Charge Distribution in the Water Molecule—
A Comparison of Methods

F. MARTIN, H. ZIPSE

Department Chemie, LMU München, Butenandstr. 13, D-81377 München, Germany

Abstract: The charge distribution in the water molecule has been analyzed using a broad variety of basis sets, four different quantum mechanical methods (Hartree–Fock, Becke3LYP, MP2, and QCISD), and six population analysis methods (Mulliken, NPA, AIM, CHELPG, Merz–Kollman, and Resp). The influence of the molecular structure on the calculated atomic charges has been studied using small perturbations of the experimentally determined structure.

© 2004 Wiley Periodicals, Inc.   J Comput Chem 26: 97–105, 2005

Key words: charge distribution; water molecule; quantum mechanical methods

Introduction

The calculation of effective atomic charges plays an important role in the application of quantum mechanical calculations to molecular systems.[1–3] This is despite all conceptual problems connected to dividing up the overall molecular charge density in atomic contributions, and all practical problems connected to finding a convenient and robust algorithm applicable to a wide range of systems. The usefulness of effective atomic charges as parameters for the calculation of electrostatic interactions in a variety of molecular mechanics simulation packages is certainly one important area of application. Partial atomic charges serve a different, but even more important, purpose in the qualitative rationalization of organic and inorganic reactivity. Despite all fundamental problems the need for reliable procedures for the calculation of atomic charges will therefore persist. Although the deficiencies of some of the established methods such as the Mulliken population analysis have meanwhile achieved textbook status,[1,2] a broad comparison of procedures is difficult to find even for small model systems. This is in particular the case when we accept that a "procedure" should cover more than the algorithm of charge density division alone. The choice of the basis set and the selection of a quantum mechanical Hamiltonian represent two other essential points, which can hardly be neglected. The choice of the molecular structure may be an additional, but slightly less important issue. Due to the absence of studies covering all three important factors we have selected here the water molecule as a model system for the comparison of several procedures for the calculation of effective atomic charges. Water is very attractive due to its fundamental role as a medium in the life sciences and many technologically important processes. The ground-state structure of the water molecule

has been studied repeatedly[4,5] and we will use here the structure suggested by Benedict and coworkers[4] with $r_{OH} = 95.72$ and $a_{HOH} = 104.52°$. The gas phase molecular dipole moment of water has been measured to 1.855 D.[6] Assuming the structure of water to be rigid at its experimental geometry, this corresponds to effective atomic charges of $q(O) = -0.66e$ and (assuming C_{2v} symmetry) $q(H) = +0.33e$. Previous studies of the electron density distribution in water include that of Lipscomb and coworkers.[7] They calculated the total electron density in water as a function of the basis set and the electron correlation treatment using the experimental gas phase geometry[4] of the water molecule. A good overview of atomic charges and the different procedures to derive them has been given by Wiberg and Rablen.[8] They calculated atomic charges for a wide range of molecules at the HF/6-31G(d,p) and HF/6-311++G(d,p) level using the HF/6-31G(d) optimized geometries. The population analysis schemes employed in this study include the Mulliken population analysis, NPA, AIM, ChelpG, Hirshfeld, and GAPT methods. Not only the atomic charges, but also the geometry of water have been investigated by Kim and Jordan.[9] They optimized the geometry of water using the MP2 and several DFT methods in combination with a selection of augmented correlation consistent basis sets. Sigfridsson and Ryde[10] compared different methods for deriving the molecular electrostatic potential and atomic charges, such as the Mulliken population analysis, NPA, AIM, Chelp, ChelpG, MK, Resp, and introduced the Chelp–Bow and Chelmo methods. The calculations have been performed with the 6-31G(d) Pople basis set in combination with the B3LYP hybrid density functional method. The water geometry used was optimized at the HF/3-21G level. Astrand,

Correspondence to: H. Zipse; e-mail: zipse@cup.uni-muenchen.de

Ruud, Mikkelsen, and Helgaker[11] investigated the Mulliken charges for the cc-pVXZ (X = 2–6) and aug-cc-pVXZ (=2–5) basis sets, a wide range of ANO-basis sets and a selection of Pople basis sets such as STO-3G, 6-31G(d), and 6-31G(d,p). These studies have been performed at the Hartree–Fock level using the experimental water geometry. Partial charges for water have also been reported by Jensen at the Hartree–Fock level using a selection of different Pople and correlation consistent basis sets in combination with the Mulliken, Löwdin, ESP, NPA, and AIM schemes.[1] The geometry used in this study has, unfortunately, not been specified. Truhlar and coworkers have reported atomic charges calculated with the Mulliken and Löwdin schemes for a wide variety of systems in their effort to develop the redistributed Löwdin population analysis (RLPA).[12] Values for water have been calculated using the MPW1PW91 hybrid density functional method in combination with the 6-31G(d) and 6-31+G(d) basis sets, using the mPW1PW91/MIDI! optimized geometry. Most recently, Baerends and coworkers reported charge parameters for a wide variety of systems using the newly formulated Voronoi deformation density (VDD) method.[13] Values for water have been calculated using the BP86 density functional method in combination with the TZ2P basis set (constructed using Slater-type basis functions), using the BP86/TZ2P optimized geometry. The VDD results compare quite favorably with those obtained from other approaches based on deformation densities (such as the Hirshfeld method), but are significantly smaller than those obtained using the Mulliken, NPA, and AIM methods. Our interest here is in the comparison of different methods to describe the electron distribution in water as broadly as possible, and assess the sensitivity of the calculated charges (and thus the dipole moment) to changes in (1) the choice of population analysis method; (2) the choice of the basis set; (3) the choice of the quantum mechanical Hamiltonian; (4) the molecular structure.

Computational Methods

If not stated otherwise, all calculations have been performed with water in its experimentally known gas phase geometry with $r_{OH} = 95.72$ pm and $a_{HOH} = 104.52°$.[4] Population analysis methods include the following choices: (1) the Mulliken population analysis as the historically most important wave function based method,[14] whose results are known to strongly depend on the choice of the molecular basis set.[1,2] Despite its known deficiencies, the Mulliken population analysis is still widely used due to its simplicity; (2) the Natural Population Analysis (NPA) is a more refined wave function-based method, that solves most of the problems of the Mulliken scheme by construction of a more appropriate set of (natural) atomic basis functions;[15,16] (3) the Atoms in Molecules (AIM) scheme developed by Bader and coworkers,[17] in which the molecular charge density is divided up in atomic contributions based on its overall topology. In practical terms, this is achieved through the construction of atomic compartments defined through (interatomic) zero-flux surfaces of the electronic density; (4) the method of electrostatic potential (ESP) derived charges, in which atomic partial charges are selected to reproduce the molecular electrostatic potential at a (large) number of grid points.[18,19] Various schemes have been defined that differ in the choice of grid points and the

actual fitting procedure. The three variants considered here include the ChelpG, the Merz–Kollman (MK), and the Resp scheme. In the ChelpG scheme,[20] the points are selected in a regularly spaced cubic grid while the MK[21,22] and the Resp schemes[23] use points located on nested Connolly surfaces. One fundamental problem of ESP methods is the prediction of unrealistically large charges for atoms that lie deeply buried inside the molecule (not connected to the surface of the van der Waals radii). The Resp scheme by Bayly et al.[23] tries to avoid this complication by restraining the charges of selected atoms to zero with a hyperbolic penalty function. For the water molecule, this should not lead to substantially different results than the MK scheme, as there are no deeply buried atomic centers. Variations in the basis set include a number of Pople basis sets [STO-3G,[24] 3-21G,[25,26] 6-31G,[27] 6-31G(d), 6-31G(d,p), 6-311G(d,p),[28] 6-311+G(2d,2p), 6-311+G(2df,2pd), and 6-311++G (3df, 3pd)] and the full range of correlation consistent basis sets (cc-pVDZ, cc-pVTZ, cc-pVQZ, cc-pV5Z, cc-pV6Z),[29] and of augmented correlation consistent basis sets (aug-cc-pVDZ, aug-cc-pVTZ, aug-cc-pVQZ, aug-cc-pV5Z, aug-cc-pV6Z) developed by Dunning and coworkers.[30,31] Four different quantum mechanical methods have been used including (a) self-consistent field theory at the Hartree–Fock level (HF); (b) density functional theory (DFT) in its Becke3LYP[32,33] incarnation as implemented in *Gaussian 03, Rev. B.03* in combination with a pruned (75/302) grid defined as "finegrid"; (c) second-order Møller–Plesset (MP2)[34–36] theory as the most economical way to treat electron correlation in a nonparameterized fashion; (d) the QCISD scheme treating electron correlation effects in a more sophisticated manner. The frozen core approximation has been used for all MP2 and QCISD calculations. All calculations have been performed with *Gaussian 03, Rev. B.03.*[37]

Results and Discussion

All results (atomic charges for the oxygen atom and the associated dipole moment) are summarized by quantum mechanical method in Tables 1–4 (Hartree–Fock, Becke3LYP, MP2, QCISD). The charge distribution could not be calculated at the QCISD level with the largest cc basis sets considered here due to technical problems. The results can, however, better be represented in graphical form as has been done in Figures 1 and 2. To reflect some information on the choice of basis set the results obtained using Pople basis sets are represented with open bars, those for the correlation consistent basis sets with gray bars, and those for the augmented correlation-consistent basis sets with black bars.

Mulliken

As must be expected from the available literature results, the Mulliken charges cover a very broad range of values from −0.30 at the low side to −0.95 at the high side at the Hartree–Fock level (Table 1). This is even true when using the very systematic family of correlation consistent basis sets. For the latter, one can observe a steady increase in negative charge at oxygen when increasing the basis set size from double to quintuple-zeta cc-basis sets. The largest member of this family does, however, break this trend, and predicts substantially smaller oxygen charges. The same very

Atomic Charges on Oxygen (H₂O)
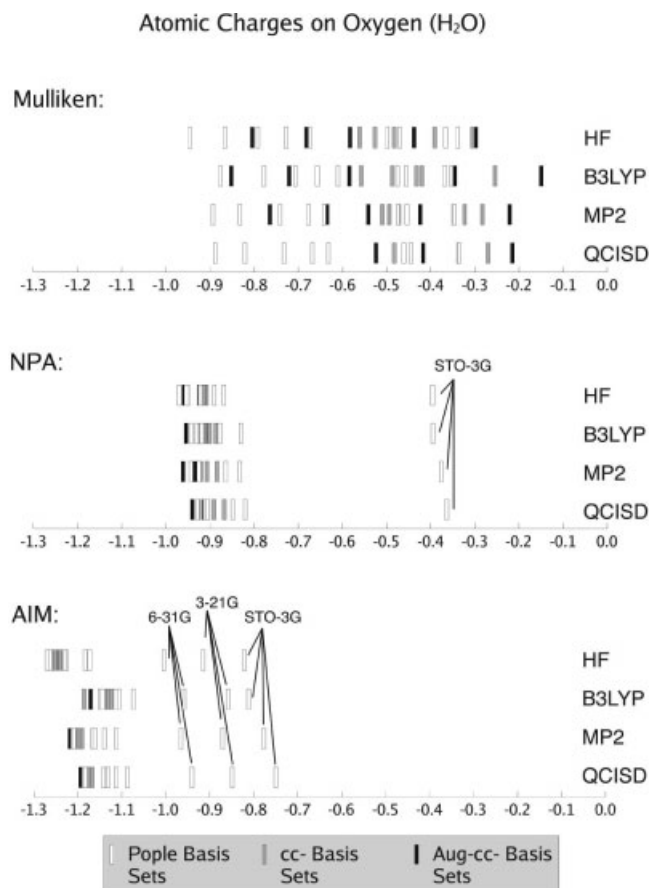
Figure 1. Oxygen partial charges calculated with the Mulliken, the NPA, and the AIM population analysis methods for water.

broad dispersion of results as a function of the basis set can also be observed for Mulliken charges calculated with the other three quantum mechanical methods (Tables 2–4). Quite in line with expectation the partial charges calculated with correlated methods are somewhat smaller than those calculated at the Hartree–Fock level. On average, the reductions amount to 7% at the Becke3LYP level, 7% at the MP2 level, and 8% at the QCISD level. There are, however, some nonsystematic exceptions to this rule, in particular at the Becke3LYP level, the partial charges being actually larger than the HF values for some of the larger basis sets. In conclusion, the results illustrated in Figure 1 show that the Mulliken charges for water cannot be brought to conversion, whatever the choice of basis set or quantum mechanical method.

NPA

The Natural Population Analysis (NPA) method shows an entirely different response to changes in the basis set and the quantum mechanical Hamiltonian (Fig. 1). Neglecting the results obtained with the smallest Pople basis set (STO-3G) all other combinations of basis sets and quantum mechanical methods predict atomic charges in a narrow window of $0.2e$ width centered at $-0.8e$. The window is even smaller when neglecting the results obtained with

the second smallest Pople basis set, 3-21G. Both the QCISD and the Becke3LYP level predict uniformly smaller oxygen charges compared to the Hartree–Fock level, the effect being slightly larger for the QCISD (3% reduction) than Becke3LYP level (1% reduction). The MP2 charges are for some basis sets smaller, but for others (in particular, the larger members of the cc-series) larger than the HF values. On average, a reduction of 1% is also predicted for the MP2 level.

AIM

The partial charges obtained using the *Atoms in Molecules* (AIM) scheme show a significantly larger dispersion as a function of basis set choice and quantum mechanical method compared to the NPA scheme. This is mainly a consequence of the results obtained with the smallest three Pople basis sets (STO-3G, 3-21G, and 6-31G). The latter provide partial charges, which are much less negative than those obtained using all other basis sets, irrespective of the quantum mechanical method used. Neglecting these latter results, the AIM charges all group in a window of $0.2e$ width centered at $-1.2e$. The AIM charges are thus, on average, $0.3e$ more negative than those obtained with the NPA scheme. The influence of the choice of the Hamiltonian is also significantly increased, the re-

Atomic Charges on Oxygen (H₂O)
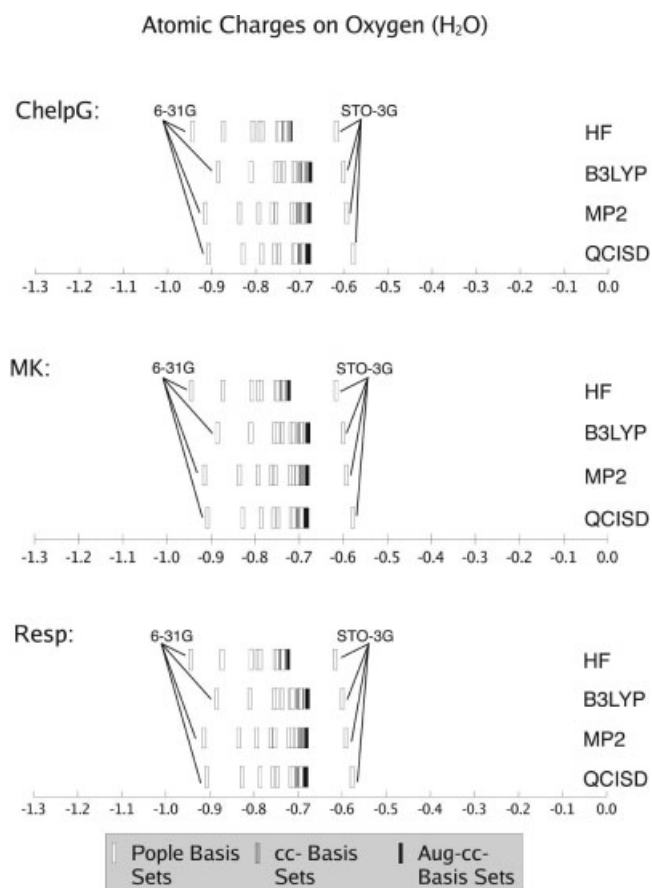
Figure 2. Oxygen partial charges calculated with the CHELPG, the MK, and the Resp schemes for water.

Table 1. Partial Charge of the Oxygen Atom in Water and the Resulting Dipole Moment as Calculated at the Hartree–Fock Level of Theory.[a,b]

Method	Mulliken[c]		ChelpG		NPA		AIM		MK		Resp	
	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ
HF/STO-3G	−0.3664	1.0309	−0.6173	1.7370	−0.3957	1.1135	−0.8209	2.3099	−0.6162	1.7340	−0.6156	1.7322
HF/3-21G	−0.7278	2.0480	−0.8734	2.4576	−0.8690	2.4451	−0.9149	2.5743	−0.8726	2.4553	−0.8720	2.4535
HF/6-31G	−0.7918	2.2280	−0.9438	2.6558	−0.9507	2.6750	−1.0041	2.8252	−0.9442	2.6569	−0.9435	2.6550
HF/6-31G(d)	−0.8663	2.4377	−0.8071	2.2709	−0.9543	2.6852	−1.1726	3.2995	−0.8071	2.2710	−0.8064	2.2692
HF/6-31G(d,p)	−0.6736	1.8953	−0.7923	2.2293	−0.9693	2.7274	−1.2254	3.4480	−0.7922	2.2292	−0.7916	2.2274
HF/6-311G(d,p)	−0.4985	1.4026	−0.7856	2.2107	−0.8900	2.5042	−1.1825	3.3274	−0.7858	2.2112	−0.7852	2.2094
HF/6-311+G(2d,2p)	−0.4705	1.3238	−0.7501	2.1107	−0.9236	2.5987	−1.2291	3.4585	−0.7518	2.1155	−0.7512	2.1137
HF/6-311+G(2df,2pd)	−0.3391	0.9543	−0.7458	2.0985	−0.9233	2.5980	−1.2693	3.5715	−0.7476	2.1035	−0.7469	2.1017
HF/6-311++G(3df,3pd)	−0.9454	2.6602	−0.7332	2.0631	−0.9215	2.5928	−1.2620	3.5511	−0.7349	2.0678	−0.7342	2.0660
HF/cc-pVDZ	−0.3054	0.8595	−0.7466	2.1008	−0.9121	2.5665	−1.2559	3.5338	−0.7464	2.1002	−0.7458	2.0984
HF/cc-pVTZ	−0.4826	1.3580	−0.7376	2.0754	−0.9088	2.5573	−1.2505	3.5186	−0.7384	2.0777	−0.7377	2.0759
HF/cc-pVQZ	−0.5261	1.4804	−0.7309	2.0567	−0.9202	2.5894	−1.2414	3.4931	−0.7324	2.0608	−0.7317	2.0590
HF/cc-pV5Z	−0.5611	1.5789	−0.7283	2.0494	−0.9125	2.5676	−1.2375	3.4821	−0.7302	2.0545	−0.7295	2.0527
HF/cc-pV6Z	−0.3912	1.1008	−0.7248	2.0394	−0.9145	2.5733	−1.2308	3.4632	−0.7267	2.0447	−0.7260	2.0429
HF/aug-cc-pVDZ	−0.2976	0.8374	−0.7283	2.0494	−0.9594	2.6995	−1.2454	3.5044	−0.7299	2.0538	−0.7293	2.0520
HF/aug-cc-pVTZ	−0.4381	1.2328	−0.7231	2.0346	−0.9224	2.5955	−1.2686	3.5697	−0.7249	2.0397	−0.7242	2.0379
HF/aug-cc-pVQZ	−0.5826	1.6394	−0.7218	2.0311	−0.9244	2.6012	−1.2446	3.5020	−0.7237	2.0363	−0.7230	2.0345
HF/aug-cc-pV5Z	−0.8239	2.3184	−0.7218	2.0310	−0.9124	2.5673	−1.2376	3.4824	−0.7236	2.0360	−0.7229	2.0342
HF/aug-cc-pV6Z	−0.6817	1.9183	−0.7218	2.0310	−0.9129	2.5687	−1.2313	3.4645	−0.7236	2.0360	−0.7229	2.0342

[a]Atomic charge in units of e.
[b]Dipole moment in units of D.
[c]The oxygen charges reported at the Hartree–Fock level using the experimental geometry in ref. 5 are: STO-3G: −0.3660; 6-31G(d): −0.8974; 6-31G(d,p): −0.6810; cc-pVDZ: −0.3058; cc-pVTZ: −0.4828; cc-pVQZ: −0.5264; cc-pV5Z: −0.5611; cc-pV6Z: −0.4408; aug-cc-pVDZ: −0.2984; aug-c-pVTZ: −0.4387; aug-cc-pVQZ: −0.5830; aug-cc-pV5Z: −0.8245.

duction of the partial oxygen charge relative to the Hartree–Fock level now amounting to −8% for the Becke3LYP method, −4% for the MP2 method, and −7% for the QCISD method.

CHELPG

The charges calculated with the CHELPG method show a significantly more systematic and predictable behavior than the methods based on an analysis of the wave function or the electron density topology. Neglecting again the smallest Pople basis sets STO-3G and 3-21G, the oxygen charges predicted at the Hartree–Fock level show a systematic decrease with increasing basis set size. Although no complete convergence can be achieved with the Pople basis sets, limiting values of $−0.72e$ are predicted for the oxygen partial charge for both subsets of the correlation consistent basis sets. This corresponds to an effective dipole moment of 2.03 D, 9% above the experimental value. Compared to the charges calculated at the Hartree–Fock level the consideration of correlation effects reduces the charges by 6, 5, and 5% at the Becke3LYP, MP2, and QCISD level, respectively. Aside from the convergence properties as a function of basis set size it should be mentioned, that the limiting values for the dipole moment calculated at the Becke3LYP and MP2 levels of 1.91 D and 1.92 D are only 3% larger than the experimentally measured one of 1.855 D. The dipole moment can also be calculated directly from the electron density as the expectation value of the dipole moment operator, and values of 1.854 D and 1.862 D, respectively, are calculated at the Becke3LYP and MP2 level in combination with the largest basis set used here (aug-cc-pV6Z). Because these latter values are very close to the experimental one, it appears that the CHELPG algorithm itself is responsible for the overestimation of 3%.

MK and Resp

The Merz–Kollman (MK) and the Resp schemes have by and large the same convergence characteristics as CHELPG (Fig. 2). This is also true for the limiting oxygen charges predicted at all four quantum mechanical levels. The additional restraint imposed by the Resp method leads to a systematic reduction of the oxygen charge by $0.0006e$ compared to the MK values, and thus a minimally smaller dipole moment. The oxygen charge calculated with the Resp scheme at the HF/6-31G(d) level is $−0.8064e$, leading to an overestimation of the dipole moment by 22% (2.27 D instead of 1.855 D). This is an important result because the charges for the AMBER 94 force field[38] have all been obtained at this level of theory. Using larger basis sets and employing correlated quantum mechanical methods leads to practically the same results as obtained with CHELPG.

Table 2. Partial Charge of the Oxygen Atom in Water and the Resulting Dipole Moment as Calculated at the B3LYP Level of Theory.[a,b]

Method	Mulliken		ChelpG		NPA		AIM		MK		Resp	
	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ
B3LYP/STO-3G	−0.3687	1.0375	−0.6014	1.6921	−0.3952	1.1121	−0.8115	2.2834	−0.6003	1.6890	−0.5996	1.6872
B3LYP/3-21G	−0.6569	1.8484	−0.8105	2.2806	−0.8298	2.3350	−0.8585	2.4158	−0.8095	2.2778	−0.8089	2.2760
B3LYP/6-31G	−0.7065	1.9880	−0.8851	2.4904	−0.9218	2.5938	−0.9577	2.6949	−0.8851	2.4904	−0.8844	2.4886
B3LYP/6-31G(d)	−0.7783	2.1900	−0.7556	2.1261	−0.9300	2.6168	−1.1058	3.1115	−0.7553	2.1253	−0.7547	2.1235
B3LYP/6-31G(d,p)	−0.6100	1.1764	−0.7374	2.0750	−0.9430	2.6535	−1.1435	3.2177	−0.7372	2.0742	−0.7365	2.0724
B3LYP/6-311G(d,p)	−0.4753	1.3375	−0.7478	2.1041	−0.8776	2.4693	−1.0730	3.0193	−0.7478	2.1042	−0.7472	2.1024
B3LYP/6-333+G(2d,2p)	−0.4557	1.2823	−0.7152	2.0124	−0.9208	2.5910	−1.1134	3.1330	−0.7178	2.0199	−0.7172	2.0181
B3LYP/6-311+G(2df,2pd)	−0.3538	0.9956	−0.7105	1.9993	−0.9209	2.5911	−1.1495	3.2344	−0.7132	2.0068	−0.7126	2.0050
B3LYP/6-311++G(3df,3pd)	−0.8765	2.4663	−0.6920	1.9472	−0.9193	2.5866	−1.1430	3.2162	−0.6944	1.9538	−0.6937	1.9520
B3LYP/cc-pVDZ	−0.2546	0.7165	−0.6944	1.9538	−0.8871	2.4961	−1.1851	3.3346	−0.6940	1.9529	−0.6934	1.9511
B3LYP/cc-pVTZ	−0.4322	1.2163	−0.6996	1.9685	−0.9008	2.5345	−1.1271	3.1713	−0.7004	1.9709	−0.6998	1.9691
B3LYP/cc-pVQZ	−0.4867	1.3696	−0.6933	1.9508	−0.9157	2.5765	−1.1357	3.1957	−0.6950	1.9556	−0.6944	1.9538
B3LYP/cc-pV5Z	−0.5578	1.5694	−0.6897	1.9406	−0.9098	2.5601	−1.1287	3.1759	−0.6921	1.9474	−0.6914	1.9456
B3LYP/cc-pV6Z	−0.4207	1.1836	−0.6836	1.9235	−0.9125	2.5676	−1.1232	3.1604	−0.6862	1.9102	−0.6856	1.9084
B3LYP/aug-cc-pVDZ	−0.1500	0.4220	−0.6766	1.9038	−0.9540	2.6845	−1.1703	3.2931	−0.6789	1.9102	−0.6782	1.9084
B3LYP/aug-cc-pVTZ	−0.3457	0.9727	−0.6779	1.9074	−0.9201	2.5890	−1.1483	3.2311	−0.6802	1.9141	−0.6796	1.9123
B3LYP/aug-cc-pVQZ	−0.5842	1.6438	−0.6773	1.9057	−0.9225	2.5957	−1.1414	3.2117	−0.6797	1.9126	−0.6791	1.9108
B3LYP/aug-cc-pV5Z	−0.8752	2.4627	−0.6774	1.9062	−0.9101	2.5609	−1.1296	3.1784	−0.6799	1.9130	−0.6792	1.9112
B3LYP/aug-cc-pV6Z	−0.7208	2.0283	−0.6775	1.9062	−0.9107	2.5625	−1.1237	3.1618	−0.6798	1.9130	−0.6792	1.9112

[a]Atomic charge in units of e.
[b]Dipole moment in units of D.

Table 3. Partial Charge of the Oxygen Atom in Water and the Resulting Dipole Moment as Calculated at the MP2 Level of Theory.[a,b]

Method	Mulliken		ChelpG		NPA		AIM		MK		Resp	
	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ
MP2/STO-3G	−0.3481	0.9794	−0.5934	1.6697	−0.3757	1.0572	−0.7770	2.1862	−0.5923	1.6666	−0.5917	1.6648
MP2/3-21G	−0.6771	1.9053	−0.8364	2.3535	−0.8320	2.3411	−0.8701	2.4483	−0.8355	2.3509	−0.8348	2.3491
MP2/6-31G	−0.7413	2.0860	−0.9142	2.5724	−0.9258	2.6050	−0.9653	2.7163	−0.9143	2.5727	−0.9137	2.5709
MP2/6-31G(d)	−0.8319	2.3408	−0.7936	2.2332	−0.9431	2.6538	−1.1385	3.2034	−0.7935	2.2328	−0.7929	2.2310
MP2/6-31G(d,p)	−0.6415	1.8050	−0.7630	2.1470	−0.9431	2.6537	−1.1656	3.2797	−0.7628	2.1465	−0.7622	2.1447
MP2/6-311G(d,p)	−0.4740	1.3336	−0.7539	2.1212	−0.8645	2.4325	−1.1124	3.1300	−0.7539	2.1214	−0.7533	2.1196
MP2/6-311+G(2d,2p)	−0.4538	1.2769	−0.7180	2.0203	−0.9232	2.5976	−1.1605	3.2655	−0.7205	2.0273	−0.7198	2.0255
MP2/6-311+G(2df,2pd)	−0.3459	0.9732	−0.7113	2.0016	−0.9255	2.6041	−1.2115	3.4089	−0.7138	2.0085	−0.7132	2.0067
MP2/6-311++G(3df,3pd)	−0.8931	2.5132	−0.6926	1.9488	−0.9255	2.6041	−1.2083	3.3998	−0.7041	1.9813	−0.7035	1.9795
MP2/cc-pVDZ	−0.2834	0.7975	−0.7140	2.0091	−0.8842	2.4881	−1.2018	3.3816	−0.7137	2.0083	−0.7131	2.0065
MP2/cc-pVTZ	−0.4720	1.3280	−0.7020	1.9752	−0.9071	2.5524	−1.1901	3.3487	−0.7028	1.9776	−0.7022	1.9758
MP2/cc-pVQZ	−0.4941	1.3903	−0.6954	1.9569	−0.9252	2.6034	−1.2012	3.3798	−0.6971	1.9614	−0.6964	1.9596
MP2/cc-pV5Z	−0.5097	1.4343	−0.6922	1.9476	−0.9209	2.5914	−1.2022	3.3827	−0.6944	1.9539	−0.6937	1.9521
MP2/cc-pV6Z	−0.3234	0.9100	−0.6870	1.9330	−0.9242	2.6004	−1.1981	3.3711	−0.6893	1.9396	−0.6887	1.9378
MP2/aug-cc-pVDZ	−0.2216	0.6235	−0.6834	1.9230	−0.9609	2.7037	−1.2026	3.3840	−0.6854	1.9287	−0.6848	1.9269
MP2/aug-cc-pVTZ	−0.4238	1.1924	−0.6792	1.9112	−0.9292	2.6147	−1.2172	3.4250	−0.6814	1.9173	−0.6808	1.9155
MP2/aug-cc-pVQZ	−0.5416	1.5240	−0.6808	1.9158	−0.9340	2.6282	−1.2084	3.4003	−0.6831	1.9221	−0.6824	1.9203
MP2/aug-cc-pV5Z	−0.7835	2.2047	−0.6818	1.9185	−0.9242	2.6005	−1.2041	3.3882	−0.6840	1.9246	−0.6834	1.9228
MP2/aug-cc-pV6Z	−0.6343	1.7847	−0.6824	1.9200	−0.9251	2.6031	−1.1994	3.3748	−0.6845	1.9261	−0.6839	1.9243

[a]Atomic charge in units of e.
[b]Dipole moment in units of D.

Table 4. Partial Charge of the Oxygen Atom in Water and the Resulting Dipole Moment as Calculated at the QCISD Level of Theory.[a,b,c]

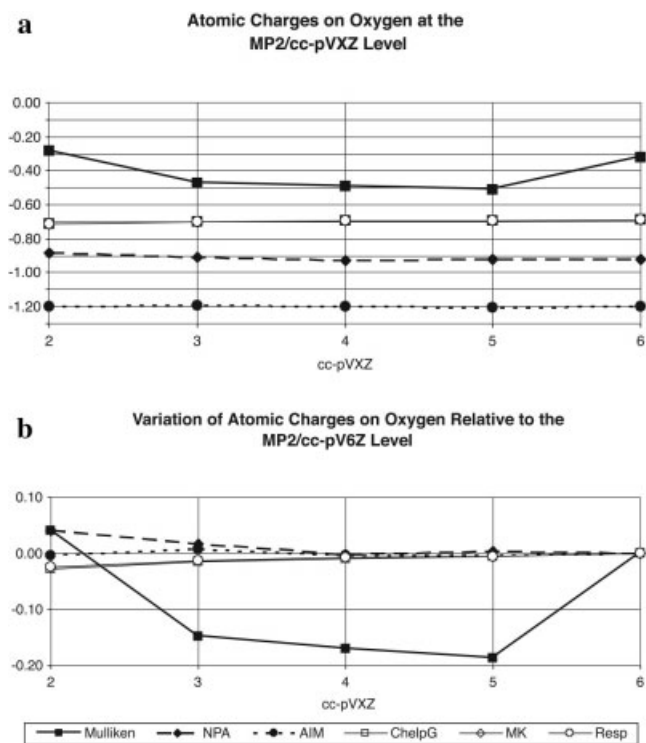| Method	Mulliken		ChelpG		NPA		AIM		MK		Resp	
	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ
QCISD/STO-3G	−0.3369	0.9479	−0.5789	1.6290	−0.3633	1.0223	−0.7493	2.1085	−0.5778	1.6258	−0.5771	1.6240
QCISD/3-21G	−0.6685	1.8810	−0.8288	2.3320	−0.8205	2.3086	−0.8485	2.3875	−0.8278	2.3294	−0.8272	2.3276
QCISD/6-31G	−0.7321	2.0600	−0.9076	2.5539	−0.9118	2.5657	−0.9395	2.6436	−0.9078	2.5543	−0.9071	2.5525
QCISD/6-31G(d)	−0.8211	2.3104	−0.7865	2.2129	−0.9271	2.6087	−1.1127	3.1309	−0.7863	2.2126	−0.7857	2.2108
QCISD/6-31G(d,p)	−0.6319	1.7781	−0.7576	2.1318	−0.9279	2.6111	−1.1407	3.2096	−0.7575	2.1314	−0.7568	2.1296
QCISD/6-311G(d,p)	−0.4611	1.2975	−0.7473	2.1028	−0.8487	2.3879	−1.0866	3.0574	−0.7474	2.1031	−0.7468	2.1013
QCISD/6-311+G(2d,2p)	−0.4446	1.2511	−0.7144	2.0103	−0.9042	2.5442	−1.1310	3.1824	−0.7167	2.0167	−0.7161	2.0149
QCISD/6-311+G(2df,2pd)	−0.3365	0.9468	−0.7091	1.9954	−0.9075	2.5535	−1.1858	3.3366	−0.7114	2.0016	−0.7107	1.9998
QCISD/6-311++G(3df,3pd)	−0.8885	2.5002	−0.6921	1.9474	−0.9066	2.5510	−1.1821	3.3261	−0.6941	1.9530	−0.6934	1.9512
QCISD/cc-pVDZ	−0.2707	0.7617	−0.7079	1.9920	−0.8681	2.4426	−1.1766	3.3108	−0.7077	1.9912	−0.7070	1.9894
QCISD/cc-pVTZ	−0.4611	1.2974	−0.6973	1.9620	−0.8909	2.5069	−1.1669	3.2833	−0.7005	1.9712	−0.6999	1.9694
QCISD/cc-pVQZ	−0.4824	1.3575	−0.6955	1.9570	−0.9083	2.5559	−1.1822	3.3265	−0.6970	1.9613	−0.6964	1.9595
QCISD/cc-pV5Z	—	—	—	—	—	—	—	—	—	—	—	—
QCISD/cc-pV6Z	—	—	—	—	—	—	—	—	—	—	—	—
QCISD/aug-cc-pVDZ	−0.2152	0.6054	−0.6811	1.9165	−0.9399	2.6446	−1.1778	3.3142	−0.6830	1.9218	−0.6824	1.9200
QCISD/aug-cc-pVTZ	−0.4172	1.1739	−0.6799	1.9131	−0.9099	2.5602	−1.1925	3.3556	−0.6819	1.9188	−0.6813	1.9170
QCISD/aug-cc-pVQZ	−0.5243	1.4752	−0.6825	1.9205	−0.9151	2.5749	−1.1884	3.3438	−0.6846	1.9263	−0.6839	1.9245
QCISD/aug-cc-pV5Z	—	—	—	—	—	—	—	—	—	—	—	—
QCISD/aug-cc-pV6Z	—	—	—	—	—	—	—	—	—	—	—	—

[a]Atomic charge in units of e.
[b]Dipole moment in units of D.
[c]The atomic charges could not be calculated at the QCISD/cc-pVXZ and QCISD/aug-cc-VXZ (X = 5, 6) levels of theory due to technical problems.

Comparison of Results at the MP2 Level

To highlight again the different convergence properties of the Pople and the correlation consistent basis sets, the results computed at the MP2 level have been compiled in Figure 3 for all five population analysis methods and all basis sets. Analysis of the QCISD data would clearly have been desirable, but the absence of results for the largest correlation consistent basis sets do not allow for a proper assessment of the convergence properties of charge derivation schemes.

In Figure 3a the results for all basis sets are included, and the rather broad dispersion of results from the Mulliken population analysis can again be compared to the distributions obtained from the other population analysis methods. Excluding again the results obtained with the STO-3G and 3-21G basis sets, it is clear that the NPA analysis offers the least dispersion, while AIM and the ESP methods are slightly less favorable in this aspect. Eliminating all results obtained with the Pople basis sets and concentrating on the correlation consistent basis sets (Fig. 3b) yields a somewhat different picture. It is now the AIM method that predicts an impressively narrow distribution of results, centered at the rather negative value of $-1.2 \pm 0.01e$. A substantially broader distribution is now predicted by the NPA method with values at around $-0.92 \pm 0.04e$. All three ESP methods predict essentially the same oxygen charges centered at around $-0.70 \pm 0.02e$. These latter values yield dipole moments in much closer agreement with the experimental value than with any of the other methods. It is clear from this comparison that the combination of any of the ESP methods

a Atomic Charges on Oxygen (H₂O) at MP2 Level

b Atomic Charges on Oxygen (H₂O) at MP2 Level
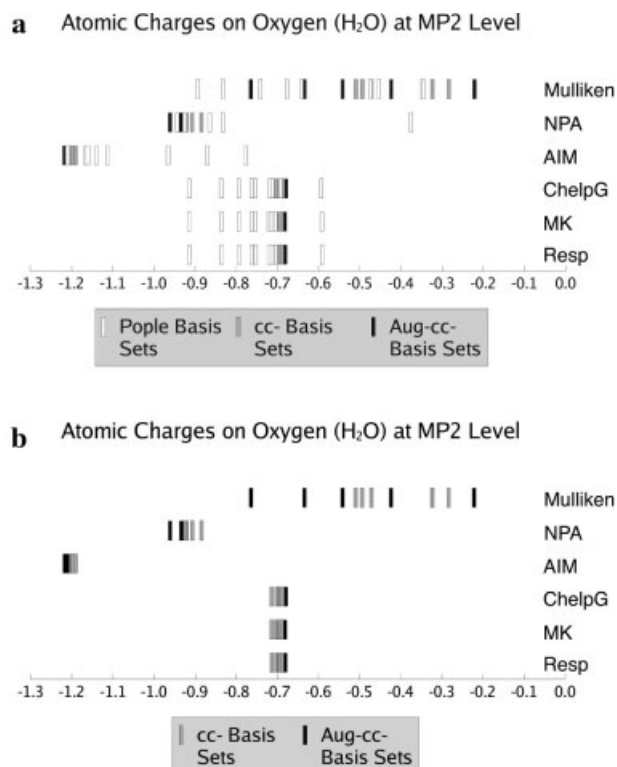
Figure 3. Oxygen partial charges in water at the MP2 level of theory (a) for all basis sets, and (b) for the correlation consistent basis sets.

Figure 4. Variation of the oxygen partial charge at the MP2/cc-pVXZ level.

The same type of analysis can be performed for the results obtained with the aug-cc-pVxZ basis sets (Fig. 5). The conclusions are practically the same as those for the cc basis sets, the performance of the smallest member of this series being considerably better (predicting charges rather close to the larger basis sets) in the aug-cc-series than in the cc-series. This finding is well in line with structural and energetic data for the water monomer and dimer obtained from cc- and aug-cc-basis sets at the correlated level.[41]

Sensitivity of the Molecular Dipole Moment to Structural Parameters

The choice of the molecular structure represents one additional variable in the calculation of atomic charge parameters. To explore the influence of structural variations on the charge distribution, we have recalculated the MP2/aug-cc-pV6Z charges with systematically perturbed geometries (Table 5). The structural perturbations include variations of ±2% in the H—O—H bond angle (resulting in values of 106.61 and 102.43°) and the O—H bond distance (resulting in values of 97.63 and 93.81 pm). Although it can hardly be expected that the electron density responds in a linear fashion to structural changes, one would hope that the dipole moments predicted by the different charge derivation methods vary proportional to the dipole moment μ_{den} predicted directly from the wave function. Variation of the latter amounts to approximately 1.6% for changes in the bond angle and to only 0.6% for changes in the bond distance. The dipole moment predicted for water is thus not

with the cc basis sets provides a very robust platform for the calculation of effective charge parameters.

Distribution and Convergence of Atomic Charges

The results shown in Figure 3b can also be used to discuss more precisely the variation of atomic charge as a function of the basis set size. It has been shown before that the cc basis set family produces (at correlated level) total energies, bond energies, vibrational frequencies, etc., that vary so systematically with increasing basis set size, that simple exponential functions can be used to extrapolate to the basis set limit.[39,40] The oxygen charges calculated at the MP2 level with the cc basis sets have therefore been plotted in Figure 4 once in absolute terms with increasing order from left to right (top) and then as the difference to the value obtained with a given population analysis method when using the cc-pV6Z basis set. In both types of presentation it can clearly be seen that the Mulliken analysis does not lead to converged results. This is due to both the rather large variation of the predicted charges as well as the unsystematic variation of these charges with basis set size. The other methods show, in comparison, a very small variation in the predicted value with increasing basis set size, and it can, in particular, be seen from the bottom representation in Figure 4 that the variations become smaller with increasing basis set size. The results obtained from the three ESP methods are so similar in both types of representation that all data fall on practically one line.

Figure 5. Variation of the oxygen partial charge at the MP2/aug-cc-pVXZ level.

Table 5. Partial Charge of the Oxygen Atom in Water and the Resulting Dipole Moment as Calculated at the MP2/aug-cc-pV6Z Level of Theory as a Function of the Molecular Structure.[a,b]

|			Mulliken		ChelpG		NPA		AIM		MK		Resp	
r_{OH}[c]	a_{HOH}[c]	μ_{den}	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ	q_o	μ
95.72	104.52	1.8619	−0.6343	1.7847	−0.6824	1.9200	−0.9251	2.6031	−1.1994	3.3748	−0.6845	1.9261	−0.6839	1.9243
95.72	106.61	1.8305	−0.6371	1.7502	−0.6909	1.8979	−0.9294	2.5531	−1.2063	3.3138	−0.6923	1.9018	−0.6917	1.9001
95.72	102.43	1.8926	−0.6318	1.8193	−0.6746	1.9426	−0.9211	2.6524	−1.1928	3.4348	−0.6777	1.9515	−0.6771	1.9498
97.63	104.52	1.8724	−0.6917	1.7785	−0.6739	1.9341	−0.9319	2.6745	−1.1810	3.3894	−0.6755	1.9387	−0.6749	1.9369
93.81	104.52	1.8510	−0.6466	1.7831	−0.6908	1.9050	−0.9174	2.5299	−1.2182	3.3594	−0.6933	1.9119	−0.6927	1.9102

[a]Atomic charge in units of e.
[b]Dipole moment in units of D.
[c]Distances in pm and angles in degrees.

highly sensitive to the choice of a particular structural parameter. For both types of perturbation considered here the direction of the changes are as expected, with larger dipole moments predicted for smaller bond angles and longer bond distances. The dipole moments predicted by the population analysis methods mirror the changes in μ_{den} quite closely in most instances. To facilitate the analysis of the results we define here the sensitivity of a given method as the ratio of the predicted change in dipole moment (induced through a given structural perturbation) and the change in μ_{den}. A sensitivity of 1 thus indicates that the dipole moment calculated with a given population analysis method responds to a given structural perturbation in exactly the same way as μ_{den}. A sensitivity of 0 indicates no change in dipole moment as a function of structural perturbation, and a negative sensitivity indicates that a given population analysis responds in the opposite direction as μ_{den} to structural changes. The sensitivity values shown in Figure 6 for changes in the bond angle of water are all fairly close to 1, and rather similar for positive or negative structural perturbations. All three ESP methods have sensitivities of 0.7, indicating slightly smaller changes of the predicted dipole moment compared to changes in μ_{den}. The situation is quite different for perturbations in the O—H bond distance, where larger variations in the sensitivities can be observed. Although the Mulliken population analysis predicts changes in the dipole moment, which are much too small and

for longer O—H bond distances also in the wrong direction, the NPA scheme is much more sensitive to the structural changes than μ_{den}. All three ESP methods predict changes in the dipole moment that are very similar to those in μ_{den}. These results imply that a comparison of results from different literature references is only meaningful when using identical geometries, and that the use of tight convergence criteria in geometry optimizations might be particularly important when using wave function-based population analysis methods such as Mulliken and NPA.

Conclusions

1. The Mulliken population analysis shows the largest variation of atomic charges with changes in the basis set or the Hamiltonian. The variations are larger when using Pople basis sets compared to the more systematically constructed cc-pVXZ or aug-cc-pVXZ basis sets. However, even with this latter choice there is no convergence of the results with increasing basis set size.
2. A very narrow dispersion of results is obtained when using ESP methods in combination with either cc-pVXZ or aug-cc-pVXZ basis sets at practically every quantum mechanical level. The results are more or less the same for the three ESP methods studied here.
3. The charges predicted by the AIM and NPA schemes are significantly larger than those predicted by the ESP schemes and overestimate the effective dipole moment by 35% (NPA) and 65% (AIM) at the MP2 level. The differences are even larger at the Hartree–Fock level. Similar observations have previously been made using Slater-type basis functions in combination with the BP86 density functional method.[13] This latter study traced the NPA results back to the far-reaching tails of atomic basis functions, while the excessive AIM charges appear to result from the particular definition of the atomic region as well as the direct integration of electron densities (instead of using electron density differences).
4. The minimal Pople basis sets STO-3G and 3-21G (and to a certain extend also the 6-31G basis set) predict a charge distribution, which is dramatically different from all other basis sets
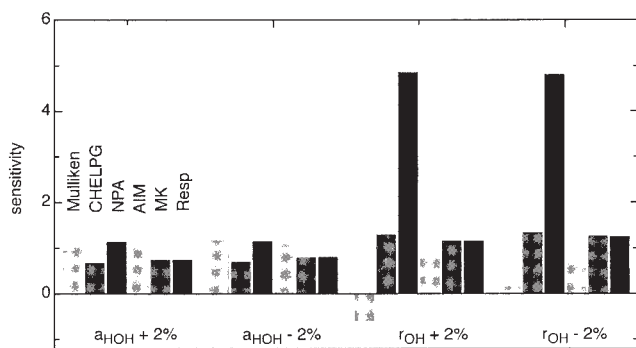
Figure 6. Sensitivity of the dipole moment to changes in the geometry of water.

used. This is true even for the most robust population analysis methods such as ChelpG and Resp.

5. In comparison to the experimentally measured dipole moment, the ESP-based methods overestimate the charge separation in water by approximately 3% using the largest aug-cc basis sets in combination with the MP2 or Becke3LYP methods. All other combinations of basis set, quantum mechanical method, and population analysis method lead to larger deviations from the experimental value.

6. The reduction of partial charges and the molecular dipole moment through the use of correlated methods is not constant, and depends on the basis set and the population analysis method. Using average values over all basis sets used here and concentrating on the results obtained at MP2 level, the largest effect is found for the Mulliken scheme (−7%), followed by the CHELPG and AIM methods (−5% and −4%), and the NPA scheme (−1%).

7. The effects of structural perturbations on the calculated charge distribution and the molecular dipole moment depend significantly on the type of population analysis, the wave function-based methods being particularly sensitive to changes in bond distances.

For the calculation of partial atomic charges in small, polar molecules it thus appears that the combination of any of the ESP methods with the larger members of the cc-basis set family and one of the three correlated quantum mechanical methods used here provides a reliable approach. This conclusion should be carefully reconsidered for much larger systems, for which the ESP methods are inherently less suitable, as well as for systems containing mainly apolar bonds, in which electron correlation effects may be more significant.

Acknowledgment

References

1. Jensen, F. Introduction to Computational Chemistry; Wiley: New York, 1999.
2. Cramer, C. J. Essentials of Computational Chemistry; Wiley: New York, 2002.
3. Cioslowski, J. Encyclopedia of Computational Chemistry; Wiley: New York, 1998, p. 892.
4. Benedict, W. S.; Gailar, N.; Plyer, E. K. J Chem Phys 1956, 24, 1139.
5. Darling, B. T.; Dennison, D. M. Phys Rev 1940, 57, 128.
6. Clough, S. A.; Beers, Y.; Klein, G. P.; Rothman, L. S. J Chem Phys 1973, 59, 2254.
7. Bicerano, J.; Marynick, D. S.; Lipscomb, W. N. J Am Chem Soc 1978, 100, 732.
8. Wiberg, K. B.; Rablen, P. R. J Comput Chem 1993, 14, 1504.
9. Kim, K.; Jordan, K. D. J Phys Chem 1994, 98, 10089.
10. Sigfridsson, E.; Ryde, U. J Comp Chem 1998, 19, 377.
11. Astrand, P.-O.; Ruud, K.; Mikkelsen, K. V.; Helgaker, T. J Phys Chem A 1998, 102, 7686.
12. Thompson, J. D.; Xidos, J. D.; Sonbuchner, M.; Cramer, C. J.; Truhlar, D. G. Phys Chem Commun 2002, 5, 117.
13. Fonseca Guerra, C.; Handgraaf, J.-W.; Baerends, E. J.; Bickelhaupt, F. M. J Comp Chem 2004, 25, 189.
14. Mulliken, R. S. J Chem Phys 1955, 23, 1833; 1962, 36, 3428.
15. Reed, A. E.; Weinstock, R. B.; Weinhold, F. A. J Chem Phys 1985, 83, 735.
16. Reed, A. E.; Curtiss, L. A.; Weinhold, F. Chem Rev 1988, 88, 899.
17. Bader, R. F. W. Atoms in Molecules, A Quantum Theory; Clarendon Press: Oxford, 1993.
18. Chipot, C.; Maigret, B.; Rivail, J.-L.; Scheraga, H. A. J Phys Chem 1992, 96, 10276.
19. Angyan, J. G.; Chipot, C. Chem Phys Lett 1995, 241, 51.
20. Breneman, C. M.; Wiberg, K. B. J Comput Chem 1990, 11, 361.
21. Sing, U. C.; Kollman, P. A. J Comput Chem 1984, 5, 129.
22. Besler, B. H.; Merz, K. M.; Kollman, P. A. J Comput Chem 1990, 11, 431.
23. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. J Phys Chem 1993, 97, 10269.
24. Hehre, W. J.; Stewart, R. F.; Pople, J. A. J Chem Phys 1996, 51, 2657.
25. Binkley, J. S.; Pople, J. A.; Hehre, W. J Am Chem Soc 1980, 102, 939.
26. Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. J Am Chem Soc 1982, 104, 2797.
27. Hehre, W. J.; Ditchfield, R.; Pople, J. A. J Chem Phys 1972, 56, 2257.
28. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. J Chem Phys 1980, 72, 650.
29. Dunning, T. H., Jr. J Chem Phys 1989, 90, 1007.
30. Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. J Chem Phys 1992, 96, 6769.
31. Wilson, A.; Mourik, T. v.; Dunning, T. H., Jr. J Mol Struct (Theochem) 1997, 388, 339.
32. Becke, A. D. J Chem Phys 1992, 96, 2155; 1993, 98, 5648.
33. Lee, C.; Yang, W.; Parr, R. G. Phys Rev B 1988, 37, 785.
34. Moeller, C.; Plesset, M. S. Phys Rev 1934, 46, 618.
35. Pople, J. A.; Seeger, R.; Krishnan, R. Int J Chem Symp 1977, 11, 149.
36. Bartlett, R. J. J Annu Rev Phys Chem 1981, 32, 359.
37. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vrevon, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Menucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomparts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Namayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, revision B.03; Gaussian, Inc.: Pittsburg, PA, 2003.
38. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179.
39. Dunning, T., Jr. In Encyclopedia of Computational Chemistry; Wiley: New York, 1998, p. 88.
40. Martin, J. M. L. In Encyclopedia of Computational Chemistry; Wiley: New York, 1998, p. 115.
41. Xantheas, S. S.; Dunning, T. H., Jr. J Chem Phys 1993, 99, 8774.