

FULL PAPER

New Multicentre Point Charge Models for Molecular Electrostatic Potentials from Semiempirical MO-Calculations

Peter Gedeck^{1,†}, Torsten Schindler¹, Alexander Alex², and Timothy Clark¹

¹Computer-Chemie-Centrum, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nögelsbachstraße 25, D-91052 Erlangen, Germany. E-mail: clark@chemie.uni-erlangen.de

²Central Research, Pfizer Limited, Sandwich, Kent CT13 9NJ, United Kingdom.

Received: 15 March 1999/ Accepted: 27 January 2000/ Published: 16 June 2000

Abstract Two quasi-multipole electrostatic models for molecular charge distributions are presented. They assign arrays of point charges to nonhydrogen atoms on the basis of hybrid orbitals or localised molecular orbitals. When used with common semiempirical MO-techniques, they reproduce natural atomic orbital derived point charge (NAO-PC) and *ab initio* molecular potentials well. The localised orbital technique (LMO-PC) is intuitively more attractive than the hybrid orbital-point charge (HO-PC) method, although the former is more CPU-intensive.

Keywords Molecular electrostatic potential, AM1, PM3, MNDO, Point charge model

Introduction

The molecular electrostatic potential (MEP) plays an important role in areas such as intermolecular interactions and chemical reactivity,[1] biological interactions,[2] solvation phenomena,[3] crystal packing,[4] and electron density studies.[5, 6] Therefore, it is of great interest to describe the molecular electrostatic potential as well as possible within

computationally efficient methods. It has recently become increasingly clear that atomic monopole models cannot describe the subtle details of the MEP in binding regions around the molecule adequately. We have therefore investigated a number of atomic multipole models with two goals in mind. The first is to be able to describe the MEP in the chemically interesting region around molecules as accurately and efficiently as possible. The second aim, however, is to investigate possible approaches that allow us to develop non-quantum mechanical quasi-multipole charge models that will allow us to describe the MEP around macromolecules accurately.

The most rigorous way to calculate the MEP of a molecule is to use the charge density derived from a quantum mechanical calculation. The MEP at a point r in space is calculated using equation (1).

$$V(r) = \sum_{i=1}^n \frac{Z_i}{|r - R_i|} - \int \frac{\rho(r')}{|r - r'|} dr' \quad (1)$$

Correspondence to: T. Clark

[†]Current address: Novartis Horsham Research Centre, Wimblehurst Road, Horsham, West Sussex, RH12 5AB, United Kingdom.

The authors like to thank Dr. Manfred Sippl for acting as guest editor

The sum is the Coulomb potential caused by the n nuclei of the molecule. The coordinates of the nuclei i are R_i , their charge Z_i . The integral gives the Coulomb potential caused by the electron charge density $\rho(r)$.

In order to calculate high quality MEPs from *ab initio* calculations, it is necessary to use large basis sets including at least polarisation functions. This has the disadvantage that for large molecules the calculation of the MEP is quite time consuming. A much faster approach is to use the charge density derived from semiempirical MO calculations. A comparison of the MEP derived from semiempirical MO calculations with those calculated at the *ab initio* HF/6-31G(d) and MP2/6-31G(d) levels showed that the semiempirical methods provide a good description of the MEP in the whole space surrounding the molecule [7–11].

An alternative approach to using the charge density lies in representing the charge distribution by a number of point charges. In this case equation (1) simplifies to a sum over n point charges.

$$V(r) = \sum_{i=1}^n \frac{q_i}{|r - R_i|} \quad (2)$$

where R_i is the coordinate of the i -th point charge and q_i its charge.

An often used ansatz are atomic monopole models. Here, the point charges are located at atomic centres (so called net atomic charges). It is common practice to determine the net atomic charges in a linear least-squares fitting procedure in order to represent *ab initio* or semiempirical potentials calculated with equation (1) at a small number of points around the molecule.[12–15] These electrostatic potential (ESP) or MEP-derived charges have become very popular for force-field applications, but it has become clear that atom-centred monopoles are not sufficient for reproducing the MEP to a high degree of accuracy.[16]

However, the number of point charges must not necessarily be equal to the number of atoms in the molecule. In order to describe the MEP better using point charges, it is necessary to use more point charges than atoms to represent the charge distribution of the molecule in a quasi-multipole approach. The NAO-PC model (natural atomic orbital derived point charges) [17,18] is one possible approach to this problem. Instead of using only one, a set of nine point charges is used to represent the charge distribution around an atom. The positions and charges of these point charges are determined from the natural atomic orbitals (NAO). We have shown [17,18] that this model provides a more accurate description of the MEP than atom-centred monopoles. However, the NAO-PC model fails for molecules with degenerate point groups or with NAOs of degenerate occupancy. To overcome this problem we have developed two modifications.

The first, HO-PC, is similar to the NAO-PC model. However, instead of using NAOs, we use a hybrid orbital (HO) basis created in such a way as to reflect the topology of the atom. The second model, LMO-PC, is based on localised molecular orbitals. Instead of using the density matrix as in

the other two models, we start with the occupied molecular orbitals and generate localised molecular orbitals (LMOs). Based on the LMOs we can create a set of point charges.

As will be shown in this paper, both models preserve the molecular symmetry. In addition, the two models reflect the chemist's view of lone-pairs, σ - and π -bonds.

A further intention in developing these models is a later use in a fragment-based approach for the assignment of charges that can be used in a force field program to give a better description of conformationally independent point charges.

All models can be used with the semiempirical methods MNDO, [19] AM1, [20] and PM3 [21] and have been implemented in the program package VAMP.[22]

Theory

Representing a hybrid orbital by two point charges

Before we describe the two models, it is necessary to explain the concept used to assign pairs of point charges to hybrid orbitals. As in the NAO-PC model, the point charges used to calculate the 'amount of charge' density of the lobes of hybrid orbitals. A typical sp^n hybrid orbital is shown in Figure 1. It is possible by numerical integration to calculate the charge located in the positive and the negative lobes of the sp^n hybrid orbital using equations (3).

$$\begin{aligned} q^+ &= \int |\psi_{sp}^+(r)|^2 dr \\ q^- &= \int |\psi_{sp}^-(r)|^2 dr \end{aligned} \quad (3)$$

Here, $\Psi_{sp}(r)$ is the wave function of the sp^n hybrid orbital determined by the semiempirical parameters. The superscripts

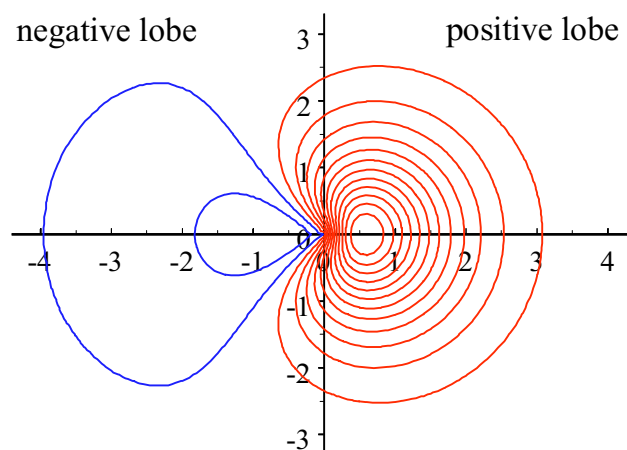


Figure 1 Electron density plot of a carbon sp^3 hybrid orbital using the standard AM1 parameters

+ and - denote the larger (+) and the smaller (-) lobes of the hybrid orbital, respectively. In a similar way the distances of the centres of charge density from the nucleus for each lobe are given by.

$$r^+ = \frac{1}{q^+} \int r |\psi_{sp}^+(r)|^2 dr$$

$$r^- = \frac{1}{q^-} \int r |\psi_{sp}^-(r)|^2 dr \quad (4)$$

For all atoms for which semiempirical parameters are available, values of q^+ , q^- , r^+ and r^- have been precalculated for a large number of degrees of hybridisation, stored and used for lookup using interpolation.

Calculation of distributed point charges based on hybrid orbitals (HO-PCs)

The calculation of hybrid orbital-based point charges (HO-PCs) is described schematically in Figure 2. To describe the MEP caused by the nuclei, we place point charges equal to the core charges at the position of each atom.

In semiempirical calculations, hydrogens are represented by a single s -orbital. Therefore, we add the part of the electron charge density located on the atom to the point charge already assigned to the nuclei. The resulting charge is actually the Coulson charge of the hydrogen atoms.

In MNDO, AM1 and PM3 calculations the atomic orbital basis for heavy atoms consists of one s - and three p -orbitals. Therefore, the electron charge distribution around an atom is in general anisotropic. To represent this asymmetry, we approximate the electron charge density of each heavy atom by eight point charges. The locations and charges of the eight point charges are determined using the following procedure.

The heavy atom and its surrounding n neighbours are treated as a XH_n . For this small system an extended-Hückel type calculation [23] is carried out. The heavy atom is described by one s - and three p -orbitals. The diagonal matrix elements of the Hückel matrix are $\alpha_X^s = \langle s_X | h | s_X \rangle$, $\alpha_X^p = \langle p_i | h | p_i \rangle$ and $\alpha_H^s = \langle s_H | h | s_H \rangle$. In the Hückel-matrix only interactions between X and H atoms are included (nearest-neighbour approximation). The matrix elements $\beta^{ss} = \langle s_X | h | s_H \rangle$ are all the same, whereas the $\beta_i^{p_x s} = \langle p_{x_X} | h | s_H \rangle$ represent the steric orientation of the H with respect to the X. We used the following expression for this

$$\vec{\beta}_i^{ps} = \begin{pmatrix} \beta_i^{p_x s} \\ \beta_i^{p_y s} \\ \beta_i^{p_z s} \end{pmatrix} = \beta^{ps} \vec{r}_{XH_i}^0 \quad (5)$$

Again, β^{ps} is constant and equal for all H atoms. The vector $\vec{r}_{XH_i}^0$ is the unit vector from X to the i -th H-atom.

To summarise, the extended-Hückel matrix used is of the following form.

$$H = \begin{pmatrix} \alpha_X^s & 0 & 0 & 0 & \beta^{ss} & \beta^{ss} & \beta^{ss} & \beta^{ss} \\ 0 & \alpha_X^p & 0 & 0 & \beta^{p_x s} & \beta^{p_x s} & \beta^{p_x s} & \beta^{p_x s} \\ 0 & 0 & \alpha_X^p & 0 & \beta^{p_y s} & \beta^{p_y s} & \beta^{p_y s} & \beta^{p_y s} \\ 0 & 0 & 0 & \alpha_X^p & \beta^{p_z s} & \beta^{p_z s} & \beta^{p_z s} & \beta^{p_z s} \\ \beta^{ss} & \beta^{p_x s} & \beta_i^{p_y s} & \beta^{p_z s} & \alpha_H^s & 0 & 0 & 0 \\ \beta^{ss} & \beta^{p_x s} & \beta_2^{p_y s} & \beta^{p_z s} & 0 & \alpha_H^s & 0 & 0 \\ \beta^{ss} & \beta^{p_x s} & \beta_3^{p_y s} & \beta^{p_z s} & 0 & 0 & \alpha_H^s & 0 \\ \beta^{ss} & \beta^{p_x s} & \beta_4^{p_y s} & \beta^{p_z s} & 0 & 0 & 0 & \alpha_H^s \end{pmatrix} \quad (6)$$

The absolute values of the constants α_X^s , α_X^p , α_H^s , β^{ss} and β^{sp} are not important for our purpose and were chosen so that reasonable molecular orbitals were obtained after diagonalisation of the extended-Hückel matrix. In the next step, the resulting molecular orbitals are localised and from the localised molecular orbitals hybrid orbitals for the atom X obtained.

Using the method described in the previous section, we can then approximate each hybrid orbital by two point charges. The amount of charge located in each of the hybrid orbitals is determined by the diagonal elements of the one-atom density matrix expressed in the basis of the four hybrid orbitals.

In total we obtain one point charge for each hydrogen and nine point charges for the remaining heavy atoms. This is the same number of point charges as in the NAO-PC model.

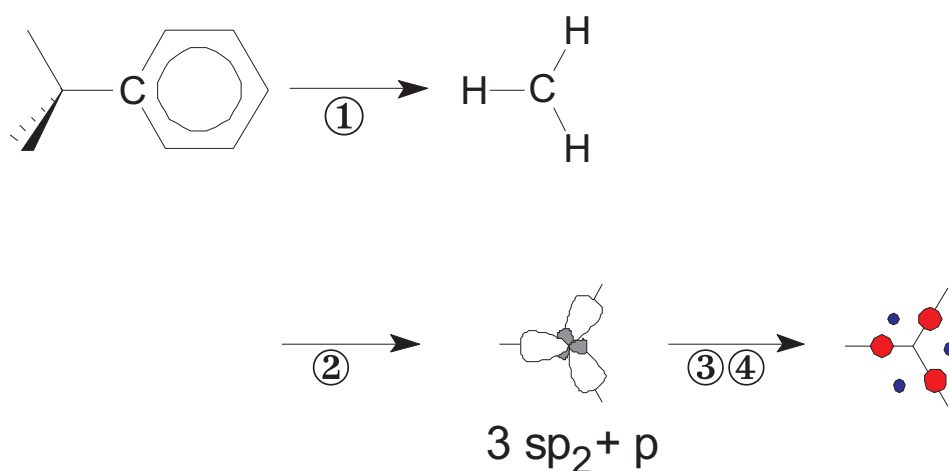
Calculation of point charges based on localised molecular orbitals (LMO-PC)

The second distributed point charge representation of the molecular charge density is based on localised molecular orbitals (LMOs). The idea behind the method is described schematically in Figure 3.

To represent the core charge, we assign point charges to the nuclei. Initially, the charge of these atom-centred point charges is the core charge of the corresponding atom. The electron density is represented by a set of point charges derived from the LMOs.

The occupied molecular orbitals, which are in general delocalised over a considerable part of the molecule, are localised according to Perkins and Stewart.[24] In general, an LMO describing a lone-pair is localised on one atom; a σ -

Figure 2 Scheme describing the determination of HO-PC. 1) Determination of molecular topology. 2) Calculation of hybrid-orbitals using a pseudo-HMO calculation. 3) Transformation of the one-atom block of the density matrix into the hybrid orbital basis. 4) Approximation of the charge density by point charges



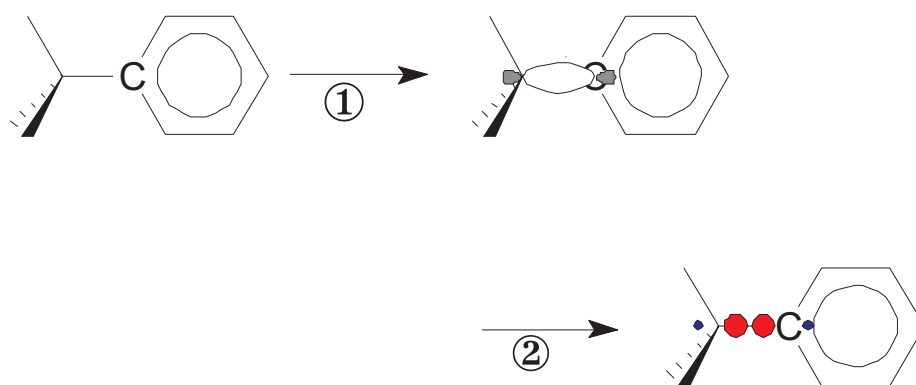
orbital is described by an LMO localised on two atoms. Only π -orbitals are sometimes localised on more than two atoms to a considerable amount. In all other cases, the MO coefficients of the other atoms are only small.

A lone pair LMO, for example, is in general localised to 98% on one atom. This also means that almost all the charge of the two electrons in the LMO is located at this atom. In addition, we know the hybridisation of the LMO from the coefficients of the s , p_x , p_y , and p_z atomic orbitals. For hydrogens, we only have an s -atomic orbital; so if the LMO is localised on a hydrogen, we just add the amount of charge on this atom to the core point charge that is already assigned to the nucleus. In the other cases, we can use the method described above to assign two point charges for the sp^x hybrid defined by the LMO coefficients of this atom. The total charge of the two point charges is determined by the degree

of localisation of the LMO on the atom and the distribution between the two point charges by the hybridisation of the sp^x hybrid. In the same way, we create pairs of point charges for all atoms where the LMO is localised to a considerable degree. As outlined above, we have in general no more than three atoms (one for lone pairs, two for σ -LMOs, and two or three for π -LMOs, see Figure 4 for an example of a σ -LMO of the C-H bond in methane).

Of course the same approach could be used to assign point charges for the LMO to all other atoms in the molecule. However, this would cause an unreasonably large number of point charges, most of them very small. To reduce the number, we add in these cases the charge located at an atom to the already assigned atom-centred point charge. As threshold a negative charge of 0.07 a.u. has been found to be appropriate. We found that neglecting the anisotropy of small point

Figure 3 Scheme describing the determination of LMO-PC. 1) Determination of localised molecular orbitals. 2) Approximation of the localised molecular orbitals by point charges



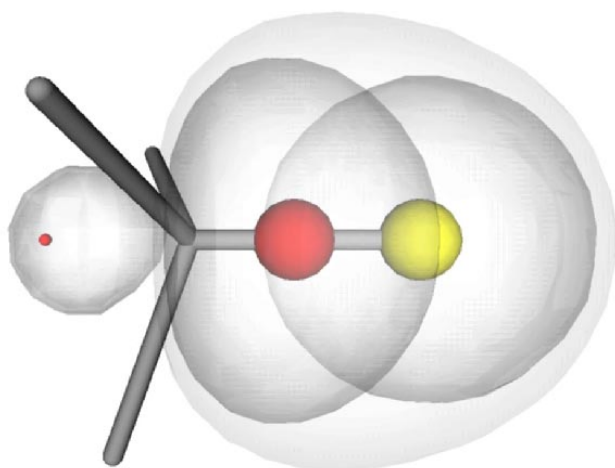


Figure 4 Approximation of the localised molecular orbital of a C-H bond by three point charges. The charge of the point charges is indicated by the size of the spheres. The red (dark) spheres approximate the carbon sp^x hybrid orbital, the yellow (light) sphere the hydrogen s-orbital. The shape of the atomic hybrid orbitals (sp^3 for carbon and s for hydrogen) as well as the resulting LMO is shown as an isosurface

charges does not affect the quality of the results. It is sufficient to add the charge to the core point charge.

The number of point charges created can be reduced even further for aromatic systems. In this case, similar pairs of point charges are created for the atoms of the aromatic system for the π -LMOs. An atom usually contributes to more than one π -LMO. However, because the point charges are located at almost the same positions, we can combine them into one pair of point charges.

Using the simplifications described above, we usually end up with about the same number of point charges as with the NAO-PC and the HO-PC methods.

Calculation of the MEP from the HO-PCs and the LMO-PCs

Using the HO-PCs or LMO-PCs the molecular electrostatic potential can now be calculated in a straightforward way using the usual equation for a set of n point charges.

$$V(r) = \sum_{i=1}^n \frac{q_i}{|r - R_i|} \quad (7)$$

Although the time required to calculate the molecular electrostatic potential from the HO-PCs or the LMO-PCs is typically five to six times higher than that using atom-centred point charges, the calculation is much faster than calculating the MEP directly from the density matrix. [17,18]

Results

Qualitative comparison of the HO-PC and the LMO-PC model

In general, the distribution and the charges of the LMO-PCs are quite similar to the HO-PCs. This can be seen quite well for the example imidazole in Figure 5. Nine point charges, one per atom, are located at the nuclei. Both methods create a pair of large point charges along each C-C and C-N bond and a corresponding pair of smaller point charges outside the C-C and C-N bonds. For the C-H and N-H bonds we have a pair of point charges for each heavy atom, the larger of the two point charges is located on the side of the hydrogen. The nitrogen lone-pair is described by a large point charge outside and a very small corresponding point charge inside the five-membered ring. All these point charges are located in the ring plane. The remaining ten point charges represent the π -system and are located symmetrically above and below the ring plane.

The two approaches are clearly very similar. However, they each have their individual advantages, so that we will consider both methods further.

The LMO-PC model is far more appealing from a theoretical point of view. Because of the way the point charges are created, they are a natural representation of the electron charge density of the molecule. In contrast, the NAO-PC and HO-PC charges use only the one-atom blocks of the density matrix and neglect the two-centre elements. However the results show that this factor is not important.

On the other hand, the LMO-PC model requires the calculation of localised molecular orbitals - a quite time-consuming step for large molecules. Not only is the time required to calculate the LMO-PCs considerably longer than for NAO-PCs or HO-PCs. But it also scales with $O(n^3)$ (n is

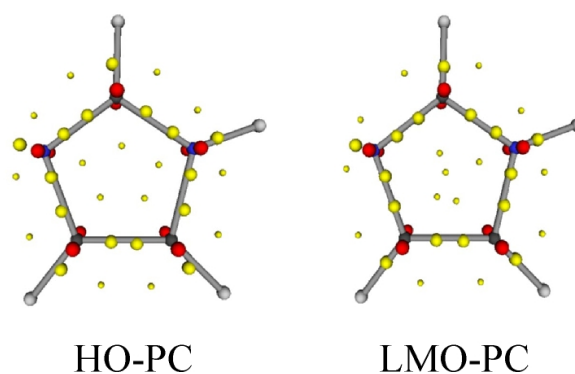


Figure 5 Comparison of point charges created by the HO-PC and LMO-PC methods for imidazole. The point charges representing the sp^2 bond orbitals and lone pairs are shown in yellow (light). The point charges approximating the p-orbital are given in red (dark). Point charges assigned to the nuclei of the atoms are not shown

the number of atoms), whereas in the other cases it is only $O(n)$.

The fact that the position of the HO-PCs is only determined by the bond topology and the relative geometrical orientation of the atoms suggests that HO-PCs can be a starting point for the development of a conformationally invariant quasi-multipole point charge model.

Qualitative comparison of the NAO-PC and the HO-PC/LMO-PC model

One of the limitations of the NAO-PC model is its inability to reproduce the molecular symmetry in the point charges created. This is due to the fact that degenerate NAOs can be represented in an infinite number of ways. However, in most cases the symmetry of the resulting molecular electrostatic potential deviates only slightly from the molecular symmetry.

One of the rare cases where the NAO-PC model gives a completely false MEP is C_{60} . Figure 6 shows the MEP of the fullerene C_{60} calculated from NAO-PCs and HO-PCs at the van der Waals surface of the molecule. The LMO-PCs give a MEP similar to HO-PCs and are therefore not shown here. Because of the spherical structure of C_{60} the sp^2 hybridised carbon atoms are bent slightly. This causes electron density to be shifted to the outside of the cage. In consequence, the outside of C_{60} should be negatively, the inside positively

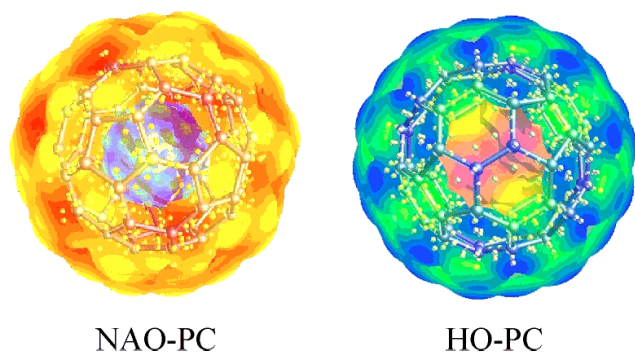
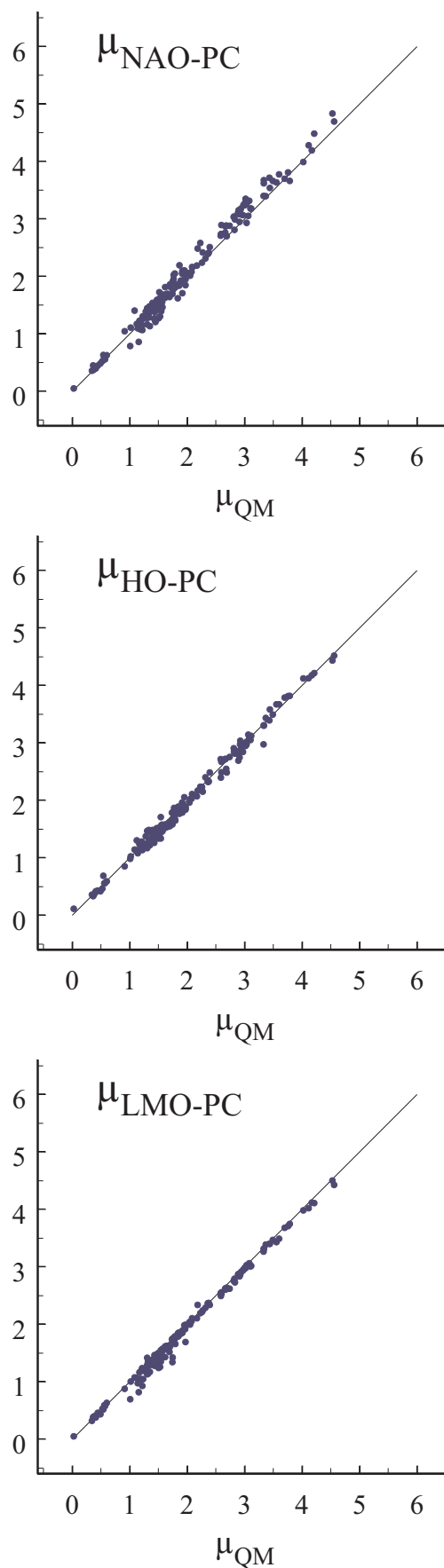


Figure 6 Comparison of point charges created by the NAO-PC and HO-PC methods for C_{60} . The point charges are shown as yellow spheres. The molecular electrostatic potential on the van der Waals surface is shown in colour coding. Red areas indicate a positive molecular electrostatic potential, blue areas a negative one. The molecular electrostatic potential on the outside should be negative and on the inside positive. The MEP calculated from NAO-PC is wrong

Figure 7 (right column) Comparison of dipole moments calculated using the methods NAO-PC (top panel), HO-PC (middle panel), and LMO-PC (bottom panel) with the dipole moments calculated using quantum mechanics. Results for AM1 calculations. For details on the studied compounds see text



charged. As can be seen in Figure 6 this is the case for the HO-PC charges. In contrast, the NAO-PCs give the opposite result and are thus wrong. We must emphasise that in general NAO-PC charges give correct MEPs and that C₆₀ is an extreme case. The MEP calculated from the HO-PCs has already been used successfully to study complexation of fullerenes. [25]

Reproduction of multipole moments by the point charge models

In order to determine how well the semiempirical charge density is represented by point charges created by the different methods, we compared the dipole moments calculated from the point charges with the usual Hamiltonian-based techniques.[26] The results of this comparison are shown in Figure 7.

The test set contained a large variety of organic compounds. The compounds were selected in order to contain all common functional groups. These comprise cyclic and acyclic, aromatic and nonaromatic compounds containing oxygen (e. g. alcohols, ethers, carboxylic acids, esters, ketones and aldehydes), nitrogen (e. g. amines, imines, cyanides, amides), and halogens (fluorine, chlorine, bromine and iodine). In addition, compounds containing sulphur and silicon were added.

CORINA [27,28] was used to generate starting geometries automatically from SMILES strings.[29] The geometries were then optimised and the dipole moment calculated by the usual Hamiltonian-based technique and from the point charges generated by the different methods. In Figure 7 the dipole moment calculated classically from the distributed point charges is plotted against the quantum mechanically calculated dipole moments.

In general, all three methods reproduce the quantum mechanically calculated dipole moment very well. The correlation coefficients are $r^2 = 0.985$ for NAO-PC, $r^2 = 0.992$ for HO-PC, and $r^2 = 0.993$ for LMO-PC. The excellent agreement is also reflected in the standard deviation of the difference between the quantum mechanically and the classically calculated dipole moments. It is largest for NAO-PC (0.132), and smaller for HO-PC (0.084) and LMO-PC (0.078). The result for LMO-PC would be even better if chlorine containing compounds were excluded from the calculation. Then the standard deviation would be only 0.047. In contrast, the dipole moments calculated for the NAO-PC and HO-PC point charges are better for chlorine containing compounds.

Additionally for a test set of 16 molecules the centre of mass molecular quadrupole moment was calculated using the Buckingham definition [30]:

$$Q_{kl} = \frac{1}{2} \left\{ \sum_{\alpha} Z_{\alpha} (3c_{k,\alpha}c_{l,\alpha} - R_{\alpha}^2\delta_{kl}) + \left\langle \Psi_{el} \left| \sum_i (3c_{k,i}c_{l,i} - r_i^2\delta_{kl}) \right| \Psi_{el} \right\rangle \right\} \quad (8)$$

where δ_{kl} is the Kronecker-delta and c_k the Cartesian coordinates of the atoms or electrons. In the point charge approximation the electronic contribution simplifies to:

$$\left\langle \Psi_{el} \left| \sum_i (3c_{k,i}c_{l,i} - r_i^2\delta_{kl}) \right| \Psi_{el} \right\rangle \approx \sum_{\alpha(\alpha \in H)} \sum_{j=1}^{2n_{\alpha}} q_j (3c_{k,j}c_{l,j} - r_j^2\delta_{kl}) + \sum_{\alpha(\alpha \in H)} q_{\alpha} (3c_{k,\alpha} - R_{\alpha}^2\delta_{kl}) \quad (9)$$

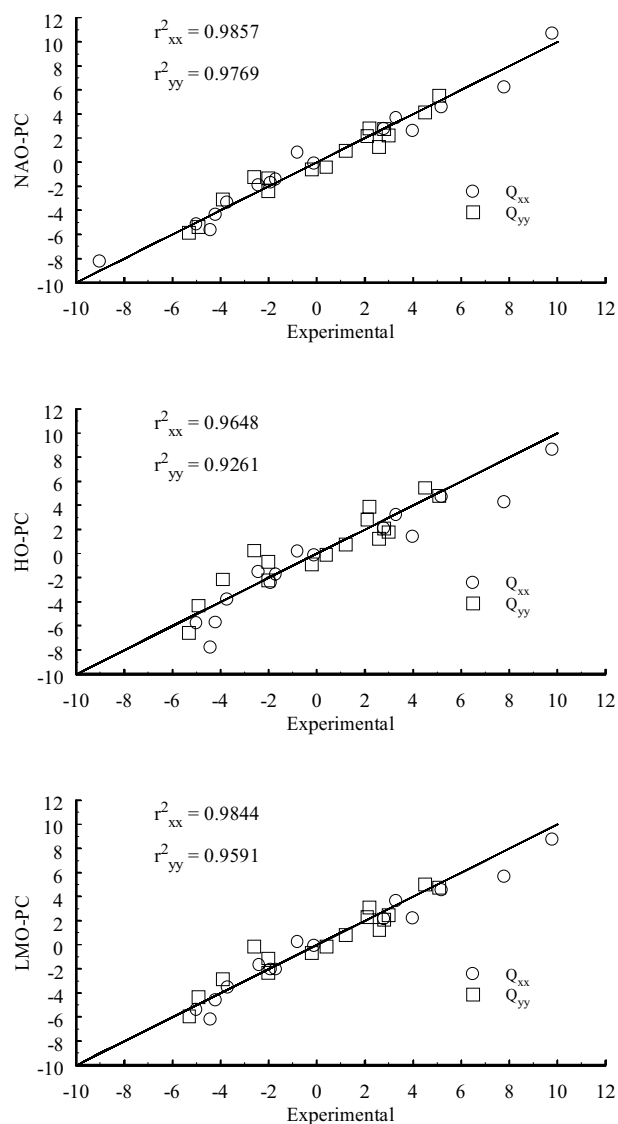


Figure 8 Correlation diagrams for the xx- and yy-elements of the quadrupole tensor of the different point charge models to experimental data in units of (Debye-Angstrom). Calculated results are from AM1 calculations and experimental data are taken from the work of Flygare and his coworkers [37]

Table 1 *Quadrupole moments (in Debye*Angstrom units) of the different point charge models compared to experimental data*

	Molecule	Point charge models			Experimental data		
		Q _{xx}	Q _{yy}	Q _{zz}	Q _{xx}	Q _{yy}	Q _{zz}
Cyanogenfluoride	NAO-PC	-4.37	2.18	2.18	-4.2±??.?	2.1±??.?	2.1±??.?
	HO-PC	-5.71	2.85	2.85			
	LMO-PC	-4.62	2.31	2.31			
Fluoroacetylene	NAO-PC	2.59	-1.30	-1.30	4.0±0.2	-2.0±0.2	-2.0±0.2
	HO-PC	1.39	-0.70	-0.70			
	LMO-PC	2.23	-1.11	-1.11			
Water	NAO-PC	-0.08	1.28	-1.20	-0.1±0.1	2.6±0.1	-2.5±0.1
	HO-PC	-0.16	1.23	-1.07			
	LMO-PC	-0.06	1.26	-1.20			
cis-1,2-Difluoroethylene	NAO-PC	-1.46	2.22	-0.76	-1.7±0.4	3.0±0.3	-1.3±0.5
	HO-PC	-1.74	1.80	-0.06			
	LMO-PC	-2.01	2.47	-0.46			
Benzene	NAO-PC	2.78	2.78	-5.55	2.8±1.4	2.8±1.4	-5.6±2.8
	HO-PC	2.07	2.07	-4.15			
	LMO-PC	2.12	2.12	-4.24			
Difluoroformaldehyde	NAO-PC	-3.34	-0.57	3.92	-3.7±0.7	-0.2±0.5	3.9±1.1
	HO-PC	-3.81	-0.90	4.71			
	LMO-PC	-3.56	-0.68	4.24			
Ammonia	NAO-PC	-1.90	0.95	0.95	-2.4±0.1	1.2±0.1	1.2±0.1
	HO-PC	-1.53	0.76	0.76			
	LMO-PC	-1.67	0.83	0.83			
Chloroactylene	NAO-PC	6.22	-3.11	-3.11	7.8±??.?	-3.9±??.?	-3.9±??.?
	HO-PC	4.29	-2.14	-2.14			
	LMO-PC	5.67	-2.83	-2.83			
Ethane	NAO-PC	0.81	-0.41	-0.41	-0.8±0.1	0.4±0.1	0.4±0.1
	HO-PC	0.20	-0.10	-0.10			
	LMO-PC	0.25	-0.13	-0.13			
Formic acid	NAO-PC	4.57	-5.83	1.26	5.2±0.4	-5.3±0.4	0.1±0.4
	HO-PC	4.72	-6.55	1.84			
	LMO-PC	4.53	-5.94	1.40			
Fluorobenzene	NAO-PC	-1.68	5.52	-3.84	-1.9±0.8	5.1±1.0	-3.2±1.0
	HO-PC	-2.44	4.79	-2.35			
	LMO-PC	-2.03	4.74	-2.71			
Dimethyl ether	NAO-PC	3.70	-2.40	-1.30	3.3±0.6	-2.0±0.5	-1.3±1.0
	HO-PC	3.22	-2.20	-1.02			
	LMO-PC	3.65	-2.33	-1.32			
Dicyanogen	NAO-PC	-8.25	4.12	4.12	-9.0±??.?	4.5±??.?	4.5±??.?
	HO-PC	-10.92	5.46	5.46			
	LMO-PC	-10.05	5.03	5.03			
Carbon dioxide	NAO-PC	-5.62	2.81	2.81	-4.4±0.2	2.2±0.2	2.2±0.2
	HO-PC	-7.81	3.90	3.90			
	LMO-PC	-6.19	3.10	3.10			
1,3-Pentadiine	NAO-PC	10.71	-5.35	-5.35	9.8±0.8	-4.9±0.8	-4.9±0.8
	HO-PC	8.62	-4.31	-4.31			
	LMO-PC	8.72	-4.36	-4.36			
1,3-Difluorobenzene	NAO-PC	-5.15	-1.24	6.39	-5.0±0.9	-2.6±1.3	7.6±1.0
	HO-PC	-5.78	0.27	5.51			
	LMO-PC	-5.41	-0.16	5.57			

Questionmark designates a experimentally undetermined value

Table 2 Statistical data of the MEP calculated on surface points using different methods. All values are in kcal·mol⁻¹


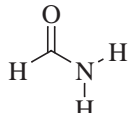
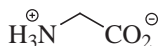
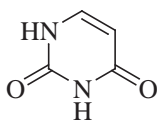
		Molecule					
		1	2	3	4	5	6
Surface points		856	1105	1570	1930	2002	2218
Maximum	MP2	34	54	99	71	70	98
	RHF 6-31G**	40	58	101	76	72	102
	NAO-PC	25	36	105	55	43	101
	HO-PC	29	36	112	55	47	109
	LMO-PC	26	39	109	54	47	105
	Coulson	26	31	98	43	34	97
Minimum	MP2	-38	-50	-91	-46	-72	-87
	RHF 6-31G**	-45	-58	-100	-53	-81	-97
	NAO-PC	-45	-55	-123	-51	-86	-117
	HO-PC	-54	-58	-134	-62	-88	-129
	LMO-PC	-43	-57	-117	-53	-88	-113
	Coulson	-30	-41	-107	-36	-55	-104
	MP2	24.302	30.718	63.050	30.179	35.328	49.966
	RHF 6-31G**	29.395	34.560	67.787	34.372	39.757	53.856
Standard Deviation	NAO-PC	22.295	27.550	65.537	26.177	32.050	51.806
	HO-PC	23.707	24.851	69.713	27.679	32.917	55.513
	LMO-PC	22.019	27.538	66.558	26.822	33.001	52.758
	Coulson	18.509	23.278	60.127	22.494	26.686	47.591
		Molecule					
		7	8	9	10	11	12
Surface points		2794	2867	3013	3385	4458	4928
Maximum	MP2	139	54	69	53	71	-
	RHF 6-31G**	143	54	78	53	79	7
	NAO-PC	141	66	105	38	35	-12
	HO-PC	141	64	103	36	66	-6
	LMO-PC	139	69	112	36	64	-7
	Coulson	129	58	92	32	33	-26
Minimum	MP2	-96	-47	-39	-62	-47	-
	RHF 6-31G**	-105	-55	-48	-64	-53	-175
	NAO-PC	-112	-71	-58	-79	-59	-187
	HO-PC	-123	-82	-67	-74	-60	-193
	LMO-PC	-113	-73	-58	-78	-58	-187
	Coulson	-105	-61	-52	-46	-36	-180
	MP2	71.639	25.957	20.411	19.436	22.029	-
	RHF 6-31G**	75.635	28.369	23.947	20.445	25.490	42.393
Standard Deviation	NAO-PC	72.699	27.334	24.377	18.911	20.125	41.389
	HO-PC	73.919	28.370	24.846	17.838	21.357	42.073
	LMO-PC	73.073	27.845	24.542	17.774	20.819	41.548
	Coulson	65.852	24.435	23.055	15.866	16.643	37.482

Quadrupole moments show a strong dependency on the quality of the wave function,[31] so that even high level *ab initio* calculations may give considerable deviations; therefore it seems more reasonable to compare these values with experimental data.[32] Most molecules of the test set contain only elements of first and second periods, because of the lack of

experimental data for compounds with heavy elements. Figure 8 shows the correlation diagrams for the *xx*- and *yy*-elements of the quadrupole tensor.

Because the *zz*-element is linearly dependent on the other two diagonal elements, it is not shown. The correlation coefficients for the *xx*- and *yy*-components are $r_{xx}^2 = 0.986$ and

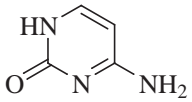
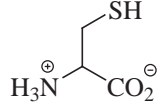
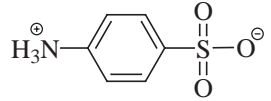
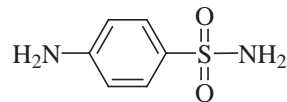
Table 3 (continues next page) *Hodgkin-Indices of the MEPs calculated from different methods*

a) Formaldehyde (1)							
	MP2	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
	1.000	1.000					
	RHF 6-31G**	0.980	1.000				
	Coulson	0.909	0.872	1.000			
	NAO-PC	0.960	0.931	0.932	1.000		
	HO-PC	0.950	0.928	0.897	0.990	1.000	
	LMO-PC	0.965	0.938	0.958	0.994	0.980	1.000
b) Formamide (2)							
	MP2	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
	1.000	1.000					
	RHF 6-31G**	0.990	1.000				
	Coulson	0.899	0.884	1.000			
	NAO-PC	0.958	0.951	0.961	1.000		
	HO-PC	0.929	0.908	0.940	0.980	1.000	
	LMO-PC	0.957	0.951	0.953	0.996	0.980	1.000
c) Glycine zwitterion (3)							
	MP2	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
	1.000	1.000					
	RHF 6-31G**	0.997	1.000				
	Coulson	0.979	0.972	1.000			
	NAO-PC	0.990	0.990	0.992	1.000		
	HO-PC	0.985	0.989	0.983	0.996	1.000	
	LMO-PC	0.990	0.991	0.989	0.999	0.998	1.000
d) Uracil (4)							
	MP2	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
	1.000	1.000					
	RHF 6-31G**	0.990	1.000				
	Coulson	0.854	0.826	1.000			
	NAO-PC	0.944	0.924	0.952	1.000		
	HO-PC	0.959	0.947	0.888	0.979	1.000	
	LMO-PC	0.960	0.945	0.908	0.986	0.997	1.000

$r_{yy}^2 = 0.977$ for NAO-PC, $r_{xx}^2 = 0.965$ and $r_{yy}^2 = 0.926$ for HO-PC and $r_{xx}^2 = 0.9844$ and $r_{yy}^2 = 0.959$ for LMO-PC. It can be seen that all three point charge models are able to reproduce higher multipole moments. The test set used is shown in Table 1. Calculated results are from AM1 calculations and

the experimental data are taken from the work of Flygare and coworkers.[32] Unfortunately experimental values of octopole and higher multipole moments are not available, so it was not possible to check the calculated molecule moments higher than third order quantitatively.

Table 3 (continued) *Hodgkin-Indices of the MEPs calculated from different methods*

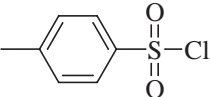
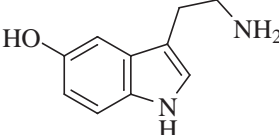
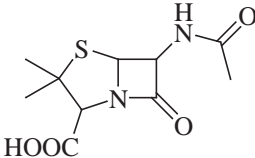
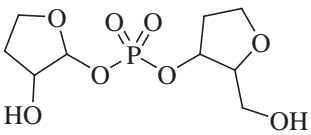
e) Cytosine (5)						
	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
MP2	1.000					
RHF 6-31G**	0.991	1.000				
Coulson	0.865	0.842	1.000			
NAO-PC	0.949	0.940	0.928	1.000		
HO-PC	0.957	0.951	0.891	0.991	1.000	
LMO-PC	0.957	0.952	0.898	0.994	0.999	1.000
f) Cysteine zwitterion (6)						
	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
MP2	1.000					
RHF 6-31G**	0.997	1.000				
Coulson	0.950	0.944	1.000			
NAO-PC	0.975	0.974	0.962	1.000		
HO-PC	0.974	0.978	0.960	0.991	1.000	
LMO-PC	0.977	0.978	0.959	0.998	0.994	1.000
g) Sulfanil zwitterion (7)						
	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
MP2	1.000					
RHF 6-31G**	0.998	1.000				
Coulson	0.974	0.971	1.000			
NAO-PC	0.986	0.988	0.988	1.000		
HO-PC	0.985	0.987	0.982	0.998	1.000	
LMO-PC	0.987	0.989	0.984	0.999	0.999	1.000
h) Sulfanilamide (8)						
	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
MP2	1.000					
RHF 6-31G**	0.993	1.000				
Coulson	0.781	0.786	1.000			
NAO-PC	0.897	0.909	0.923	1.000		
HO-PC	0.882	0.896	0.895	0.987	1.000	
LMO-PC	0.895	0.908	0.900	0.992	0.995	1.000

Comparison of the MEP calculated using different methods

To evaluate the quality of the molecular electrostatic potential resulting from the new distributed point charge models, we compared the MEPs calculated by different methods for a set of molecules. Beside the two charge models described

here, we looked at the MEP calculated from NAO-PC and Coulson charges and at the MEP obtained from *ab initio* calculations. The *ab initio* calculations used the program package Gaussian 94.[33] In all cases, we calculated the molecular electrostatic potential on a set of points on the solvent excluded surface [34] created by Vamp.

Table 3 (continued) Hodgkin-Indices of the MEPs calculated from different methods

i) Tosyl chloride (9)						
	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
	MP2	1.000				
	RHF 6-31G**	0.987	1.000			
	Coulson	0.758	0.765	1.000		
	NAO-PC	0.842	0.852	0.937	1.000	
	HO-PC	0.808	0.816	0.912	0.976	1.000
	LMO-PC	0.840	0.849	0.909	0.992	0.986
j) 5-Hydroxytryptamine (10)						
	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
	MP2	1.000				
	RHF 6-31G**	0.995	1.000			
	Coulson	0.660	0.690	1.000		
	NAO-PC	0.861	0.877	0.849	1.000	
	HO-PC	0.867	0.879	0.798	0.976	1.000
	LMO-PC	0.870	0.878	0.781	0.981	0.988
k) Penicillin G (11)						
	MP2	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC
	MP2	1.000				
	RHF 6-31G**	0.988	1.000			
	Coulson	0.667	0.659	1.000		
	NAO-PC	0.837	0.826	0.838	1.000	
	HO-PC	0.882	0.873	0.734	0.951	1.000
	LMO-PC	0.890	0.882	0.752	0.973	0.980
l) 3-1',2'-Desoxyribose (12)						
	RHF 6-31G**	Coulson	NAO-PC	HO-PC	LMO-PC	
	RHF 6-31G**	1.000				
	Coulson	0.987	1.000			
	NAO-PC	0.992	0.997	1.000		
	HO-PC	0.994	0.995	0.998	1.000	
LMO-PC	0.994	0.996	0.999	1.000	1.000	

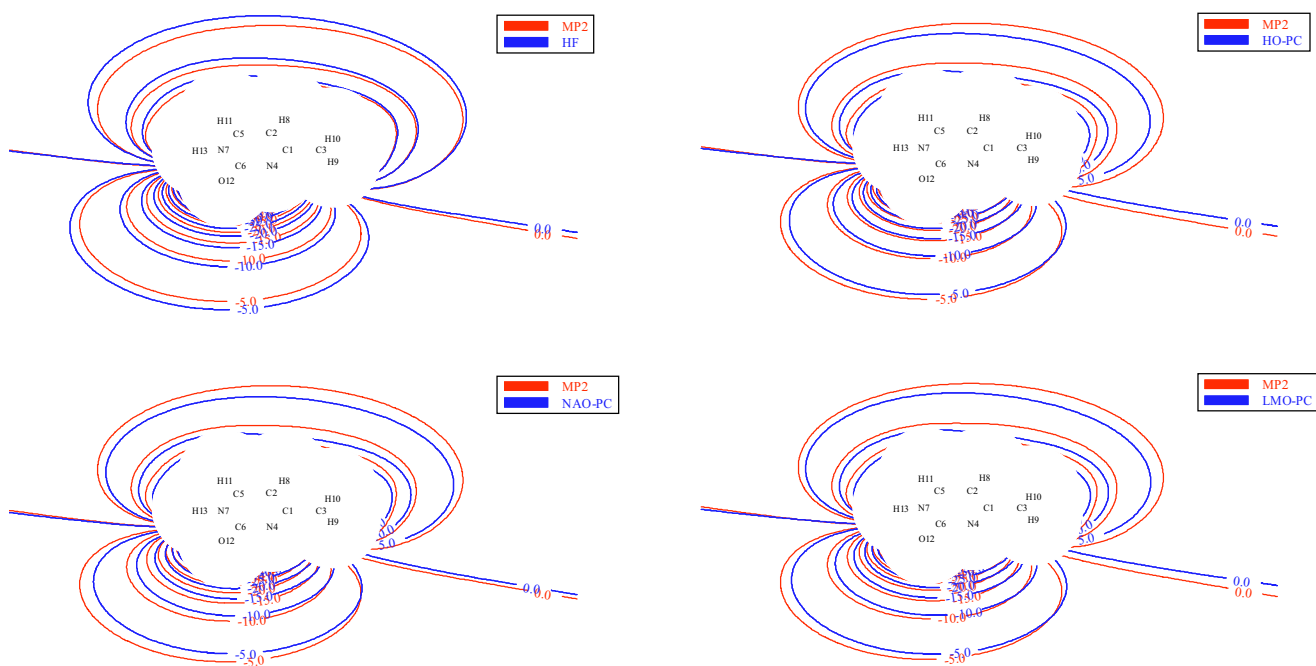


Figure 9 Isopotential maps in the plane of cytosine (kcal mol^{-1})

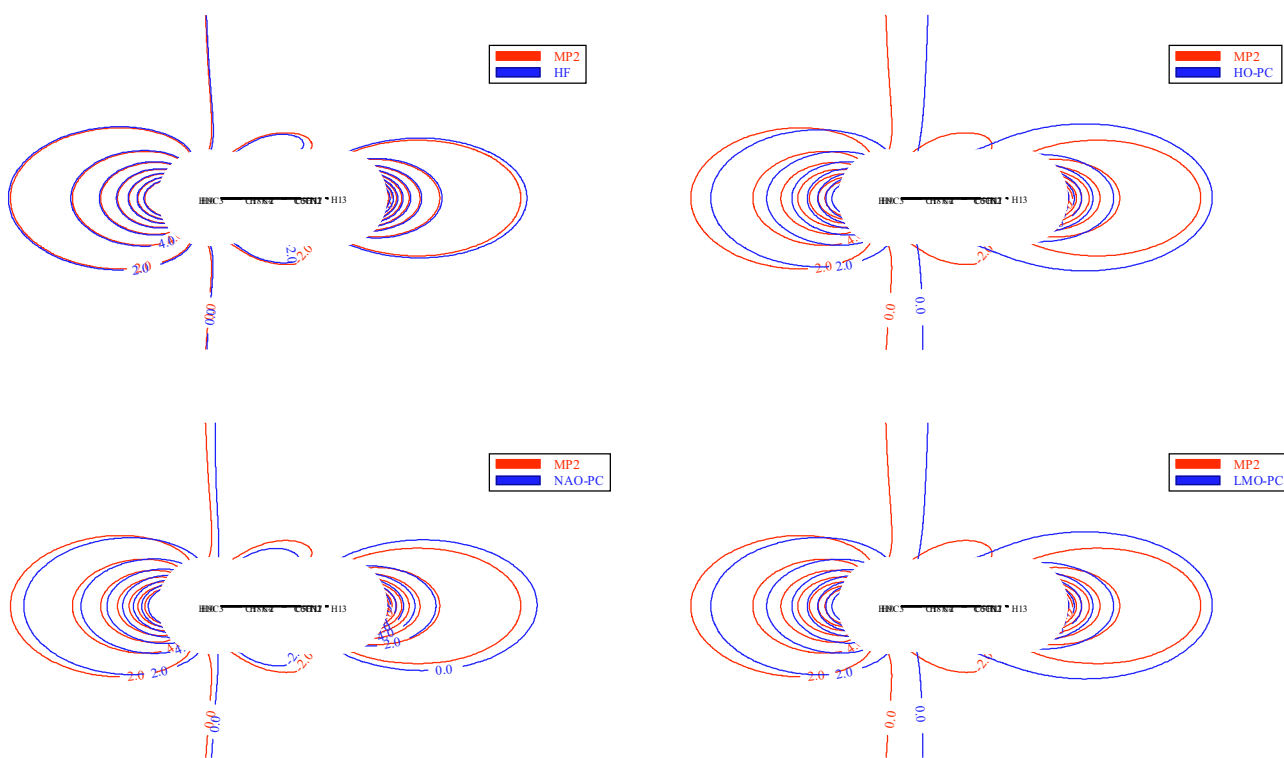


Figure 10 Isopotential maps in a plane perpendicular to cytosine (kcal mol^{-1})

The test molecules were selected in order to represent structural features found in biologically active compounds.

In Table 2 we show the results of a standard statistical analysis of the distribution. In addition to the standard statistical criteria, we also looked at the similarity index according to Hodgkin and Richards' modification [35] of the Carbo-Index.[36] This Carbo-Hodgkin-index, H_{rv} , is given by

$$H_{rv} = \frac{2\sum V_r V_v}{\sqrt{\sum V_r^2} + \sqrt{\sum V_v^2}} \quad (10)$$

where, V_r is the reference potential and V_v is the potential being compared. The sums run over all grid points. In contrast to the regression coefficient, the Carbo-Hodgkin-Index H_{rv} , was initially developed to compare molecular electron densities and describes the similarity of the distribution around 0. It is shown in Table 3 together with the structures.

As in our previous study of the NAO-PC method, [17,18] the statistical properties of the molecular electrostatic potential derived from LMO-PCs and HO-PCs is similar to the other methods. The differences to the previous work are due to two facts. On the one hand, solvent excluded surfaces and not van der Waals surfaces were used for the determination of the grid points used in the potential calculation. On the other hand, we used the larger basis set 6-31+G(d,p) [37] instead of 6-31G(d) for the MP2 and HF *ab initio* calculations.

The statistical properties of the distribution of the molecular electrostatic potential on the surface points are very similar for the three multicentre point charge methods NAO-PC, HO-PC and LMO-PC. The average of the Hodgkin-indices calculated between the different distributions are 0.99 for NAO-PC vs LMO-PC and LMO-PC vs HO-PC. For NAO-PC vs HO-PC it is only slightly smaller, 0.98. The molecular electrostatic potentials obtained with the three methods are very similar. This indicates that the HO-PC and the LMO-PC methods reproduce the electrostatic potential of the semiempirical wave function as well as the NAO-PC method.

In contrast, the MEP calculated from Coulson charges deviates significantly. The average of the calculated Hodgkin-indices comparing the MEP derived from Coulson charges to that derived from the multicentre point charge models is 0.92.

The semiempirical results compare reasonably well with those from the *ab initio* calculations. The maxima of the electrostatic potentials calculated with the semiempirical methods are generally smaller than those calculated with *ab initio* methods, but they are still in an acceptable range. The deviations of the minima are less distinct. The standard deviations are also comparable between the different methods. Thus, the distribution of the molecular electrostatic potential on the surface is comparable for *ab initio* and the semiempirical methods. The Hodgkin-indices confirm this conclusion; their average value is 0.93. The atom centred Coulson charges result in a molecular electrostatic potential that is less similar to the *ab initio* MEP.

As a further illustration of the strength of the point charge models, the isopotential maps of the MEP in the plane of

cytosine are compared to the MP2 *ab initio* calculated MEPs in Figure 9. The potential deviation is seen to vary between 1-2 kcal mol⁻¹ for the different models. To check the representation of the conjugated π -system, potentials were also calculated for cytosine in a plane perpendicular to the plane of the base. The NAO-PC model gives the best result, whereas the deviation in the ring plane for the HO-PC and LMO-PC models is greater, as shown in Figure 10. Again the errors were less than 2 kcal mol⁻¹ everywhere.

Conclusions

The two quasi-multipole techniques presented here are able to give a good qualitative representation of the molecular electrostatic potential. They also provide a bridge between classical and quantum mechanical techniques for the assignment of molecular charge models. The LMO-PC approach is conceptually particularly attractive in this respect, but suffers from the extra CPU-time needed to form the LMOs when used in its present form.

Acknowledgements Financial support by Deutsche Forschungsgemeinschaft and Pfizer Ltd. is gratefully acknowledged.

Supplementary material available statement VRML files for figure 4 and 5 are attached to this publication.

References

1. Politzer, P.; Murray J. S. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH: Weinheim, 1991; Vol 2, pp 273-312.
2. Warshel, A. *Computer Modelling of Chemical Reactions in Enzymes and Solutions*; J. Wiley & Sons, Inc.: New York, 1991.
3. Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.
4. Williams, J. H. *Acc. Chem. Res.* **1993**, *26*, 593.
5. Politzer, P. *J. Chem. Phys.* **1980**, *72*, 3027.
6. Gadre, S. R.; Bendale, R. D. *Chem. Phys. Lett.* **1986**, *130*, 515.
7. Luque, F. J.; Illas, F.; Orozco, M. *J. Comput. Chem.* **1990**, *11*, 416.
8. Orozco, M.; Luque, F. J. *J. Comput. Chem.* **1990**, *11*, 909.
9. Alemán, C.; Luque, F. J.; Orozco, M. *J. Comput. Chem.* **1993**, *14*, 799.
10. Alkorta, I.; Villar, H. O.; Artca, G. A. *J. Comput. Chem.* **1993**, *14*, 530.
11. Bonati, L.; Cosentino, U.; Frascini, E.; Moro, G.; Pitea, D. *J. Comput. Chem.* **1992**, *13*, 842.
12. Williams, D. E. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH: Weinheim, 1991; p 219.
13. Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.

14. Beck, B.; Clark, T.; Glen, R. C. *J. Mol. Model.* **1995**, *1*, 176.
15. Beck, B.; Clark, T.; Glen, R. C. *J. Comput. Chem.* **1997**, *19*, 744.
16. Tomasi, J.; Bonaccorsi, R.; Cammi, R. In *Theoretical Models of Chemical Bonding, Part IV*, Maksiæ, Z. B., Ed.; Springer-Verlag: Berlin, 1991; p 229.
17. Rauhut, G.; Clark, T. *J. Comput. Chem.* **1993**, *14*, 503.
18. Beck, B.; Rauhut, G.; Clark, T. *J. Comput. Chem.* **1994**, *15*, 1064.
19. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1975**, *97*, 1285.
20. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
21. Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.
22. Clark, T.; Alex, A.; Beck, B.; Chandrasekhar, J.; Gedeck, P.; Horn, A.; Hutter, M.; Rauhut, G.; Sauer, W.; T. Steinke, *Vamp 7.0*, Oxford Molecular, Medawar Centre, Oxford Science Park, Oxford, OX4 4GA, England, 1998.
23. Kutzelnigg W. *Einführung in die theoretische Chemie*, 2nd ed.; VCH: Weinheim, 1993; Vol 2, p 199.
24. Perkins, P. G.; Stewart, J. J. P. *J. Chem. Soc., Faraday Trans. 2* **1982**, *78*, 285.
25. Mauser, H.; Clark, T.; Hirsch, A. In *Electronic Properties of Novel Materials-Progress in Molecular Nanostructures*; Kuzmany, H.; Fink, J.; Mehring, M. Eds.; American Institute of Physics: Woodbury, 1998, pp 392–395.
26. MacGlynn, S. P. *Introduction to applied quantum chemistry*; Rinehard and Winston: New York Holt, 1972.
27. Gasteiger, J.; Rudolph, C.; Sadowski, J. *Tetrahedron Comput. Methodol.* **1990**, *3*, 537.
28. Sadowski, J.; Rudolph, C.; Gasteiger, J., *Corina 1.8*, Oxford Molecular, Medawar Centre, Oxford Science Park, Oxford, OX4 4GA, England, 1998.
29. Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.
30. Buckingham, A. D. *Quart. Rev.* **1959**, *13*, 189.
31. Senhuesa, J.E. *Theor. Chim. Acta* **1981**, *60*, 143.
32. Gierke, T. D.; Tiglaar, H. L.; Flygare, W. H. *Mol. Phys.* **1971**, *20*, 225.
33. Gaussian 94 Revision C.3; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; 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.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh PA, 1995.
34. Pascual-Ahuir, J. L.; Silla, E.; Tuñon, I. *J. Comput. Chem.* **1994**, *15*, 1127.
35. Hodgkin, E. E.; Richards W. G. *Int. J. Quant. Chem.* **1978**, *14*, 105.
36. Carbo, R.; Leyda, L.; Arnau, M. *Int. J. Quant. Chem.* **1980**, *17*, 1185.
37. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.*, **1983**, *4*, 294.