REGULAR ARTICLE

Solvatochromic probe in molecular solvents: implicit versus explicit solvent model

Andrzej Eilmes

Abstract Solvent-induced shifts in the absorption spectrum of *N,N*-diethyl-4-nitroaniline were studied by quantum-chemical methods in water, dimethylsulfoxide, acetonitrile and acetone. TDDFT methodology and semiempirical ZINDO/S and PM6-CIS approaches were used to calculate excitation energies. Solvent effect was modeled in implicit solvent model by different variants of the PCM approach. Classical molecular dynamics was applied to obtain solute–solvent geometries used in explicit solvent modeling. Most implicit solvent models fail to reproduce the sequence of solvatochromic shifts for four studied solvents, usually yielding too small effect for water. The best result of the PCM method was obtained with SMD atomic radii. Semiempirical quantum-chemical methods in explicit solvent model did not provide satisfactory description of solvatochromic shifts with the largest disagreement to experiment observed for water. TDDFT explicit solvent calculations performed the best in modeling of spectral shifts. Problems with reproduction of experimental data were attributed to specific interactions.

Keywords Solvent effects · Absorption spectra · Solvatochromic shifts · Solvent modeling · Explicit solvent

1 Introduction

It is well known that electronic spectra of organic molecules in gas phase and in solution are different. Interactions of the solute with the solvent change the energies of the ground and of the excited states of the molecule, and therefore, they induce shifts of the maxima in absorption or emission spectra. Two important factors contributing to the overall effect are the long-range electrostatic interactions with the solvent polarizing the solute and the short-range contributions such as dispersion or specific solute–solvent interactions, e.g., hydrogen bonding. Properties of the solvent which affect the spectra of the solute are commonly described by empirical parameters quantifying solvent polarity/polarizability or its ability to act as a hydrogen bond donor or acceptor. Examples of such scale are the Kamlet–Taft π^*, α and β parameters [1–3] derived experimentally from the values of solvent-induced shifts in absorption spectra of molecules serving as solvatochromic probes [4].

Two extreme approaches used to account for the solvent effect in quantum-chemical calculations are implicit and explicit solvent models. In the implicit approach, the solute is inserted into a cavity in a continuous solvent treated as an effective medium. In such a way, the electrostatic component of the solvent effect can be estimated from the interactions of the solute with surrounding polarizable medium; corrections to the energy resulting from van der Waals interactions can be also parametrized. Common examples of implicit solvent models are polarizable continuum model (PCM) [5] or conductor-like screening model (COSMO) [6]. Such methods can provide corrections to the energy of the solute in the solvent or values of solvatochromic shifts; they are computationally cheap and therefore widely used. Significant drawback of implicit approaches is that they do not account for specific interactions and the dispersion and repulsion contributions are treated via an effective parameterization.

At the other end, there are the approaches in which solvent molecules are explicitly included in calculations.

A. Eilmes (✉)
Faculty of Chemistry, Jagiellonian University, Ingardena 3,
30-060 Kraków, Poland
e-mail: eilmes@chemia.uj.edu.pl

To generate structures of the system at molecular resolution simulations of molecular dynamics (MD) are usually performed. Depending on the needs and/or available resources classical MD, Monte Carlo (MC) methods or first-principles MD may be used to sample the solvent configurations, another choice is combined QM/MM approaches treating part of the system classically while the other part is considered at quantum-chemical level. Snapshots of the MD trajectory are then used to extract the geometries of the solute and its solvation shell, which are subsequently passed as input structures to QC calculations. Quantum-chemical computations may be performed for the whole solute–solvent system, or the solvent molecules may be replaced by point charges in QM/MM approach or by polarizable multipoles (polarizable embedding methods [7, 8]). Some recent examples include works using classical MD or MC methods [7–17] or ab initio molecular dynamics [18–23].

Both implicit and explicit solvent models may be combined: few solvent molecules closest to the solute are treated explicitly and the whole system is embedded in a continuous solvent. Specific interactions between solute and explicit solvent molecules may be therefore accounted for and the implicit solvent provides corrections originating from the bulk medium with computational cost significantly lower than in fully explicit models. For these reasons, the mixed approach is quite often used in quantum-chemical calculations [15, 24–26].

Choice of the method applied to modeling of solvent effect is usually based on a balance between the speed of calculations and the requested accuracy; data sets comparing predictions of different approaches are necessary to make such decision. Such information is available for several systems, however, studies of the performance of different models for series of solvents are even more helpful in this context because allow one to verify presence or lack of systematic trend.

In this paper, we wanted to check how well popular implicit solvent approaches and the explicit model describe solvent-induced shifts of one of typical solvatochromic probes in common molecular solvents. Methods available in standard quantum-chemical software were used, in order to test the performance of routine approaches. Molecule of N,N-diethyl-4-nitroaniline (DEPNA) solvated in acetonitrile, acetone, dimethylsulfoxide and water were chosen as model systems. DEPNA is a solvatochromic probe used, e.g., in determination of Kamlet–Taft parameters. The molecule is able to form hydrogen bonds, and therefore, it may be used to assess how specific interactions affect the results of different approaches. Different computational methods, including cheap semiempirical approaches, were used to find the excitation energies. Results are compared to experimental data.

2 Calculations

Gaussian 09 rev. A.02 [27] was used in all quantum-chemical calculations in this work. Geometry of DEPNA was optimized in vacuum using Density Functional Theory methodology with the B3LYP functional and the 6-311+G* basis set. Resulting structure was subsequently used in calculations of excited states in vacuum and in solution with implicit solvents modeled by different variants of the polarizable continuum model (PCM) implemented in the software: integral equation formalism model of the polarizable continuum model IEFPCM [28], the polarizable conductor solvent model (C-PCM) [29] (which may be viewed as an implementation of COSMO model in PCM formalism) and the self-consistent isodensity PCM (SCIPCM) model [30]. The IEFPCM calculations were also performed with atomic radii and non-electrostatic terms originating from the SMD universal solvation model [31].

The TDDFT (Random Phase Approximation) 6-311+G* calculations were applied to obtain the energies and oscillator strengths of electronic transitions. Hybrid B3LYP functional was used in most calculations. As will be shown in the next section, the lowest electronic transition in DEPNA involves significant intramolecular charge transfer, therefore three functionals with long-range corrections: CAM-B3LYP [32], LC-ωPBE [33] and wB97X [34] were tested with default PCM parameters. In addition to TDDFT method, two semiempirical approaches were used: the ZINDO/S method [35] and the semiempirical PM6 [36] Hamiltonian combined with configuration interaction with single excitations (CIS) calculations.

Classical molecular dynamics simulations were performed using Tinker v. 5.1.08 [37] software for a single N,N-diethyl-4-nitroaniline molecule solvated in four liquids considered in this work. Force field parameterization was based on the MM3 force field parameters distributed with Tinker. For dimethylsulfoxide (DMSO), additional parameters were taken from the MM3 study of sulfoxides [38]. Partial atomic charges were obtained from the fit to the electrostatic potential of individual molecules calculated at the B3LYP/6-311++G** level.

System sizes used in MD simulations ranged from about 500 molecules of acetone or dimethylsulfoxide, 1,000 molecules of acetonitrile to about 1,500 molecules of water. DEPNA molecule was put in the center of the simulation box, and its structure was frozen at the vacuum B3LYP/6-311+G* geometry. With frozen geometry of the solute analysis of calculated spectra is easier, because solvent-induced shifts are not obscured by broadening of the spectrum caused by different conformations of the molecule. Periodic boundary conditions were used, and the densities of the systems DEPNA/solvent were set to the densities of pure liquids. Cut-off of 9 Å was applied to the van

der Waals interactions, and Ewald summation was used for electrostatic interactions. All MD simulations were performed in the NVT ensemble at 300 K. After 0.2 ns of equilibration, the MD trajectory was collected for each system for 1 ns.

Accumulated MD trajectories were used to prepare the input data for quantum-chemical calculations of excited states in explicit solvent. First, 100 equally spaced frames were selected from each trajectory. Next, the DEPNA molecule and N solvent molecules closest to it were extracted from each frame; the Trajectory Sculptor tool of the InSilicoLab [39] was used in manipulations with the MD trajectory. In this way, sets of 100 structures with the solvatochromic probe solvated in increasing amount of explicit solvent were prepared. The number of explicit solvent molecules N ranged from 1 to 250 depending on the solvent and the quantum-chemical method used in calculations (larger values were used for smaller solvent molecules, i.e., water and acetonitrile or/and semiempirical methods). Quantum-chemical calculations were performed to obtain transition energies for each structure, and the results were averaged over all 100 configurations used for given solvent and N value. The same QC methods were used as applied previously in the calculations for DEPNA in vacuum. To reduce the computational cost of B3LYP or CAM-B3LYP TDDFT calculations only the solute molecule was described by the 6-311+G* basis set; for the solvent the 3-21G basis was used. With increasing number of explicit solvent molecules increasing number of non-absorbing (with very small oscillator strength) excited states appearing below the first intense state was observed in B3LYP calculations; the number of such states was greatly reduced when the CAM-B3LYP functional was used. Therefore, the number of computed excited states had to be adjusted so that the absorbing state was always included in calculations. In practice, the number of calculated excited states ranged from 10 to 50 depending on the number of solvent molecules and the functional.

3 Results and discussion

3.1 Quantum-chemical calculations in vacuum and in implicit solvation models

Vertical transition energies of the lowest observable optical transition calculated for DEPNA in vacuum within ZINDO/S and PM6-CIS methodologies are 372.4 and 394.3 nm, respectively. From TDDFT/6-311+G* calculations, the energies of 345.8, 307.7, 283.9 and 289.7 nm were obtained for B3LYP, CAM-B3LYP, LC-ωPBE and wB97X functional, respectively. In TDDFT/B3LYP and PM6-CIS results, this energy corresponds to the lowest
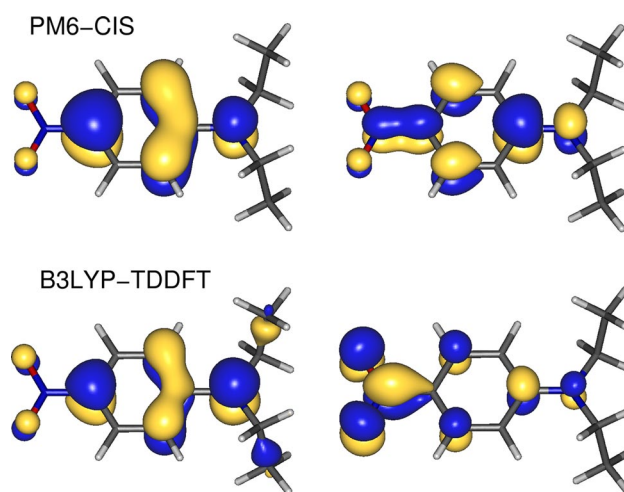
Fig. 1 Dominant natural transition orbital pairs for the lowest observable transition calculated for *N,N*-diethyl-4-nitroaniline in vacuum

calculated transition; in TDDFT with other functionals and in ZINDO/S computations, there are one (TDDFT) or two (ZINDO/S) transitions with zero oscillator strength below the first intense transition. In all cases, the first observable transition originates mainly from HOMO → LUMO excitations.

Spurious low-lying states found for range-separated functionals are not observed in B3LYP calculation in vacuum. Apparently, the rather small range of charge transfer in DEPNA does not require long-range corrections. We will, however, see later that these corrections greatly reduce the problem of spurious states in explicit solvent. Based on vacuum calculations, one may expect that long-range-corrected functionals may overestimate the transition energies in solution; nevertheless, as we are interested mainly in differences between the energies in vacuum and in the solution, selected results for TDDFT calculations with these functionals will be presented.

In Fig. 1 the dominant Natural Transition Orbital [40] pairs obtained in vacuum for the lowest observable transition are displayed. Apparently, both orbitals have a nodal plane in the plane of the molecule; therefore, the transition is of the $\pi \rightarrow \pi^*$ type. Upon excitation, the electron density shifts from the central ring of the molecule toward the NO_2 group. This shift in TDDFT results is slightly larger than in the semiempirical approach, but otherwise there are no major differences between the two methods in the shape of the NTOs involved in the transition.

Experimental values [1, 41] of the absorption maximum of DEPNA molecule in different solvents are listed in Table 1. The largest solvent-induced red-shift is measured for water, the shift induced in DMSO is 20 nm smaller and the smallest increase of λ_{max} is observed for acetonitrile

Table 1 Wavelengths (in nm) of the lowest observable transition calculated for DEPNA molecules in implicit solvation models

	Water	DMSO	Acetonitrile	Acetone
ε_{static}, ε_{∞}	78.4, 1.78	46.8, 2.01	35.7, 1.81	20.5, 1.85
Exp.	430.5[a]	411.5[a]	397.9[b]	396.5[a]
TDDFT(B3LYP)/PCM	392.8	394.3	392.1	391.2
TDDFT(B3LYP)/PCM(equil. solv.)	417.4	416.2	415.2	412.4
TDDFT(B3LYP)/PCM(st. spec. solv.)	426.3	427.8	425.0	423.3
TDDFT(CAM-B3LYP)/PCM	348.5	349.7	347.8	347.0
TDDFT(CAM-B3LYP)/PCM(st. sp. solv.)	390.2	391.7	388.8	386.9
TDDFT(LC-ωPBE)/PCM	317.3	318.2	316.7	316.1
TDDFT(wB97X)/PCM	324.5	325.6	324.0	323.3
TDDFT(B3LYP)/PCM(SAS)	365.6	357.2	358.6	357.1
TDDFT(B3LYP)/CPCM	394.5	396.3	394.0	393.3
TDDFT(B3LYP)/SCIPCM	378.1	377.7	377.4	376.4
TDDFT(B3LYP)/SMD	411.4	394.8	395.2	393.2
TDDFT(B3LYP)/SMD(st. spec. solv.)	445.8	429.3	430.6	427.4
ZINDO/PCM	502.8	513.6	501.2	499.1
PM6-CIS/PCM	430.9	436.8	431.1	431.4

[a] Ref. [1]

[b] Ref. [41]

and acetone. Wavelength of the maximum apparently correlates with the static dielectric constant of the solvent. On the other hand, there is no correlation between position of the maximum and the dielectric constant at infinite frequency. The latter parameter does not change significantly between four solvents considered in this work.

The first test of the implicit solvent approach for DEPNA was performed with standard methodology used in Gaussian 09 for integral equation formalism model of the Polarizable Continuum Model (IEFPCM; for brevity denoted in Table 1 as PCM) [28]. Program defaults were used to build the molecular cavity. There are large differences in predicted transition energies depending on the functional used, with long-range-corrected functionals yielding much too high transition energies. However, it may be easily verified that these differences reflect just the differences in the vacuum transition energies, and therefore, it is still reasonable to compare the vacuum-to-solvent shift values obtained for different solvents for given functional. A striking feature of the TDDFT PCM solvent-induced shifts is that they are not much dependent on the solvent—values calculated for water, acetonitrile and acetone are almost the same and the shift for DMSO is only 2 nm larger. Noticeably, the largest red-shift (with respect to the vacuum value) is observed for DMSO, not for water. Therefore, the sequence of solvatochromic shifts increasing from acetone to water is not reproduced at all. Similar is the case of relative shifts obtained within the PCM method applied to ZINDO/S and PM6-CIS calculations. Although absorption maxima in both semiempirical approaches are calculated at significantly lower energies than in TDDFT (especially the ZINDO/S values), the λ_{max} value for DMSO is remarkably larger than for other three

solvents (for which the shifts are similar). Transition orbitals obtained in implicit solvent in TDDFT or PM6-CIS approach do not differ from the corresponding plots presented for vacuum results in Fig. 1. It should be noted that in all cases (regardless of the quantum-chemical method and the solvent) there is no low spurious state and the intense HOMO → LUMO transition is always the lowest root obtained in continuous solvent.

Two more variants of the PCM method were tested: polarizable conductor solvent model (C-PCM) [29] and the self-consistent isodensity PCM (SCIPCM) model [30] in which the cavity is self-consistently determined from an isodensity surface. Results of the C-PCM calculations are essentially the same as the PCM data apart from small (2–3 nm) shift to longer wavelengths. Transition energies obtained from the SCIPCM are higher and they do not differ much between solvents, like in other models. In this case, at least, the sequence of solvatochromic shifts increasing from acetone to water agrees with the experiment. With differences between solvents less than 1 nm, this result is still very far from reproduction of experimental data where positions of absorption maxima differ by 10–20 nm.

In addition to the above-described models, the IEFPCM B3LYP calculations were performed with atomic radii and non-electrostatic terms parameters from the SMD universal solvation model [31]. As readily seen, the result for water has improved: λ_{max} is about 15 nm larger than for the three other solvents. Although large solvatochromic shift for water has been at least partially captured, and the values for two least polar solvent (acetone and acetonitrile) are fairy well reproduced, description of DMSO is still poor. Position of the maximum expected from the experimental data should be between those obtained for water and

acetonitrile, but the value calculated for DMSO is even smaller than the λ_{max} for acetonitrile.

The default Gaussian 09 PCM method constructs the molecular cavity based on the van der Waals molecular surface and the atomic radii from Universal Force Field scaled by a factor of 1.1. As a test, several B3LYP calculations with other available options were performed including different scaling factors, different choices of atomic radii and other definitions of molecular surface—Solvent Excluding Surface and Solvent Accessible Surface. It turned out that in most cases the effect of modified cavity is merely an uniform shift of all values obtained for standard PCM, depending on whether given set of parameters yields larger (resulting in smaller λ_{max}) or smaller cavity (reduced distance from the molecule to the dielectric continuum increases the effect of the medium and leads to increasing wavelengths of absorption maxima). In such cases, the calculated relative shifts were practically the same as for PCM results, with small differences of λ_{max} between solvents and the largest value obtained for DMSO. The only exception was observed for the results obtained with the cavity based on Solvent Accessible Surface (SAS). In this case, the λ_{max} for water is apparently larger (by about 7–8 nm) than for other solvents (Cf. Table 1). For all solvents, absorption maxima appear at shorter wavelengths, because SAS-based cavity is the largest from all used in calculations.

The results based on the SMD-based parameters and the SAS surface show that the choice of atomic radii or the cavity construction may affect the relative values of solvatochromic shifts. Nevertheless, in either case discussed so far, the sequence of measured λ_{max} and its correlation with static dielectric constant of the solvent are not reproduced.

Let us recall at this point, the difference between equilibrium and non-equilibrium solvation. Polarization of the dielectric medium arises from the reorientation of the solvent and from the changes in electron distribution on solvent molecules. The first process occurs at the timescale of molecular motions, and it is much slower than the latter. During vertical excitation, the solvent can respond to the changes of the electron density at the solute molecule via rapid polarization of the electron cloud, therefore non-equilibrium solvation is used in PCM to model the electronic excitation of molecules. The parameter of the solvent relevant to this process is the value of the dielectric constant corresponding to infinite frequency. As seen from Table 1, ε_{∞} only very little differs between four investigated solvents; interestingly, it is the largest for DMSO, which corresponds with the fact that in most models absorption maximum for DMSO exhibited the largest red-shift.

Notwithstanding the fact that the optical dielectric constant governs the solvent response to the changes of electron density between the two states, static dielectric constant still affects the vertical transition energies through

determining the polarization field for the ground state of the molecule. The relative importance of both effects depends on the difference between dipole moment in the ground and in the excited state and may be estimated, e.g., using Eq. 19 of Ref. [42]:

$$\Delta G_1^{neq} - \Delta G_0^{eq} = -g(\varepsilon_{static})\mu_0(\mu_1 - \mu_0) - \tfrac{1}{2}g(\varepsilon_{\infty})(\mu_1 - \mu_0)^2,$$
(1)

where the left-hand side is the solvent shift between the ground and the excited state with dipole moments μ_0 and μ_1, respectively. The factor $g(\varepsilon)$ relates the reaction field to the field generated by the dipole and for simple model case of a spherical cavity of radius r in a medium described by dielectric constant ε reads [42]:

$$g(\varepsilon) = \frac{2(\varepsilon - 1)}{2\varepsilon + 1}\frac{1}{r^3}.$$
(2)

Such a model of an effective cavity is certainly oversimplified for most molecules (except for small and spherical solutes) but the ratio of the two contributions in (1) does not depend on the particular value of r; therefore, the above approximation can be conveniently used as a rough estimate of the effect.

B3LYP DFT and TDDFT calculations in 6-311+G* basis yielded 8.94 and 14.96 D for the dipole moment of DEPNA in the ground and in the excited state, respectively. Inserting these data and ε_{static} and ε_{∞} values listed in Table 1 into (1) yields for the solvents considered here the ratio of the first to the second term in the RHS of (1) between 7.1 and 8.5, showing that most of the shift comes from the first term involving static dielectric constant. However, it should be noted that this estimate of both contributions refers to the case when the state-specific solvation is considered (see below); in common TDDFT calculations applying linear response approach the solvent shift is proportional to ε_{∞} [42].

Although optical dielectric constant is more adequate to describe solvent response to the rapid process of vertical excitation, one may try to check what the predictions of the PCM model are when the equilibrium solvation for both states is used instead. In Table 1, such results are presented for default IEFPCM model. Apparently, calculated position of the absorption maximum increases monotonically with increasing value of ε_{static}; nevertheless, the difference between DMSO and water is only 1 nm, despite large increase in the dielectric constant. Therefore, we may conclude that applying static dielectric constant to solvation in excited state calculation (i.e., using equilibrium solvation scheme) only qualitatively reproduced the measured sequence of solvatochromic shifts, but still did not improve quantitative predictions. Moreover, for the reason discussed above, it is not a valid approach when the concept of fast and slow contribution to the polarization of the medium

is considered. We should also note that that the observed correlation of λ_{max} with static dielectric constant may be indirect. Small difference of equilibrium solvation results between water and DMSO suggest that the increase in ε_{static} from 47 to 78 leads to rather small changes and therefore electrostatic part of the effect has saturated. Large solvatochromic shift for water originates not only from long-range electrostatics but also from hydrogen bonding. The latter phenomenon is also responsible for large dielectric constant of water; therefore, large value of λ_{max} seems to correlate with large permittivity, although both effects have common origin in specific interactions.

Finally, the state-specific solvation [42] was tested. In this formalism, the iterative procedure is applied to make the electrostatic potential of the excited state self-consistent with the solvent reaction field. Transition energy is then calculated as the difference between the energy of the solvated excited state (with self consistency achieved between excited state density and the response of the solvent) and the ground state energy calculated within equilibrium solvation. Results of the state-specific method are presented for IEFPCM and SMD solvation models. This approach yields lower energies of the excited state and therefore longer transition wavelengths. Apart from this feature, relative differences in solvatochromic shifts between solvents are the same as for standard linear response formalism. Therefore, the mere effect of state-specific method for our set of solvents is the decrease of excitation energies without any systematic improvement with respect to the experimental solvatochromic shifts.

From the above results, we can conclude that neither continuous solvent model tested in this work reproduces correctly the sequence and the size of solvatochromic shifts of DEPNA molecule in four common solvents. While one cannot expect that the difference of 1 nm in the positions of absorption maxima between acetonitrile and acetone will be captured, large differences (10–30 nm) between water, DMSO and other solvents should be at least semi-quantitatively reproduced. In most cases, however, we see the opposite: calculated λ_{max} for all solvents are within the range of less than 5 nm and DMSO, not water, leads to the largest red-shift in the spectrum. Only in two cases (SMD solvation model or IEFPCM with SAS molecular surface), the solvatochromic shift for water was remarkably larger than for other solvents, however, simultaneously, DMSO shift is smaller than for acetonitrile in clear disagreement with experiment.

It has been shown in the literature that SMD atomic radii used in PCM calculations give better description of solvent effect on absorption spectra [24, 43]. Results for DEPNA shown here agree, therefore, with these earlier findings. Likewise, remarkable difference in λ_{max} between water and other solvents was also obtained when SAS surface

was used to construct the molecular cavity. These results suggest that proper description of atomic radii and molecular surface (affecting the strength of electrostatic solute–solvent interactions) is one of key factors responsible for reproduction of experimental data. Ability of SMD radii to capture the effect for water is related to the fact that some SMD radii effectively depend on the hydrogen bond donor properties of the solvent, thus account partially for specific interactions. Another factor which presumably improved the performance of the SMD model is the effective treatment of dispersion and repulsion contributions to the solvation energy which was not included in other PCM calculations. These effects are not well described in implicit solvents but can be in principle captured in explicit models (depending on the capability of the quantum-chemical method).

It is worthwhile to mention that the overall disagreement of implicit solvation results with experimental data becomes apparent only when several solvents are taken into account. In fact, selected individual results, e.g., TDDFT/IEFPCM, CPCM or SMD values with B3LYP for acetone and acetonitrile or PM6-CIS/PCM result for water, examined alone would be considered as a good reproduction of measured absorption maximum. In such cases, it is necessary to look at the whole set of solvents to realize that the trend exhibited by the experimental data is not reproduced.

3.2 Explicit solvation

Turning our attention to explicit solvation models we analyzed the results of sequential MD/QC modeling described in Sect. 2. To get some insight into the solvation structure of DEPNA, solute–solvent radial distribution functions (RDF's) are presented in Fig. 2. The first $\pi \rightarrow \pi^*$ electronic transition involves a change of the electron density at NO_2 group, therefore the structure of the solvation shell close to this part of DEPNA molecule should be most important for the solvatochromic shifts. Accordingly, RDF's shown in Fig. 2 were calculated between the nitrogen atom from the NO_2 group and the oxygen atom of the water molecule or the carbon atom from the CH_3 groups for the three other solvents. Several maxima related to the solvation shells in DEPNA-water system are clearly visible in the plot; the first and highest appears at 4 Å, two other at 7 and 10 Å and the last diffuse maximum is noticeable at 13 Å. The first maximum for $N(O_2)$–$C(H_3)$ RDF's is found at 4.3–4.5 Å, the second at about 8.5 Å (acetonitrile) or 9 Å (acetone and DMSO), the third, weak and diffuse at about 13 Å. Integrating the RDF's we found the number of solvating molecules up to the distance above the first peak, i.e., 5 Å for water, 6 Å for acetonitrile and 6.5 Å for DMSO and acetone. The average values were 14 water, nine acetonitrile molecules
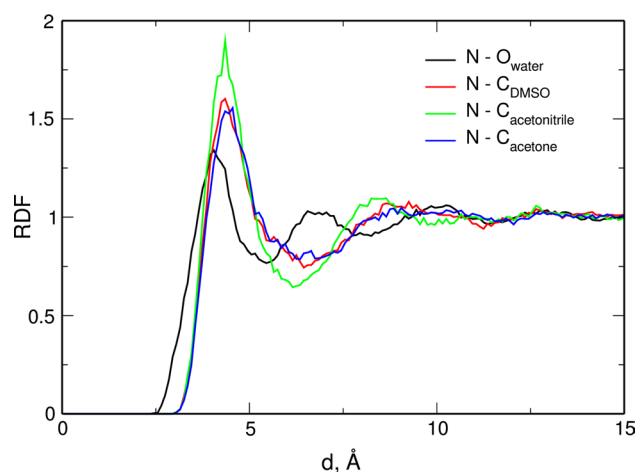
Fig. 2 Radial distribution functions obtained from MD trajectories for DEPNA and different solvents

Fig. 3 Absorption spectra simulated for DEPNA in DMSO for increasing number N_{mol} of explicit solvent molecules

and 16–17 CH_3 groups from acetone or DMSO, i.e., 8–8.5 molecules in the first solvation shell of the NO_2 group.

To produce plots of simulated absorption band, the spectrum was calculated as a superposition of normal peaks placed at calculated transition energies with heights corresponding to computed oscillator strengths and standard deviations $\sigma = 500$ cm^{-1}.

In Fig. 3, an example of changes in simulated spectrum is presented for DMSO; only the part corresponding to the first intense transition is displayed. For an isolated DEPNA molecule (no solvating DMSO molecules), the shape and the width of the band corresponds to the width of the Gaussian curve used to convolute individual transitions in the simulation of the spectrum. When the number of solvent molecules is small ($N = 1$–10), the Gaussian shape of the band is preserved, although its width increases. Simultaneously, the maximum shifts to longer wavelengths. For larger values of N, the absorption band widens even more and flattens, but its overall shape still resembles more/less a normal curve. Although increasing N leads to further red-shift of the peak, this increase gets smaller for sufficiently large N. The largest changes in the position of the band are observed for N between 5 and 20. This is not surprising: when there are only few (less than 5) solvent molecules, their effect is negligible and the spectrum of the solute is barely affected. About 15 DMSO molecules are necessary to complete the solvation shell around the whole DEPNA molecule inducing most prominent changes in the spectrum. Further solvent molecules are placed in the next solvation shell, therefore, do not interact directly with the solute and their effect is smaller. Similar dependence is observed for other solvents with the major difference that for smaller solvent molecules (water and acetonitrile) much larger N is necessary to effectively solvate the probe molecule.
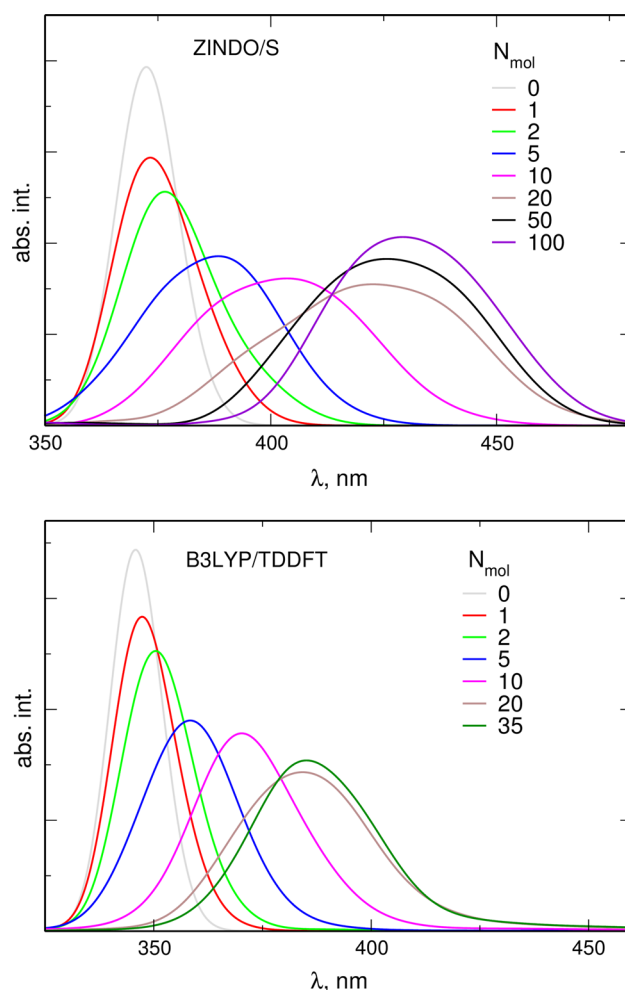
Changes in the position of the maximum in simulated spectra are displayed in Fig. 4. Because solvent molecules differ in size, total number of solvating atoms N_{atoms} rather than the number of solvent molecules N is given as the measure of the system size. The three approaches used to calculate excitation energies differ greatly in the computational cost; therefore, the largest system sizes were reached in cheapest ZINDO/S calculations, whereas cost of the TDDFT computations imposed much stronger upper bound on attainable N_{atoms} values.

Features already discussed for DMSO results are readily seen: at about 100–200 atoms in the solvation shell the slope of the curves decreases and the increase of λ_{max} tends to saturate. Unfortunately, only in few cases (acetone/ZINDO, acetone/B3LYP/TDDFT), the regime in which the increase in the number of solvent atoms does not change the position of the absorption maximum has been reached. In particular, shifts of the spectrum are still well pronounced for water.
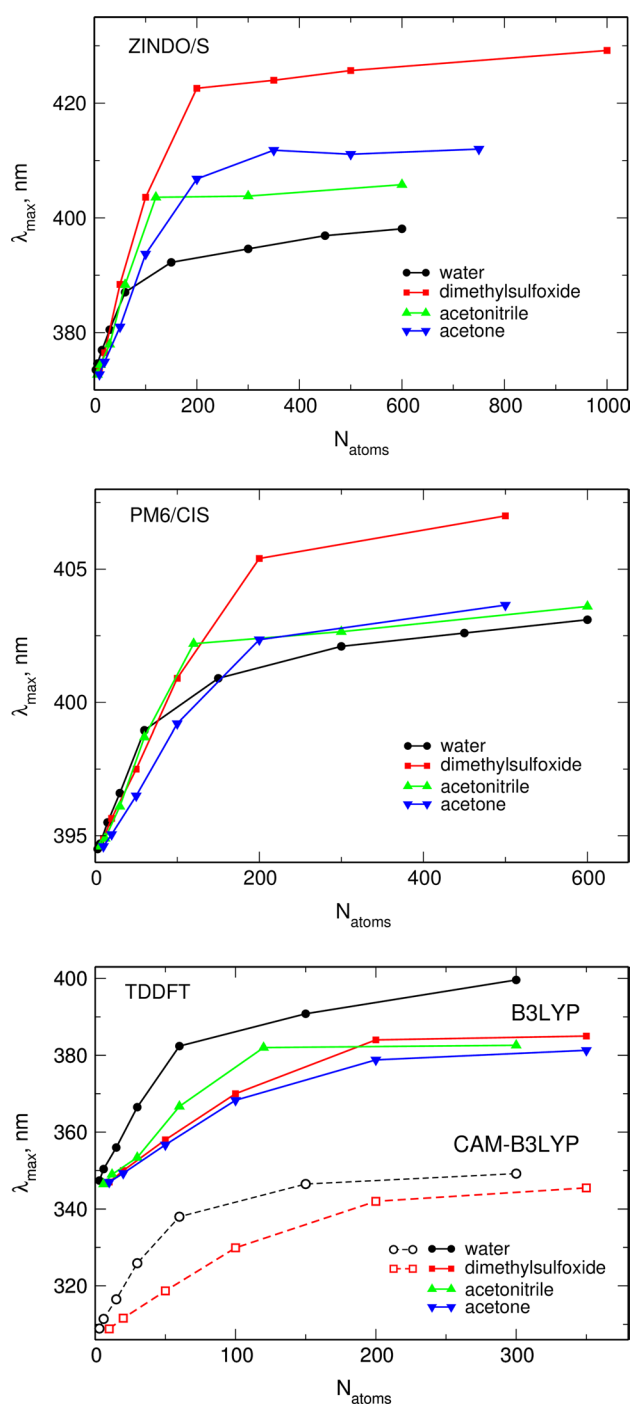
Fig. 4 Wavelengths of the maximum of the absorption spectrum of DEPNA calculated in explicit solvents for increasing number of atoms N_{atoms} in solvating molecules

Even a short glimpse at Fig. 4 reveals that semiempirical approaches combined with explicit solvent model yield wrong predictions of the transition energies of DEPNA in solution. ZINDO/S calculations give the largest difference between solvents (up to 30 nm); however, the results disagree with experiment: the largest shift is computed for

DMSO and the smallest for water; the effect of two other solvents is intermediate. The solvent with highest dielectric constant induces, therefore, the smallest solvatochromic shift in clear opposition to measured data (Table 1). Qualitatively similar is the picture for PM6-CIS calculations, although in such cases relative shifts are smaller: λ_{max} for acetone and acetonitrile is only little (1 nm) larger than for water, and the absorption maximum in DMSO is about 5 nm red-shifted with respect to water.

In the results of B3LYP TDDFT method, λ_{max} obtained for acetone, acetonitrile and DMSO is similar, slightly increasing in this sequence, with about 3 nm difference between acetone and DMSO calculated for largest system sizes. Maximum of the absorption band for water appears at about 15 nm longer wavelengths than for DMSO. Although this difference is smaller than the experimental value, and, unlike the experiment, position of the maximum for DMSO is close to those for acetonitrile and acetone, explicit solvation in B3LYP/TDDFT calculations is the sole model which at least qualitatively reproduces the sequence of solvatochromic shifts and captures significant increase of λ_{max} for water. CAM-B3LYP functional yields too high excitation energies in solution, which is the consequence of high excitation energy calculated in vacuum. The difference between water and DMSO in CAM-B3LYP calculations is smaller than in B3LYP results, though their sequence agrees with the experimental data.

Comparing whole sets of results to the experimental data one can note that in semiempirical approaches maxima of absorption for DMSO, acetonitrile and acetone are predicted at wavelengths longer than measured, in general agreement with PCM results where ZINDO/S and PM6-CIS approaches overestimated λ_{max} for all solvents. In explicit solvent, however, this overestimate is much smaller (and the position of the maximum for DMSO calculated in PM6-CIS is even underestimated). The major difference is observed for DEPNA in water, in which case semiempirical spectra are shifted to lower wavelengths. It appears therefore that the problem with reproduction of the sequence of solvatochromic shifts in explicit solvents is aggravated by the failure of semiempirical methods to yield significant red-shift in water.

On the other hand, in the TDDFT calculations with explicit solvent, all spectra are computed at higher energies compared to experimental data (especially for CAM-B3LYP as a result of high transition energy calculated in vacuum). Better agreement is found for B3LYP functional in which case λ_{max} calculated for largest numbers of solvent molecules are 400, 385, 383 and 381 nm for water, DMSO, acetonitrile and acetone, respectively, i.e., 30–15 nm lower than in experiment (Table 1). It should be, however, taken into account that some results are rather far from saturation (e.g., λ_{max} would probably increase

quite significantly when more explicit H_2O molecules were included in calculations). The systematic difference to the experiment may be also caused by too small basis set (reduced in this work to make computations tractable) and presumably the problem might be alleviated if some other functional or better method were used to calculate energies of excited states. The solute–solvent configurations used in quantum-chemical calculations resulted from classical molecular dynamics simulations with frozen solute. The geometries of systems used to calculate absorption transition are affected by the quality of the force field, being another factor which may lead to the differences with respect to the experiment. Fully quantum-chemical approach is of course possible (with the use of ab initio Molecular Dynamics) but incurs much larger computational cost of dynamics simulations.

The major disagreement between experiment and predictions of semiempirical models, much better behavior of the TDDFT methodology together with the results of implicit solvent approaches suggest that the large solvatochromic shift in water is partially caused by solute–water interactions which are not well captured in implicit solvent or semiempirical calculations. The possible reason of better stabilization of the excited state versus the ground state may be specific interactions. Indeed, the NO_2 group of the DEPNA molecule may act as an acceptor in hydrogen bond in the solvent capable of being a hydrogen donor, e.g., in water. The exited state has the electron density on oxygen atoms of NO_2 group larger than the ground state (Cf. Fig. 1), therefore, should be more sensitive to the hydrogen bond formation. Such specific interactions are not properly accounted for in implicit models or semiempirical approaches, but are better described in TDDFT calculations. It is therefore expectable that TDDFT computations based on ab initio MD structures could yield even better description of the absorption spectrum, because of improved distribution of hydrogen-bonded water molecules around DEPNA.

To check how large may be the effect of direct interaction between NO_2 group and H_2O the transition energies for a system of DEPNA interacting with a single water molecule located close to the NO_2 group were calculated using B3LYP/TDDFT. The lowest transition energy shifted from the value of 345.8 nm (for isolated DEPNA molecule) to 357.6 nm. In the explicit approach with H_2O molecules spaced around the whole DEPNA molecule, such size of the average red-shift was observed for five water molecules; therefore, even a single molecule interacting with solute through hydrogen bond contributes significantly to the overall effect of the solvent. Inspection of frontier orbitals reveals that the effect of interaction with water molecule is much larger for HOMO (shift of 0.44 eV to lower energies) than for LUMO (shift of 0.27 eV).
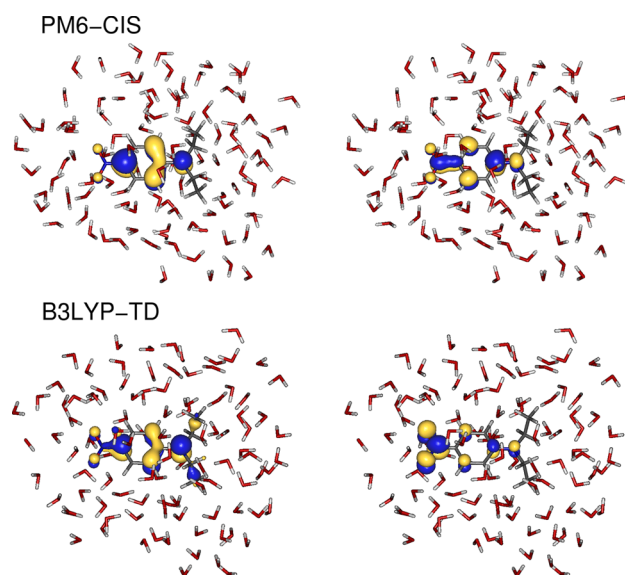
Fig. 5 Dominant natural transition orbital pairs for the lowest observable transition calculated for DEPNA in explicit water

In Fig. 5, plots of dominant Natural Transition Orbital pairs for DEPNA in 100 explicit water molecules calculated by PM6-CIS and B3LYP/TDDFT methods are displayed for the lowest transition with significant oscillator strength. In both cases, the shape of NTOs does not differ from corresponding orbitals obtained in vacuum or continuous PCM solvent. To make the plots, a particular snapshot of the MD trajectory was used, yielding transition energy practically equal to the value averaged over all configurations (i.e., to the position of the maximum in simulated spectrum), but even for configurations corresponding to significantly higher or lower transition energies the picture is very similar, therefore the arrangement of solvent molecules does not affect the shape of molecular orbitals involved in the observed transition.

In all approaches used to calculate excitation energies in continuous solvent model, the first transition with non-negligible intensity corresponds to the HOMO → LUMO excitation. In explicit solvent, this situation changes depending on the solvent and the method. Frontier orbitals may have the electron density distributed not on the dye molecule but on the solvent and the dominant contribution to the observable transition comes from the orbitals lower than HOMO or/and higher than LUMO. In TDDFT results obtained for increasing N, an increasing number of low-lying states with practically no absorption intensity is observed. Appearance of these states is partially a consequence of unbalanced basis set (6-311+G* for the solute, 3-21G for the solvent) as confirmed by the reduction of the number of spurious states in test calculations employing 6-31+G* set for the solvent. On the other hand, in CAM-B3LYP calculations

using 3-21G basis for the solvent molecules spurious states practically disappear in systems with water and their number was greatly reduced for DMSO. Long-range-corrected functionals alleviate the problem with low-lying dark states and allow one to decrease the number of states sought in TDDFT, reducing the cost of calculations. In a recent paper [44], problem of spurious states in TDDFT in explicit water was traced to the "edge" water molecules at the boundary of the cluster. Possible ways of ameliorating the problem described in [44] included the use of range-separated functionals, increasing basis set or embedding the system in external environment. Above we have discussed the two first possibilities; we will comment on the third in the next section.

3.3 Combined PCM/explicit solvent model

A quite frequently used approach combining implicit and explicit solvation models is to perform calculations for a solute and several nearest solvent molecules surrounded by continuous solvent modeled, e.g., as PCM dielectric medium. To test the performance of such method series of B3LYP TDDFT calculations for $N = 5$–50 explicit water or $N = 5$–20 DMSO molecules were repeated with the system (DEPNA $+ N$ solvent molecules) embedded in implicit PCM solvent with either default or SMD-based parameters. Such large number of explicit solvent molecules is not used in routine calculations (usually only a few are included), but here we wanted to check the trend observed when the amount of explicit solvent increases. The simulated spectra for default PCM are shown in Fig. 6 along with the corresponding results obtained from the model with purely explicit solvent.

Even for small number of water molecules, the maximum of simulated absorption peak appears at longer wavelengths than in pure PCM approach and its position is comparable to the results of explicit solvent modeling for systems with largest N. Such result may be attributed to the combined effect of both explicit (providing short-range interactions) and implicit (contributing via long-range electrostatic polarization) parts of the solvent. With N increasing to 20 for water or to 10 for DMSO, the λ_{max} is even more red-shifted and exceeds the value obtained in explicit solvent but for larger N (50 for water, 20 for DMSO) a decrease of λ_{max} is observed. This is not surprising, because for large number of explicit solvent molecules the effect of surrounding continuum becomes small and the results should converge toward the limit obtained in explicit modeling. Results of mixed approach for PCM and SMD models are summarized in Fig. 7. Changes in λ_{max} described above are readily seen. It may be easily noted that the data for DMSO obtained in PCM and SMD models are very similar, in accordance with the results of implicit model
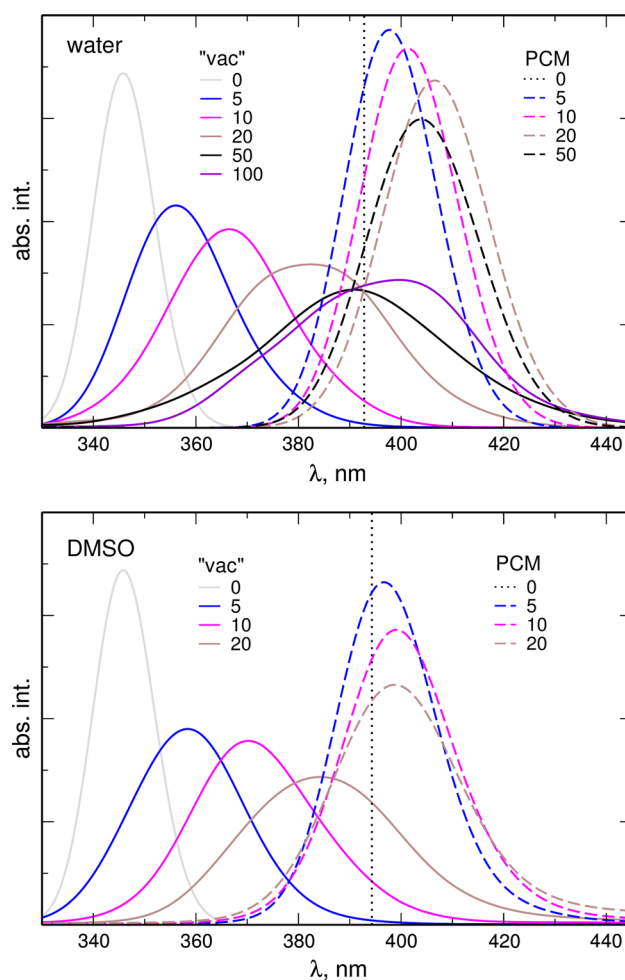
Fig. 6 Absorption spectra of DEPNA simulated for increasing number of explicit solvent molecules in purely explicit ("vac", solid lines) and combined explicit/continuous (default PCM, *broken lines*) approach. *Vertical dotted lines* mark the wavelengths obtained in purely implicit solvent (IEFPCM model)

(Sect. 3.1). On the other hand, purely implicit SMD model for water predicts absorption maximum at almost 20 nm longer wavelengths than default PCM calculations and difference persists even for large number of explicit solvent molecules.

In the hybrid approach, the implicit part of the solvent provides long-range electrostatic corrections to the effects of short-range interactions accounted for by the explicit part of the system. This embedding in external field reduces the problem with spurious states (as suggested in [44])—such states were completely eliminated in explicit/PCM calculations for DEPNA/water even for B3LYP functional and small 3-21G basis used for H_2O molecules. For large N partitioning of the system becomes less clear: disjoint cavities may appear leading to the presence of dielectric medium between explicit molecules, and for large explicit part the results tend toward the values obtainable for largest
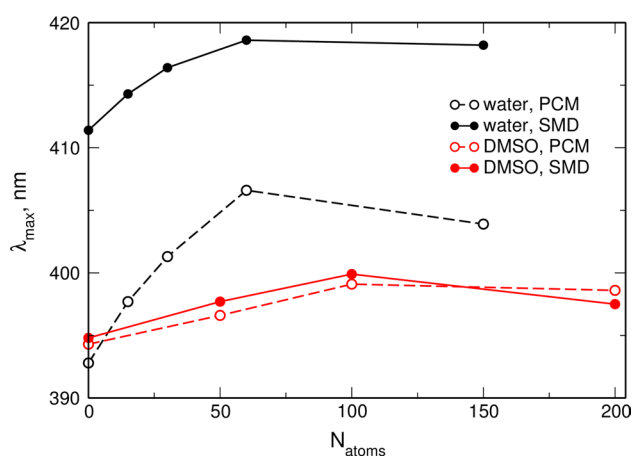
Fig. 7 Wavelengths of the maximum of the absorption spectrum of DEPNA calculated in mixed implicit/explicit solvent for increasing number of atoms (N_{atoms}) in explicit part of the system for default PCM cavity settings and for PCM with SMD-derived radii. Note that $N_{atoms} = 0$ corresponds to the result obtained in purely implicit solvent model

Fig. 8 Full widths at half maximum obtained for simulated absorption bands in explicit and mixed solvent models

systems in purely explicit approach. Nevertheless, one may try to use the largest λ_{max} values obtained in the combined approach as an estimate of the solvatochromic shift for infinitely large system (bulk solvent) provided that the implicit model gives reasonably good description of solvation (as suggested by the difference between PCM and SMD results in Fig. 7), and the quantum-chemical method used appropriately treats the interactions in the explicit part. On the other hand, results presented in this section for small number of H_2O or DMSO molecules in SMD solvent confirm the usual assumption that combining continuous and explicit model is a cheap way to improve the description of solvent effects.

Finally, let us briefly discuss the intensity and the width of absorption band simulated in explicit or combined solvent models. With increasing N, the total absorption intensity decreases a little (less than 10 %) in explicit model and such change is even smaller in combined approach. From the height of the bands in Fig. 6, it is clear that the total intensity obtained in combined model is larger than in explicit solvent. This stems from the values obtained for isolated DEPNA molecule in vacuum and in purely implicit solvent; the latter calculations yield the oscillator strength of the transition about 40 % larger than the former. Changes of the full width at half maximum (FWHM) obtained for simulated band are displayed in Fig. 8. In explicit solvent, FWHM increases rapidly up to N_{atoms} about 50 and then saturates. This is not a surprising result, because increasing the number of solvent molecules initially increases the scatter of energies through varying configurations of the solvent, but above a certain number of explicit molecules they are placed far away from the solute (thus inducing
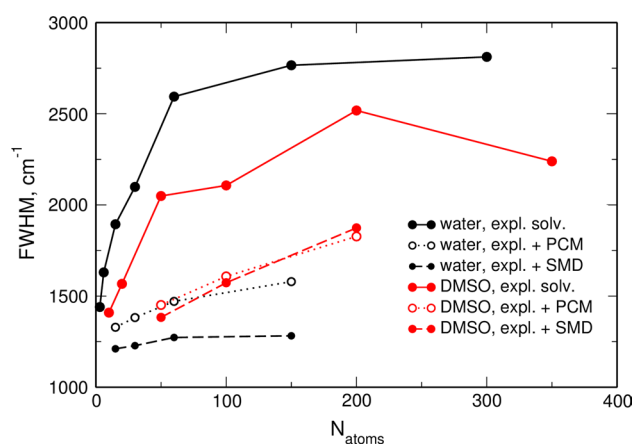
smaller changes) and the effect of fluctuations averages. On the other hand, FWHM for combined solvent model is significantly smaller than for the same N in purely explicit model showing that the reaction field from the implicit part effectively acts like a large number of explicit molecules, suppressing the fluctuations experienced by the solute.

4 Conclusions

Quantum-chemical calculations have been performed for N,N-diethyl-4-nitroaniline in four common solvents using implicit and explicit solvents as well as combined approach. The TDDFT calculations with B3LYP of CAM-B3LYP functionals and two semiempirical methods PM6-CIS and ZINDO/S have been used to obtain the energies and oscillator strengths for absorption transitions.

Most implicit solvent parameterizations (variants of the PCM method) have failed in reproduction of the experimental sequence of solvatochromic shifts correlating with static dielectric constant of the solvent. The most important difference to the experiment is significant underestimate of the absorption shift in water. Some models with atomic radii or molecular surface modified with respect to default PCM cavity settings yield larger differences between water and the three other solvents suggesting that the size of the solute's molecular cavity affects to some extent relative solvent shifts in absorption spectra. Nevertheless, neither implicit solvation model provides satisfactory reproduction of experimental data. The best results have been obtained in SMD calculations, and this fact may be attributed to the set of atomic radii improving description of electrostatic interactions and partially accounting in an effective way for hydrogen bond donating ability of water.

Both semiempirical methods used in explicit solvent model predict much too small effect for water. On the other hand, description of solvent effects in B3LYP/TDDFT results is qualitatively or even semiquantitatively correct. Because the largest disagreement with experiment is observed for DEPNA-water system where hydrogen bonds may be formed between the NO_2 group of DEPNA and the solvent molecules, it seems reasonably to attribute, at lest partially, problems encountered in semiempirical quantum-chemical calculations in explicit solvent to the same origin as in PCM model, i.e., to the deficiencies in description of specific interactions.

Results of combined implicit/explicit solvent model for water seem to converge toward values obtained in explicit model. For small number of explicit water molecules, they are quite close the explicit solvent limit for large N, confirming that combined approach may improve convergence of absorption shifts toward the bulk solvent value at moderate computational cost. Another benefit of mixed solvent approach is the decrease of the number of spurious states in TDDFT calculations.

The important conclusion which can be drawn from the above results is that one should be careful applying implicit solvation models to calculate solvatochromic shifts in solvents possessing ability to form hydrogen bonds. PCM-calculated differences between solvents are likely to be much too small in such case, and the solvent effect for water is not reproduced. Explicitly including solvent molecules, in purely explicit or in mixed solvation model, improves the agreement with experiment. Similar results regarding the improvements brought by the explicit part of the model were recently reported for the solvatochromism of acetone [8] or protein chromophore [26]. If the purely implicit solvent model has to be used, the PCM approach with SMD radii seems to be a suggested choice, yielding the best reproduction of experimental data for solvents with hydrogen-bonding ability.

References

1. Kamlet MJ, Taft RW (1976) J Am Chem Soc 98:377–383
2. Taft RW, Kamlet MJ (1976) J Am Chem Soc 98:2886–2894
3. Kamlet MJ, Abboud JL, Taft RW (1977) J Am Chem Soc 99:6027–6038
4. Reichardt C (1994) Chem Rev 94:2319–2358
5. Miertuš S, Scrocco E, Tomasi J (1981) Chem Phys 55:117–129
6. Klamt A, Schüürmann G (1993) J Chem Soc Perkin Trans 2:799–805
7. Olsen JM, Aidas K, Kongsted J (2010) J Chem Theory Comput 6:3721–3734
8. Schwabe T, Olsen JMH, Sneskov K, Kongsted J, Christiansen O (2011) J Chem Theory Comput 7:2209–2217
9. Brancato G, Rega N, Barone V (2006) J Chem Phys 125:164515
10. Olsen JM, Aidas K, Mikkelsen KV, Kongsted J (2010) J Chem Theory Comput 6:249–256
11. Murugan NA, Kongsted J, Rinkevicius Z, Aidas K, Ågren H (2010) J Phys Chem B 114:13349–13357
12. Eilmes A (2010) Theor Chem Acc 127:743–750
13. Jaramillo P, Coutinho K, Cabral BJC, Canuto S (2011) Chem Phys Lett 516:250–253
14. Sakata T, Kawashima Y, Nakano H (2011) J Chem Phys 134:014501
15. Barone V, Bloino J, Monti S, Pedone A, Prampolini G (2011) Phys Chem Chem Phys 13:2160–2166
16. Manzoni V, Lyra ML, Coutinho K, Canuto S (2011) J Chem Phys 135:144103
17. Ma H, Ma Y (2013) J Chem Phys 138:224505
18. Sok S, Willow SY, Zahariev F, Gordon MS (2011) J Phys Chem A 115:9801–9809
19. Murugan NA, Chakrabarti S, Ågren H (2011) J Phys Chem B 115:4025–4032
20. Isborn CI, Götz A, Clark MA, Walker RC, Martinez TJ (2012) J Chem Theory Comput 8:5092–5106
21. Douma DH, M'Passi-Mabiala M, Gebauer R (2012) J Chem Phys 137:154314
22. Garbulo V, Cascella M, Del Sole R, Marsili M, Pulci O (2012) J Chem Phys 137:164317
23. Eilmes A (2013) J Phys Chem A 117:2629–2635
24. Marenich AV, Cramer CJ, Truhlar DG (2010) J Chem Theory Comput 6:2829–2844
25. Caricato M, Lipparini F, Scalmani G, Capelli C, Barone V (2013) J Chem Theory Comput 9:3035–3042
26. Petrone A, Caruso P, Tenuta S, Rega N (2013) Phys Chem Chem Phys 15:20536–20544
27. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery Jr JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, revision A.02. Gaussian, Inc., Wallingford
28. Cances E, Mennucci B, Tomasi J (1997) J Chem Phys 107:3032–3041
29. Cossi M, Rega N, Scalmani G, Barone V (2003) J Comp Chem 24:669–681
30. Foresman JB, Keith TA, Wiberg KB, Snoonian J, Frisch MJ (1996) J Phys Chem 100:16098–16104
31. Marenich AV, Cramer CJ, Truhlar DG (2009) J Phys Chem B 113:6378–6396
32. Yanai T, Tew D, Handy N (2004) Chem Phys Lett 393:51–57

33. Vydrov OA, Scuseria GE (2006) J Chem Phys 125:234109
34. Chai JD, Head-Gordon M (2008) J Chem Phys 128:084106
35. Ridley JE, Zerner MC (1973) Theor Chem Acc 32:111–134
36. Stewart JJP (2007) J Mol Model 13:1173–1213
37. Tinker Molecular Modeling Package v. 5.1.08. http://dasher.wustl.edu/tinker/
38. Allinger NL, Fan Y, Varnali T (1996) J Phys Org Chem 9:159–167
39. InSilicoLab. http://insilicolab.cyfronet.pl
40. Martin RL (2003) J Chem Phys 118:4775–4777
41. Kamlet MJ, Abboud JL, Abraham MH, Taft RW (1983) J Org Chem 48:2877–2887
42. Improta R, Barone V, Scalmani G, Frisch MJ (2006) J Chem Phys 125:054103
43. Caricato M (2012) J Chem Theory Comput 8:4494–4502
44. Isborn CM, Mar BD, Curchod BFE, Tavernelli I, Martinez TJ (2013) J Phys Chem B 117:12189–12201