ORIGINAL PAPER

Alexandr V. Yatsenko · Ksenia A. Paseshnichenko

On the performance of semiempirical MO theory for dipole moments of dye molecules

Abstract The dipole moments of a set of 71 simple dye molecules calculated at the ab initio, DFT, and semi-empirical levels have been compared. The DFT dipole moments are on average 16% larger than those obtained by MP2/6-31G**. AM1 and PM3 modified with an empirical correction procedure yield dipole moments essentially at the same level of accuracy as the results of non-empirical calculations. INDO/S and CNDO/S are considerably less accurate. Among different versions of spectral methods, the CISD scheme gives the best performance.

Keywords Dipole moments · Dyes · Semiempirical methods

Introduction

In the last decade spectacular progress has been made in modeling the electronic and spatial structure of organic molecules in condensed phases. This progress was achieved using procedures based on the incorporation of electrostatic interactions between a molecule and its environment into the Hamiltonian of the molecule. Several models exploiting this approach have been introduced and were applied to solvation energies, solvatochromic shifts, and solvent-induced conformational changes and tautomerism. [1, 2, 3, 4, 5] In order to account for the first-solvation-shell effects, the electrostatic contribution can be accompanied by a term responsible for the solvent–solute dispersion interaction. [2]

A very similar approach has been developed for molecular crystals. The total energy of the crystal packing is divided into two terms: the dispersion energy calculated by summation of empirical atom–atom potentials, [6, 7] and the Coulomb energy of intermolecular attraction. Polarization of the molecules under the effect of the crystal electrostatic (Madelung) potential is taken into account via the SCF scheme. [8, 9, 10]

It is obvious that the success of such a procedure depends to a large extent on the accuracy of the representation of the molecular electrostatic potential for the ground and excited states of a molecule. As an illustration, let us consider a non-polarizable point dipole placed into a spherical cavity within a polarizable dielectric (Eq. (28) in [11]):

$$\nu_{solution} - \nu_{gas} = \frac{22679}{V} \left[\frac{\varepsilon - 1}{2\varepsilon + 1} \left(\mu_g^2 - \mu_g \mu_e \right) \right. $$
$$\left. + \frac{n^2 - 1}{2n^2 + 1} \left(\mu_g \mu_e - \mu_e^2 \right) \right] cm^{-1},$$

where $\Delta\nu = \nu_{solution} - \nu_{gas}$ is the solvatochromic shift, V is the cavity volume, μ_g and μ_e are molecular dipole moments for the ground and excited states, ε and n are the dielectric constant and the refractive index of the solvent, respectively. Let $\varepsilon = n^2$ (as for an apolar solvent) and $\mu_e = 1.4\mu_g$ (which is typical of dye molecules). If the calculations overestimate both μ_g and μ_e by 10%, $\Delta\nu$ is overestimated by 21%, but if the calculations underestimate μ_g by 10% and overestimate μ_e by 10%, the error in $\Delta\nu$ rises to 62.5%.

Correlation-corrected ab initio methods using sufficiently large basis sets (BSs) provide precise results for large organic molecules at the cost of enormous computational time. Semiempirical methods are much faster, but their accuracy normally does not exceed the accuracy of the ab initio calculations with minimal BSs. [12] Recently a procedure of empirical correction of dipole moments and atomic charges evaluated by the AM1, PM3, and ZINDO/S methods for an extensive set of simple organic molecules has been developed [13, 14] and applied to solvation energies [13] and electronic excitations. [15] This procedure is based on the idea that the error in bond

A.V. Yatsenko (✉) · K.A. Paseshnichenko
Department of Chemistry, Moscow State University,
119899, Moscow, Russia
e-mail: yatsenko@biocryst.phys.msu.su

Fig. 1 Chemical diagrams and reference numbers of discussed molecules
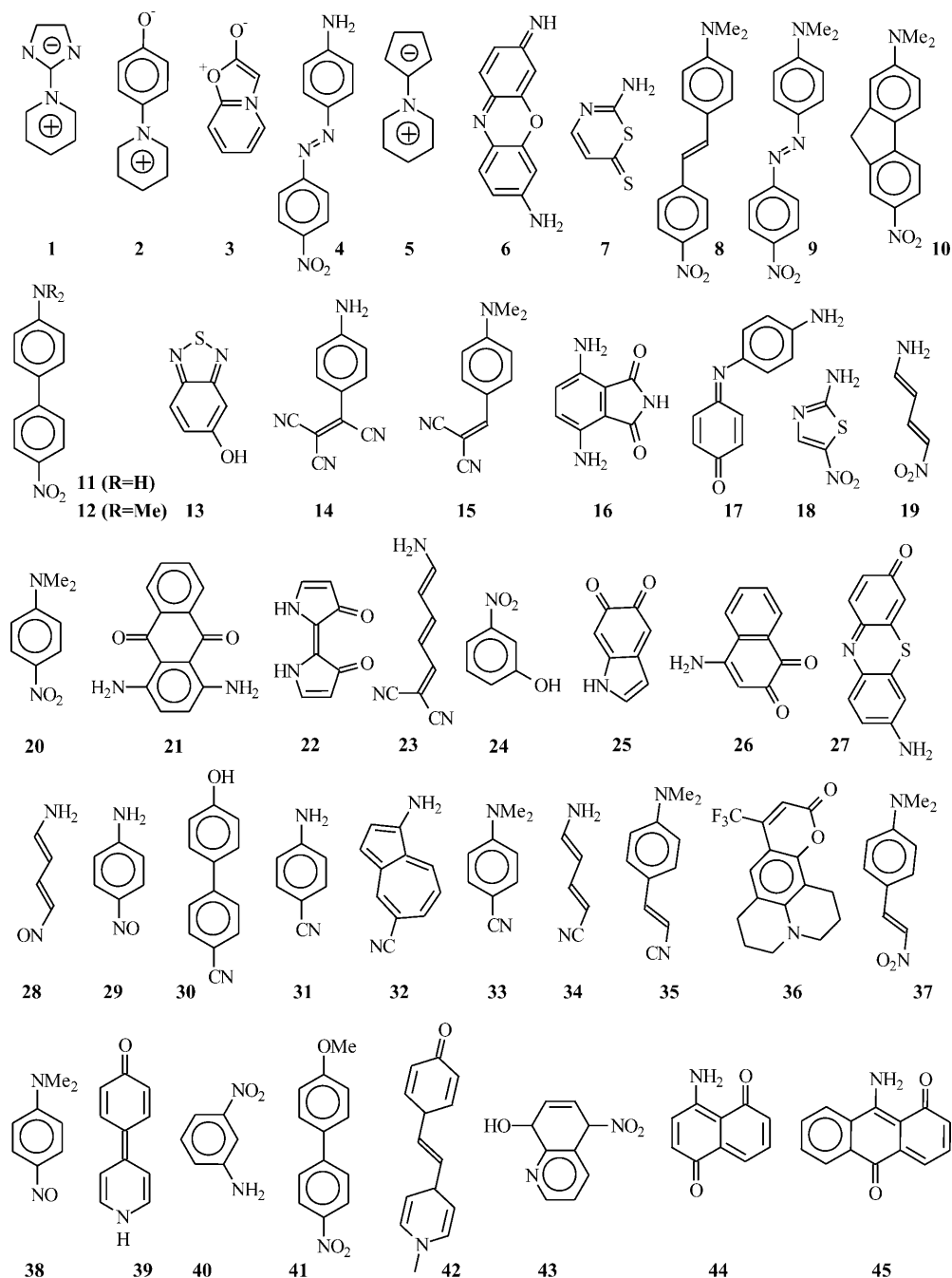
dipole moment is approximately a constant for each particular bond type.

Calculations on dye molecules can reveal some details that are not evident when dealing with simple molecules. Since the molecules of dyes have low-lying vacant MOs, the correlation correction, implicitly included in the semiempirical parameterization, may be inadequate for reproducing their electronic structure.

This work is devoted to the reliability of molecular dipole moments of organic dyes and related molecules (shown in Fig. 1) obtained from semiempirical calculations at the NDO and NDDO levels. Unfortunately, it is impossible to determine the gas-phase dipole moments

for a great majority of dyes, whereas the dipole moments measured in solution correlate poorly with those obtained from gas-phase experiments. [16] Hence the results of high-quality ab initio and DFT calculations were used as reference data.

Computational technique

The ab initio calculations were carried out with the GAMESS program [17] using the following basis sets: Dunning's TZV [18] extended with $2d1f$ and $2p$ polarization functions for heavy atoms and hydrogens, respec-

tively, and with diffuse *sp* and *s* shells for heavy atoms and hydrogens, respectively (TZV+2P), Dunning's DZV extended with 1*d* and 1*p* polarization functions for heavy atoms and hydrogens, respectively (DZVP), and standard 6-31G** and 3-21G*. All exponents, contraction schemes, and splitting factors were the GAMESS defaults. The DFT calculations were performed with the program provided by Laikov [19] employing the BLYP [20, 21] and PBE [22] exchange-correlation functionals. For the representation of the Kohn–Sham one-electron wave functions, sets of contracted Gaussian-type orbitals, {3111/1} for H, {611111/411/11} for the second-row elements, and {6111111111/5111111/11} for S, were used. For expansion of the electron density, auxiliary basis sets of uncontracted functions (5*s*1*p*) for H, (10*s*3*p*3*d*1*f*) for the second-row elements, and (14*s*7*p*7*d*1*f*1*g*) for S were employed.

The semiempirical calculations at the AM1 [23] and PM3 [24, 25] levels were run using MOPAC7.2. [26] The INDO/S and CNDO/S calculations were performed using different implementations of these methods. The popular ZINDO/S scheme was used in its original parameterization [27, 28] and with the readjusted oxygen parameters (ZINDO/S2, [14]) using the traditional scaling factors f_σ=1.267 and f_π=0.585. The configuration space was built of the 225 singly excited configurations generated from the 15 highest filled MOs and the 15 lowest empty MOs. The scheme known as GRINDOL [29] was used with the basis of 200 energy-selected singles. GRINDOL employs the Linderberg–Seamans equation [30] for the resonance integrals with the empirical scaling parameter equal to 1.1. The program for INDO-CISD and CNDO-CISD calculations provided by Dick [31, 32] was run using 200 energy-selected singles and doubles with the Pariser–Parr formula [33] for the two-center Coulomb integrals and with the parameter set proposed by Ellis et al. [34]

All molecular models were optimized at the DFT/BLYP level, and the resulting molecular geometries were used in the ab initio and semiempirical calculations.

Results and discussion

Ground-state dipole moments

The performance and basis set dependence of the ab initio and DFT dipole moments of small organic molecules have been discussed in a number of papers. The results seem to be essentially converged employing aug-cc-pVTZ at the DFT or MP2 levels. The HF dipole moments are normally too large (by 5–15%). Smaller basis sets (lacking diffuse and polarization functions) provide worse results. For example, DFT-BPW91 with DZVP, 6-31G*, and MIDI1 reproduce the experimental dipole moments for a set of 102 polar molecules containing H, C, N, and O with absolute RMS errors of 0.19, 0.20, and 0.40 D and relative errors of 7.7, 8.1, and 16.2%, respectively. [13]

Table 1 Experimental and calculated ground-state dipole moments of simple molecules: correlation summary ($Y=A \cdot X$)

X	Y	A^a	SD^b	N^c
PBE	BLYP	1.007(2)	0.040	34
MP2/TZV+2P	DFT	1.011(9)	0.097	13
MP2/TZV+2P	Experimental[d]	0.985(10)	0.107	13
MP2/TZV+2P	MP2/DZVP	1.006(14)	0.156	13
MP2/TZV+2P	MP2/6-31G**	0.980(13)	0.138	13
Experimental	DFT	1.015(10)	0.169	28
Experimental	MP2/DZVP	1.016(11)	0.197	28
Experimental	MP2/6-31G**	0.990(10)	0.170	28
MP2/DZVP	MP2/6-31G**	0.973(4)	0.083	33
MP2/6-31G**	DFT	1.026(8)	0.151	33
MP2/DZVP	DFT	0.998(9)	0.183	33

[a] In parentheses: standard deviation for slope
[b] Standard deviation from the regression, D
[c] Number of points in the correlation
[d] Data from [13, 25, 45]

Dipole moments calculated at various levels of theory for a few simple molecules are compared in Table 1. The DFT-BLYP and DFT-PBE calculations at the same molecular geometry give almost identical results (also see Table 2); thus separate discussions of the BLYP and PBE results would make no sense, and only the averaged values referred to as "DFT" are considered below. The MP2/TZV+2P results are in excellent agreement with the results of DFT calculations and with experimental data, whereas the data obtained with basis sets of double-zeta quality are slightly worse (6-31G** seems to be slightly better than DZVP). The largest distinctions between the double-zeta and triple-zeta basis sets are observed for molecules such as water, ammonia, and formaldehyde; their dipole moments are determined by contributions from sterically active lone pairs, which are very sensitive to the quality of basis set used. Overall, with respect to the experimental data, the DFT dipole moments are slightly too large, whereas those obtained by MP2/6-31G** are slightly too small, but these distinctions are essentially within the limits of standard deviations.

The calculations were carried out for a set of organic molecules containing π-systems with electron donor and acceptor centers that can be considered as prototypes for real dye molecules (Table 2). The results obtained by 6-31G** and DZVP are very close. It is known that, when calculating molecular properties, MP2 typically overshoots the correlation effects; thus one could expect that the DFT dipole moments should be somewhere between the HF and MP2 results, closer to MP2. The DFT results are, on average, closer to HF, but correlation with the MP2 data is distinctly better. It is noteworthy that the standard deviation (SD) for the regression HF/6-31G**–DFT is almost equal to the sum of SDs for the regressions HF/6-31G**–MP2/6-31G** and MP2/6-31G**–DFT, thus the difference between the DFT and ab initio results is mostly attributable to the correlation effects. The largest distinctions are exhibited by the mesoionic molecules **1**, **2**, and **3**, poorly described by the single-determinant wave function. However, the DFT–MP2

Table 2 Experimental and calculated ground-state dipole moments of dye molecules: correlation summary ($Y = A \cdot X$)

X	Y	A[a]	SD[b]	N[c]	Principal outliers[d]
PBE	BLYP	1.0012(6)	0.004	71	
MP2/DZVP	MP2/6-31G**	0.976(2)	0.11	49	
HF/DZVP	HF/6-31G**	0.988(2)	0.09	49	
MP2/6-31G**	DFT	1.155(12)	0.62	63	1 (–1.82); 2 (–1.70); 4 (1.56)
HF/6-31G**	DFT	0.976(17)	1.03	63	2 (–3.91); 1 (–2.70); 5 (–2.23)
HF/6-31G**	MP2/6-31G**	0.847(8)	0.49	63	2 (–1.95); 5 (–1.23); 6 (0.91)
MP2/3-21G*	MP2/6-31G**	1.050(5)	0.25	63	1 (–0.89); 5 (–0.73); 7 (–0.56)
MP2/3-21G*	DFT	1.238(16)	0.85	71	1 (–3.01); 8 (2.04); 9 (2.00)
Experimental[e]	DFT	1.464(46)	1.32	24	8 (2.28); 10 (2.24); 4 (2.22)
Experimental	MP2/6-31G**	1.185(27)	0.79	24	11 (1.65); 12 (1.40); 10 (1.40)
Experimental	Non-empirical	1.325(36)	1.02	24	11 (1.86); 10 (1.82); 12 (1.75)
AM1	PM3	0.993(5)	0.26	71	1 (–0.72); 13 (0.68); 7 (0.60)
AM1	Non-empirical	1.079(17)	0.96	71	14 (2.31); 1 (–2.16); 15 (1.98)
AM1	DFT	1.168(21)	1.15	71	
AM1	MP2/6-31G**	0.988(16)	0.88	71	
AM1+CI corrected	Non-empirical	1.110(8)	0.45	71	1 (–1.31); 16 (–0.99); 2 (0.93)
AM1+CI corrected	DFT	1.203(11)	0.62	71	
AM1+CI corrected	MP2	1.015(9)	0.50	71	
PM3+CI corrected	Non-empirical	1.104(9)	0.47	71	18 (–1.16); 19 (–1.12); 20 (–0.93)
ZINDO/S	Non-empirical	0.879(17)	1.18	71	2 (–4.04); 3 (–2.93); 13 (–2.20)
ZINDO/S2	Non-empirical	0.934(19)	1.18	71	13 (–3.77); 2 (–3.00); 21 (2.56)
GRINDOL	Non-empirical	1.007(25)	1.43	71	2 (–3.82); 15 (3.63); 7 (–3.40)
INDO-CISD	Non-empirical	0.962(15)	0.92	64[f]	22 (–2.04); 23 (1.97); 24 (–1.89)
CNDO-CISD	Non-empirical	0.981(17)	1.00	64	3 (–2.14); 23 (1.93); 24 (–1.92)
ZINDO/S2 corrected	Non-empirical	0.953(17)	1.03	64	6 (–3.28); 21 (2.60); 2 (2.20)
INDO-CISD corrected	Non-empirical	0.983(13)	0.79	64	22 (–2.26); 25 (–1.71); 26 (–1.60)
GRINDOL corrected	Non-empirical	1.000(19)	1.12	64	2 (–3.73); 3 (–2.12); 9 (2.03)

[a] In parentheses: standard deviation for slope
[b] Standard deviation from the regression, D
[c] Number of points in the correlation
[d] In parentheses: deviation from the regression ($Y–A \cdot X$, D)

[e] Solution data from [46]
[f] INDO-CISD and CNDO-CISD calculations were not performed for seven S-containing molecules since these schemes have not been parameterized for sulfur

Table 3 Examples of DFT>HF/6-31G**>MP2/6-31G** ordering of dipole moments (D)

Molecule	DFT	HF[a]	MP2[a]
14	11.10	10.38	9.35
4	10.72	9.67	7.93
28	9.72	8.91 (9.06)	7.94 (8.28)
27	9.47	8.19	7.71
17	8.05	6.18	5.82
29	7.27	6.18 (6.27)	5.85 (6.05)

[a] In parentheses: the TZV+2P data

correlation is not as good as that observed for simple molecules (Table 1, [35]), and for some dye molecules MP2 increases the difference between the DFT and HF dipole moments (for examples see Table 3). This indicates that the difference between the DFT and ab initio electron density distribution is sometimes more complex. In other words, it is not obvious which data are preferable. For this reason, the semiempirical dipole moments are compared not only with the DFT and MP2/6-31G** results, but also with the averaged values referred to from here on as "non-empirical".

The 3-21G* basis set is often believed to be too small for the MP2 correction. However, the MP2/3-21G* dipole moments are in much better agreement with the

MP2/6-31G** and DFT data than with the HF/6-31G** results. Thus, the set of dyes was supplemented by several larger molecules with dipole moments calculated by MP2/3-21G* and then extrapolated to MP2/6-31G**.

The dipole moments of dyes measured in solution are significantly smaller than the calculated values and, in contrast to the gas-phase data for small molecules, correlation between the calculated and experimental values is rather poor.

AM1 usually gives more realistic atomic charges for compounds involving nitrogen than does PM3. [36] However, dipole moments computed with these methods compare well, but do not correlate very well with the non-empirical results, with relative SDs of 15–17%. It is well known that the AM1 and PM3 dipole moments of the majority of simple organic molecules are in reasonably good agreement with the experimental data, with the exception of nitro compounds (dipole moments systematically overestimated by 0.86 and 0.65 D, respectively) and nitriles (systematically underestimated by 0.90 and 0.70 D, respectively). After the empirical correction for the NO_2 and CN groups [10] the SDs decrease by a factor of 1.3–2.0. The alternative mapping procedure CM2 [13] improves dipole moments of nitriles, but is unsuccessful for nitro compounds since it was parameterized using no such molecules. At this stage the largest discrepancies were observed for mole-

cules **1**, **2**, **5**, and **17**. We have found that the use of configuration interaction including only HOMO and LUMO in the configuration space considerably improves these outliers and decreases the deviations to such an extent that the empirically corrected AM1+CI and PM3+CI dipole moments are essentially of the same quality as the non-empirical results, with a relative SD of 7%. The corrected AM1+CI data correlate with the MP2 and DFT results even better than MP2 correlates with DFT: correlation coefficients are 0.977, 0.980, and 0.970 for the MP2-AM1, DFT-AM1, and MP2-DFT regressions, respectively.

On first sight, the accuracy of the INDO/S and CNDO/S dipole moments is only slightly worse than the uncorrected AM1 and PM3 data. GRINDOL and ZINDO/S significantly overestimate the dipole moments of mesoionic molecules **1**–**3** and **5**, whereas INDO-CISD involving the ground state in CI is superior to the CIS-based schemes in modeling the electronic structure of these molecules. Similarly to AM1 and PM3, all spectral methods tend to overestimate the dipole moments of nitro compounds and underestimate the dipole moments of nitriles. Fitting the INDO/S dipole moments to the non-empirical results, the following empirical correction factors were derived for the nitro and cyano groups: –0.48 and +0.34 D (ZINDO/S2), –0.77 and +0.55 D (INDO-CISD), and –0.28 and +1.83 D (GRINDOL), respectively.

Once the mesoionic and S-containing molecules have been excluded from consideration and the empirical correction for the NO_2 and CN groups has been applied, all three INDO/S schemes demonstrate essentially the same level of accuracy, being, however, distinctly inferior to the empirically corrected AM1-CI and PM3-CI results. There are three obvious reasons for lower reliability of the INDO/S dipole moments, viz.:

1. As a result of improper representation of lone pairs, the NDO/S schemes have problems with reproducing the dipole moments of many simple compounds such as alcohols, ketones, acids, esters, and amines. [14, 37] These errors are, however, systematic. Thus they can be compensated for using an empirical correction for particular groups (as for the NO_2 and CN) or for bond dipoles (CM2).
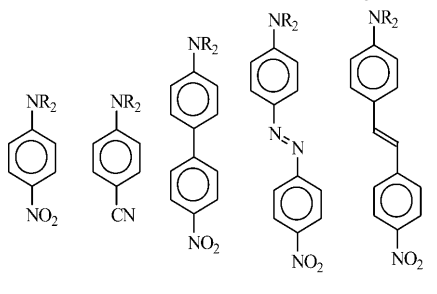2. ZINDO/S, GRINDOL, and, to a lesser degree, INDO-CISD underestimate the σ-effects that arise from replacement of the amino hydrogen atoms by the methyl groups (Table 4). This problem can be fixed by readjusting the H atom parameters, e.g., running ZINDO/S2 with $\beta_H=19$ eV instead of the traditional 12 eV.
3. The empirical correction for the nitro and cyano groups is less efficient for INDO/S than for AM1 and PM3. After this correction, the dipole moments of simple nitriles and nitro compounds are systematically overestimated by 1.0–1.2 D, whereas the dipole moments of highly polar molecules are underestimated by 1.0–1.5 D (Table 5). In other words, all three INDO/S schemes underestimate the π-acceptor capacity of the NO_2 and CN groups. Since the observed error in dipole moment is not a constant, but depends on the electronic structure of the whole molecule, there is little sense in attempting to compensate for it with an empirical procedure.

Excited-state dipole moments

High-level non-empirical calculations with a full CI expansion become unmanageably time-consuming even for

Table 4 Effect of amino group methylation on ground-state molecular dipole moment, $\Delta\mu=\mu(R=CH_3)-\mu(R=H)$ (D)

X	A	B	C	D	E
Non-empirical	1.12	0.95	0.56	1.44	1.24
AM1	1.16	1.02	0.50	1.41	1.18
ZINDO/S2					
β_H=12 eV	0.65	0.54	0.29	0.95	0.70
β_H=19 eV	0.99	0.88	0.69	1.38	1.07
GRINDOL	0.62	0.53	–0.17	0.64	0.74
INDO-CISD	0.84	0.76	0.40	1.08	0.77

Table 5 Non-empirical and semiempirical ground-state dipole moments (D) of CN-containing compounds (semiempirical data include empirical correction and are scaled to non-empirical data using regression coefficients from Table 2)

Molecule	Non-emp.	AM1	ZINDO/S	GRIN-DOL	INDO-CISD
CH_3CN	3.83	4.27	5.01	5.52	4.75
C_6H_5CN	4.50	4.72	5.62	5.68	5.68
30	5.67	5.11	5.84	5.47	6.51
31	6.60	6.64	6.95	6.91	7.06
32	6.61	6.01	6.51	6.18	7.13
33	7.55	7.77	7.48	7.50	7.81
34	7.89	7.80	7.91	8.06	7.75
14	10.23	9.66	9.79	9.44	9.67
15	10.88	10.56	10.17	9.30	10.15
23	11.96	11.82	11.55	10.97	10.70

Table 6 Experimental and calculated excited-state dipole moments of dye molecules: correlation summary ($Y=A\cdot X$)

X	Y	A[a]	SD[b]	N[c]	Principal outliers[d]
ZINDO/S	Experimental[e]	0.853(32)	2.66	14	9 (5.71); 35 (4.87); 20 (–3.04)
ZINDO/S2	Experimental	0.986(43)	3.08	14	9 (5.60); 20 (–4.74); 36 (+4.00)
INDO-CISD	Experimental	1.114(31)	1.97	14	37 (–3.96); 10 (3.39); 38 (–2.60)
CNDO-CISD	Experimental	1.265(40)	2.21	14	37 (–4.06); 38 (–4.05); 20 (–2.69)
GRINDOL	Experimental	1.437(61)	3.00	14	35 (5.67); 20 (–5.10); 37 (–3.81)
AM1	Experimental	1.568(51)	2.31	14	37 (–4.46); 35 (3.94); 20 (–3.19)
ZINDO/S	Averaged	0.772(12)	1.52 (1.33)	64[f]	39 (3.41); 2 (3.38); 40 (–3.09)
ZINDO/S2	Averaged	0.873(13)	1.40 (1.16)	64	30 (–4.11); 41 (–3.08); 36 (3.07)
INDO-CISD	Averaged	0.971(15)	1.47 (1.11)	64	2 (–5.48); 5 (–4.14); 1 (3.26)
CNDO-CISD	Averaged	1.072(15)	1.29 (0.83)	64	2 (–5.87); 5 (–4.25); 1 (3.07)
GRINDOL	Averaged	1.158(18)	1.49 (1.35)	64	30 (3.02); 2 (2.93); 8 (2.82)
AM1	Averaged	1.315(19)	1.34 (1.13)	64	42 (–3.55); 41 (3.33); 2 (2.83)

[a] In parentheses: standard deviation for slope
[b] Standard deviation from the regression, D. In parentheses: SD on exclusion of four principal outliers
[c] Number of point in the correlation
[d] In parentheses: deviation from the regression ($Y–A\cdot X$, D)
[e] Experimental data from [47, 48, 49, 50]
[f] Molecules containing S atoms are excluded

the simplest dye molecules. Experience with the ab initio calculations suggests that CIS gives wave functions of roughly HF quality, and inclusion of at least doubles and triples is required. [12, 37] For example, the DFT calculations of Coumarin C153, **36**, employing CIS and CISD schemes yield μ_e equal to 23.4 and 17.7 D, respectively, [38] whereas the experimental value is 14.5 D. [39] Alternatively, the gas-phase measurements on the excited-state dipole moments are restricted to very simple molecules, for none of them the excitation being of the charge-transfer type. [14] Thus the solvatochromic and electrochromic measurements in solutions or in crystalline matrices are the only available source of reference data. Since these techniques employ some approximations and arbitrarily chosen parameters, their results have not been considered as pure experimental values, but they are suitable to rate the results of semiempirical calculations.

For a limited set of substituted anilines, azobenzenes, stilbenes, and biphenyl derivatives, ZINDO/S systematically overestimates, whereas GRINDOL and AM1 systematically underestimate the molecular dipole moments, and the best correlation with the experimental data is achieved using INDO-CISD and CNDO-CISD (Table 6).

In order to reveal the specificity of each of the semiempirical techniques used, the calculated data were compared with the values obtained by averaging over all of these methods, referred to from here on as "averaged" (Table 6). With respect to the averaged data, ZINDO/S overestimates the dipole moments of nitro compounds. The readjusting of the O atom parameters (ZINDO/S2) eliminates this shortcoming, but, as a side effect, the dipole moments of the molecules containing the OH and OR groups, **30** and **41**, become too large. In contrast to ZINDO/S2, AM1 underestimates the dipole moments of the two latter molecules, but overestimates the dipole moment of molecule **42**. In the case of GRINDOL, the obvious outliers are absent.

The INDO-CISD and CNDO-CISD data for four mesoionic molecules **1–3**, **5** deserve special consideration. They differ greatly from the averaged data, but as shown

Table 7 Excited-state dipole moments (D) of some molecules calculated with various semiempirical schemes

Molecule	9	39	5	2
Program INDO [31, 32]:				
CISD,CI = 200	20.72	15.34	6.83	12.56
CI=300	20.58	14.87	6.70	12.29
CIS, CI=200	17.16	12.88	0.28	4.61
CI=200, MN[a]	21.77	12.46	0.13	3.95
CI=300, MN	21.67	12.37	0.12	3.98
ZINDO/S	22.61	12.96	0.16	4.33
ZINDO/S, $f_{\pi\pi}$=0.64	22.72	12.67	0.31	3.98
ZINDO/S2	19.68	12.21	0.16	4.28
GRINDOL	15.40	12.00	0.18	3.28
AM1-CIS, CI=225	15.59	11.19	0.57	2.96
AM1, C.I.=(7,3) OPEN(2,2)	19.15			6.60

[a] Mataga–Nishimoto formula [51] for the two-center Coulomb integrals was used
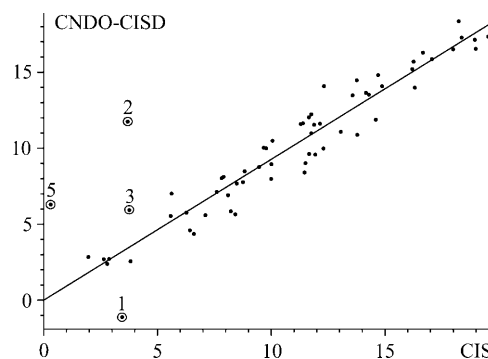
Fig. 2 Excited-state CNDO-CISD molecular dipole moments plotted against averaged CIS data. The CNDO-CISD dipole moment of molecule **1** is directed oppositely to the averaged CIS dipole moment

in Fig. 2, the correlation between the CISD and CIS data becomes perfect if the points corresponding to these zwitterions are eliminated. The data given in Table 7 suggest that for moderately polar molecules (e.g., **9**) the effect of the inclusion of doubles is not prominent. For

Table 8 Ground- and excited-state dipole moments (D) of betaine dye

	Ground	Excited
ZINDO/S	18.8	2.9
ZINDO/S2	16.7	2.9
GRINDOL	16.4	1.6
INDO-CISD	13.6	7.6
CNDO-CISD	15.5	5.8
AM1-CIS	12.4	0.2
Non-empirical	12.9	
Experimental [52, 53]	14.8	6.2

highly polar molecule **39**, which is close to the isoelectronic point, this effect exceeds the effects of all other possible changes in the computational scheme, and for the zwitterions **2** and **5** the CISD and CIS data are incompatible. What scheme – CIS or CISD – provides preferable results for such compounds? Let us consider as a test example a molecule of a betaine dye containing the pyridinium phenolate fragment (analog of molecule **2**). This dye demonstrates a pronounced hypsochromic shift of its absorption band on transfer from an apolar to a polar solvent, [40] and its dipole moment decreases by 8.6 D on excitation. The calculations predict heavy dependence of both (ground and excited) dipole moments on the interplanar angle between the pyridinium and phenolate rings, whereas the effect of peripheral phenyl groups is insignificant. In crystals this angle is equal to 68°, [41] and the DFT optimization of molecular geometry yields 56°; thus the value of 65° was used for the semiempirical calculations (Table 8). All CIS schemes considerably underestimate the excited-state dipole moment and overestimate the difference between μ_g and μ_e, whereas the CISD results are in better agreement with the experiment.

The molecule of Brooker's merocyanine **42** exhibits a modest bathochromic shift on transfer from apolar to moderately polar solvent, and a pronounced hypsochromic shift on further increase in solvent polarity. Therefore, for an isolated molecule μ_g is expected to be smaller than μ_e, and in highly polar solvents μ_g becomes larger than μ_e, because the molecular structure reverts to the zwitterion. [42] In line with the experiment, all six semiempirical schemes yield μ_e being 2–14% larger than μ_g.

The ana-quinonoid molecules **44** and **45** exhibit hypsochromic shifts on increase in solvent polarity. [43, 44] For **45**, all methods properly predict $\mu_e < \mu_g$, whereas for **44** only ZINDO/S2 and CNDO-CISD are successful: others yield $\mu_e > \mu_g$, although the calculated difference between μ_e and μ_g does not exceed 0.6 D.

Conclusions

The NDDO methods (AM1, PM3) extended with the empirical correction procedure are as good as the non-empirical (DFT and ab initio) methods for the ground-state dipole moments of a set of simple dye molecules.

Therefore, these methods are expected to be capable of reproducing the higher-order electrostatic moments of extended multi-functional dye molecules to a high degree of accuracy. In particular, they seem to be suitable for the modeling of intermolecular electrostatic interactions.

Various versions of INDO/S and CNDO/S are approximately two times less accurate for the ground-state dipole moments, and there are no grounds to expect that their reliability can be markedly improved without the thorough re-parameterization of these methods. The distinctions between the experimental and calculated excited-state dipole moments are rather large, and thus the modeling of solvatochromism based on the INDO/S electrostatic potential is restricted to only semi-quantitative level.

The CISD scheme is superior to CIS for the ground- and excited-state dipole moments of molecules with large intramolecular charge transfer. The use of more efficient schemes than the CIS technique to account for the correlation effects seems to be a crucial point in development of semiempirical methods capable of modeling the solvatochromic effects.

Supplementary material. Chemical diagrams of all molecules used in calculations, and a full list of non-empirical and semiempirical dipole moments are available.

References

1. Tomasi J, Persico M (1994) Chem Rev 94:2027–2094
2. Cramer CJ, Truhlar DG (1999) Chem Rev 99:2161–2200
3. Luzhkov V, Warshel A (1991) J Am Chem Soc 113:4491–4499
4. Lipiński J, Bartkowiak W (1997) J Phys Chem A 101:2159-2165
5. Gao J (1996) Acc Chem Res 94:298–305
6. Filippini G, Gavezzotti A (1993) Acta Crystallogr, Sect B 49:868–880
7. Willock DJ, Price SL, Leslie M, Catlow, CRA (1995) J Comput Chem 16:628–647
8. Popelier P, Lenstra ATH, Van Alsenoy C, Geise H (1989) J Am Chem Soc 111:5658–5660
9. Ferenczy GG, Csonka GI, Náray-Szabó G, Ángyán JG (1998) J Comput Chem 19:38–50
10. Yatsenko AV, Paseshnichenko KA (2000) Chem Phys 262:293–301
11. Karelson MM, Zerner MC (1992) J Am Chem Soc 96:6949–6957
12. Jensen F (1999) Introduction to computational chemistry. Wiley, Chichester
13. Li J, Zhu T, Cramer CJ, Truhlar DG (1998) J Phys Chem A 102:1820–1831
14. Li J, Williams B, Cramer CJ, Truhlar DG (1999) J Chem Phys 110:724–733
15. Li J, Cramer CJ, Truhlar DG (2000) Int J Quantum Chem 77:264–280
16. Exner O (1975) Dipole moments in organic chemistry. Georg Thieme, Stuttgart, p 74
17. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su SJ, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347–1363
18. Dunning TH (1971) J Chem Phys 55:716–723

19. Laikov DN (1997) Chem Phys Lett 281:151–154
20. Becke AD (1988) Phys Rev A 38:3098–3100
21. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789
22. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865–3868
23. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107:3902–3909
24. Stewart JJP (1989) J Comput Chem 10:209–220
25. Stewart JJP (1989) J Comput Chem 10:221–264
26. Stewart JJP (1993) MOPAC7 program package, QCPE No. 455
27. Ridley J, Zerner M (1973) Theor Chim Acta 32:111–134
28. Ridley JE, Zerner MC (1976) Theor Chim Acta 42:223–236
29. Lipiński J (1988) Int J Quantum Chem 34:423-435
30. Linderberg J, Seamans L (1974) Int J Quantum Chem 8:925–940
31. Dick B, Hohlneicher G (1979) Theor Chim Acta 53:221–251
32. Dick B, Nickel B (1983) Chem Phys 78:1–16
33. Pariser R, Parr RG (1953) J Chem Phys 21:767–776
34. Ellis RL, Kuehnlenz G, Jaffé HH (1972) Theor Chim Acta 26:131–140
35. De Proft F, Tielens F, Geerlings P (2000) J Mol Struct (Theochem) 506:1–8
36. Rzepa HS, Yi MY (1990) J Chem Soc Perkin Trans 2:943–951
37. Zerner MC (1991) Semiempirical molecular orbital methods. In: Lipkovits KB, Boid DB (eds) Reviews in computational chemistry, vol. 2. VCH, New York, pp 313–365
38. Mühlpfordt A, Schanz R, Ernsting NP, Farztdinov V, Grimme S (1999) Phys Chem Chem Phys 1:3209–3218
39. Baumann W, Nagy Z (1993) Pure Appl Chem 65:1729–1734
40. Reichardt C (1988) Solvents and solvent effects in organic chemistry. VCH, Weinheim, p 168
41. Allmann R (1969) Z Kristallogr 128:115–132
42. Morley JO, Morley RM, Docherty R, Charlton MH (1997) J Am Chem Soc 119:10192–10202
43. Gorelik MV, Titova SP, Trdatyan VA (1979) Zh Org Khim 15:167–171
44. Ektova LV, Shishkina RP, Fokin EP (1988) Izv Akad Nauk, SSSR Ser Khim 1672–1676
45. McClellan KAL (1974) Experimental dipole moments, vol. 2. Rahara Enterprises, San Francisco, Calif.
46. Matsunaga N, Dixon DA (1992) J Phys Chem 96:6232–6241
47. Liptay W (1974) Dipole moments and polarizabilities of molecules in excited electronic states. In: Lim EC (ed) Excited states, vol. 1. Academic Press, New York, pp 129–229
48. Liptay W (1965) Dipole moments of molecules in excited states and the effect of external electric field on the optical absorption of molecules in solutions. In: Sinanoǧlu O (ed) Modern quantum chemistry. Part III: action of light and organic crystals. Academic Press, New York, pp 45–66
49. Suppan P (1990) J Photochem Photobiol A Chem 50:293–330
50. Ravi M, Samanta A, Radhakrishnan TP (1994) J Phys Chem 98:9133–9136
51. Mataga N, Nishimoto K (1957) Z Phys Chem (Frankfurt) 13:140–157
52. Schweig A, Reichardt C (1966) Z Naturforsch Part A 21A:1373–1376
53. Liptay W (1966) Z Naturforsch Part A 21A:1605–1618