation involves around 5% of the CPU cost of the pure QM one.

The QM/MM strategy can be compared with simple cutoff approaches utilized to partition the grid between QM and MM regions. This has been done for selected molecules and the results are given in Table 4. Clearly, the searching algorithm presented here is nec-

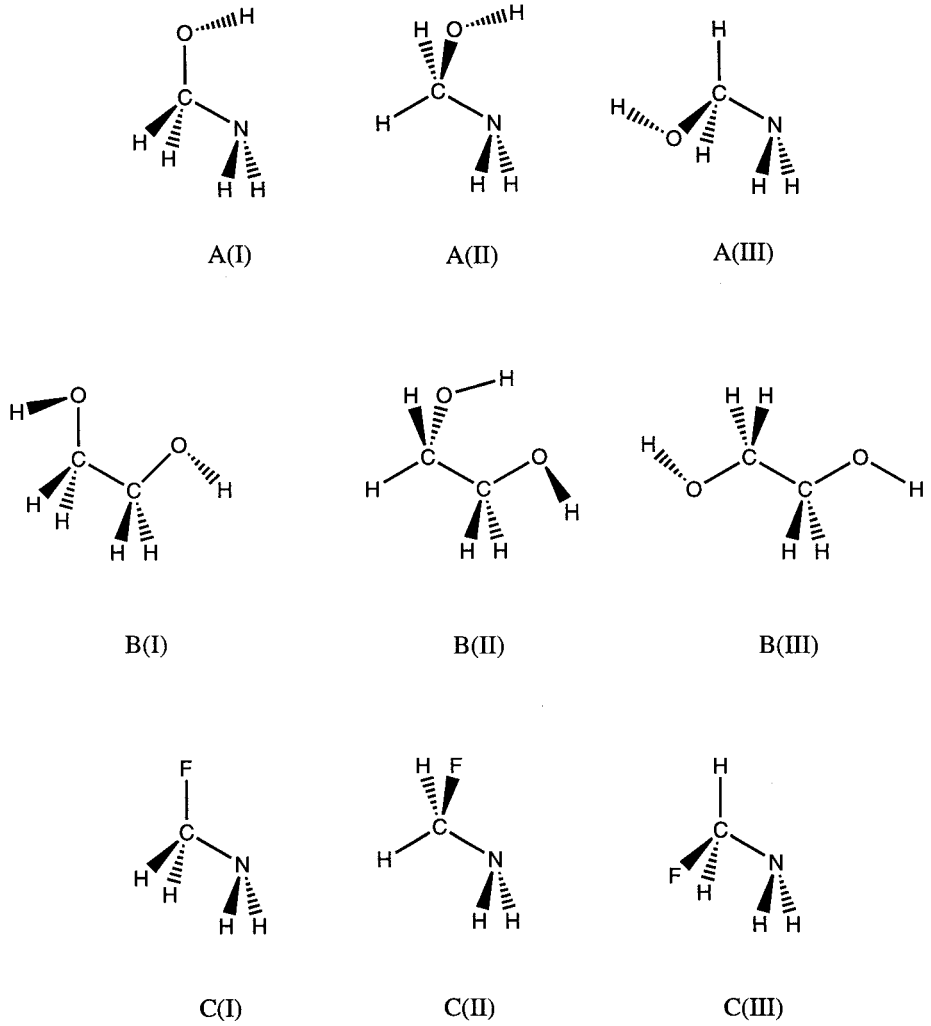

Figure 5. Selected conformers of molecules considered to analyze the performance of QM/MM MEPs.

essary to guarantee the accuracy of the QM/MM MEP, especially in terms of maximum RMS error. Thus, the use of a simple cutoff radius leads to maximum errors which are clearly above the selected threshold value owing to the anisotropy of the MEP distribution. However, the QM/MM mixed approach greatly reduces the errors, at the expense of a very moderate increase in CPU time, since the QM calculation of the MEP is performed selectively for those points having the largest errors. In fact, the results are more accurate than those obtained using a simple cutoff CUT1=0.1 kcal mol^{-1} with even less computational effort. Accordingly, the QM/MM method guarantees that no errors greater than CUT2 exist, while the single cutoff procedure does not.

We should note that the searching strategy presented here to optimize the region of partition between the QM and MM MEPs does not lead to any dramatic increase in the number of points to be considered at the QM level. This is clearly shown in Figure 4, which demonstrates that: (i) the optimum boundary between QM and MM MEPs is not isotropic, and accordingly cannot be easily defined using atomic-based cutoffs, and (ii) there are few extra points introduced by the searching algorithms.

As a final test, we have examined the ability of the QM/MM approach to reproduce the MEP of different conformers compared to the description provided by using a common set of Boltzmann-averaged charges at the MM level. Three molecules in selected conformations were considered for this study (Figure 5). In all

Table 4. Root mean square deviation (RMS in roman), and maximum error (in italic) between QM/MM and QM MEPs when a simple cutoff (CUT1) is used to partition QM and MM regions. Values of CUT1 equal to 0.1, 0.2 and 0.5 kcal mol^{-1} (corresponding approximately to spherical cutoffs around 6.0, 5.0 and 4.0 Å) were used. The results obtained with the QM/MM standard procedure are displayed for comparison. In all the cases the CPU times (relative to the pure QM calculations) are shown in bold. Error values are in kcal mol^{-1}

Molecule	CUT1=0.5 radii ≈ 4.0 Å	CUT1=0.2 radii ≈ 5.0 Å	CUT1 = 0.1 radii ≈ 6.0 Å	Standard CUT1=0.5 CUT2=0.2
H_2O	0.06	0.04	0.04	0.04
	0.68	*0.35*	*0.21*	*0.17*
	0.05	**0.08**	**0.13**	**0.11**
NH_3	0.08	0.06	0.06	0.06
	0.63	*0.32*	*0.21*	*0.18*
	0.05	**0.08**	**0.13**	**0.12**
CH_3F	0.02	0.02	0.01	0.02
	0.23	*0.15*	*0.09*	*0.15*
	0.05	**0.08**	**0.13**	**0.05**
$HCONH_2$	0.05	0.04	0.04	0.03
	0.58	*0.32*	*0.18*	*0.17*
	0.05	**0.08**	**0.13**	**0.07**
HCOOH	0.04	0.03	0.03	0.03
	0.56	*0.27*	*0.16*	*0.17*
	0.05	**0.08**	**0.13**	**0.06**
Imidazole	0.08	0.06	0.04	0.04
	0.96	*0.47*	*0.26*	*0.18*
	0.05	**0.09**	**0.14**	**0.10**
Pyridine	0.09	0.06	0.05	0.04
	1.07	*0.64*	*0.41*	*0.17*
	0.06	**0.10**	**0.15**	**0.12**
Pyrimidine	0.07	0.05	0.05	0.04
	0.78	*0.46*	*0.29*	*0.17*
	0.05	**0.09**	**0.13**	**0.09**

cases a common set of ESP charges was determined for each molecule by Boltzmann-weighted averaging of ESP charges fitted for every conformer. Such a common set of charges greatly resembles those obtained for the most stable conformer, as expected from the relative energies between conformers (see Table 5). Furthermore, charges for rotationally equivalent atoms were averaged following the standard ESP strategy.

The results in Table 5 demonstrate the excellent agreement between QM and QM/MM MEPs for all the conformers. Thus, the average RMS deviations are typically around 0.05 kcal mol^{-1}, while sensibly

Table 5. Root mean square deviation (RMS) and maximum error in QM/MM and classical MEPs for conformers of molecules A, B and C (see Figure 5). Classical and QM MEPs were compared only at grid points located at more than 1.4 times the van der Waals radii, while all regions were considered for comparison of QM/MM and QM MEPs. The CPU times relative to the full QM calculation are shown for QM/MM calculations (they are below 0.005 for classical calculations). The energies of each conformer relative to the most stable one are displayed for each molecule (see Figure 5)

Conformer	ΔE (kcal mol^{-1})	QM/MM (CUT2=0.2)	Classical
(A)			
A(I)	3.2	0.07	0.34
		0.19	*9.3*
		0.25	
A(II)	0.0	0.05	0.21
		0.18	*8.1*
		0.11	
A(III)	6.9	0.08	0.39
		0.19	*9.7*
		0.39	
(B)			
B(I)	5.2	0.08	0.13
		0.19	*4.5*
		0.36	
B(II)	0.0	0.03	0.15
		0.17	*5.9*
		0.06	
B(III)	1.8	0.08	0.13
		0.19	*5.9*
		0.27	
(C)			
C(I)	0.0	0.04	0.25
		0.18	*12.3*
		0.11	
C(II)	2.3	0.06	0.26
		0.18	*10.4*
		0.14	
C(III)	3.4	0.06	0.28
		0.19	*10.1*
		0.17	

larger values are obtained when classical MEPs are used, even when QM and MM MEPs are compared only in outer regions (outside 1.4 times the van der Waals surface). It is worth noting the large reduction

in the maximum errors in chemically relevant regions when the QM/MM procedure is used. Thus, whereas the maximum error at the MM level ranges from 5 to 12 kcal mol^{-1} (as a consequence of the intrinsic errors derived from the use of a single set of Boltzmann-averaged charges), it is less than 0.2 kcal mol^{-1} (the value of CUT2) in the QM/MM approach. It is also worth noting the large reduction in computational effort, typically around 70–90% of the cost in the QM calculation. As expected, QM/MM calculations are more expensive when conformers other than the more stable one are considered, owing to the problems associated with the use of a common set of point charges for the different conformers. However, the computational cost can be modulated by appropriate choice of the cutoff CUT2. Thus, for 1,2-ethanediol the CPU ratio between QM-MM and QM calculations is only 5%, 10% and 8% for conformers I–III, respectively, when CUT2=0.5 kcal mol^{-1}. Accordingly, the QM/MM approach can reproduce nicely the QM MEPs of flexible molecules, even when a single set of point charges is used for the different conformers.

Conclusions

Classical (MM) potentials are able to reproduce quantum mechanical (QM) molecular electrostatic potentials (MEPs) at large distances from the nuclei, but there are important errors in regions relevant for understanding chemical reactivity and for intermolecular interactions. The QM/MM algorithm allows us to compute the MEP quantum mechanically only in those regions where the MM MEP presents large deviations, whereas the MEP for the rest of the space surrounding the molecule is computed at the classical level. A compromise between computational cost and accuracy of the computed MEP can be easily adjusted by choosing appropriate cutoff values.

The test calculations reported in this study demonstrate that the QM/MM approach provides MEP maps fully comparable to those obtained from a pure QM computation, but at an extremely small fraction of the computational cost. The accuracy of the QM/MM MEP method and the computational efficiency of the method make it very suitable for structure-activity relationship studies, like similarity analysis, or pharmacophore search, where the MEP has to be computed in very large grids of points around molecules of medium or larger size.

Acknowledgements

We thank Dr F. Sanz for valuable discussions. This work has been supported by the NATO (grant CRG972068) and by DGICYT (projects PB961005 and PB97-0908). The Centre de Supercomputació de Catalunya (CESCA; Molecular Recognition Project) is acknowledged for computational facilities. This is a contribution from the Centre Especial de Recerca en Química Teòrica (CERQT) de la Universitat de Barcelona.

References

1. Bonaccorsi, R., Petrongolo, C., Scrocco, E. and Orozco, M., Theoret. Chim. Acta, 20 (1971) 331.
2. Bonaccorsi, R., Pullman, A., Scrocco, E. and Tomasi, J., Theoret. Chim. Acta, 24 (1972) 51.
3. Berthier, G., Bonaccorsi, R., Scrocco, E. and Tomasi, J., Theoret. Chim. Acta, 26 (1972) 101.
4. Scrocco, E. and Tomasi, J., Top. Curr. Chem., 42 (1973) 95.
5. Giessner-Prettre, C. and Pullman, A., Theoret. Chim. Acta, 25 (1972) 83.
6. Srebrenik, A., Weinstein, H. and Pauncz, R., Chem. Phys. Lett., 20 (1973) 419.
7. Politzer, P., J. Am. Chem. Soc., 207 (1980) 3027.
8. Ritchie, J.P., J. Am. Chem. Soc., 107 (1985) 1829.
9. Politzer, P. and Murray, J.S., In Murray, J.S. and Sen, K. (Eds) Molecular Electrostatic Potentials. Concepts and Applications, Elsevier, Amsterdam, 1996, pp. 649–660.
10. Oliveira Neto, M., J. Comput. Chem., 7 (1986) 617.
11. Gilson, M.K. and Honig, B., Proteins, 4 (1988) 7.
12. Murray, J.S. and Politzer, P., J. Org. Chem., 56 (1991) 6715.
13. Tomasi, J., Mennucci, B. and Cammi, R., In Murray, J.S. and Ken, K. (Eds) Molecular Electrostatic Potentials. Concepts and Applications, Elsevier, Amsterdam, 1996, pp. 1–85.
14. Luque, F.J., Barril, X. and Orozco, M., J. Comput.-Aided Mol. Design, 13 (1999) 139.
15. Spackman, M.A. and Stewart, R.F., In Politzer, P. and Truhlar, D.G. (Eds) Chemical Applications of Atomic and Molecular Electrostatic Potentials, Plenum, New York, NY, 1981, pp. 407–415.
16. Weber, H.P. and Craven, B.M., Acta Crystallogr., B46 (1990) 532.
17. White, J.C. and Hess, A.C., J. Phys. Chem., 97 (1993) 6398.
18. Gadre, S.R., Bhadane, P.K., Pundlik, S.S. and Pingale, S.S., In Murray, J.S. and Ken, K. (Eds) Molecular Electrostatic Potentials. Concepts and Applications, Elsevier, Amsterdam, 1996, pp. 219–256.
19. Mishra, P.C. and Kumar, A., In Murray, J.S. and Sen, K. (Eds) Molecular Electrostatic Potentials. Concepts and Applications, Elsevier, Amsterdam, 1996, pp. 257–296.
20. Náray-Szabó, G. and Surjan, P.R., In Náray-Szabó, G. (Ed.) Theoretical Chemistry of Biological Systems, Elsevier, Amsterdam, 1986, pp. 1–30.
21. March, N.H., In Murray, J.S. and Sen, K. (Eds) Molecular Electrostatic Potentials. Concepts and Applications, Elsevier, Amsterdam, 1996, pp. 619–648.

22. Baeten, A., In Murray, J.S. and Sen, K. (Eds) Molecular Electrostatic Potentials. Concepts and Applications, Elsevier, Amsterdam, 1996, pp. 587–618.
23. Loew, G.H. and Berkowitz, D.S., J. Med. Chem., 18 (1975) 656.
24. Hayes, D.M. and Kollman, P.A., J. Am. Chem. Soc., 98 (1976) 335.
25. Hayes, D.M. and Kollman, P.A., J. Am. Chem. Soc., 98 (1976) 7861.
26. Weinstein, H., Osman, R., Topiol, S. and Green, J.P., Ann. N.Y. Acad. Sci., 367 (1981) 434.
27. Lavery, R. and Pullman, B., Int. J. Quantum Chem., 20 (1981) 259.
28. Kollman, P.A., McKelvey, J., Johansson, A. and Rothenberg, S., J. Am. Chem. Soc., 97 (1975) 955.
29. Scrocco, E. and Tomasi, J., Adv. Quant. Chem., 11 (1978) 115.
30. Tomasi, J., In Ratajczak, H. and Orville-Thomas, W.T. (Eds) Molecular Interactions, Vol. 3, Wiley, New York, NY, 1982, pp. 119–166.
31. Momany, F.A., J. Phys. Chem., 82 (1978) 592.
32. Williams, D.E., J. Comput. Chem., 9 (1988) 745.
33. Ferenzcy, G.G., Reynolds, C.A. and Richards, W.G., J. Comput. Chem., 11 (1990) 159.
34. Besler, B.H., Merz, K.M. and Kollman, P.A., J. Comput. Chem., 11 (1990) 431.
35. Cornell, W.D., Cieplak, P., Bayly, C.I. and Kollman, P.A., J. Am. Chem. Soc., 115 (1993) 9620.
36. Bonaccorsi, R., Scrocco, E., Petrongolo, C. and Tomasi, J., Theor. Chim. Acta., 20 (1984) 331.
37. Petrongolo, C. and Tomasi, J., Int. J. Quantum Chem., Quantum Biol. Symp., 2 (1975) 181.
38. Orozco, M., Canela, E.I. and Franco, R., Mol. Pharmacol., 35 (1989) 257.
39. Orozco, M., Canela, E.I. and Franco, R., Eur. J. Biochem., 188 (1990) 155.
40. Warshel, A. and Åqvist, J., Annu. Rev. Biophys. Biophys. Chem., 20 (1991) 267.
41. Náray-Szabó, G. and Nagy, P., Int. J. Quantum Chem., 35 (1989) 215.
42. Breneman, C.M. and Martinov, M., In Murray, J.S. and Sen, K. (Eds) Molecular Electrostatic Potentials. Concepts and Applications, Elsevier, Amsterdam, 1996, pp. 143–180.
43. Burt, C., Huxley, P. and Richards, W.G., J. Comput. Chem., 11 (1990) 1139.
44. Besalú, E., Carbó, R., Mestres, J. and Solà, M., In Sen, K. (Ed.), Topics in Current Chemistry: Molecular Similarity I, Vol. 173, Springer-Verlag, Berlin, 1995, p. 31.
45. Rogríguez, J., Manault, F. and Sanz, F., J. Comput. Chem., 14 (1993) 922.
46. Sanz, F., López, E., Rodríguez, J. and Manault, F., Quant. Struct.-Act. Relat., 13 (1994) 281.
47. Sanz, F., Manaut, F., Rodríguez, J., Lozoya, E. and López, J., J. Comput.-Aided Mol. Design, 7 (1993) 337.
48. Charlton, M.K. and Thomson, C., J. Chem. Soc., Faraday Trans., 90 (1994) 3533.
49. Orozco, M. and Luque, F.J., J. Comput. Chem., 14 (1993) 587.
50. Orozco, M. and Luque, F.J., Biopolymers, 33 (1993) 1851.
51. Alhambra, C., Luque, F.J. and Orozco, M., J. Phys. Chem., 99 (1995) 3084.
52. Orozco, M. and Luque, F.J., In Murray, J.S. and Sen, K. (Eds) Molecular Electrostatic Potentials. Concepts and Applications, Elsevier, Amsterdam, 1996, pp. 181–218.
53. Francl, M.M., J. Phys. Chem., 89 (1985) 428.
54. Luque, F.J. and Orozco, M., J. Comput. Chem., 19 (1998) 866.
55. Cubero, E., Luque, F.J. and Orozco, M., Proc. Natl. Acad. Sci. USA, 95 (1998) 5976.
56. Colominas, C., Orozco, M., Luque, F.J., Borrell, J.I. and Teixidó, J., J. Org. Chem., 63 (1998) 4947.
57. Gao, J., Luque, F.J. and Orozco, M., to be published.
58. Luque, F.J., Illas, F. and Orozco, M., J. Comput. Chem., 11 (1990) 446.
59. Ferenczy, G.G., Reynolds, C.A. and Richards, W.G., J. Comput. Chem., 11 (1990) 159.
60. Besler, B.H., Merz, K.M. and Kollman, P.A., J. Comput. Chem. 11, (1990) 431.
61. Wang, B. and Ford, G.P., J. Comput. Chem., 12 (1994) 200.
62. Alhambra, C., Luque, F.J. and Orozco, M., J. Comput. Chem., 15 (1994) 12.
63. Soliva, R., Orozco, M. and Luque, F.J., J. Comput. Chem., 18 (1997) 980.
64. Soliva, R., Luque, F.J. and Orozco, M., Theor. Chem. Acc., 98 (1997) 42.
65. Luque, F.J., Orozco, M., Illas, F. and Rubio, J., J. Am. Chem. Soc., 113 (1991) 5203.
66. Williams, D.E., J. Comput. Chem., 9 (1988) 745.
67. Williams, D.E., Biopolymers, 29 (1990) 1367.
68. Gadre, S.R. and Shrivastava, I.H., Chem. Phys. Lett., 204 (1993) 350.
69. Singh, U.C. and Kollman, P.A., J. Comput. Chem., 5 (1984) 129.
70. Chipot, C., Angyán, J.G., Ferenczy, G.G. and Scheraga, H.A., J. Phys. Chem., 97 (1993) 6628.
71. Williams, D.E., J. Comput. Chem., 15 (1994) 719.
72. Alemán, C., Orozco, M. and Luque, F.J., Chem. Phys., 189 (1994) 573.
73. Orozco, M. and Luque, F.J., J. Comput.-Aided Mol. Design, 4 (1990) 411.
74. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Gill, P.M.W., Johnson, B.G., Robb, M.A., Cheeseman, J.R., Keith, T.A., Petersson, G.A., Montgomery, J.A., Raghavachari, K., Al-Laham, M.A., Zakrzewski, V.G., Ortiz, J.V., Foresman, J.B., Cioslowski, J., Stefanov, B.B., Nanayakkara, A., Challacombe, M., Peng, C.Y., Ayala, P.Y., Chen, W., Wong, M.W., Anfres, J.L., Replogle, E.S., Gomperts, R., Martin, R.L., Fox, D.J., Binkley, J.S., Defress, D.J., Baker, J., Stewart, J.J.P., Head-Gordon, M., Gonzalez, C. and Pople, J.A., GAUSSIAN 94 (Rev. A.1), GAUSSIAN Inc., Pittsburgh, PA, 1995.
75. Luque, F.J., Alhambra, C. and Orozco, M., MOPETE-99, Universitat de Barcelona, 1999.