Theor Chem Acc (2009) 123:113–118
DOI 10.1007/s00214-009-0554-8

REGULAR ARTICLE

Analysis of the basis set and correlation effects on the computation of molecular polarizabilities using molecular polarization maps

Lurdes Roset · Jaime Rubio-Martinez · Juan J. Perez

Received: 25 February 2009 / Accepted: 26 February 2009 / Published online: 17 March 2009

Abstract In the present report, we analyze the effect of augmenting the basis set and the level of the theory in the computation of the molecular polarization potential and it is used as a tool to understand their spatial effects in the calculation of molecular polarizabilities. This is analyzed in the present report through the results obtained for three reference molecules: hydrogen cyanide, formaldehyde and urea. Accordingly, different basis sets were used for the calculations, including: 6-31G(D), 6-31G(D,P), 6-31G(2D,2P) and 6-31++G(2D,2P). On the other hand, calculations at different levels of the theory were considered, including: Hartree–Fock, Moller–Plesset second order (MP2) and fourth order (MP4) as well as the functionals BLYP and B3LYP at the density functional theory. The results of the polarization maps reveal similar trends observed with the calculation of the different components of the polarizability tensor. However, the use of the maps permits to understand spatial effects specific for each level of calculation and each basis set.

Keywords Polarization maps · Polarizability · Polarization · Basis set effect

---

Dedicated to Professor Santiago Olivella on the occasion of his 65th birthday and published as part of the Olivella Festschrift Issue.

L. Roset · J. J. Perez (✉)
Department of Chemical Engineering,
Technical University of Catalonia, ETS d'Enginyeria Industrial,
Av. Diagonal, 647, 08028 Barcelona, Spain
e-mail: juan.jesus.perez@upc.edu

J. Rubio-Martinez
Department of Physical Chemistry,
The Institut de Recerca en Química Teòrica i Computacional
(IQTCUB), University of Barcelona,
Martí i Franqués 1, 08028 Barcelona, Spain

1 Introduction

Molecular polarization potential maps (MPPs) [1–6] account for the spatial distribution of deformability of the electronic cloud of a molecule and can be considered as a correction to the molecular electrostatic potential maps (MEPs) [1, 7–9]. However, whereas the latter are graphical representations of the first order (coulombic) contribution of the interaction energy between a point charge and a molecule, the former represent the second order (induced) contribution to the interaction energy of a molecule in the same conditions. The information provided by MPPs at the van der Waals surface of the molecule and longer distances complements the picture provided by MEPs [1, 6] and may reveal new features in the description of the interaction energy between two molecules.

Molecular polarization potential maps were introduced as a tool for analyzing the early stages of the drug–receptor interaction [1, 8] and since then a large number of papers have been written devoted to describe its use in studies concerning different aspects of molecular recognition [9]. The success of these maps as a predictive tool is based in two factors: on the one hand, the importance of the electrostatic contribution to the interaction energy between two molecular systems and on the other, to the fact that they provide a picture of the interactions not attainable from a multipole expansion analysis. Despite their use underscoring the most important features involving molecular recognition between two molecules at long distances, polarization contributions at short distances may become equally important [9] and these effects can be properly accounted through MPPs [1–6]. Accordingly, MPPs have been used in the past to explain specific features of molecular interactions, besides being used to generate new force field parameters [10].

Alternatively, MPPs can also be used as a tool for understanding the effects of the basis set and method used in the evaluation of molecular polarizabilities. In fact, despite being known that in order to compute reasonable estimates of molecular polarizabilities a set of well tempered basis set including polarization and diffuse functions is necessary [11], the spatial distribution due to the increase of the basis set or due to the use of a different evaluation procedure has never been analyzed. Accordingly, we report in this work the analysis carried out on three small molecules: hydrogen cyanide, formaldehyde, and urea used as benchmark in a previous study [4].

2 Methods

Molecular polarizabilities were estimated by computing the analytical derivatives of the energy as described elsewhere [12]. For the calculation of the MPPs, molecules were placed in the *YZ* plane and polarization maps were computed at $x = 4$ bohr from the molecular plane. Grid points were computed every bohr and extended approximately 6 bohr beyond the van der Waals surface of the molecules. Since the molecules selected have a symmetry plane $\sigma_{s,}$ only one half of the points of the grid needed to be computed. The total number of points calculated was 231 points for hydrogen cyanide, 242 for formaldehyde and 312 for urea. The value of the polarization at every point of the grid was computed using the finite field method with a probe charge of 0.1 a.u. For this purpose at each point, the scaled value of the electrostatic potential (V) was subtracted from the interaction energy of the molecule with the probe charge (ΔE). From these values the polarization potential (P) was easily obtained, from the following expression:

$$\Delta E(r) = eV(r) + e^2 P(r)$$

Molecular polarization potential maps were computed with three different basis sets and using different approximations. The molecular geometry was optimized for each basis set and method used. Basis sets selected for the present study include the standard 6-31G(d) [13], 6-31G(d,p) [13], 6-31 (2d,2p) [14] basis sets with polarization functions and the 6-311++G(2d,2p) [15] that includes diffuse functions. Calculation were done at the Hartree–Fock, second order Moller–Plesset (MP2), fourth order Moller–Plesset (MP4) levels [16] and using the functional density theory, with the BLYP [17] and B3LYP [18] functional. All the calculations were performed with the Gaussian94 package [19].

3 Results and discussion

In Table 1 are listed the estimates of the components of the polarizability tensor for the three molecules studied in the present work, together with the experimental results available. As can be seen in Table 1a–c, a second set of polarization functions improves all the components of the tensor around 10%, but only for the out-of-the *YZ* plane components. In contrast, the use of diffuse functions is indispensable to obtain reliable estimates with increases of around 30% in regard to the calculations carried out with the basis sets augmented with two sets of polarization functions, although this improvement is smaller for the YZ plane components for formaldehyde and urea. Present results confirm also that DFT reproduces well the results obtained at the ab initio level at the MP4 level, although they are about 5% overestimated in the case of the basis set augmented with diffuse functions (Fig. 1).

Molecular polarization potential maps provide the spatial distribution of these trends. MPPs of these molecules were described in a previous report [4]. Figure 2a–c show the MPP maps of the three molecules studied in the present study at the MP4 level with the 6-311++(2d,2p) basis set. In contrast to MEPs, MPPs do not present any stationary point, but the separation between level curves inform of the regions where the molecule is more polarizable. Thus, looking at the level curves, we can see that in formaldehyde the oxygen is more polarizable in the directions of its non-bonding pair and in the path that intersects the two hydrogens. In the case of hydrogen cyanide, the map shows that the molecule exhibits a similar polarizable pattern in any direction. Finally, in the case of urea it exhibits as more polarizability paths in the regions between the carbon and the oxygen as well as the hydrogen bisecting directions. The different basis sets and methods used in the present work reproduce well the main features of these maps.

Whereas qualitative differences found among different computing schemes carried out in this work are small, quantitative differences are noticeable. We have analyzed the effect of correlation in the description of the maps and the effect of the augmentation of the basis set. Improvement on the calculations follows the same trends as found on the estimates of the polarizability tensor shown in Table 1. Figure 2 shows the correlation between the MPPs of urea computed with the basis set at the MP4 level with other levels of calculation. Compared with the Hartree–Fock and MP2 all the points of the map behave similarly, independent of their value. Moreover, it can be seen that differences with MP2 are not noticeable and with the calculations carried out at the Hartree–Fock level, differences are around 2%. Interestingly, it had been previously

Table 1 Estimates of the polarizability tensor for the different molecules reported in the present work

	6-31G(d)	6-31G(d,p)	6-31G(2d,2p)	6-311 ++ G(2d,2p)
(a) *Results for hydrogen cyanide, with an average experimental value of 17.5 a.u.* [21]				
RHF	*xx* 7.47	7.53	8.70	11.36
	yy 7.47	7.53	8.70	11.36
	zz 19.52	19.64	19.68	26.12
	<α> 10.48	**11.56**	**12.36**	**16.28**
MP2	*xx* 7.49	7.55	9.35	11.51
	yy 7.49	7.55	9.35	11.51
	zz 18.28	18.41	18.61	26.34
	<α> 11.09	**11.17**	**12.44**	**16.45**
MP4	*xx* 7.62	7.70	9.42	11.44
	yy 7.62	7.70	9.42	11.44
	zz 18.81	18.89	18.85	26.99
	<α> 11.34	**11.43**	**12.63**	**16.62**
BLYP	*xx* 7.73	7.79	9.52	12.22
	yy 7.73	7.79	9.52	12.22
	zz 19.22	19.36	19.41	27.01
	<α> 11.56	**11.64**	**12.82**	**17.15**
B3LYP	*xx* 7.66	7.71	9.47	11.88
	yy 7.66	7.71	9.47	11.88
	zz 19.15	19.28	19.30	26.49
	<α> 11.49	**11.57**	**12.75**	**16.75**
(b) *Results for formaldehyde with an average experimental value of 16.5 a.u.* [22]				
RHF	*xx* 6.48	6.77	8.73	10.48
	yy 12.92	13.04	14.04	15.23
	zz 17.64	17.81	17.93	21.51
	<α> 12.35	**12.54**	**13.57**	**15.74**
MP2	*xx* 6.45	6.65	8.71	10.71
	yy 13.56	13.66	14.77	15.24
	zz 17.40	17.52	17.84	21.56
	<α> 12.48	**12.61**	**13.78**	**15.89**
MP4	*xx* 7.50	6.92	8.80	10.86
	yy 13.71	14.02	14.82	16.54
	zz 17.42	17.62	17.98	22.09
	<α> 12.88	**12.85**	**13.87**	**16.50**
BLYP	*xx* 6.76	6.96	8.84	11.17
	yy 14.16	14.31	15.43	17.90
	zz 17.95	18.09	18.32	23.58
	<α> 12.96	**13.12**	**14.20**	**17.55**
B3LYP	*xx* 6.69	6.89	8.80	10.99
	yy 13.74	13.84	14.97	16.96
	zz 17.74	17.69	18.10	22.72
	<α> 12.72	**12.89**	**13.96**	**16.87**
(c) *Results for urea, with an experimental value of 31.9 a.u* [23]				
RHF	*xx* 11.30	12.16	13.01	22.61
	yy 28.12	28.66	28.92	32.47
	zz 30.90	31.42	32.60	35.37
	<α> 23.44	**24.08**	**24.84**	**30.15**

Table 1 continued

	6-31G(d)	6-31G(d,p)	6-31G(2d,2p)	6-311 ++ G(2d,2p)
MP2	*xx* 12.06	12.31	13.69	25.14
	yy 29.11	29.52	29.62	33.45
	zz 30.84	31.24	31.82	36.72
	<α> 24.01	**24.36**	**25.04**	**31.77**
MP4	*xx* 12.16	12.42	13.70	25.17
	yy 30.24	30.67	30.91	34.50
	zz 30.82	31.82	31.85	36.98
	<α> 24.41	**24.97**	**25.48**	**32.22**
BLYP	*xx* 12.25	12.51	14.71	26.45
	yy 31.89	32.30	30.73	40.77
	zz 32.15	32.49	32.28	42.10
	<α> 25.43	**25.77**	**25.91**	**36.44**
B3LYP	*xx* 12.16	12.41	14.70	25.13
	yy 30.26	30.67	28.99	38.01
	zz 31.05	31.40	32.19	39.91
	<α> 24.49	**24.83**	**25.29**	**34.35**

Values in every cell correspond to the different tensor components (*xx*, *yy*, *zz*) followed by the average value (in bold), computed as $<\alpha> = (\alpha_{xx} + \alpha_{yy} + \alpha_{zz})/3$. All values are in atomic units ($1e^{2a^{2}}{}_{0}E_{h}^{-1} = 1.648778 \times 10^{-41}\mathrm{C}^{2}\,\mathrm{m}^{2}\,\mathrm{J}^{-1}$)

reported that the effect of correlation does not affect the MEPs considerably [20]. In contrast, in the case of the DFT calculations, it can be seen that the correlation with MP4 calculations is poor. Specifically, in the case of using the B3LYP functional, there are a series of points that are not properly computed and are distributed at different values. In contrast, the BLYP functional exhibits a better behavior with points that differ in regard to the reference calculation found at values close to zero.

Spatial distributions of differences provides information about the areas that are described differentially. For this purpose, relative difference maps of the polarization computed with two different methods or basis sets were computed. Each point of the grid was computed using the formula:

$$\varepsilon_i = \frac{\{P(\text{method A})i - P(\text{method B})i\}}{P(\text{method A})i} \times 100$$

where P is the polarization at point i of the grid computed with one of the basis sets and with a specific method. The extent of the relative deviation between the two different methods is measured by the average values of ε for all points of the grid. Figure 3a–c shows the effect of augmenting the basis set in a series of calculations carried out con urea at the MP4 level and using as reference the 6-311++(2d,2p) basis set. As can be seen in Fig. 3a, differences with the basis that includes two sets of polarization functions are around 10% in the areas close to the nuclei and increase up to a 20% in the outermost regions of the molecule. Interestingly, when we compare the differences in the computation of the polarizability as shown in Table 1, it can be clearly concluded that there is the need of a proper description of this region for a proper calculation of some of the components of the polarizability specially the out-of-the *YZ* plane component. However, when we compare the reference results with the basis set that includes one set of polarization functions (Fig. 3b, c), differences are larger, but more interestingly, the area close to the nuclei is not properly computed. Indeed, the map shows positive deviations in the region close to the nuclei, like if they were overestimated and negative differences in the outermost regions of the molecule with nearly zero deviations on top of the middle of the different bonds. The same trends are observed in the map obtained with the smaller basis set.

4 Conclusions

The aim of the present report is to assess the spatial distribution of the effects observed when using different basis sets and at different levels of the theory for the computation of molecular polarizabilities. For this purpose, we computed the molecular polarization potential of three small molecules: hydrogen cyanide, formaldehyde, and urea. The results show that the inclusion of a second set of polarization functions in the basis set increases about 10% the values of the polarizabilities and this improvement is

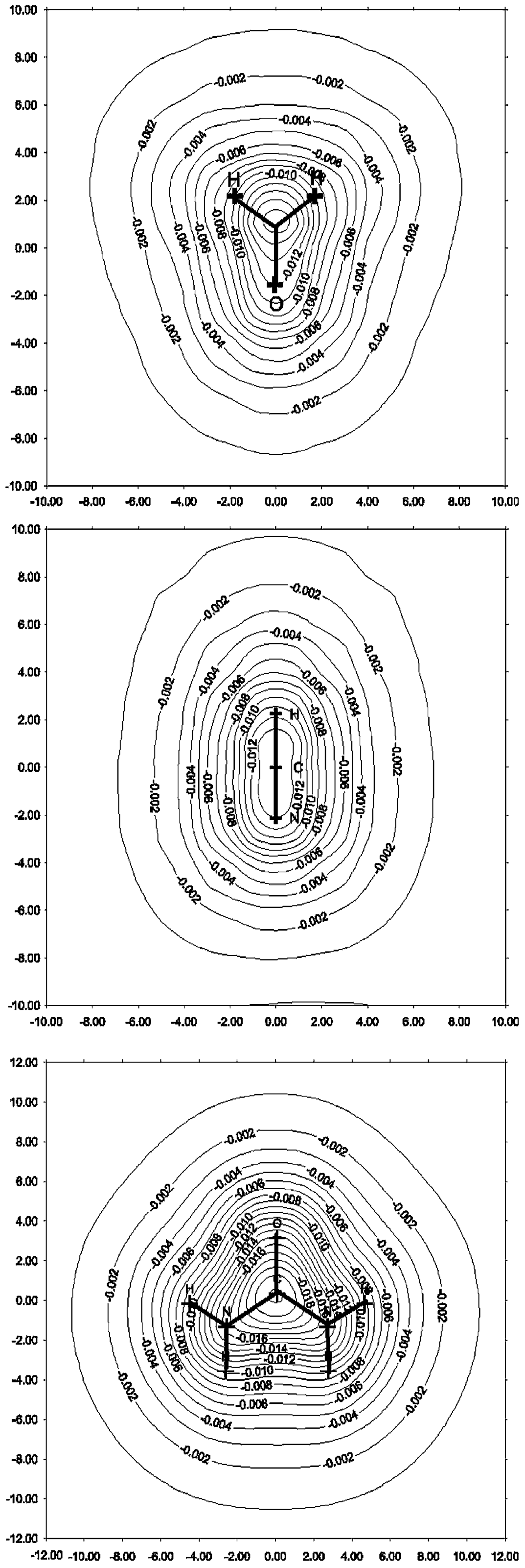

Fig. 1 Molecular polarization maps computed with a 6-311++G(2D,2P) basis set at the MP4 level of calculation. All the results are in a.u.: **a** formaldehyde, **b** hydrogen cyanide, **c** urea

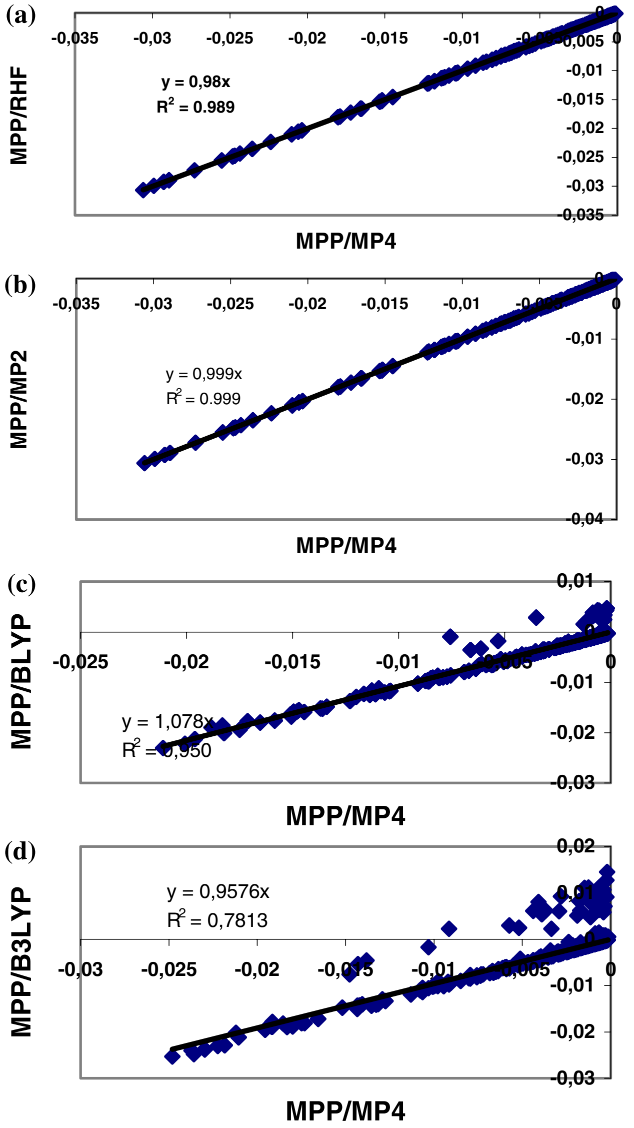

Fig. 2 Comparison of the values of the molecular polarization maps computed at different levels of theory using a 6-311++G(2D,2P) basis set and using the MP4 as reference. All the results are in a.u.: **a** Hartree–Fock, **b** MP2, **c** BLYP, **d** B3LYP

basically located in the most polarizable regions like lone pairs and bonds. This has a parallel in the evaluation of the different components of the tensor. On the other hand, basis augmentation with diffuse functions is critical for an accurate description of the less polarizable directions of the molecule. Increments of about 30% are obtained and mostly located in the outermost regions of the molecule. Inclusion of the electronic correlation does not alter much the spatial distribution of the molecular polarization maps, although there is a difference of around 2% in the values. In contrast, calculations carried out at the DFT level exhibit a poor correlation compared to MP4 calculations, with

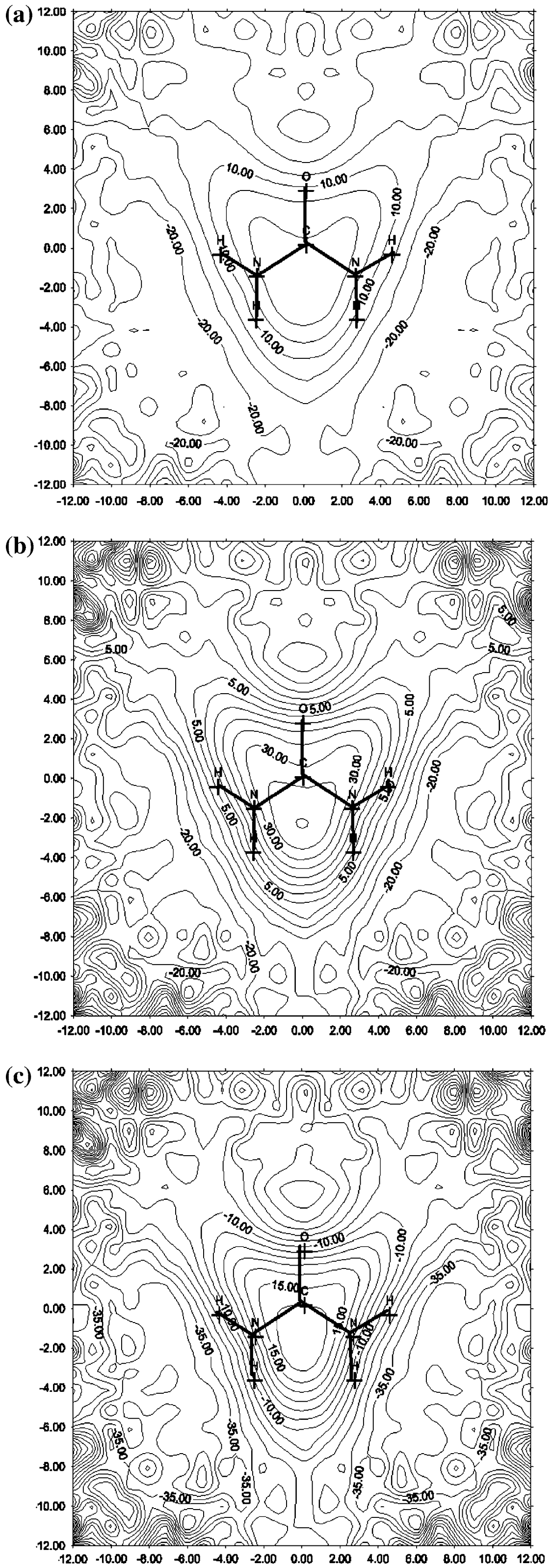

Fig. 3 Maps of the difference (in percentage) of molecular polarization maps for urea, using the 6-311++G(2D,2P) basis at the MP4 level of the theory as reference calculation: **a** comparison with 6-31G(2D,2P), **b** comparison with 6-31G(D,P), **c** comparison with 6-31G(D)

the BLYP functional exhibiting a better behavior. Interestingly, deviations are preferentially located in regions where the polarization values gets close to zero.

Acknowledgments The Barcelona Supercomputer Center is acknowledged for a generous allocation of computer time. The Spanish Ministry of Science and Technology supported this work through the projects CTQ2006-06588/BQU to J.R.M. and SAF2005-08148-C04-01 to J.J.P.

References

1. Bonaccorsi R, Scrocco E, Tomasi J (1976) Theor Chim Acta 43:63. doi:10.1007/BF00548957
2. Francl MM (1985) J Phys Chem 89:428. doi:10.1021/j100249a012
3. Alkorta I, Perez JJ, Villar HO (1994) J Mol Graph 12:3. doi:10.1016/0263-7855(94)80002-2
4. Alkorta I, Villar HO, Perez JJ (1993) J Phys Chem 97:9113. doi:10.1021/j100138a008
5. Alkorta I, Perez JJ (1996) Int J Quantum Chem 57:123. doi:10.1002/(SICI)1097-461X(1996)57:1<123::AID-QUA14>3.0.CO;2-9
6. Alkorta I, Perez JJ (1995) Electron J Theor Chem 1:26. doi:10.1002/ejtc.6
7. Scrocco E, Tomasi J (1973) In: Topics in current chemistry, vol 42, Springer, Berlin
8. Weinstein H (1975) Int J Quantum Chem QBS2:5969
9. Politzer P, Murray JS (1991) In: Lipskowitz KB, Boyd DB (eds) Reviews of computational chemistry vol 2. VCH Publishers, New York
10. Alkorta I, Bachs M, Perez JJ (1994) Chem Phys Lett 224:160. doi:10.1016/0009-2614(94)00512-5
11. Sadlej AJ (1991) Theor Chim Acta 81:339. doi:10.1007/BF01118573
12. Rice JE, Amos RD, Colwell SM, Handy NC, Sanz J (1990) J Chem Phys 93:8828. doi:10.1063/1.459221
13. Hariharan PC, Pople JA (1973) Theor Chim Acta 28:213. doi:10.1007/BF00533485
14. Frisch MJ, Pople JA, Binkley JS (1984) J Chem Phys 80:3265. doi:10.1063/1.447079
15. Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PR (1983) J Comput Chem 4:294. doi:10.1002/jcc.540040303
16. Hinchliffe A (1996) Modelling Molecular Structures. Wiley and Sons, Chichester
17. Becke AD (1988) Phys Rev A 38:3098. doi:10.1103/PhysRevA.38.3098
18. Becke AD (1993) J Chem Phys 98:5648. doi:10.1063/1.464913
19. Frisch MJ, Trucks G W, Schlegel HB, Gill P MW, Johnson BG, Robb MA, Cheeseman JR, Keith T, Peterson GA, Montgomery JA, Raghavachari K, Al-Laham MA, Zakrzewski VG, Ortiz JV, Foresman JB, Peng CY, Ayala PY, Chen W, Wong MW, Andres JL, Replogle ES, Gomperts R, Martin RL, Fox J, Binkley JS, Defrees DJ, Baker J, Stewart JJP, Head-Gordon M, Gonzales C, Pople JA (1995) Gaussian94, Revision B 2. Gaussian Inc, Pittsburgh PA
20. Luque FJ, Orozco M, Illas F, Rubio J (1991) J Am Chem Soc 113:5203. doi:10.1021/ja00014a010
21. McClellan AL (1963) Tables of experimental dipole moments. Freeman, San Francisco
22. Landolt-Bornstein (1951) Atom und Molecularphysik, vol 1 part 3, Springer Berlin
23. Applequist J, Carl JR, Fung KK (1972) J Am Chem Soc 94:2953