

Reliability of MEP and MEP-derived properties computed from DFT methods for molecules containing P, S and Cl

Robert Soliva¹, Francisco J. Luque², Modesto Orozco¹

¹ Departament de Bioquímica i Biologia Molecular. Facultat de Química. Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona Spain

² Departament de Farmàcia, Unitat Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avda Diagonal s/n, E-08028 Barcelona, Spain

Received: 4 March 1997 / Accepted: 27 June 1997

Abstract. We present a systematic study on the reliability of different theoretical methods to represent the molecular electrostatic potential (MEP), and MEP-derived properties of prototypical compounds containing phosphorus, sulfur and chlorine. Calculations at the Hartree-Fock and Møller-Plesset up to fourth-order level of theory, as well as local, non-local and hybrid density functional computations were performed for a representative set of neutral molecules. The study was carried out using different basis sets ranging from the medium-sized 6-31G(*d*) to the large 6-31G(2*d*,2*p*) basis set, but in some test calculations more extended basis sets were also considered. The analysis of the results was performed discussing separately the effect of the basis set and of the level of theory used to determine the molecular wavefunction on the reliability of the MEP and MEP-derived properties.

Key words: Molecular electrostatic potential – Density functional methods – Electrostatic properties – Second-row atoms

1 Introduction

The molecular electrostatic potential (MEP; [1]) is the electrostatic interaction energy between the molecular charge distribution and the positive unit charge. Within the quantum mechanical framework the MEP is defined as the expectation value of the \mathbf{r}^{-1} operator. In the context of the molecular orbital-linear combination of atomic orbitals (MO-LCAO) approach, a simple expression actually implemented in many computational packages is derived for the MEP (Eq. 1),

$$V(\mathbf{r}_1) = \sum_A \frac{Z_A}{|\mathbf{r}_1 - \mathbf{R}_A|} - \sum_{\mu} \sum_{\nu} P_{\mu\nu} \int \frac{\chi_{\mu}(\mathbf{r})\chi_{\nu}(\mathbf{r})}{|\mathbf{r}_1 - \mathbf{r}|} d\mathbf{r}, \quad (1)$$

where Z_A denotes the nuclear charge of atom A , placed at \mathbf{R}_A , $P_{\mu\nu}$ is the element $\mu\nu$ of the first-order density matrix, and χ stands for the basis of atomic orbitals.

The MEP has largely been used as a molecular descriptor of chemical reactivity in different research fields, which include the study of reactivity patterns, non-covalent complexes and biological interactions [2–12], the analysis of the molecular electronic structure [13–21], and the description of phenomena in condensed phases [22–30]. Moreover, the MEP has become a widely used tool in force-field parametrization as a source for the derivation of partial charges, which are used to evaluate electrostatic interactions in classical calculations [31–37].

The calculation of the MEP is not very expensive from a computational point of view, specially with the advent of efficient algorithms for the calculation of the mono-electronic integrals [38–40]. Nevertheless, knowledge of the first-order density matrix typically requires the molecular wavefunction to be determined previously, and this can be very expensive if the highest levels of the quantum mechanical theory are used. Therefore, the a priori knowledge of the expected accuracy in the description of the MEP would be very valuable to establish the minimum level of theory necessary to determine the essential features of the MEP and MEP-derived properties.

Previous studies performed by our group and others examined the influence of the basis set and the electron correlation on the quality of the MEPs determined at the ab initio level, and the accuracy of the MEPs derived from semiempirical wavefunctions [41–52]. Furthermore, these studies have been recently extended to determine the suitability of DFT methods for providing an accurate description of the MEP and MEP-related properties [53–55]. These studies were basically carried out using a variety of molecules containing first-row atoms (H, C, N, O, F), and less attention was paid to second-row atoms like phosphorus, sulfur, and chlorine, which are of importance in chemistry and biochemistry.

In this paper a systematic study on the MEP and MEP-derived properties of molecules containing P, S and Cl is presented. The influence of the basis set and electron correlation is discussed in the framework of

SCF, MP2, and MP4 levels of theory. The suitability of different DFT methods (including local, non-local and hybrid functionals) is also critically analyzed. Discussion of the results is made in comparison with those obtained in previous studies for first-row elements.

2 Methods

The molecules considered in the study are H₂S, HCl, PH₃, CH₃Cl, S(CH₃)₂, PH₂CH₃, SO₂, SCl₂, PF₃, HCCCl, HCOCl, PF₂NH₂, CH₂ClF, CH₃CH₂Cl, PHF₂, CNCl, PClF₂, thiophene, and CH₃NCS. The geometries were optimized at the HF/6-31G(*d*) level, and the optimized geometries were then used for single-point calculations with the other formalisms.

Owing to the expensiveness of the MP4 calculations with the largest basis sets, the influence of the basis set on the quality of the results was examined considering a limited number of molecules, which is referred to in the text as “subset”: H₂S, HCl, PH₃, CH₃Cl, S(CH₃)₂, PH₂CH₃, SO₂, SCl₂, PF₃, HCCCl, and P(CH₃)₃. The following basis sets [56–58] were considered in all cases: 6-31G(*d*), 6-31G(*d,p*), 6-31G(2*d*), and 6-31(2*d,2p*). In addition, selected test calculations (data not shown) were performed with the basis 6-311G(*d*), 6-311G(*d,p*), 6-311+G(*d,p*) and 6-311++G(*d*) [59–61]. In all these cases the differences in the MEP distributions with respect to those derived from the corresponding split-valence basis set without diffuse functions were negligible, which led us to ignore these extensions to the basis set in the systematic study of the basis set dependence.

Calculations were performed at the HF, MP2, and MP4(SDQ) levels [62], as well as using two local functionals: SVWN [63, 64] and SPL [63, 65], one non-local functional: BP86 [66, 67], and three hybrid non-local functionals: B3P86 [67, 68], B3LYP [69], and B3PW91 [68, 70]. Test calculations at the QCISD level were also performed with different basis sets for selected molecules. In all cases the QCISD results were almost identical to those derived at the MP4 level for the same basis set, which led us to exclude the very expensive QCISD calculations from the study. The use of MP4(SDQ) calculations as a source of reference data is also supported by our previous studies [46, 53], which revealed a perfect agreement between the MEPs determined at the MP4(SDQ), CIPSI(G+M) [71] and Full-CI calculations. Let us note, nevertheless, that the choice of MP4(SDQ) results as reference values does not imply a priori any assumption on the quality of the MEPs determined at other levels of theory. Møller-Plesset calculations were performed using the frozen-core approximation, which is justified by our previous findings [53] that the contribution owing to excitations of inner electrons has a negligible effect on the MEP.

In order to examine the influence of the basis set and the treatment of electron correlation, the MEP was computed in layers around the molecule using Connolly’s approach [72]. The layers were placed at different distances from the molecule ranging from 0.6 to 2.0 times the van der Waals radii of atoms (in Å). In addition,

electrostatic potential-derived (ESP) atomic charges and dipoles were determined using Momany’s strategy [73] and the optimization protocol noted in detail elsewhere [35, 43, 74]. Finally, inspection of the molecular electrostatic properties was extended to the dipole moments determined rigorously from the molecular wavefunction, which were compared with the gas phase experimental values.

All the wavefunctions were determined with Gaussian-94 [75]. The MEPs were computed with either Gaussian-94 or MOPETE/MOPFIT [76] computer programs. ESP charges were calculated with MOPETE/MOPFIT. Calculations were carried out on the IBM-SP2 computer of the Centre de Supercomputació de Catalunya (CESCA), as well as in workstations in our laboratory.

3 Results and discussion

In order to clarify the discussion we shall firstly examine the dependence of the MEP and MEP-derived properties determined at different computational levels (HF, DFT, MPx) on the basis set, and subsequently the dependence of the results on the level of theory used in the treatment of electron correlation effects.

3.1 Basis set dependence

As noted before, we limited the study to the following basis sets: 6-31G(*d*), 6-31G(*d,p*), 6-31G(2*d*), and 6-31(2*d,2p*), since a series of test calculations (data not shown) revealed that the use of triple-zeta basis for valence electrons or the inclusion of diffuse functions was found to have little influence on the main features of the MEP. On the other hand, the use of more extended basis sets including another set of *d* orbitals, or even *f* orbitals, which is computationally very demanding, does not seem justifiable in view of the results obtained for the largest basis sets (see below).

Comparison of ESP atomic charges and dipole moments determined with the different basis sets at a given level of theory allowed us to analyze the importance of: (1) polarization functions on hydrogen atoms, and (2) a second set of *d* functions on heavy atoms. Furthermore, comparison with experimental gas phase dipole moments allowed us to examine the ability of theoretical calculations to reproduce experimental measures of the molecular charge distribution.

Tables 1 and 2 report the statistical results of the comparison for ESP charges (only charges over P, S or Cl atoms were considered in the analysis) and for the QM dipoles (similar results were obtained for ESP dipoles, but are omitted for simplicity), respectively. The values derived at the corresponding computational level using the 6-31G(2*d,2p*) basis set were used as reference in the comparison. Inclusion of *p* orbitals on hydrogen atoms has a negligible influence at the HF and DFT levels, as noted by the close similarity between the root-mean square (RMS) deviations and the optimum ratios, while they have only a moderate effect in MPx calcula-

Table 1. Results of the statistical analysis of electrostatic potential-derived charges for P, S and Cl atoms in neutral molecules. The atomic charges derived at the corresponding level of theory using the 6-31G(2*d*,2*p*) basis set are used as reference

Method	Basis	r ^a	c ^b	RMS ^c
HF	6-31G(<i>d</i>)	0.986	0.947	0.037
	6-31G(<i>d,p</i>)	0.988	0.945	0.034
	6-31G(2 <i>d</i>)	0.999	0.993	0.004
SVWN	6-31G(<i>d</i>)	0.991	0.930	0.030
	6-31G(<i>d,p</i>)	0.991	0.941	0.029
	6-31G(2 <i>d</i>)	0.999	0.978	0.006
SPL	6-31G(<i>d</i>)	0.991	0.930	0.029
	6-31G(<i>d,p</i>)	0.991	0.941	0.028
	6-31G(2 <i>d</i>)	0.999	0.978	0.006
BP86	6-31G(<i>d</i>)	0.991	0.923	0.029
	6-31G(<i>d,p</i>)	0.991	0.933	0.028
	6-31G(2 <i>d</i>)	0.999	0.977	0.006
B3P86	6-31G(<i>d</i>)	0.994	0.913	0.028
	6-31G(<i>d,p</i>)	0.991	0.941	0.029
	6-31G(2 <i>d</i>)	0.999	0.982	0.006
B3LYP	6-31G(<i>d</i>)	0.991	0.931	0.028
	6-31G(<i>d,p</i>)	0.990	0.940	0.030
	6-31G(2 <i>d</i>)	0.999	0.985	0.005
B3PW91	6-31G(<i>d</i>)	0.990	0.932	0.030
	6-31G(<i>d,p</i>)	0.991	0.940	0.029
	6-31G(2 <i>d</i>)	0.999	0.982	0.006
MP2	6-31G(<i>d</i>)	0.985	0.921	0.037
	6-31G(<i>d,p</i>)	0.984	0.951	0.035
	6-31G(2 <i>d</i>)	0.998	0.940	0.014
MP4	6-31G(<i>d</i>)	0.986	0.908	0.037
	6-31G(<i>d,p</i>)	0.985	0.943	0.036
	6-31G(2 <i>d</i>)	0.998	0.931	0.016

^a Pearson correlation coefficient

^b The optimum (6-31G(2*d*,2*p*) basis/other basis) ratio

^c The root mean square deviation (in units of electron)

tions. On the contrary, a second set of *d* orbitals on heavy atoms has a major influence on the MEP-derived properties at all levels of theory. Interestingly, the dependence of the results on extension of the basis set is very similar in all cases (HF, DFT, or MPx), which reveals that the effect of the basis set on the MEP and MEP-derived properties is quite independent of the influence played by the electron correlation. This is also shown in Tables 3 and 4, which give the statistical results of the comparison of ESP charges (Table 3) and QM dipoles (Table 4) obtained at the MP4 level with regard to the values derived from HF, DFT and MP2 calculations using the same basis set. It is clear that the difference between the MP4 results and the values determined from other methods does not change greatly for different basis sets.

The ability of the different basis sets in HF, DFT or MPx calculations to reproduce experimental measures can be seen in the comparison between theoretical and experimental dipole moments. The statistical results are given in Table 5. In general, inclusion of a second set of *d* functions on heavy atoms is more important than addition of *p* orbitals on hydrogen atoms, even though these latter orbitals are important in MPx calculations (see also Tables 1, 2). As expected, the enlargement of the basis set improves the agreement between computed and experimental dipoles at all the levels of calculation, as noted in the changes of the experimental/theoretical

Table 2. Results of the statistical analysis of dipole moments for neutral molecules. The dipoles derived at the corresponding level of theory using the 6-31G(2*d*,2*p*) basis set are used as reference

Method	Basis	r ^a	c ^b	RMS ^c
HF	6-31G(<i>d</i>)	0.977	0.901	0.184
	6-31G(<i>d,p</i>)	0.979	0.900	0.183
	6-31G(2 <i>d</i>)	0.999	0.999	0.020
SVWN	6-31G(<i>d</i>)	0.982	0.877	0.201
	6-31G(<i>d,p</i>)	0.984	0.888	0.180
	6-31G(2 <i>d</i>)	0.999	0.989	0.026
SPL	6-31G(<i>d</i>)	0.982	0.877	0.201
	6-31G(<i>d,p</i>)	0.984	0.889	0.182
	6-31G(2 <i>d</i>)	0.999	0.989	0.026
BP86	6-31G(<i>d</i>)	0.982	0.881	0.190
	6-31G(<i>d,p</i>)	0.984	0.889	0.173
	6-31G(2 <i>d</i>)	0.999	0.991	0.025
B3P86	6-31G(<i>d</i>)	0.966	0.899	0.187
	6-31G(<i>d,p</i>)	0.983	0.892	0.174
	6-31G(2 <i>d</i>)	0.999	0.992	0.024
B3LYP	6-31G(<i>d</i>)	0.982	0.891	0.174
	6-31G(<i>d,p</i>)	0.979	0.892	0.177
	6-31G(2 <i>d</i>)	0.999	0.994	0.021
B3PW91	6-31G(<i>d</i>)	0.981	0.882	0.191
	6-31G(<i>d,p</i>)	0.983	0.891	0.177
	6-31G(2 <i>d</i>)	0.999	0.992	0.023
MP2	6-31G(<i>d</i>)	0.977	0.867	0.210
	6-31G(<i>d,p</i>)	0.977	0.904	0.162
	6-31G(2 <i>d</i>)	0.993	0.939	0.094
MP4	6-31G(<i>d</i>)	0.977	0.861	0.218
	6-31G(<i>d,p</i>)	0.971	0.890	0.184
	6-31G(2 <i>d</i>)	0.990	0.930	0.119

^a Pearson correlation coefficient

^b The optimum (6-31G(2*d*,2*p*) basis/other basis) ratio

^c The root mean square deviation (in Debye)

ratio, which indicate that HF, DFT, and MPx calculations overestimate the polarity of the charge distribution, and that the overestimation is reduced as the basis set becomes larger. The reduction of the molecular polarity is quite constant at all the levels of theory. Thus, the change from 6-31G(*d*) to 6-31G(2*d*,2*p*) increases the experimental/theoretical ratio by ca 0.1 units, and the RMS deviation decreases by around 0.12 Debye. This finding also supports the separate modulation of the basis set and electron correlation effects on the determination of molecular electrostatic properties.

Another item of interest is the relevance of the basis set effect on the MEP in different regions of the space. In order to analyze this point, the dependence of the MEP computed in layers around the molecule with the basis set was examined. Layers were placed at 0.6, 0.8, 1.0, 1.2, 1.4, 1.6 and 2.0 times the van der Waals atomic radii. Calculations were performed only at the MP2 level, since the conclusions can be extrapolated to the other methods owing to the large separability between basis set and electron correlation effects. Figure 1 shows the profiles of the RMS deviation, averaged for the subset of molecules, between the MEP at the MP2/6-31G(2*d*,2*p*) level with regard to the MP2 values computed with the smaller basis sets. Inspection of Fig. 1 shows the close similarity of the profiles for the basis 6-31G(*d*) and 6-31G(*d,p*), which only differ in the innermost layer, and the reduction in the RMS error obtained for the

Table 3. Results of the statistical analysis of electrostatic potential-derived charges for P, S and Cl atoms in neutral molecules. The atomic charges derived at the MP4 level of theory using the corresponding basis set are used as reference

Method	Basis	r^a	c^b	RMS ^c
HF	6-31G(<i>d</i>)	0.994	0.905	0.040
	6-31G(<i>d,p</i>)	0.997	0.884	0.040
	6-31G(2 <i>d</i>)	0.992	0.925	0.040
	6-31G(2 <i>d,2p</i>)	0.998	0.866	0.038
SVWN	6-31G(<i>d</i>)	0.982	1.044	0.052
	6-31G(<i>d,p</i>)	0.974	0.980	0.060
	6-31G(2 <i>d</i>)	0.991	1.056	0.030
	6-31G(2 <i>d,2p</i>)	0.982	0.992	0.043
SPL	6-31G(<i>d</i>)	0.981	1.042	0.053
	6-31G(<i>d,p</i>)	0.972	0.979	0.062
	6-31G(2 <i>d</i>)	0.990	1.052	0.040
	6-31G(2 <i>d,2p</i>)	0.980	0.990	0.045
BP86	6-31G(<i>d</i>)	0.987	1.073	0.047
	6-31G(<i>d,p</i>)	0.980	1.020	0.053
	6-31G(2 <i>d</i>)	0.992	1.082	0.030
	6-31G(2 <i>d,2p</i>)	0.985	1.028	0.040
B3P86	6-31G(<i>d</i>)	0.995	1.042	0.028
	6-31G(<i>d,p</i>)	0.991	0.992	0.033
	6-31G(2 <i>d</i>)	0.999	1.053	0.020
	6-31G(2 <i>d,2p</i>)	0.995	0.990	0.022
B3LYP	6-31G(<i>d</i>)	0.997	1.041	0.023
	6-31G(<i>d,p</i>)	0.997	0.992	0.033
	6-31G(2 <i>d</i>)	0.998	1.044	0.020
	6-31G(2 <i>d,2p</i>)	0.996	0.990	0.020
B3PW91	6-31G(<i>d</i>)	0.996	1.040	0.027
	6-31G(<i>d,p</i>)	0.991	0.992	0.033
	6-31G(2 <i>d</i>)	0.999	1.056	0.020
	6-31G(2 <i>d,2p</i>)	0.994	0.993	0.024
MP2	6-31G(<i>d</i>)	0.998	1.023	0.017
	6-31G(<i>d,p</i>)	0.998	1.017	0.018
	6-31G(2 <i>d</i>)	0.998	1.021	0.010
	6-31G(2 <i>d,2p</i>)	0.997	1.013	0.016

^a Pearson correlation coefficient

^b The optimum (MP4/other method) ratio

^c The root mean square deviation (in units of electron)

6-31G(2*d*) basis in all the layers, which confirms the larger influence of the second set of *d* orbitals.

As a summary, present results indicate that the accurate representation of electrostatic properties for these molecules requires the use of large basis sets, including at least a double set of *d* polarization functions on heavy atoms. It is worth noting that these results do not agree with our previous studies for molecules containing first-row atoms, for which the 6-31G(*d*) basis set was found to provide a reasonable description of the MEP [41, 46, 53]. For instance, for 27 molecules containing only first-row atoms the experimental/MP4 ratio between dipole moments was 0.996, the RMS deviation being only 0.16 Debye (53). These results are in contrast with those reported here for the same method and basis set, which are 0.822 and 0.29 Debye (Table 5), respectively, for the subset of molecules, and 0.878 and 0.34 Debye for the entire set of molecules.

3.2 Dependence on correlation effects

The influence of the level of theory on the MEP and MEP-derived properties can be examined from the

Table 4. Results of the statistical analysis of dipole moments for neutral molecules. The dipoles derived at the MP4 level of theory using the corresponding basis set are used as reference

Method	Basis	r^a	c^b	RMS ^c
HF	6-31G(<i>d</i>)	0.956	0.930	0.16
	6-31G(<i>d,p</i>)	0.980	0.897	0.18
	6-31G(2 <i>d</i>)	0.958	0.956	0.14
	6-31G(2 <i>d,2p</i>)	0.990	0.886	0.17
SVWN	6-31G(<i>d</i>)	0.863	0.975	0.22
	6-31G(<i>d,p</i>)	0.763	0.946	0.27
	6-31G(2 <i>d</i>)	0.951	1.028	0.14
	6-31G(2 <i>d,2p</i>)	0.917	0.970	0.19
SPL	6-31G(<i>d</i>)	0.851	0.975	0.23
	6-31G(<i>d,p</i>)	0.746	0.945	0.28
	6-31G(2 <i>d</i>)	0.945	1.027	0.15
	6-31G(2 <i>d,2p</i>)	0.908	0.969	0.19
BP86	6-31G(<i>d</i>)	0.892	1.013	0.20
	6-31G(<i>d,p</i>)	0.805	0.983	0.24
	6-31G(2 <i>d</i>)	0.957	1.064	0.15
	6-31G(2 <i>d,2p</i>)	0.927	1.003	0.17
B3P86	6-31G(<i>d</i>)	0.963	1.009	0.12
	6-31G(<i>d,p</i>)	0.922	0.962	0.17
	6-31G(2 <i>d</i>)	0.991	1.036	0.08
	6-31G(2 <i>d,2p</i>)	0.978	0.973	0.10
B3LYP	6-31G(<i>d</i>)	0.976	1.015	0.10
	6-31G(<i>d,p</i>)	0.956	0.979	0.12
	6-31G(2 <i>d</i>)	0.993	1.053	0.09
	6-31G(2 <i>d,2p</i>)	0.983	0.985	0.09
B3PW91	6-31G(<i>d</i>)	0.965	0.992	0.12
	6-31G(<i>d,p</i>)	0.914	0.961	0.17
	6-31G(2 <i>d</i>)	0.989	1.038	0.08
	6-31G(2 <i>d,2p</i>)	0.975	0.975	0.11
MP2	6-31G(<i>d</i>)	0.987	0.994	0.07
	6-31G(<i>d,p</i>)	0.970	0.999	0.10
	6-31G(2 <i>d</i>)	0.990	1.001	0.06
	6-31G(2 <i>d,2p</i>)	0.990	0.992	0.07

^a Pearson correlation coefficient

^b The optimum (MP4/other method) ratio

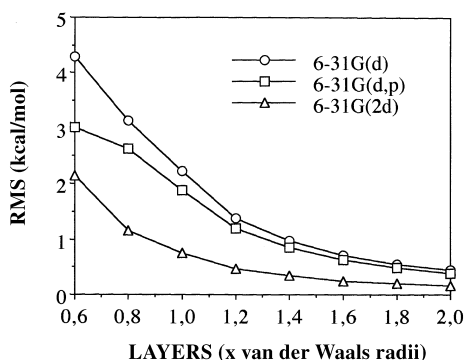
^c The root mean square deviation (in Debye)

results in Tables 3 and 4. HF values overestimate the molecular polarity with respect to MP4 results, as reflected in the enlarged atomic charges and dipoles. The use of local and BP86 DFT methods does not lead to a marked improvement over the HF values, since the corresponding charges and dipoles do not correlate well with MP4 values, and the RMS errors are quite large. Hybrid non-local DFT methods perform noticeably better, and the results are only slightly worse than those determined at the MP2 level. No significant differences are found between the three hybrid non-local functionals considered in the study.

In order to verify the trends observed from the preceding analysis, we extended the study to the entire set of 19 molecules. Taking advantage of the separability between basis set and electron correlation effects (see above), the study was limited to the 6-31G(*d*) basis set. Results of the statistical analysis for ESP charges and dipoles with regard to the MP4 values is given in Table 6, which also includes the results for the subset of molecules for comparison. There is general agreement between the results for the whole set of compounds and for the subset. Some differences are found in the correlation coefficients, since the enlargement in the number

Table 5. Results of the statistical analysis of dipole moments for neutral molecules. The gas phase experimental values are used as reference

Method	Basis	r ^a	c ^b	RMS ^c
HF	6-31G(<i>d</i>)	0.958	0.768	0.37
	6-31G(<i>d,p</i>)	0.956	0.766	0.38
	6-31G(2 <i>d</i>)	0.982	0.843	0.23
SVWN	6-31G(2 <i>d,2p</i>)	0.983	0.843	0.23
	6-31G(<i>d</i>)	0.836	0.798	0.38
	6-31G(<i>d,p</i>)	0.846	0.809	0.37
SPL	6-31G(2 <i>d</i>)	0.853	0.890	0.28
	6-31G(2 <i>d,2p</i>)	0.857	0.899	0.27
	6-31G(<i>d</i>)	0.827	0.798	0.39
BP86	6-31G(<i>d,p</i>)	0.837	0.809	0.37
	6-31G(2 <i>d</i>)	0.842	0.889	0.29
	6-31G(2 <i>d,2p</i>)	0.846	0.898	0.28
B3P86	6-31G(<i>d</i>)	0.869	0.831	0.32
	6-31G(<i>d,p</i>)	0.878	0.842	0.31
	6-31G(2 <i>d</i>)	0.867	0.922	0.33
B3P86	6-31G(2 <i>d,2p</i>)	0.871	0.931	0.24
	6-31G(<i>d</i>)	0.908	0.827	0.31
	6-31G(<i>d,p</i>)	0.933	0.823	0.30
B3LYP	6-31G(2 <i>d</i>)	0.932	0.904	0.25
	6-31G(2 <i>d,2p</i>)	0.935	0.910	0.20
	6-31G(<i>d</i>)	0.937	0.835	0.28
B3LYP	6-31G(<i>d,p</i>)	0.934	0.835	0.27
	6-31G(2 <i>d</i>)	0.943	0.920	0.21
	6-31G(2 <i>d,2p</i>)	0.944	0.924	0.18
B3PW91	6-31G(<i>d</i>)	0.923	0.814	0.32
	6-31G(<i>d,p</i>)	0.929	0.822	0.30
	6-31G(2 <i>d</i>)	0.927	0.904	0.18
MP2	6-31G(2 <i>d,2p</i>)	0.930	0.911	0.20
	6-31G(<i>d</i>)	0.937	0.816	0.31
	6-31G(<i>d,p</i>)	0.956	0.854	0.24
MP2	6-31G(2 <i>d</i>)	0.917	0.871	0.25
	6-31G(2 <i>d,2p</i>)	0.950	0.932	0.17
	6-31G(<i>d</i>)	0.955	0.822	0.29
MP4	6-31G(<i>d,p</i>)	0.952	0.852	0.25
	6-31G(2 <i>d</i>)	0.944	0.872	0.23
	6-31G(2 <i>d,2p</i>)	0.978	0.950	0.12

^a Pearson correlation coefficient^b The optimum (experimental/theoretical) ratio^c The root mean square deviation (in Debye)**Fig. 1.** Representation of the average RMS deviation (kcal/mol) between MP2/6-31G(2*d,2p*) MEPs in different layers around the molecule and MEP estimates obtained from MP2/6-31G(2*d*), MP2/6-31G(*d,p*), and MP2/6-31G(*d*) calculations

of compounds improves, as expected, the goodness of the correlation. This is particularly remarkable for local DFT methods, which exhibited the poorest correlations for the subset. In spite of these changes, the trends

Table 6. Results of the statistical analysis of electrostatic potential-derived charges for P, S and Cl atoms and exact dipole moments (italics) determined using the 6-31G(*d*) basis set in neutral molecules. The atomic charges at the MP4 level of theory are used as reference. Values are given for the total set and for the subset (in parenthesis) of molecules

Method	r ^a	c ^b	RMS ^c
HF	0.995 (0.994)	0.904 (0.905)	0.035 (0.040)
SVWN	0.992 (0.956)	0.925 (0.930)	0.19 (0.16)
	0.975 (0.982)	1.062 (1.044)	0.054 (0.052)
SPL	0.974 (0.863)	1.030 (0.975)	0.21 (0.22)
	0.973 (0.981)	1.060 (1.042)	0.056 (0.053)
BP86	0.972 (0.851)	1.030 (0.975)	0.22 (0.23)
	0.981 (0.987)	1.088 (1.073)	0.050 (0.047)
B3P86	0.980 (0.892)	1.052 (1.013)	0.20 (0.20)
	0.994 (0.995)	1.053 (1.042)	0.028 (0.028)
B3LYP	0.993 (0.963)	1.023 (1.009)	0.11 (0.12)
	0.996 (0.997)	1.053 (1.041)	0.025 (0.023)
B3PW91	0.995 (0.976)	1.032 (1.015)	0.11 (0.10)
	0.994 (0.996)	1.051 (1.040)	0.029 (0.027)
MP2	0.993 (0.965)	1.018 (0.992)	0.11 (0.12)
	0.998 (0.998)	1.023 (1.023)	0.015 (0.017)
	0.998 (0.987)	1.001 (0.994)	0.06 (0.07)

^a Pearson correlation coefficient^b The optimum (MP4/other method) ratio^c The root mean square deviation (in units of electron)**Table 7.** Results of the statistical analysis of MEP minima near P, S and Cl atoms determined using the 6-31G(*d*) basis set in the subset of neutral molecules. The depth of MEP minima determined from MP4 calculations are used as reference

Method	r ^a	c ^b	RMS ^c
HF	0.998	0.948	1.5
SVWN	0.996	0.943	1.8
SPL	0.996	0.943	1.8
BP86	0.997	0.967	1.3
B3P86	0.998	0.962	1.2
B3LYP	0.999	0.974	0.8
B3PW91	0.999	0.956	1.3
MP2	1.000	0.973	0.7

^a Pearson correlation coefficient^b The optimum (MP4/other method) ratio^c The root mean square deviation (in kcal/mol)

discussed for the subset of compounds are also valid for the entire set of molecules: only the hybrid non-local DFT methods, which provide very similar results, represent a significant improvement with respect to HF calculations for determining the MEP and MEP-related properties.

Results in Table 5 show that the accuracy of the dipole moment increases as the level of theory changes from HF to MPx. Thus, the experimental/theoretical ratio increases around 0.08 units and the RMS error decreases by around 0.1 Debye upon changing from HF to MP4 irrespectively of the basis set. As expected, the best agreement with experimental values are obtained at the MP4 level, and for the largest basis set. Let us note, however, that even though inclusion of electron correlation is important, the effect of basis set extension is also relevant, as stated by the fact that the HF/6-31G(2*d,2p*) dipoles are slightly better than the MP4/

Table 8. Average root-mean square deviations (in kcal/mol) between the MEPs computed at different levels of theory using the 6-31G(*d*) basis set with regard to the values derived from MP4 calculations at selected layers from the molecules. See text for details

Layer	HF	SVWN	SPL	BP86	B3P86	B3LYP	B3PW91	MP2
0.6	4.7	7.1	6.6	6.1	5.6	4.4	5.5	1.1
0.8	2.8	4.4	4.2	3.6	3.0	2.3	3.0	0.8
1.0	1.9	2.7	2.7	2.2	1.7	1.3	1.7	0.5
1.2	1.3	1.8	1.8	1.5	1.0	0.8	1.0	0.4
1.4	1.0	1.3	1.3	1.1	0.7	0.5	0.7	0.3
1.6	0.8	1.0	1.0	0.8	0.5	0.4	0.5	0.2
1.8	0.6	0.7	0.8	0.7	0.4	0.3	0.4	0.2
2.0	0.5	0.6	0.6	0.5	0.3	0.3	0.3	0.1

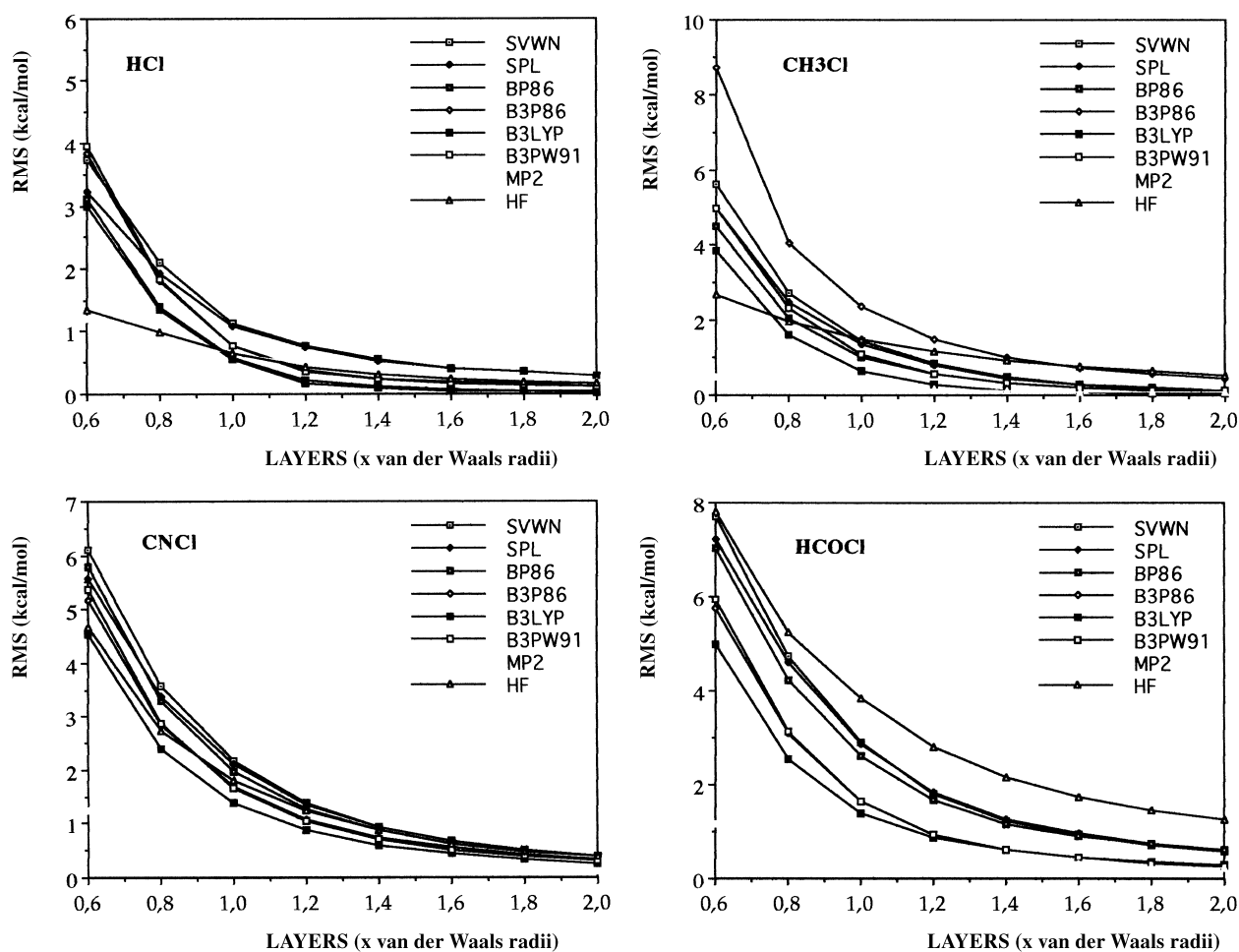


Fig. 2. Representation of the RMS deviation (kcal/mol) between MP4/6-31G(*d*) MEPs in different layers around the selected molecules and MEP estimates obtained from other QM methods using in all cases the 6-31G(*d*) basis set

6-31G(*d*) ones. For medium sized molecules like those considered here the cost of the HF/6-31G(2*d*,2*p*) calculation involves around half of the expense required for the MP2/6-31G(*d*) calculation. This leads us to recommend the use of large basis sets for the calculation of the MEP in molecules containing second-row elements, even when this implies the need to use a poorer description of correlation effects in the wavefunction.

The DFT dipoles are in all cases larger than the experimental values, i.e., the molecular polarity is overestimated. This finding is opposite to the trend observed

for molecules containing exclusively first-row atoms [53]. The hybrid non-local methods provide results comparable to the values determined at the MP2, or even MP4, levels. Even though the poor correlations obtained for the BP86 functional can improve if the whole set of molecules is considered, it does not appear to be the best alternative to compute the electrostatic properties for molecules containing second-row atoms.

In order to complete the analysis, the influence of electron correlation on the MEP minima and on the MEP in layers around the molecule was examined. The analysis was performed using only the 6-31G(*d*) basis set, and the MP4 values were taken as reference for comparison. Table 7 gives the results of the statistical analysis for the MEP minima. HF results are overestimated (in absolute value) with regard to the MP4 values,

the RMS error being 1.5 kcal/mol. As expected from previous studies ([46, 53] and references therein), at the MP2 level the agreement with the MP4 results improves remarkably, as noted in the decrease of the RMS deviation, which amounts to 0.7 kcal/mol. The DFT methods overestimate in all cases the depth of MEP minima. The results provided from local DFT methods are worse than the HF values, as noted in the larger RMS deviation (around 1.8 kcal/mol). The statistical results for non-local BP86 and hybrid DFT methods are slightly worse than the values determined at the MP2 level. The best agreement with the MP4 results is found for the B3LYP functional, for which the RMS error amounts only to 0.8 kcal/mol.

Table 8 gives the average RMS deviation for the MEPs computed at different layers, and Fig. 2 shows the RMS profile for selected molecules. In all cases the largest discrepancies with regard to the MP4 values are in layers inside the van der Waals sphere, where correlation effects are very important. Comparison with our previous results [46, 53] shows that these effects are more important for molecules containing second-row elements, owing probably to the greater electron density in inner regions. The influence of electron correlation diminishes very fast outside the van der Waals sphere. Thus, the average RMS deviations are lower than 1 kcal/mol at 1.6 times the van der Waals radii. The MP2 results almost reproduce the MP4 values in all the regions of space. The behavior of DFT methods is correct in outer regions, where all of them, specially the hybrid non-local functionals, give average RMS errors smaller than those obtained at the HF level. However, the performance of DFT methods is notably worse in inner regions, where the HF results are even closer to the MP4 values.

4 Conclusions

The MEP and MEP-derived properties of neutral molecules containing second-row atoms (P, S and Cl) are more difficult to represent than the MEP of molecules containing only first-row atoms (H, C, N, O, F).

The 6-31G(*d*) basis set, which provides good results for molecules containing only first-row atoms, is not flexible enough to represent the MEP in molecules with second-row atoms. Results presented here show that a second set of *d* orbitals is necessary to obtain an accurate representation of the MEP, whereas inclusion of *p* orbitals on hydrogen atoms provides a moderate improvement in the description of electrostatic properties.

It is found that electron correlation and basis set effects are quite independent, and that accurate values of the MEP and MEP-derived properties are very sensitive to the proper choice of the basis set. Electron correlation effects are more important here than for molecules containing only first-row atoms. The results indicate that calculations at the MP2 level capture properly most of the electron correlation effect, and it is not necessary to use higher level methods.

Local and non-local BP86 DFT methods lead to MEP distributions which, at a given basis set, do not

noticeably improve the MEP representation derived from HF calculations. A more remarkable improvement is obtained using hybrid non-local DFT methods, even though no significant differences are found in general between the three hybrid non-local functionals considered in the study.

Acknowledgements. We would like to thank Prof. C.A. Reynolds for discussions and for a preprint of his work prior to publication. We also thank the Centre de Supercomputació de Catalunya (CESCA, Mol. Recog. Project) for computational facilities, and the Direcció General de Investigació Científica y Tècnica (DGICYT; grants PB93-0779 and PB94-0940) for financial support. R. S. is recipient of a pre-doctoral grant from the Comissió Interministerial de Recerca i Innovació Tecnològica.

References

- Scrocco E, Tomasi J (1973) *J Comp Chem* 42:95
- Scrocco E, Tomasi J (1978) *Adv Quant Chem* 11:115
- Politzer P, Daiker KC (1981) In: Debb DM (ed) *The Force concept in chemistry*. Van Nostrand Reinhold, Amsterdam, p. 294
- Politzer P, Truhlar DG (eds) (1981) *Chemical applications of atomic and molecular electrostatic potentials*. Plenum, New York
- Politzer P, Murray JS (1991) In: Lipkowitz KB, Boyd DB (eds) *Reviews in computational chemistry*. VCH, New York, p. 273
- Tomasi J, Alagona G, Bonaccorsi R, Ghio C, Cammi R (1991) In: Maksić ZB (ed) *Theoretical models in chemical bonding, Part IV*. Springer, Berlin Heidelberg New York, p. 229
- Kollman P, McKelvey J, Johansson A, Rothenberg S (1975) *J Am Chem Soc* 97:955
- Murray JS, Zilles BA, Jayaruriya K, Politzer P (1986) *J Am Chem Soc* 108:915
- Orozco M, Luque FJ (1993) *J Comp Chem* 14:587
- Náray-Szabó G, Ferenczy GG (1995) *Chem Rev* 95:829
- Orozco M, Luque FJ (1996) In: Murray JS, Sen K (eds) *Molecular electrostatic potentials: concepts and applications, theoretical and computational chemistry, vol 3*, Elsevier, Amsterdam, pp 181–210
- Tomasi J, Menucci B, Cammi R (1996) In: Murray JS, Sen K (eds) *Molecular electrostatic potentials: concepts and applications, theoretical and computational chemistry, vol 3*, Elsevier, Amsterdam, pp 1–85
- Richards NGJ, Price SL (1989) *Int J Quantum Chem, Quantum Biol Symp* 16:73
- Burt C, Richards WG, Huxley P (1990) *J Comp Chem* 11:1139
- Richard AM (1991) *J Comp Chem* 12:959
- Rodríguez J, Manaut F, Sanz F (1993) *J Comp Chem* 14:922
- Petke JD (1993) *J Comp Chem* 14:928
- Besalú E, Carbó R, Mestres J, Solà M (1995) In: Sen K (ed) *Topics in current chemistry: molecular similarity, vol 173*. Springer, Berlin Heidelberg New York, p 31
- Sen KD, Politzer P (1989) *J Chem Phys* 90:4370
- Gadre SR, Kulkarni SA, Shrivastava IH (1992) *J Chem Phys* 96:5253
- Gadre SR, Kulkarni SA, Suresh CH, Shrivastava IH (1995) *Chem Phys Lett* 239:273
- Lecomte C, Ghermani N, Pichon-Pesme V, Souhasson M (1992) *J Mol Struct (THEOCHEM)*, 255:241
- White JC, Hess AC (1993) *J Phys Chem* 97:6398
- Stewart RF, Craven BM (1993) *Biophys J* 65:998
- Gilson MK, Honig B (1988) *Proteins* 4:7
- Zauhar RJ, Morgan RS (1988) *J Comp Chem* 9:171
- Politzer P, Lane P, Murray JS, Brinck T (1992) *J Phys Chem* 96:7938
- Tomasi J, Persico M (1994) *Chem Rev* 94:2027

29. Alhambra C, Luque FJ, Orozco M (1995) *J Phys Chem* 99:3084
30. Rivail JL, Rinaldi D (1996) In: Leszczynski J (ed) *Computational chemistry, review of current trends*. World Scientific, Singapore (in press)
31. Chirlian LE, Francl MM (1987) *J Comp Chem* 8:894
32. Williams DE (1988) *J Comp Chem* 9:745
33. Woods RJ, Khalil M, Pell W, Moffat SH, Smith VH (1990) *J Comp Chem* 11:297
34. Breneman CM, Wiberg KB (1990) *J Comp Chem* 11:361
35. Orozco M, Luque FJ (1990) *J Comp Chem* 11:909
36. Alemán C, Luque FJ, Orozco M (1993) *J Comput Aided Mol Des* 7:721
37. Cornell WD, Cieplak P, Bayly CI, Kollman PA (1991) *J Am Chem Soc* 113:5203
38. Gadre SR, Shrivastava IH, Kulkarni SA (1990) *Chem Phys Lett* 170:271
39. Gadre SR, Kulkarni SA, Pathak RK (1989) *J Chem Phys* 91:3596
40. Johnson BG, Gill PMW, Pople JA, Fox DJ (1993) *Chem Phys Lett* 206:239
41. Luque FJ, Illas F, Orozco M (1990) *J Comp Chem* 11:416
42. Ferenczy GG, Reynolds CA, Richards WG (1990) *J Comp Chem* 11:159
43. Besler BH, Merz KM Jr., Kollman PA (1990) *J Comp Chem* 11:431
44. Wang B, Ford GP (1994) *J Comp Chem* 15:200
45. Alhambra C, Luque FJ, Orozco M (1994) *J Comp Chem* 15:12
46. Luque FJ, Orozco M, Illas F, Rubio J (1991) *J Am Chem Soc* 113:5203
47. Petrongolo C (1978) *Gazz Chim Ital* 108:445
48. Culberson JC, Zerner MC (1985) *Chem Phys Lett* 122:436
49. Edwards WD, Weinstein H (1973) *Chem Phys Lett* 56:213
50. Mó O, Yañez M (1973) *Theor Chim Acta* 28:213
51. Ramos MJ, Webster B (1983) *J Chem Soc Faraday Trans 2*, 79:1389
52. Bebbster R, Hilczer M, Ramos MJ, Carmichael J (1985) *J Chem Soc Faraday Trans II* 81:1761
53. Soliva R, Orozco M, Luque FJ (1997) *J Comp Chem* 18:980
54. Winn PJ, Ferenczy GG, Reynolds CA *J Phys Chem* (in press)
55. Köstr AM, Leboeuf M, Salahub DR (1996) In: Murray JS, Sen K (eds) *Molecular electrostatic potentials: concepts and applications, theoretical and computational chemistry, vol 3*. Elsevier, Amsterdam, pp 105–137
56. Ditchfield R, Hehre WJ, Pople JA (1971) *J Chem Phys* 54:724
57. Hariharan PC, Pople JA (1973) *Theor Chim Acta* 28:213
58. Frisch MJ, Pople JA, Binkley JS (1984) *J Chem Phys* 80:3265
59. McLean AD, Chandler GS (1980) *J Chem Phys* 72:5639
60. Krishnan R, Binkley JS, Seeger R, Pople JA (1980) *J Chem Phys* 72:650
61. Clark T, Chandraskhar J, Spitznagel GW, Shleyer PvR (1983) *J Comp Chem* 4:294
62. Møller C, Plesset MS (1934) *Phys Rev* 46:618
63. Kohn W, Sham LJ (1965) *Phys Rev* 140:A1133
64. Vosko SH, Wilk L, Nusair M (1980) *Can J Phys* 58:1200
65. Perdew JP, Zunger A (1981) *Phys Rev B* 23:5048
66. Becke AD (1988) *Phys Rev A* 38:3098
67. Perdew JP (1986) *Phys Rev B* 33:8822
68. Becke AD (1993) *J Chem Phys* 98:5648
69. Stevens PJ, Devlin FJ, Chablowski CF, Frisch MJ (1994) *J Phys Chem* 98:11623
70. Perdew JP, Wang Y (1992) *Phys Rev B* 45:13244
71. Huron B, Malrieu JP, Rancurel P (1973) *J Chem Phys* 58:5745
72. Connolly M (1981) *QCPE Bull* 1:75
73. Momany FA (1978) *J Phys Chem* 82:592
74. Orozco M, Luque FJ (1990) *J Comput Aided Mol Des* 4:411
75. Gaussian 94 (Rev. A.1) (1995) Frisch MJ, Trucks GW, Schlegel HB, Gill PMW, Johnson BG, Robb MA, Cheeseman JR, Keith TA, Petersson GA, Montgomery JA, Raghavachari K, Al-Laham MA, Zakrzewski VG, Ortiz JV, Foresman JB, Cioslowski J, Stefanov BB, Nanayakkara A, Challacombe M, Peng CY, Ayala PY, Chen W, Wong MW, Andres JL, Replogle ES, Gomperts R, Martín RL, Fox DJ, Binkley JS, Defrees DJ, Baker J, Stewart JP, Head-Gordon M, Gonzalez C, Pople JA. Gaussian, Pittsburgh, Pa
76. Luque FJ, Orozco M (1995) Unpublished version of MO-PETE/MOPFIT program. University of Barcelona