Journal of Computer-Aided Molecular Design, 4 (1990) 411–426
ESCOM

J-CAMD 097

A practical procedure for the determination of electrostatic charges of large molecules

M. Orozco[a] and F.J. Luque[b]

[a]*Departamento de Bioquímica y Fisiología, Facultad de Química, Universidad de Barcelona, C/. Martí i Franqués 1, 08028 Barcelona, Spain*

[b]*Departamento de Farmacia, Unidad Físico-Química, Facultad de Farmacia, Universidad de Barcelona, Av. Diagonal s/n, 08028 Barcelona, Spain*

Received 2 May 1990
Accepted 14 June 1990

Key words: Electrostatic charges; Molecular electrostatic potential; Fractional models; MNDO wavefunction

SUMMARY

A practical procedure for the precise determination of electrostatic charges, which are evaluated by fitting the rigorous quantum mechanical molecular electrostatic potential to a monopole-monopole expression, is presented. The proposal of this procedure arises from the study of the minimum requirements necessary to obtain reliable electrostatic charges. Such a study is focused on: (i) the dependence of the electrostatic charges on the set of points where the quantum mechanical and the monopole-monopole molecular electrostatic potentials are fitted; thus, both the influence of the number of points and their distribution in layers located out of the van der Waals radii of the atoms are examined, and (ii) the reliability of the use of fractional models for the evaluation of electrostatic charges of large molecules. Results point out that the optimum number of points is defined by a density of points ranging from 0.45 to 0.60 points/Å^2 when four layers (separated by 0.2 Å) are considered. Nevertheless, the use of only two layers (separated by 0.4 Å) for large molecules is recommended, thus enabling one to obtain reliable charges at a reduced computational cost. Moreover, results justify the use of fractional models for the determination of electrostatic charges of extremely large molecules, even when aromatic structures exist.

INTRODUCTION

In the last years molecular mechanic (MM) and dynamic (MD) methods have been largely used for the study of biological macromolecules and of their interactions [1–3]. Both methodologies are based on the assumption that the energy of a molecule can be expressed as the addition of the 'bonded' interactions, which include the bond stretching and bending terms as well as the rotation around bonds, and the 'non-bonded' interactions, which refer to the van der Waals and electrostatic energy terms.

The correct representation of the electrostatic energy is of major importance for obtaining reliable results in MM and MD studies. The 'force fields' used to define the molecular energy in such studies approximate the electrostatic energy by a simple monopole-monopole coulombic expression:

$$V_{EL} = 332{,}4 \sum \frac{q_i\, q_j}{\varepsilon\, R_{ij}} \text{ (in kcal/mol)} \tag{1}$$

where ε is the dielectric constant, q_i and q_j are localized charges, and R_{ij} is the distance separating both charges.

The representation of the electrostatic energy by a simple coulombic expression is a drastic simplification, whose goodness depends mainly on the quality of the set of localized charges. Two strategies are currently used for determining localized charges: (i) the evaluation of charges from empirical values or from Hückel calculations [4–8], and (ii) the computation of charges from the quantum mechanical molecular wavefunction [9–12]. The computational cost of the first strategy is very small, but the quality of the charges derived from quantum mechanical calculations is fairly superior.

The most widely used method for obtaining localized charges from quantum mechanical wavefunctions was suggested by Mulliken [9]. The popularity of Mulliken analysis is mainly due to its conceptual and computational simplicity. Nevertheless, the shortcomings of this procedure for reproducing the SCF and experimental charge distributions are well known [11, 13–15].

In the seventies, Momany proposed a new strategy to compute localized charges [11], which has proved its superiority with respect to other methods, like Mulliken population analysis [11–15]. Actually, charges derived from versions of Momany's procedure (named 'electrostatic charges') are used in MM and MD studies. Especially noticeable is the use of such charges in Kollman's 'force field' [16–18], one of the most used for MM and MD studies of proteins and nucleic acids. Following Momany's strategy [11], the charges are determined by imposing the condition that the monopole-monopole electrostatic potential computed by using such charges has to be equal to the rigorous quantum mechanical molecular electrostatic potential. In practice, according to the popular algorithm proposed by Singh and Kollman [19], the quantum mechanical molecular electrostatic potentials are evaluated in a set of Connolly layers [20] defined out of the van der Waals radii of the atoms. Then, the electrostatic charges are obtained by a Levenberg–Marquardt [21] non-linear fitting of the rigorous quantum mechanical and the monopole-monopole electrostatic potentials.

The goodness of the electrostatic charges depends on two factors: (i) the quality of the basis set used to compute the SCF wavefunction, and (ii) the set of points where the quantum mechanical and monopole-monopole molecular electrostatic potentials are fitted.

Within the ab initio framework, a *split-valence* basis set including polarization functions at least for the non-hydrogen atoms is necessary to obtain accurate molecular electrostatic potentials [22, 23], the 6-31G* [24] basis set being recommended [23]. Even though, approximate electrostatic charges can be derived from the less sophisticated and costly ab initio 3-21G [25] *split-valence* wavefunction [14]. Nevertheless, the computation of ab initio wavefunctions is in practice a hard and often impossible task for biologically relevant molecules, not only at the *split-valence* level, but also when the minimal STO-3G [26] basis set is used. Thus, fractional models, whose reliabili-

ty has not been well demonstrated, must be employed for computing electrostatic charges of large molecules.

Recently, we have proposed [23, 27] a new strategy for the calculation of accurate quantum mechanical molecular electrostatic potentials from MNDO [28] and AM1 [29] wavefunctions. Moreover, we have demonstrated [15] the goodness of the electrostatic charges derived from both semiempirical molecular electrostatic potentials. The reliability of this strategy has also been examined in two recent papers [30, 31], which point out the usefulness of the semiempirical electrostatic potentials for deriving atomic charges. Such a methodology appears promising for the determination of electrostatic charges of large molecules at a reduced computational cost, avoiding in most cases the use of fractional models.

With regard to the set of points where the quantum mechanical and monopole-monopole molecular electrostatic potentials should be fitted, different works [14, 15, 19] have suggested the use of 200–300 points defined in four Connolly layers (the inner one being placed at 1.4 times the van der Waals radii of the atoms, and the distance between layers being equal to 0.2 Å) for evaluating electrostatic charges. Nevertheless, those works were performed for small molecules, and consequently it is not clear whether such recommendations are also valid for large molecules.

Following our recent developed methodology, in this paper we try to determine for large biologically relevant molecules: (i) the minimum number of points where molecular electrostatic potentials should be fitted, (ii) the optimum number of Connolly layers to be used, and (iii) the reliability of the fractional models.

METHODS

The quantum mechanical molecular electrostatic potential at a point r is defined according to Eq. 2 [32], where Z_A is the nuclear charge of atom A, located at R_A, and $\rho(r_1)$ is the molecular electron density.

$$V(r) = \sum_A \frac{Z_A}{|r - R_A|} - \int \frac{\rho(r_1)}{|r - r_1|} dr_1 \tag{2}$$

Within the framework of the MO-LCAO approximation, Eq. 2 can be rewritten in terms of the basis set of atomic orbitals Φ.

$$V(r) = \sum_A \frac{Z_A}{|r - R_A|} - \sum_\mu \sum_\upsilon P_{\mu\upsilon} \int \frac{\Phi^*_\mu(r_1)\Phi_\upsilon(r_1)}{|r - r_1|} dr_1 \tag{3}$$

where $P_{\mu\upsilon}$ stands for the element $\mu\upsilon$ of the first-order density matrix.

The computation of molecular electrostatic potentials from semiempirical wavefunctions implies three steps previous to the application of Eq. 3 [33–35]:

(i) The semiempirical wavefunction is deorthogonalized by means of Löwdin inverse transformation [36].

(ii) Each Slater type orbital of the minimal basis set used in Dewar's methods is expressed as a set of four gaussian type orbitals to facilitate the computation of atomic integrals in Eq. 3.

(iii) Since Dewar's methods are based on the *All Valence Electron* approximation, the nuclear charge is set equal to the 'core' effective charge, which is calculated by subtracting the inner electrons from the atomic number of the atom.

All the quantum mechanical molecular electrostatic potentials have been calculated from the semiempirical MNDO wavefunction, since our previous studies [15] and those of Besler et al. [30] have demonstrated that the electrostatic charges derived from such a wavefunction are more reliable than those derived from the AM1 calculations. At this point, it must be stressed that the MNDO wavefunction, unlike the CNDO method, which fails in reproducing for aromatic compounds the electrostatic potential above the molecular plane, correctly reflects the characteristics of the ab initio electrostatic potential originated from the π-electron distribution [23].

Electrostatic charges were determined according to the procedure suggested by Singh and Kollman [19]. Thus, quantum mechanical and monopole-monopole molecular electrostatic potentials calculated in points located in a set of Connolly layers [20] placed out of the van der Waals radii of the atoms are fitted by using a Levenberg–Marquardt non-linear optimization procedure [21]. The distance from the molecule to the inner layer was taken equal to 1.4 times the van der Waals radii of the atoms in all the calculations, since the short-range nuclear repulsion is not negligible for those points closest to the nuclei.

The statistical quality of the fitting between the rigorous quantum mechanical and the monopole-monopole molecular electrostatic potentials is measured by means of the relative root-mean-square deviation (rel RMS) defined according to Eq. 4,

$$\text{rel RMS} = \sqrt{\frac{\sum(\text{MEP}_{ik} - \text{MEP}_{ij})^2}{\sum(\text{MEP}_{ik})^2}} \tag{4}$$

where k and j denote the methods (k: quantum mechanical; j: monopole-monopole) used to evaluate the molecular electrostatic potential at the point i.

The variation of the electrostatic charges and dipoles due to the consideration of the fractional models, as well as the dependence of charges and dipoles on the number of points and layers used in the fitting is measured by means of the mean deviation defined according to Eq. 5a–b,

$$\text{M.D.q} = \sum \frac{|q_{A\alpha} - q_{A\beta}|}{n} \tag{5a}$$

$$\text{M.D.}\mu = |\mu_\alpha - \mu_\beta| \tag{5b}$$

where $(q_{A\alpha},\mu_\alpha)$ and $(q_{A\beta},\mu_\beta)$ represent the values of the charge A and the dipole moment evaluated by considering either the whole molecule or the fractional model, or by defining a different number of points and layers, n being the total number of charges.

Three biologically relevant molecules have been studied (Fig. 1): the natural nucleoside adenosine and two DNA-interacting drugs: proflavine, which is an intercalative drug [37], and psoralen, which is a cross-linker [37] that causes a large unwinding in the DNA structure. The molecular geometries have been optimized at the MNDO level with the only restriction of the planarity of

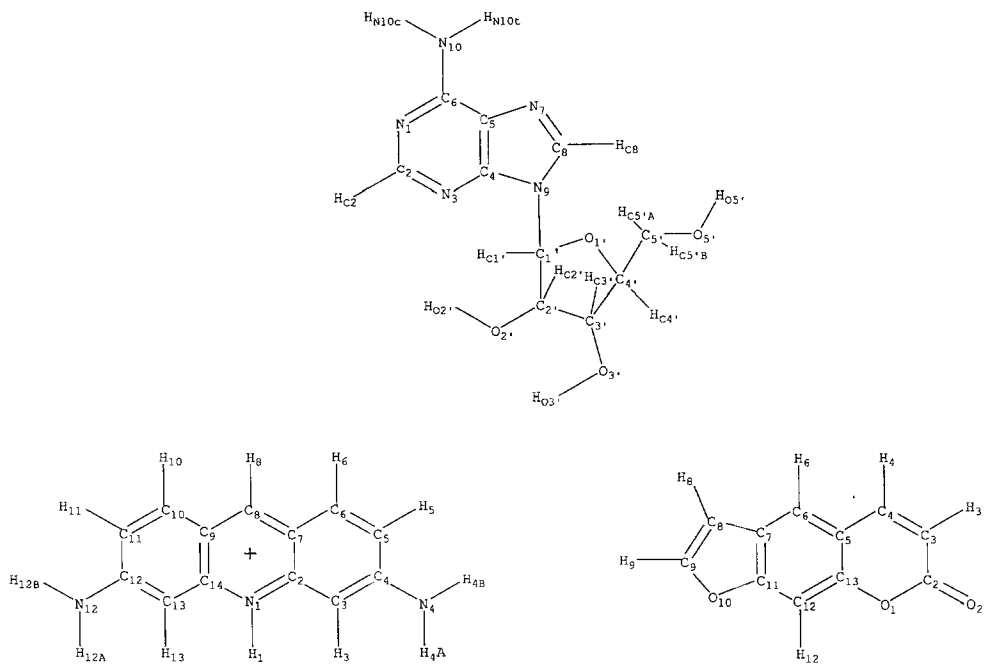

Fig. 1. Schematic representation and numbering of adenosine, proflavine and psoralen.

the aromatic ring. No symmetry considerations have been used in the computation of the electrostatic charges of proflavine.

MNDO calculations have been carried out from the standard parameters [26, 38] by using a locally modified [39] version of the MOPAC [40] computer program. The molecular electrostatic potentials and the electrostatic charges and dipoles have been computed with a largely modified version of the molecular electrostatic potential computer program developed by the Pisa group. All calculations have been performed with the IBM-3090 of the Centre de Informàtica de la Universitat de Barcelona.

RESULTS

In a first calculation, accurate electrostatic charges and dipoles for adenosine, proflavine and psoralen have been determined in order to have a reference calculation, and to provide reliable charges that may be useful for research groups working in drug–DNA interactions. For this purpose, electrostatic charges and dipoles were evaluated by fitting the quantum mechanical and monopole-monopole molecular electrostatic potentials in around 4500 points placed in four Connolly layers, the distance between layers being equal to 0.2 Å. Results are displayed in Tables 1–3, where the charges scaled by a factor equal to 1.32 in order to obtain the sophisticated ab initio 6-31G* charges are also included. This scaling factor, which was defined from the comparison of the MNDO and the ab initio 6-31G* electrostatic charges by means of a regression analysis of

the type $y = cx$ [15], permits to reproduce the charge distributions computed at the 6-31G* level at a small computational cost. It must be noted that the MNDO/6-31G* scaling factor derived from the work of Besler et al. [30] closely agrees with the value found by us.

Number of points

The electrostatic charges and dipoles of adenosine, proflavine and psoralen have been calculated considering 100, 200, 400, 500, 800, 1000, 1400, 1700 and 2000 points distributed in all cases in

TABLE 1
ELECTROSTATIC CHARGES FOR ADENOSINE COMPUTED FROM THE MNDO WAVEFUNCTION CONSIDERING EITHER THE WHOLE MOLECULE OR THE FRACTIONAL MODEL[a]

Atoms	Whole molecule		Model
	MNDO	Scaled MNDO[b]	MNDO
N1	−0.708	−0.935	−0.706
C2	0.446	0.589	0.486
N3	−0.695	−0.917	−0.644
C4	0.539	0.711	0.402
C5	−0.373	−0.492	−0.319
C6	0.828	1.093	0.749
N7	−0.367	−0.484	−0.340
C8	0.235	0.310	0.150
N9	−0.395	−0.521	−0.244
HC2	0.157	0.207	0.129
HC8	0.176	0.232	0.171
N10	−0.935	−1.234	−0.841
HN10c	0.440	0.581	0.407
HN10t	0.431	0.569	0.400
O1′	−0.261	−0.345	−0.309
C1′	0.073	0.096	0.349
C2′	0.045	0.059	−0.046
C3′	−0.068	−0.090	0.163
C4′	0.010	0.013	−0.018
C5′	−0.170	−0.224	−0.003
HC1′	0.181	0.239	0.078
HC2′	0.095	0.125	0.082
O2′	−0.439	−0.579	−0.491
HO2′	0.343	0.519	0.349
HC3′	0.157	0.207	0.076
O3′	−0.418	−0.552	−0.487
HO3′	0.356	0.470	0.351
HC4′	0.162	0.214	0.121
HC5′A	0.095	0.125	0.047
HC5′B	0.112	0.148	0.062
O5′	−0.417	−0.550	−0.430
HO5′	0.364	0.480	0.329

[a]See text for more details.
[b]Charges scaled to reproduce the ab initio 6-31G* values are also included.

TABLE 2
ELECTROSTATIC CHARGES FOR PSORALEN COMPUTED FROM THE MNDO WAVEFUNCTION CONSIDERING EITHER THE WHOLE MOLECULE OR THE FRACTIONAL MODEL[a]

Atoms	Whole molecule		Model
	MNDO	Scaled MNDO[b]	MNDO
O1	−0.454	−0.599	−0.388
C2	0.693	0.915	0.596
C3	−0.489	−0.645	−0.552
C4	0.144	0.190	0.426
C5	−0.290	−0.383	−0.533
C6	−0.142	−0.187	−0.077
C7	−0.013	−0.017	0.119
C8	−0.342	−0.451	−0.507
C9	−0.046	−0.061	−0.056
O10	−0.175	−0.231	−0.194
C11	0.310	0.409	0.320
C12	−0.533	−0.704	−0.578
C13	0.531	0.701	0.570
O2	−0.373	−0.492	−0.321
H3	0.200	0.264	0.222
H4	0.107	0.141	0.036
H6	0.173	0.228	0.104
H8	0.216	0.285	0.270
H9	0.219	0.289	0.256
H12	0.263	0.347	0.287

[a]See text for more details.
[b]Charges scaled to reproduce the ab initio 6-31G* values are also included.

four Connolly layers, the distance between them being equal to 0.2 Å. The charges and dipoles obtained from these calculations are compared with their respective reference calculations.

Figure 2 shows the variation of the relative root-mean-square deviation of the fitting between the quantum mechanical and the monopole-monopole molecular electrostatic potentials on the number of points for adenosine, proflavine and psoralen. From Fig. 2, it is evident that the three molecules share a common relative RMS profile. The smallest deviation appears when the fitting is done on 100 points. The deviation increases strongly when the number of points varies from 100 to 200 and more slightly when the number of points increases from 200 to 400. The relative RMS remains almost constant when the number of points used in the fitting is greater than 400.

The mean differences between the electrostatic charges obtained from the fitting with different points for adenosine, proflavine and psoralen with respect to the reference charges are displayed in Fig. 3. The differences (in absolute values) between the electrostatic dipoles with regard to the dipoles derived from the reference calculations for the three molecules are shown in Fig. 4.

Both the electrostatic charges and dipoles strongly depend on the number of points used in the fitting. Results clearly state the general similarity between the charge and dipole-deviation profiles. As could be expected, the general trend is that the charges and dipoles come closer to the ref-

TABLE 3
ELECTROSTATIC CHARGES AND DIPOLE FOR PROFLAVINE COMPUTED FROM THE MNDO WAVE-FUNCTION CONSIDERING THE WHOLE MOLECULE

Atoms	Whole molecule		Atoms	Whole molecule	
	MNDO	Scaled MNDO[a]		MNDO	Scaled MNDO[a]
N1	−0.580	−0.766	H1	0.391	0.516
C2	0.500	0.660	H3	0.200	0.264
C3	−0.610	−0.805	N4	−0.844	−1.114
C4	0.614	0.810	H4A	0.431	0.569
C5	−0.364	−0.480	H4B	0.421	0.556
C6	0.005	0.007	H5	0.189	0.249
C7	−0.244	−0.322	H6	0.169	0.223
C8	0.106	0.140	H8	0.144	0.190
C9	−0.247	−0.326	H10	0.167	0.220
C10	0.009	0.012	H11	0.190	0.251
C11	−0.363	−0.479	N12	−0.838	−1.106
C12	0.607	0.801	N12A	0.435	0.574
C13	−0.608	−0.803	N12B	0.417	0.550
C14	0.504	0.665	H13	0.198	0.261

[a]Charges scaled to reproduce the ab initio 6-31G* values are also included.

erence ones as the number of points considered in the fitting increases. Nevertheless, three points must be emphasized:

(i) The approach of the charges and dipoles to the reference values is not absolutely regular. Thus, in some cases a small increase in the number of points used in the fitting does not improve the quality of the charges and dipoles.

(ii) Large errors can appear when few points are selected. For instance, the mean difference in the charges of proflavine when 100 and 4500 points are used is 2.5 unit charges. Moreover, the difference in the dipole for this molecule calculated from 100 and 4500 points is equal to 9 Debyes.

(iii) The profiles exhibit two well-differentiated areas separated by a threshold value of the number of points used in the fitting, which ranges from 400 points for psoralen to 800 for adenosine. When the number of points is higher than the threshold value, the accuracy of the charges and dipoles notably improves, and a plateau appears indicating a slight dependence of the charges and dipoles on the number of points considered.

Number of layers

All the computations performed so far have been carried out by considering four layers and a separation between layers equal to 0.2 Å. The molecular electrostatic potential at both short and long distance from the molecule is well represented by using these four layers. Nevertheless, the use of four layers when a small number of points is considered implies a low density of points in the layers, which can lead to poor results (see below). A possibility to increase the density keeping constant the number of points lies in the use of only two layers, the inner and the outer one describing the molecular electrostatic potential in regions close and far from the molecule, respectively.

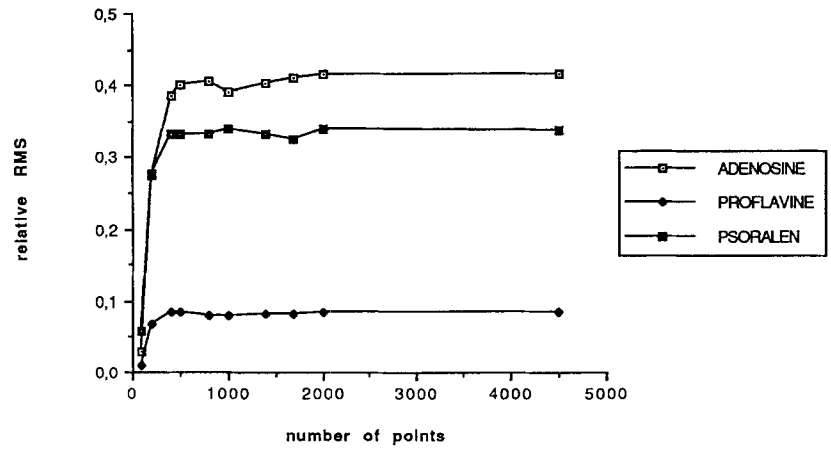

Fig. 2. Variation of the relative root-mean-square deviation of the fitting between the quantum mechanical and the monopole-monopole molecular electrostatic potentials on the number of points for adenosine, proflavine and psoralen.

Charges and dipoles have been calculated by considering 100, 200, 400, 500, 800 and 1000 points distributed in two layers separated by 0.4 Å, the inner one being located at 1.4 times the van der Waals radii of the atoms. The differences in the electrostatic charges and dipoles of the three molecules with regard to the reference values (determined from four layers and 4500 points) are shown in Figs. 5 and 6, respectively, which also display for comparison purposes the results derived from calculations with the same number of points, but distributed in four layers.

The deviation profiles corresponding to two layers notably differ with respect to those obtained when four layers are considered, the differences being especially noticeable when few points are considered. For all compounds the deviations in the charges and dipoles with respect to the refer-

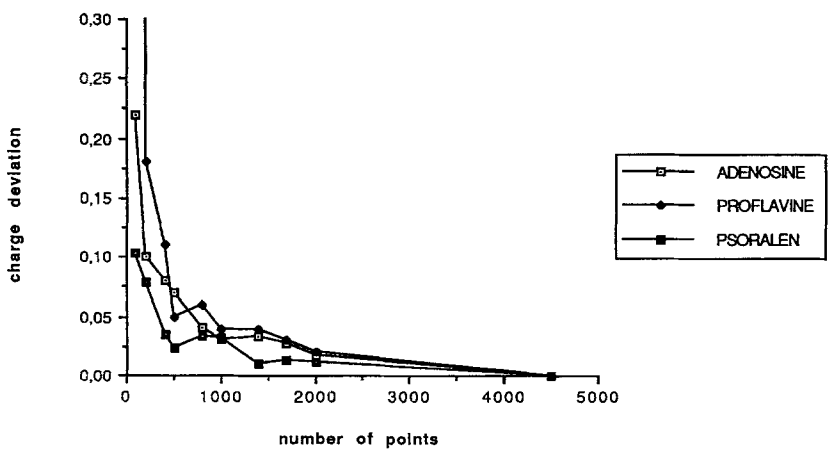

Fig. 3. Variation of the mean difference between the electrostatic charges derived from the fitting between the quantum mechanical and the monopole-monopole molecular electrostatic potentials on the number of points for adenosine, proflavine and psoralen.

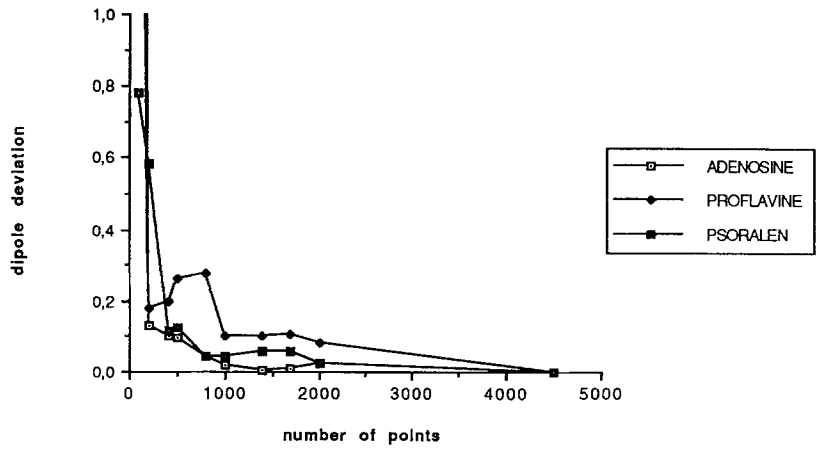

Fig. 4. Variation of the mean difference between the electrostatic dipoles derived from the fitting between the quantum mechanical and the monopole-monopole molecular electrostatic potentials on the number of points for adenosine, proflavine and psoralen.

ence ones are smaller when only two layers are defined. The inspection of the difference profiles indicates that the use of only 100–200 points in two layers leads to results of similar quality to those obtained when defining 500–800 points in four layers. In fact, the increase in the number of points when two layers are considered leads only to a small improvement in the quality of the electrostatic charges and dipoles.

Fractional models

The use of the semiempirical methodology drastically reduces the computational cost of the determination of electrostatic charges and dipoles, and consequently they can be evaluated for large biologically relevant structures like nucleosides. Nevertheless, the calculation of charges and dipoles for very large molecules is prohibitive even at the semiempirical level, and consequently the use of fractional models remains necessary in these cases.

Fractional models are usually applied for large molecules with two or more moieties clearly defined and joined by single bonds. Nevertheless, the use of fractional models is also necessary for the determination of charges of large aromatic structures. In this paper, we have studied the reliability of the fractional models for both kinds of compounds. Adenosine has been used as an example of a molecule with two subunits (the purine and the ribose ring) joined by a single bond, and psoralen has been used as an example of a large aromatic molecule.

The electrostatic charges of adenosine computed considering the whole molecule, four layers and 4500 points (Table 1) have been compared with those calculated taking separately the adenine and the ribose moieties. A methyl group is added to the N9 position of adenine in order to mimick the effect of the ribose, whereas an amino group simulates the effect of the purine on the ribose electronic charge distribution. The fitting of the quantum mechanical and monopole-monopole molecular electrostatic potentials has been done defining four layers (separated by 0.2 Å), and around 2500 points for the ribose and 2000 points for the adenine. Therefore, the total number of

points used is similar to that employed for the computation of the electrostatic charges of the whole adenosine molecule.

The results shown in Table 1 demonstrate the goodness of the fractional model for evaluating the charges of adenosine. Thus, the mean difference between the charges computed considering

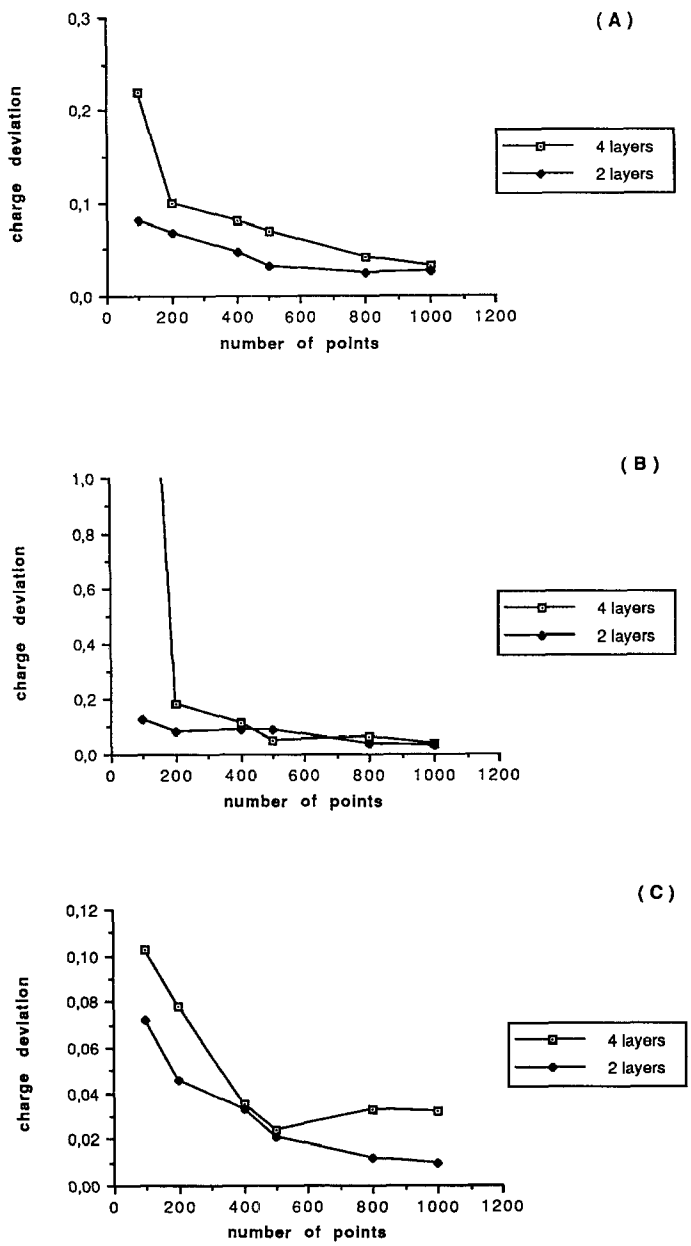

Fig. 5. Variation of the mean differences between the electrostatic charges computed by defining two or four layers on the number of points for adenosine (A), proflavine (B) and psoralen (C).

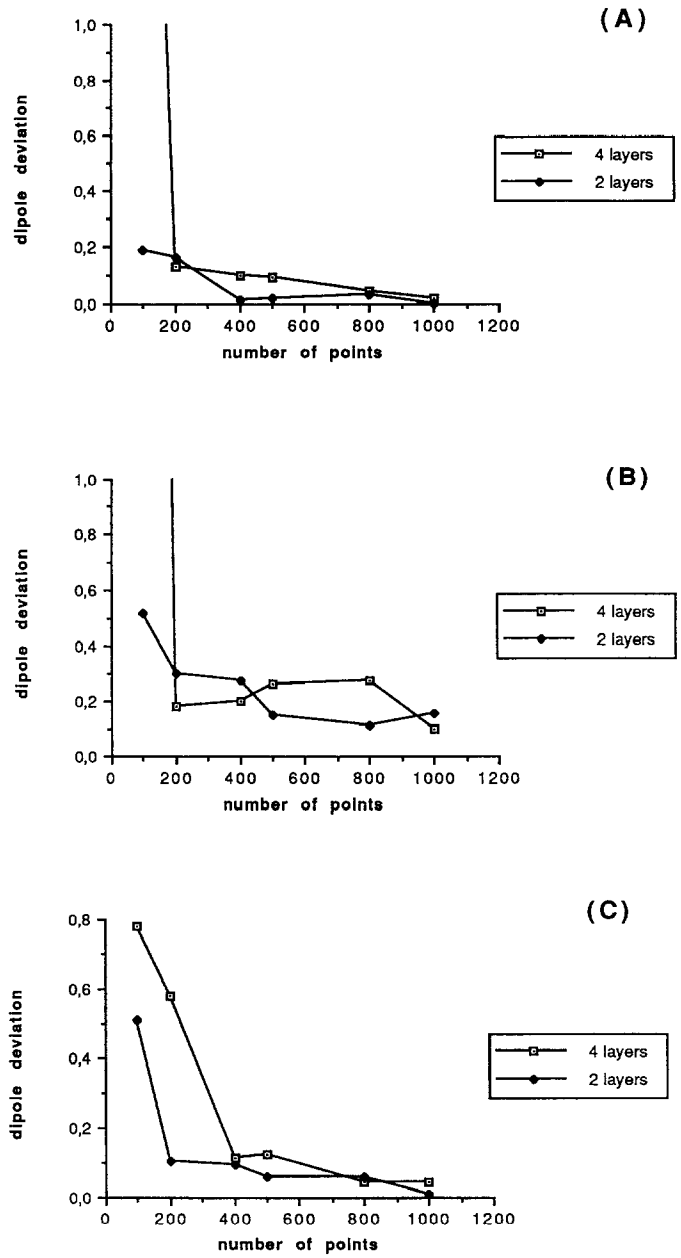

Fig. 6. Variation of the mean differences between the electrostatic dipoles computed by defining two or four layers on the number of points for adenosine (A), proflavine (B) and psoralen (C).

the whole molecule and the fractional model is only 0.064 charge units. As could be expected, the maximum differences appear for the atoms closest to the glycosidic bond. Nevertheless, it is worthwhile to note the quality of the charges derived from the fractional model for the most 'ex-

ternal' atoms, which are the most relevant ones in determining the electrostatic characteristics of the molecule. Finally, it must be emphasized that the molecule maintains its neutrality (± 0.03 charge units), the net charge for the purine in the fractional model being -0.20, and that of the ribose $+0.23$, values very similar to those of the adenine (-0.22) and ribose ($+0.22$) moieties of adenosine.

The fractional model used for psoralen divides the molecule in two aromatic moieties (see Fig. 7). The first moiety (structure A) contains the terminal 5-membered ring, and the second (structure B) possesses the terminal 6-membered ring, the central 6-membered ring appearing in both moieties (the central 6-membered ring has the same geometry in both structures). The molecular electrostatic potential of structure A is computed in 1900 points and that of structure B in 2100 points, these points being distributed in both cases in four layers. Charges of the psoralen molecule are evaluated by fitting simultaneously the charges of both structures A and B. To carry out this fitting, the central 6-membered ring and the Connolly surfaces of each structure are superposed, thus enabling to 'rebuild' the whole molecule and its corresponding Connolly surfaces*. Moreover, the molecular electrostatic potential is evaluated by combining the molecular electrostatic potentials derived from both structures (in our calculations the total number of points after the junction of the two structures is 3400). By using this procedure, the molecular electrostatic potentials in the surroundings of the 5- and 6-membered rings are determined from the

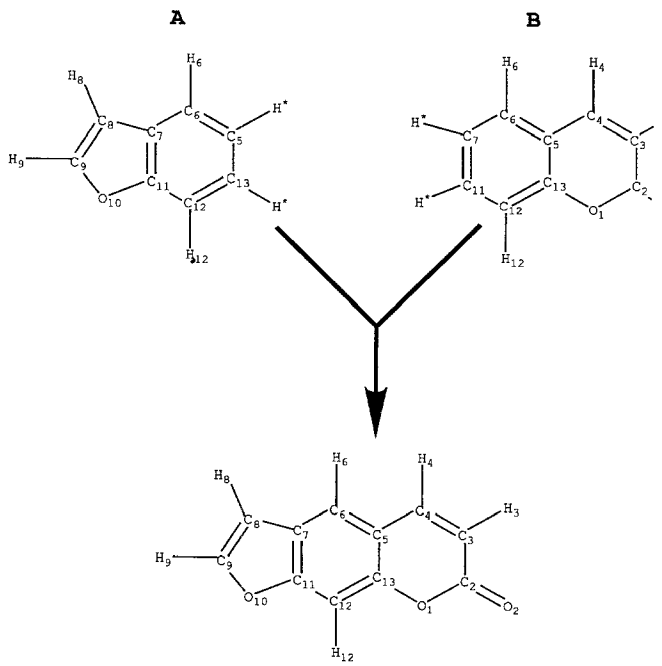

Fig. 7. Representation of the fractional model used to compute the electrostatic charges of psoralen (see text for more details).

*When the psoralen molecule is 'rebuilt', the hydrogen atoms quoted by an asterisk in Fig. 7 and the points of the Connolly surfaces closest to such atoms are eliminated.

structures A and B, respectively, whereas the molecular electrostatic potential in the surroundings of the central 6-membered ring is determined by mixing the molecular electrostatic potentials of the structures A and B.

Results displayed in Table 2 demonstrate the reliability of the fractional method here employed. Thus, the mean difference in the charges obtained considering the whole molecule and the fractional model is only 0.0782 charge units. As is expected, the maximum differences appear for the atoms of the central 6-membered ring.

DISCUSSION

The results presented in this paper clearly point out that reliable electrostatic charges can be obtained only when the quantum mechanical and the monopole-monopole molecular electrostatic potentials are fitted in a large number of points. It must be emphasized that the assumption that correct charges are obtained when there are small deviations of the RMS or relative RMS is absolutely erroneous, since these parameters give information only about the statistical significance of the fitting between both electrostatic potentials, but they do not reflect the quality of the charges and dipoles, as is demonstrated by the fact that the relative RMS deviation obtained when 100 points are considered is very small, whereas the charges and dipoles obtained are absolutely erroneous.

In practice, the evaluation of the quantum mechanical molecular electrostatic potential in thousands of points is not possible due to the computational cost of such a calculation. Therefore, it is necessary to define a precise criterion for determining the minimum number of points where the quantum mechanical and monopole-monopole molecular electrostatic potentials have to be fitted in order to provide reliable charges. Our results point out that when four layers are considered, the minimum number of points is 800 for adenosine, 500 for proflavine and 400 for psoralen. If the molecular size of these three molecules is taken into account, it can be concluded that the optimum number of points when four layers are employed is defined by a density of points ranging from 0.45 to 0.60 points/Å^2.

The use of four Connolly layers guarantees that the molecular electrostatic potential in both the inner and the outer regions is well represented. Nevertheless, when few points are considered, the use of four layers implies a small density of points around the molecule, and consequently the fitting can lead to erroneous results. Here we suggest to use only two Connolly layers separated by 0.4 Å instead of four layers separated by 0.2 Å. This strategy leads to the definition of Connolly layers with a greater density of points and has the advantage that the fitting of the molecular electrostatic potentials can be performed in a reduced number of points, since reasonable charges are obtained considering only 100–200 points for proflavine and psoralen and 400 for adenosine. It should be emphasized that the densities of points around the molecules considering only two layers are similar to the values reported before (density of 0.5 points/Å^2).

Accordingly, from the results presented in this paper, it can be concluded that the number of points used to fit the quantum mechanical and monopole-monopole molecular electrostatic potentials must be selected depending upon the size of the molecule, in such a way that the density of points around the molecule has to be at least 0.5 points/Å^2. For large molecules, where the calculation of quantum mechanical molecular electrostatic potentials is very expensive, the use of

only two layers instead of the usual four permits to drastically reduce the number of points needed to fit the molecular electrostatic potentials with the subsequent computational saving.

The use of fractional models should be a priori avoided. Nevertheless, they become necessary for the calculation of the electrostatic charges for extremely large molecules. There are two different cases where fractional models are used: the first one is for molecules with different structures joined by single bonds, and the second case corresponds to molecules having large aromatic systems. In the first case, the electrostatic charges can be easily obtained with notable accuracy following these steps:

(i) The large molecule is divided (disconnected) into several moieties; (ii) a set of structures is defined by adding to each moiety small groups, which simulate the rest of the moieties of the molecule; (iii) the electrostatic charges of the whole molecule are defined once the electrostatic charges of the different structures are obtained; and (iv) the global charge of the molecule is determined by the addition of the electrostatic charges of all the atoms. If such a charge is equal to the real net charge of the molecule, the calculation is finished. Otherwise, step (iii) is repeated, but imposing the condition that the charge of every small group added to each moiety possesses the same global charge as the subunit which it simulates. This procedure is repeated until the global charge and the net charge are equal.

The use of fractional models for large aromatic structures is more complex, since it is not possible to divide the molecule into moieties by disconnecting single bonds. Results have demonstrated that reasonable charges can be obtained from the use of fractional models in aromatic structures when the following methodology is used:

(i) The aromatic structure is divided into several moieties, each one having at least one aromatic ring in common with the rest of the moieties. The geometry of such common rings must be kept frozen during the geometry optimization of all the fragments; (ii) the quantum mechanical molecular electrostatic potentials of the different structures are computed on Connolly surfaces around the corresponding structures; (iii) the different structures are joined by fitting the corresponding common rings (during this process, the Connolly surfaces of the different structures are superposed in such a way that the whole molecule surrounded by its Connolly layers is rebuilt); (iv) the electrostatic charges of the whole molecule are directly computed.

If the Z-matrix used as input in the MOPAC computation is adequately built, this methodology is only slightly more complex than the simple fractional strategy used for non-aromatic molecules. The increase in the complexity is compensated by the fact that if the electrostatic charges are computed simultaneously for the whole molecule, then the global charge is always identical to the net charge, and consequently, the electrostatic charges are always calculated in only one iteration.

CONCLUSION

The results presented in this work permit to define the minimum requirements necessary for the precise determination of electrostatic charges. The number of points where the quantum mechanical and the monopole-monopole molecular electrostatic potentials are fitted is defined by a density of points ranging from 0.45 to 0.60 points/Å^2 when four layers separated by 0.2 Å are considered. Nevertheless, the evaluation of electrostatic charges for large molecules can be performed by using only two layers separated by 0.4 Å, the density of points being in the range noted above. Finally, when the electrostatic charges have to be determined for extremely large molecules, the re-

sults justify the application of fractional models, even when aromatic moieties exist in the molecules.

REFERENCES

1 Tapia, O. and Åqvist, J., Computer-Assisted Modeling of Receptor–Ligand Interactions: Theoretical Aspects and Applications to Drug Design, Alan R. Liss, New York, 1989.
2 McCammon, J.A., Science, 238 (1987) 486.
3 Kollman, P., Annu. Rev. Phys. Chem., 38 (1987) 303.
4 Smith, R.P., Ree, T., Magee, J.L. and Eyring, H., J. Am. Chem. Soc., 73 (1951) 2263.
5 Del Re, G., J. Chem. Soc., (1958) 4031.
6 Allinger, N.L. and Wuesthoff, T., Tetrahedron, 33 (1977) 3.
7 Abraham, R.J., Griffiths, L. and Loftus, P., J. Comput. Chem., 3 (1982) 407.
8 Mullay, J., J. Comput. Chem., 9 (1988) 399.
9 Mulliken, R.S., J. Chem. Phys., 36 (1962) 3428.
10 Wiberg, K.B., J. Am. Chem. Soc., 101 (1979) 1718.
11 Momany, F.A., J. Phys. Chem., 82 (1978) 592.
12 Reed, A.E., Weinstock, R.B. and Weinhold, F., J. Chem. Phys., 83 (1985) 735.
13 Cox, S.R. and Williams, D.E., J. Comput. Chem., 2 (1981) 304.
14 Chirlian, L.E. and Francl, M.E., J. Comput. Chem., 8 (1987) 894.
15 Orozco, M. and Luque, F.J., J. Comput. Chem., 11 (1990) 909.
16 Weiner, P.K. and Kollman, P.A., J. Comput. Chem., 2 (1981) 287.
17 Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, Jr., S. and Weiner, P., J. Am. Chem. Soc., 106 (1984) 765.
18 Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A., J. Comput. Chem., 7 (1986) 230.
19 Singh, U.C. and Kollman, P.A., J. Comput. Chem., 5 (1984) 129.
20 Connolly, M., QCPE Bull., 1 (1981) 75.
21 Fletcher, R., Practical Methods of Optimization, Wiley, New York, 1980.
22 Orozco, M. and Luque, F.J., Chem. Phys. Lett., 160 (1989) 305.
23 Luque, F.J., Illas, F. and Orozco, M., J. Comput. Chem., 11 (1990) 416.
24 Hariharan, P.C. and Pople, J.A., Theor. Chim. Acta, 28 (1973) 213.
25 Binkley, J.S., Pople, J.A. and Hehre, W.J., J. Am. Chem. Soc., 102 (1980) 939.
26 Hehre, W.J., Stewart, R.F. and Pople, J.A., J. Chem. Phys., 51 (1969) 2657.
27 Luque, F.J. and Orozco, M., Chem. Phys. Lett., 168 (1990) 269.
28 Dewar, M.J.S. and Thiel, W., J. Am. Chem. Soc., 99 (1977) 4899.
29 Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. and Stewart, J.J.P., J. Am. Chem. Soc., 107 (1985) 3902.
30 Besler, B.H., Merz, K.M. and Kollman, P.A., J. Comput. Chem., 11 (1990) 431.
31 Ferenczy, G.G., Reynolds, C.A. and Richards, W.G., J. Comput. Chem., 11 (1990) 159.
32 Scrocco, E. and Tomasi, J., Topics Curr. Chem., 42 (1973) 95.
33 Giessner-Prettre, C. and Pullman, A., Theoret. Chim. Acta, 25 (1972) 83.
34 Duben, A.J., Theoret. Chim. Acta, 59 (1981) 81.
35 Culberson, J.C. and Zerner, M.C., Chem. Phys. Lett., 122 (1985) 436.
36 Löwdin, P.O., J. Chem. Phys., 18 (1950) 365.
37 Gale, E.F., Gundliffe, T., Reynolds, P.E., Richmond, M.H. and Waring, M.J., The Molecular Basis of Antibiotic Action, Wiley, New York, 1981.
38 Dewar, M.J.S. and Thiel, W., J. Am. Chem. Soc., 99 (1977) 4907.
39 Olivella, S., QCPE Bull., 4 (1984) 109 (modified by S. Olivella and J.M. Bofill, 1987).
40 Stewart, J.J.P., QCPE Bull., 3 (1983) 101.