Theor Chem Acc (2000) 104:417–425
DOI 10.1007/s002140000166

Theoretical
Chemistry Accounts

Regular article

Dipole moments and molecular electrostatic potentials from MSINDO

Karl Jug, Lars Kunert, Andreas M. Köster

Theoretische Chemie, Universität Hannover, Am Kleinen Felde 30, 30167 Hannover, Germany

Received: 1 October 1999 / Accepted: 15 February 2000 /

Abstract. To achieve an improved description of solvation effects in the newly derived version MSINDO of the SINDO semiempirical formalism an appropriate parameterization of dipole moments and electrostatic potentials was undertaken. The mean error of the dipole moment could be reduced by an appropriate choice of orbital exponents. It is also shown that the approximation of the molecular electrostatic potential (MESP) in the asymptotic density model (ADM) implemented in MSINDO results in a superior description compared to the previous SINDO1 implementation. The accuracy is demonstrated for a selected number of small molecules with carbon, nitrogen and oxygen atoms.

Key words: Dipole moments – Molecular electrostatic potential

1 Introduction

For a realistic description of organic systems the inclusion of solvation effects is desirable, if not unavoidable. Whereas early quantum chemical descriptions frequently neglected such solvation effects, more and more approaches and methods have evolved in the past two decades which explicitly include such effects in the study of molecular structure and reactions. For such systems experiments in solution are the standard and gas-phase data are frequently not available. For the description of solution effects a variety of methods are available: Monte Carlo methods [1, 2], molecular mechanics methods [3] and continuum models [4] are frequently used. In continuum models the solvent is considered as a homogeneous dielectric. Only macroscopic properties of the solvent are introduced into calculations. The solute molecule is located in a cavity of the dielectric. The various continuum models differ in the form of the cavities and the charge distribution of the molecules. Depending on the model the resulting electrostatic equations can be solved analytically [5–8] or must be solved numerically [9]. We found the latter approach, implemented in the polarizable continuum model (PCM) attractive, because it had been introduced to semiempirical methods [10–12] and because we intended to introduce isodensity surfaces as cavity surfaces which were to be calculated numerically.

The basis for our structure and energy calculations was the semiempirical molecular orbital method SINDO1, which was developed for first-row [13], second-row [14] and third-row [15] elements. The PCM model was implemented into SINDO1 in combination with the asymptotic density model (ADM) [16–18] for the molecular electrostatic potential (MESP). The latter reduces the amount of computer time considerably, while retaining the quality of the MESP. Several applications evolved, notably the calculation of solvation energies [19], which was achieved based on isodensity surfaces for the cavity surface and on the ADM. Further applications included solvation effects on diradicals and zwitterions of tetramethylenes and trimethylenes [20] and treatment of reactions in solutions with isodensity surfaces [21]. In the implementation of the PCM in SINDO1 a parameterization of the MESP and a scaling of the atomic charges for the generation of the MESP was necessary to achieve high accuracy of the solvation energies. We have recently modified the method SINDO1. The new modified version is called MSINDO [22–24] and has a substantially improved accuracy for geometries, heats of formation, ionization energies and dipole moments of compounds with first- and second-row elements. The goal of the present article is to present an additional parameterization of MSINDO with respect to solvation effects. We want to demonstrate that this parameterization is capable of reproducing the basic features of the MESP of some compounds with first-row elements better than SINDO1 despite the fact that SINDO1 was parameterized to give good solvation energies. A similar philosophy underlies the charge models of the SM4 formalism by Giesen et al. [25]. These au-

---

Present address: A. M. Köster, Departamento de Quimica, CINVESTAV, 7000 Mexico DF, Mexico
Correspondence to: K. Jug

thors have fitted their charge model to experimental dipole moments and high-level ab initio partial charges.

2 Parameterization of dipole moments

For our generation of the MESP the idea of cumulative atomic multipole moments (CAMMs) [26, 27] is essential. It is therefore desirable to generate improved dipole moments by an effective and simple parameterization in order to keep the expansion in the CAMMs as short as possible, i.e. up to the cumulative dipole moments.

The dipole moment, $\boldsymbol{\mu}$, of a molecule can be written as the sum of nuclear and electronic contributions

$$\boldsymbol{\mu} = \sum_A Z_A \mathbf{R}_A - \sum_{\mu\nu} \mathbf{P}_{\mu\nu}\, \mathbf{r}_{\mu\nu} \ . \tag{1}$$

Here Z_A and $\mathbf{R}_A$ are the nuclear charges and coordinates, respectively. $\mathbf{P}$ and $\mathbf{r}$ are the density matrix and the dipole matrix over atomic orbitals (AOs) μ and ν. In semiempirical methods the AOs are usually considered as an orthogonalized set $\{\lambda\}$ [28]. In this way Eq. (1) can be rewritten as

$$\boldsymbol{\mu} = \sum_A Z_A \mathbf{R}_A - \sum_A \sum_{\mu\nu}^{A} \mathbf{P}^{\lambda}_{\mu\nu}\, \mathbf{r}_{\mu\nu} = \sum_A \boldsymbol{\mu}^{\lambda}_A \ , \tag{2}$$

with $\boldsymbol{\mu}^{\lambda}_A = Z_A \mathbf{R}_A - \sum_{\mu\nu}^{A} \mathbf{P}^{\lambda}_{\mu\nu} \mathbf{r}_{\mu\nu}$.

Here $\mathbf{r}_{\mu\nu}$ is calculated over the usual AO's of s, p and d type. The two-center terms are neglected due to the zero differential overlap (ZDO) assumption. Alternatively, the dipole moment could be written as the sum over contributions from nonorthogonal orbitals $\{\chi\}$ with the Mulliken partitioning [29].

$$\boldsymbol{\mu} = \sum_A Z_A \mathbf{R}_A - \sum_{\mu\nu} \mathbf{P}^{\chi}_{\mu\nu}\, \mathbf{r}_{\mu\nu} = \sum_A \boldsymbol{\mu}^{\chi}_A \ , \tag{3}$$

with $\boldsymbol{\mu}^{\chi}_A = Z_A \mathbf{R}_A - \sum_{\mu}^{A} \sum_{\nu} \mathbf{P}^{\chi}_{\mu\nu}\, \mathbf{r}_{\mu\nu}$.

Here two-center terms of $\mathbf{r}_{\mu\nu}$ are present. In SINDO1 the form of Eq. (2) was used and in MSINDO Eq. (2) was also the preferred choice; however, MSINDO uses, as SINDO1 did, two sets of exponents, one (ζ^{U}) for the one-center atomic integrals and one (ζ) for the two-center molecular integrals. In order to avoid a conflict in the calculations of dipole moments over nonorthogonal orbitals which contain one-center and two-center terms, a new third set of exponents was generated for the dipole moments. For first- and second-row elements the exponents $\zeta^{\mathrm{D}}_s, \zeta^{\mathrm{D}}_{p\sigma}, \zeta^{\mathrm{D}}_{p\pi}$ are introduced for each element. The exponents were chosen in such a way that they would minimize the mean error of the calculated dipole moments with respect to experimental dipole moments for a set of test molecules. The transformation from the orthogonal to the nonorthogonal basis is performed with the $\mathbf{S}^{-1/2}$ matrix.

$$\mathbf{P}^{\chi} = \mathbf{S}^{-1/2}\, \mathbf{P}^{\lambda}\, \mathbf{S}^{-1/2} \ . \tag{4}$$

$\mathbf{S}^{-1/2}$ and $\mathbf{r}$ are calculated with the ζ^{D} exponents.

The exponents for the elements H, C, N and O are collected in Table 1. Since the contributions to the dipole moment are initially calculated in a local coordinate system for pairs of atoms, a distinction between $p\sigma$ and $p\pi$ is necessary. The exponents listed in Table 1 were optimized with respect to the mean error of the dipole moments of the compounds listed in Table 2. The comparison of the calculated and the experimental values [30, 31] shows that the MSINDO values are much better than the SINDO1 values. The MSINDO values with an orthogonal basis, i.e. the ZDO assumption, can be improved if a transformation to the nonorthogonal basis is made via the density matrix $\mathbf{P}^{\chi}$ according to Eq. (4) and if Eq. (3) is used instead of Eq. (2). In gen-

Table 1. Parameters for dipole moments

	ζ^{D}_s	$\zeta^{\mathrm{D}}_{p\sigma}$	$\zeta^{\mathrm{D}}_{p\pi}$
H	1.080	–	–
C	1.595	1.925	1.258
N	2.065	2.092	1.904
O	2.352	2.431	2.168

Table 2. Calculated and experimental dipole moments (debye)

Compound	SINDO1	MSINDO[a]	MSINDO[b]	Exp[c]
Propane	0.00	0.21	0.17	0.08
Propene	0.50	0.68	0.62	0.37
Propyne	0.52	1.09	0.84	0.78
Cyclopentadiene	0.31	0.65	0.64	0.42
Toluene	0.09	0.78	0.68	0.36[d]
Ammonia	2.20	2.36	1.67	1.47
Methylamine	2.01	1.98	1.29	1.31
Dimethylamine	1.83	1.74	1.21	1.03
Trimethylamine	1.63	1.41	0.92	0.63
Aniline	2.25	1.94	1.32	1.53[d]
Diazomethane	0.41	1.35	1.04	1.5
Formamide	2.50	3.63	3.32	3.73[d]
Pyrrole	1.73	1.81	1.64	1.74
Pyridine	1.87	1.96	1.53	2.22
Hydrogen cyanide	2.36	2.49	2.65	2.98
Acetonitrile	0.88	3.68	3.51	3.92[d]
Nitrous oxide	1.47	0.59	0.81	0.16
Nitrous acid	1.68	2.22	1.96	1.43
Nitric acid	2.51	2.67	2.20	2.17
Water	1.94	2.42	1.80	1.85
Methanol	1.67	1.92	1.46	1.70
Phenol	2.37	1.99	1.19	1.22[d]
Dimethyl ether	0.85	1.53	1.17	1.30
Furan	0.44	0.71	0.30	0.66
Carbon monoxide	0.58	0.26	0.25	0.11
Formaldehyde	1.52	2.09	2.12	2.33
Ketene	0.13	1.45	1.49	1.42
Acetaldehyde	2.14	3.12	2.95	2.75
Acetone	2.43	3.65	3.34	2.88
Acrolein	2.03	3.03	2.99	3.12
Propynal	1.99	2.66	2.73	2.74
Formic acid	0.79	1.06	1.20	1.41
Acetic acid	1.08	1.85	2.06	1.74
Methyl formate	1.09	1.65	1.65	1.77
Mean error	0.66	0.35	0.24	

[a] Orthogonal basis
[b] Nonorthogonal basis, parameters from Table 1
[c] Ref. [30]
[d] Ref. [31]

eral, we observe that the calculated MSINDO dipole moments of compounds with a lone pair, for example, ammonia, aniline or water are too large with the ZDO approximation and are reasonably reduced in the nonorthogonal basis. In contrast, compounds with a lone pair and a multiple bond at the same center, for example cyanic acid, formaldehyde or pyridine are too small in both variants of the MSINDO dipole moments. We have tried other combinations of exponents and scaling, but the improvements were not substantial.

Our results compare favorably with recent results by Li et al. [32, 33]. The class IV charge model, called charge model 1 (CM1) involves a semiempirical mapping of Mulliken charges with the parameters optimized to reproduce experimental dipole moments. The new class IV charge model, charge model 2 (CM2) has been successfully implemented for Hartree–Fock calculations and some semiempirical methods.

In the first article the accuracy of ground-state dipole moments calculated using INDO/S and INDO/S2 wavefunctions is presented. For a large number of compounds containing H, C, N, O, F, Si, P, S, Cl, Br and I atoms the root-mean-square errors of the CM2 method were 0.24 and 0.26 D respectively. This is the accuracy achieved in Table 2 of our work. The dipole moments calculated by the Mulliken analysis and density showed errors between 0.62 to 0.74. This is the same magnitude as for SINDO1 in Table 2. The second article reports dipole moments from Hartree–Fock calculations with correlation-consistent polarized-valence double-zeta basis sets, which lower the errors in the CM2 parameterization to 0.18 D.

3 Parameterization of the MESP

The MESP is an important indicator for intra- and intermolecular interactions. In particular, for the latter it can help to describe the perturbation of a molecular system by another molecular system. In the past this aspect had been emphasized in the literature. For an approximation of the MESP a multipole expansion was usually chosen, which resulted in a reasonable description of the MESP in the outer region around a molecule. A few years ago we suggested an alternative description called the ADM [16], which was based on the idea that an approximation of the MESP should be quite accurate close to the nuclei and far away from the nuclei. This would also help to maintain an acceptable approximation in the bonding region of a molecule. We repeat here those portions of the theory related to the parameterization. Since the CAMMs of the molecule are necessary components for the generation of the MESP, we suggested an approximate density for each atom A in the molecule in the following form

$$\rho_A^{\mathrm{ADM}}(\mathbf{r}_A) = c_0 \exp(-\alpha_0 r_A) Y_{00}(\vartheta_A, \varphi_A) + \sum_{m=-1}^{1} c_{1m} r_A \exp(-\alpha_1 r_A) Y_{1m}(\vartheta_A, \varphi_A) \ . \quad (5)$$

The first term describes the spherical part and the second simulates the polarization of an atom in a molecule. In principle, they reflect the monopole and dipole terms of a charge distribution. This atomic density, ρ_A, is related to the corresponding atomic portion, U_A, of the MESP by the Poisson equation

$$\nabla^2 U_A^{\mathrm{ADM}}(\mathbf{r}_A) = -4\pi \rho_A^{\mathrm{ADM}}(\mathbf{r}_A) \ . \quad (6)$$

The total MESP is the sum of all atomic MESPs.

$$U^{\mathrm{ADM}} = \sum_A U_A^{\mathrm{ADM}} \ . \quad (7)$$

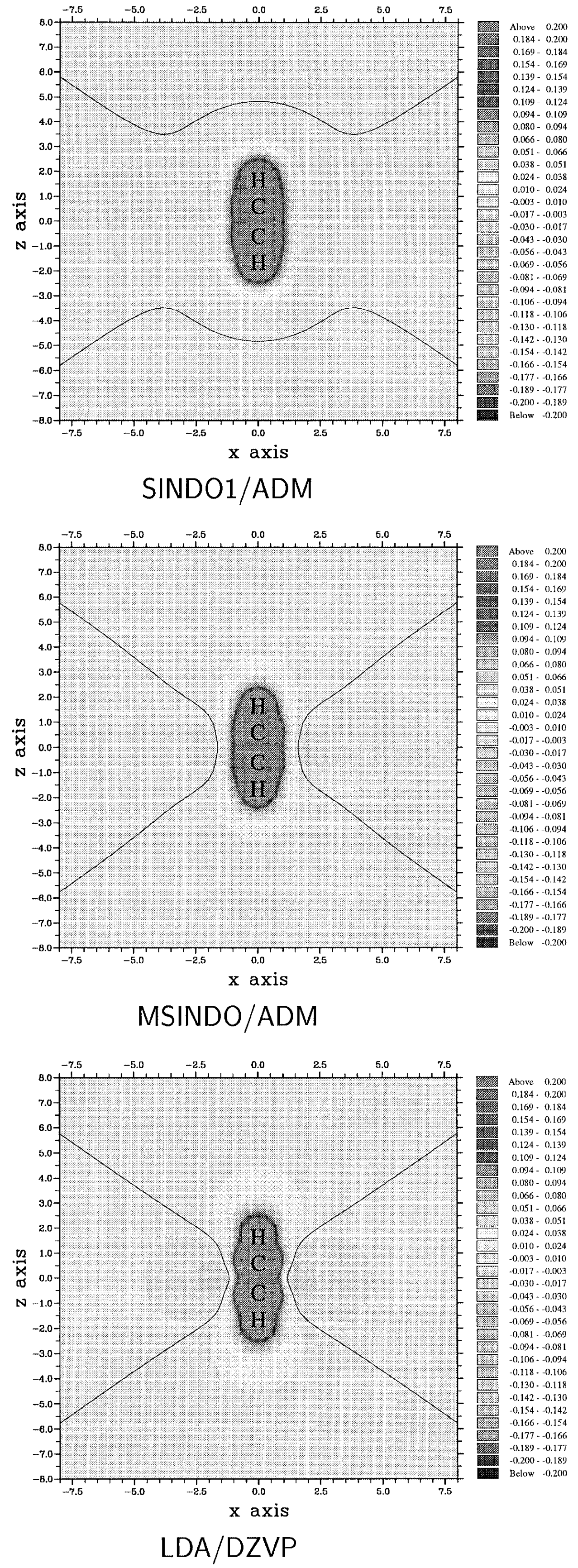
The exponents α_0 and α_1 of Eq. (5) are fixed in such a way that the approximate MESP of a series of test molecules reproduces the accurate MESP as well as possible. The implementation of the ADM in MSINDO follows the same lines as published previously [34]. The ADM parameters for MSINDO are collected in Table 3. Here α'_0 corresponds to α'_{00} and α_0 is calculated from Eq. (12) of the earlier article [34]. The CAMMs over nonorthogonal orbitals which refer to Eqs. (13) and (14) of the previous article [34] are calculated with the ζ^{D} exponents, whereas the corresponding expressions in Eqs. (15) and (16) over orthogonalized orbitals are calculated with the ζ^{U} exponents because only atomic integrals occur.

The MESP of a molecule can be characterized by its critical points. These are minima and saddle points. Much work was done on this topic [35–37]. We have contributed to this development by semiempirical [38] and density functional theory (DFT) [39] calculations.

In the present work we want to demonstrate that a highly accurate approximation of the essential features of the MESP can be achieved based on MSINDO molecular orbitals. For this purpose we selected the following test molecules: acetylene, ethylene, butadiene, benzene, ammonia, hydrazine, cyanic acid, water and formaldehyde. We took the dipole moments of column 3 of Table 2. The dipole moments of column 2 of Table 2 were not considered because of the lesser accuracy. We present the results in Figs. 1–9. SINDO1 and MSINDO calculations based on the ADM are compared with DFT calculations with the program AllChem [40]. In the DFT calculations the local density approximation (LDA) [41] is used with a double-zeta valence polarization basis set [42]. Figure 1 shows the MESP for acetylene: the critical points consist of a ring of minima around the triple bond. In the figure we see these minima to the left and to the right of the CC bond for the MSINDO and the LDA. A similar shape of the MESP is found in Fig. 2 for ethylene. Here there are two minima for the MSINDO and the LDA to the left and right of the CC bond. They describe the π system. In Fig. 3 we see the saddle point for the C_2C_3 bond of butadiene, again only for the MSINDO and the LDA. The π system

Table 3. Parameters for the asymptotic density model

	α'_0	α_1
H	1.40	0.80
C	1.35	0.95
N	1.35	0.60
O	1.20	1.00

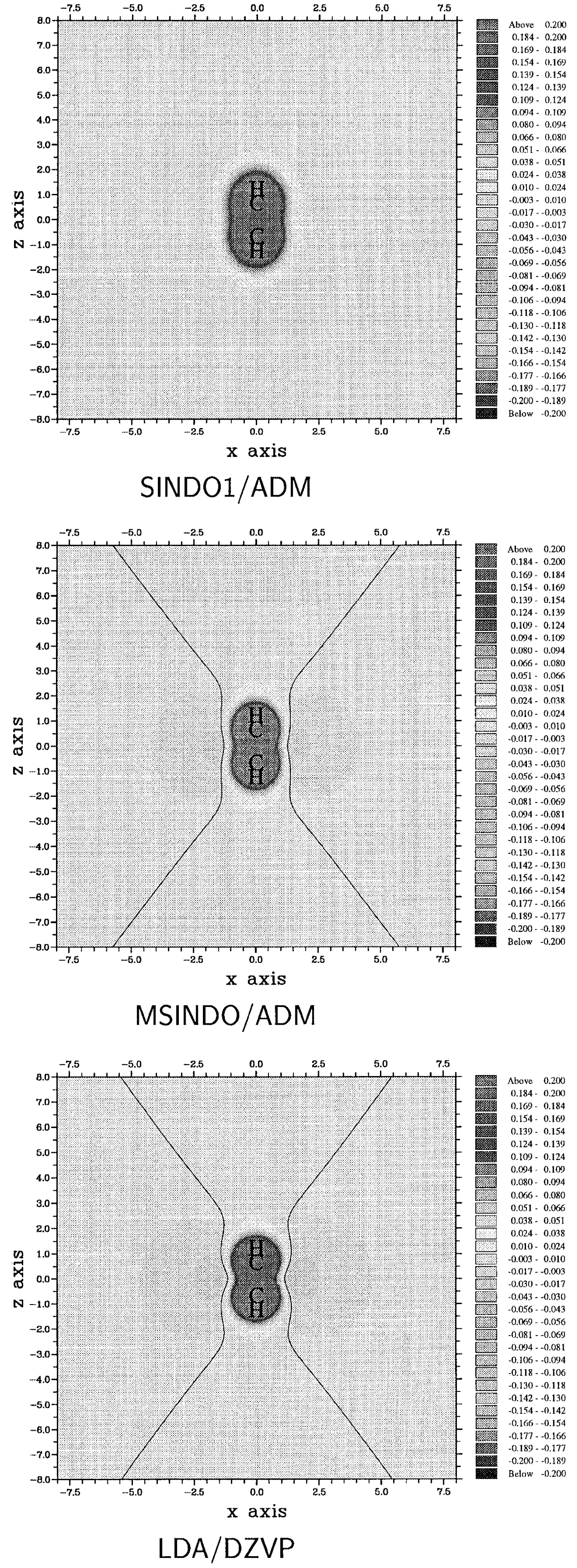


SINDO1/ADM

MSINDO/ADM

LDA/DZVP

Fig. 1. Molecular electrostatic potential (MESP) of acetylene

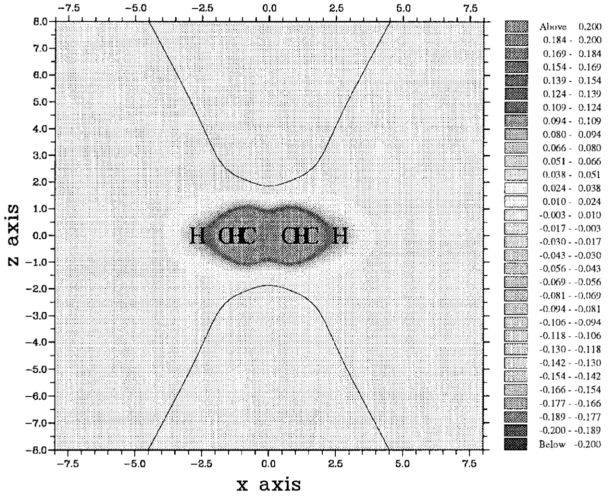


SINDO1/ADM

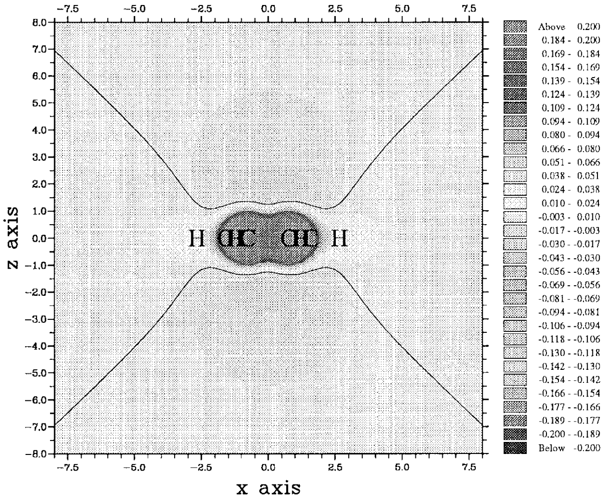
MSINDO/ADM

LDA/DZVP

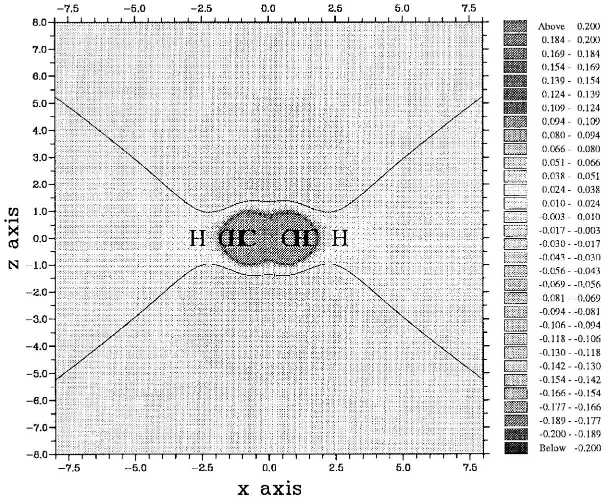
Fig. 2. Molecular electrostatic potential (MESP) of ethylene



SINDO1/ADM

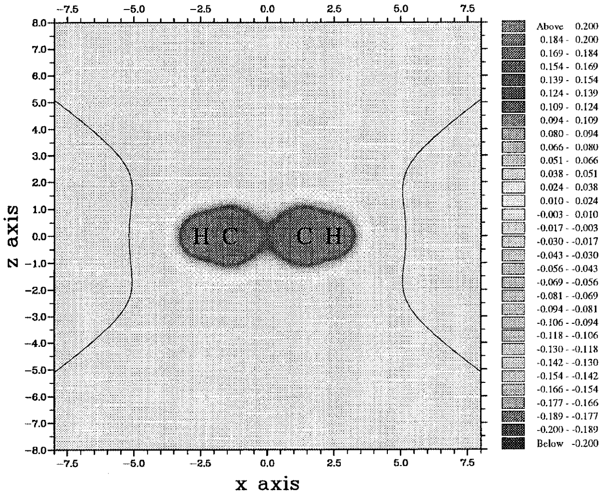


MSINDO/ADM

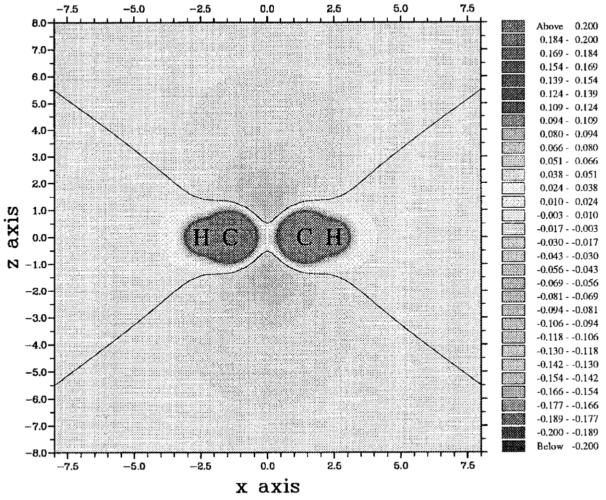


LDA/DZVP

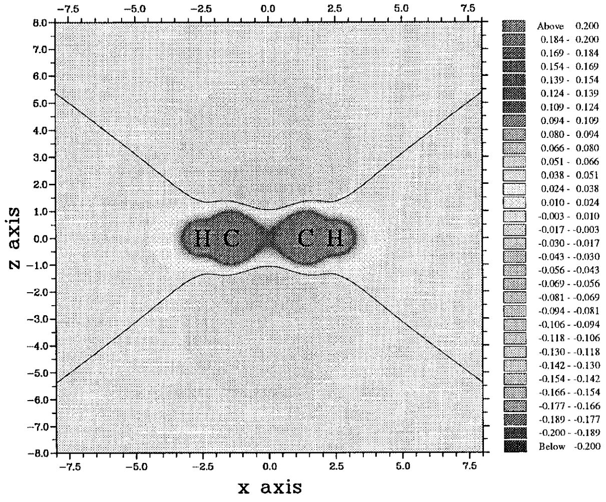
Fig. 3. Molecular electrostatic potential (MESP) of *t*-butadiene



SINDO1/ADM

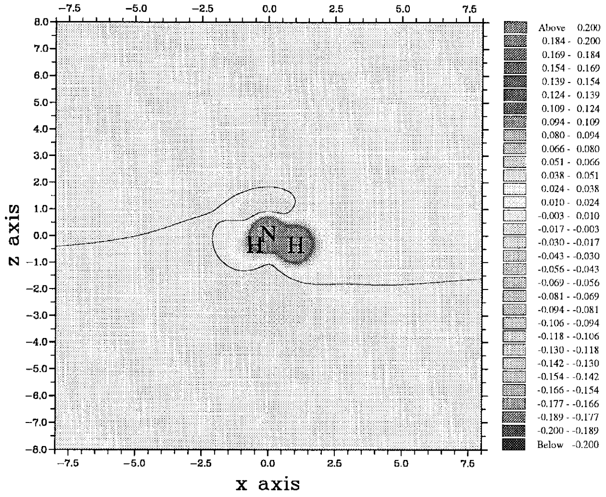


MSINDO/ADM

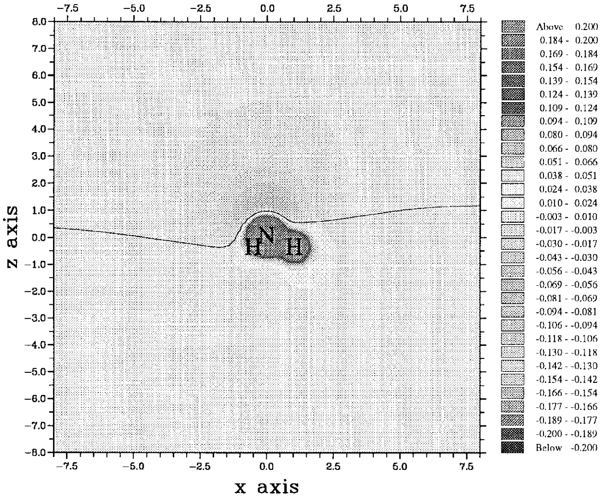


LDA/DZVP

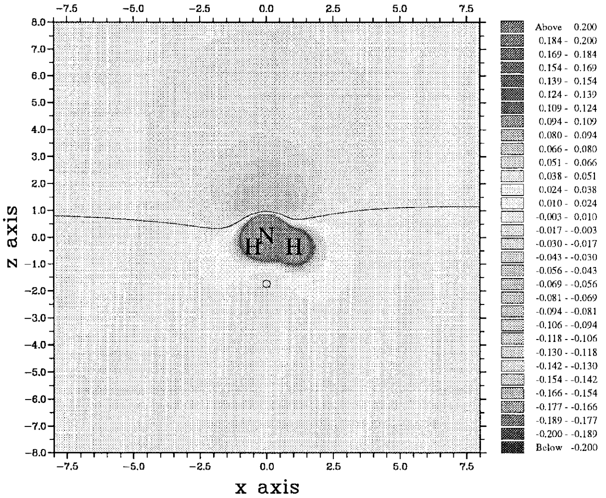
Fig. 4. Molecular Pelectrostatic potential (MESP) of benzene



SINDO1/ADM

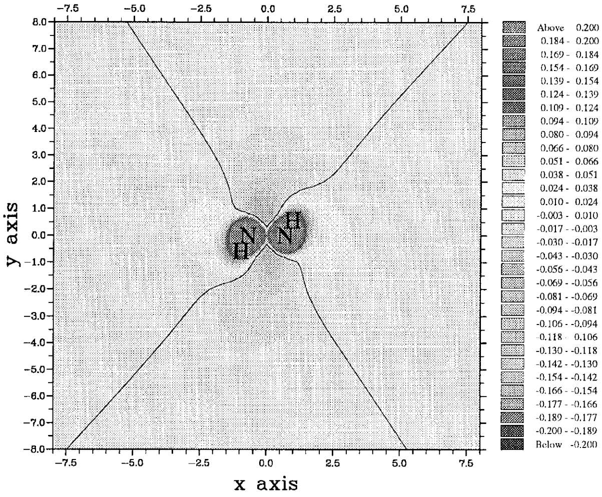


MSINDO/ADM

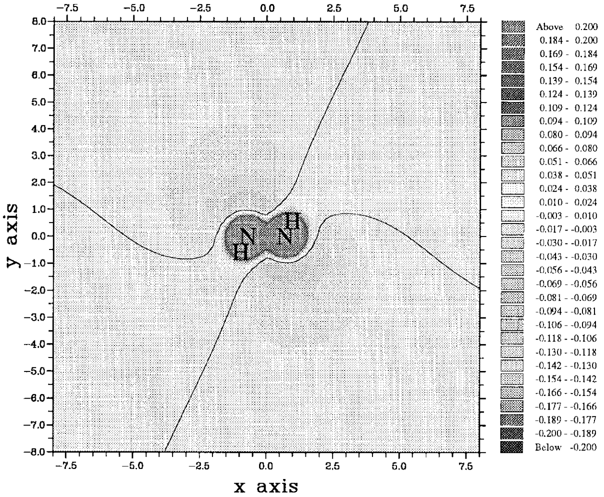


LDA/DZVP

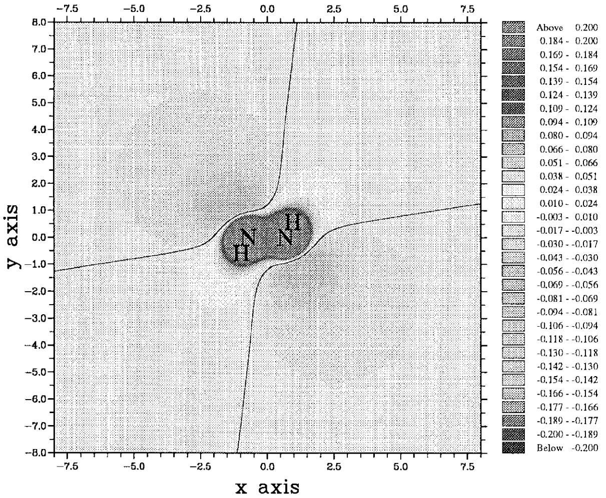
Fig. 5. Molecular electrostatic potential (MESP) of ammonia



SINDO1/ADM

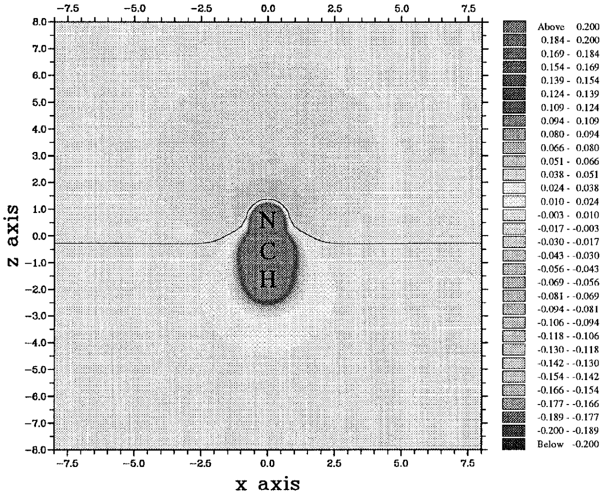


MSINDO/ADM

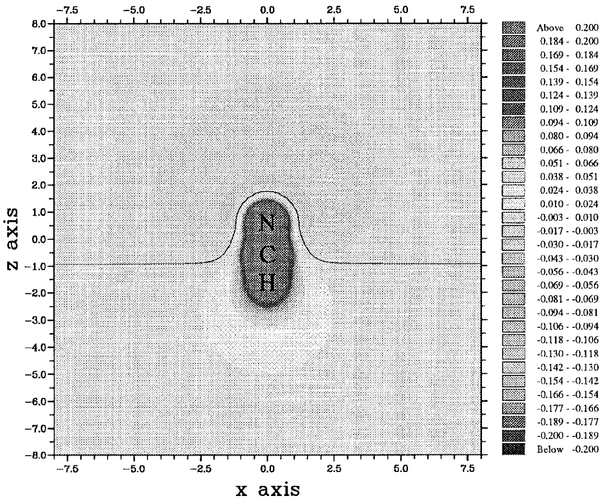


LDA/DZVP

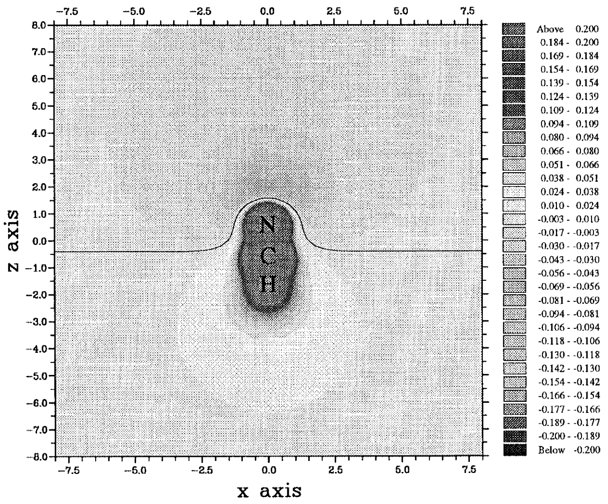
Fig. 6. Molecular electrostatic potential (MESP) of hydrazine



SINDO1/ADM

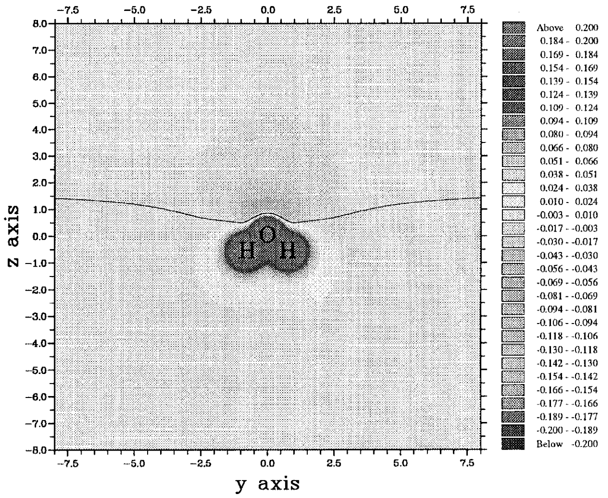


MSINDO/ADM

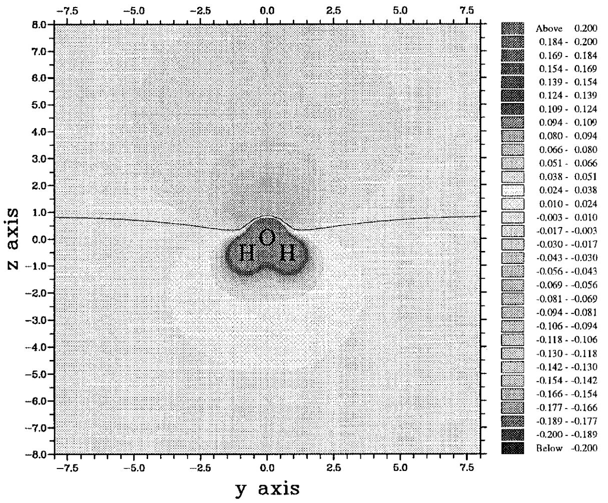


LDA/DZVP

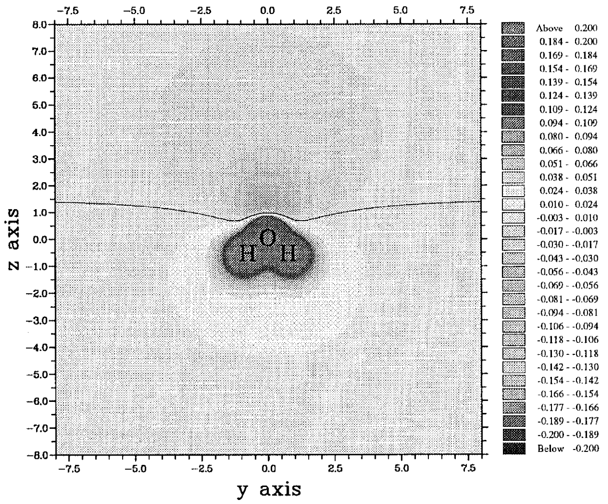
Fig. 7. Molecular electrostatic potential (MESP) of hydrogen cyanide



SINDO1/ADM

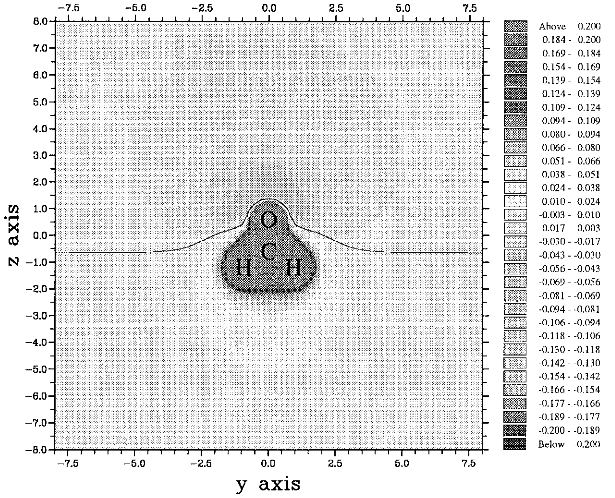


MSINDO/ADM

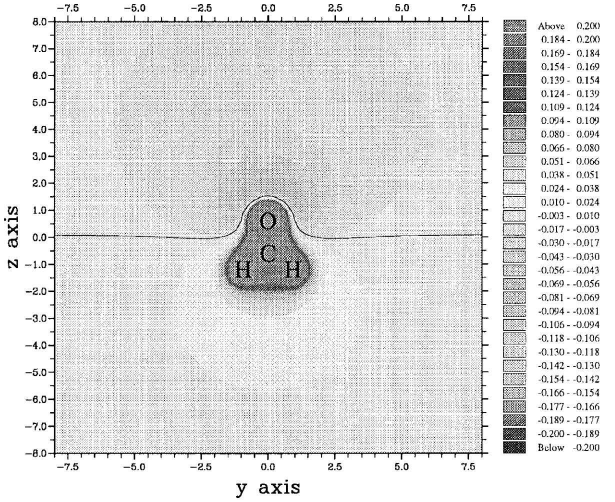


LDA/DZVP

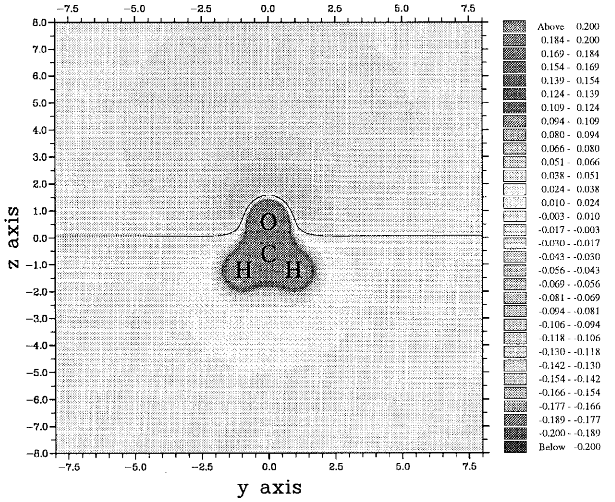
Fig. 8. Molecular electrostatic potential (MESP) of water



SINDO1/ADM



MSINDO/ADM



LDA/DZVP

Fig. 9. Molecular electrostatic potential (MESP) of formaldehyde

of benzene is represented in Fig. 4. Here the MSINDO and the LDA show a corresponding minimum. Lone pairs are seen at the nitrogen of ammonia (Fig. 5), hydrazine (Fig. 6) and hydrogen cyanide (Fig. 7). Again the MSINDO and the LDA features of the MESP are very similar. Only for hydrogen cyanide is the SINDO1 presentation very good. For water (Fig. 8) and formaldehyde (Fig. 9) the lone pairs on oxygen are clearly visible in all three presentations. Whereas the SINDO1 features of the MESP are accurate usually only at larger distances, the MSINDO features are also accurate in the bonding region. All critical points of the MESP, i.e. minima and saddle points are correctly reproduced by the MSINDO.

4 Conclusions

We have shown that an extra set of exponents, ζ^D, for the calculation of dipole moments with the MSINDO brings about a substantial reduction of the mean error of this quantity calculated for a large number of small molecules consisting of hydrogen, carbon, nitrogen and oxygen. With these rather accurate dipole moments an ADM-based approximation of the MESP is possible which retains all important features of the MESP characterized by its critical points. This means that we can hope for a further improvement of solvation effects based on the qualitative and quantitative accuracy of the MESP.

Acknowledgement. We thank the Deutsche Forschungsgemeinschaft for partial support of this work.

References

1. Valleau JP, Whittington SG (1977) In: Berne BJ (ed) Statistical mechanics part A. Plenum, New York, p 137
2. Valleau JP, Torrie GM (1977) In: Berne BJ (ed) Statistical mechanics part A. Plenum, New York, p 169
3. Kushik J, Berne BJ (1977) In: Berne BJ (ed) Statistical mechanics part B. Plenum, New York, p 41
4. Tomasi J, Persico M (1994) Chem Rev 94: 2027
5. Böttcher CJF (1973) Theory of electric polarization, vol 1. Elsevier, Amsterdam
6. Rinaldi D, Rivail J-L (1973) Theor Chim Acta 32: 57
7. Tapia O, Goscinski O (1975) Mol Phys 29: 1653
8. Wong MW, Frisch MJ, Wiberg KB (1991) J Am Chem Soc 113: 4776
9. Miertus S, Scrocco E, Tomasi J (1981) Chem Phys 55: 117
10. Ford GP, Wang B (1992) J Am Chem Soc 114: 10563
11. Rauhut G, Clark T, Steinke T (1993) J Am Chem Soc 115: 9174
12. Luque FJ, Bachs M, Orozco M (1994) J Comput Chem 15: 847
13. Nanda DN, Jug K (1989) Theor Chim Acta 57: 95
14. Jug K, Iffert R, Schulz J (1987) Int J Quantum Chem 32: 265
15. Li J, Correa de Mello P, Jug K (1992) J Comput Chem 13: 85
16. Köster AM, Kölle C, Jug K (1993) J Chem Phys 99: 1224
17. Köster AM, Leboeuf M, Salahub DR (1996) In: Murray JS, Sen K (eds) Molecular electrostatic potentials: concepts and applications. Theoretical and computational chemistry 3. Elsevier, Amsterdam, p 105
18. Leboeuf M, Köster AM, Salahub DR (1997) Theor Chem Acc 96: 23
19. Kölle C, Jug K (1997) J Comput Chem 18: 1
20. Jug K, Kölle C (1998) J Phys Chem B 102: 6605

21. Jug K, Kölle C, Neumann F (1999) J Comput Chem 20: 301
22. Ahlswede B, Jug K (1999) J Comput Chem 20: 563
23. Ahlswede B, Jug K (1999) J Comput Chem 20: 572
24. Ahlswede B (1998) PhD thesis. Hannover
25. Giesen DJ, Storer JW, Cramer CJ, Truhlar DG (1995) J Am Chem Soc 117: 1057
26. Sokalski WA, Poirier RA (1983) Chem Phys Lett 98: 86
27. Sokalski WA, Shibata M, Ornstein RL, Rein R (1992) J Comput Chem 13: 883
28. Löwdin PO (1950) J Chem Phys 18: 365
29. Mulliken RS (1955) J Chem Phys 23: 1833
30. Kolb M (1991) PhD thesis. Wuppertal
31. McClellan AL (1963) Tables of experimental dipole moments. Freeman, San Francisco
32. Li J, Williams B, Cramer CJ, Truhlar DG (1999) J Chem Phys 110: 724
33. Li J, Ching J, Cramer CJ, Truhlar DG (1999) J Chem Phys 111: 885
34. Krack M, Köster AM, Jug K (1997) J Comput Chem 18: 301
35. Sen KD, Politzer P (1988) J Chem Phys 90: 4370
36. Pathak RK, Gadre SR (1990) J Chem Phys 93: 1770
37. Gadre SR, Kulkarni SA, Shrivastava IH (1992) J Chem Phys 96: 5253
38. Jug K, Kölle C (1994) Proc Indian Acad Sci 106: 283
39. Leboeuf M, Köster AM, Jug K, Salahub D (1999) J Chem Phys 111: 4893
40. Köster AM, Krack M, Leboeuf M, Zimmermann B (1998) ALLCHEM. Universität Hannover
41. Vosko SH, Wilk L, Nusair M (1980) Can J Phys 58: 1200
42. Godbout N, Salahub DR, Andzelm J, Wimmer E (1992) Can J Chem 70: 560