Perturbative Multi-Reference Configuration Interaction (CI-MRPT2) Calculations in a Focused Dynamical Approach: a Computational Study of Solvatochromism in Pyrimidine

Ivo Cacelli[a,b], Alessandro Ferretti[b], Giacomo Prampolini[b]

30 marzo 2019

a) Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via G. Moruzzi 3, I-56124 Pisa, Italy
b) Istituto di Chimica dei Composti OrganoMetallici (ICCOM-CNR), Area della Ricerca, via G. Moruzzi 1, I-56124 Pisa, Italy

Abstract

We have investigated solvatochromic effects over a solvent series of increasing polarity on the prototype molecule pyrimidine as solute species. The line shape profiles, obtained by a time dependent approach based on quantum mechanical calculations performed over frames sampled from classical molecular dynamics trajectories, were directly compared to the available experimental bands. The multi-reference configuration interaction–second order perturbation (CI-MRPT2) calculations are in quantitative agreement with the experiment. The results also confirm how non protic solvents can be confidently modeled by continuous solvation models as the Polarizable Continuum Model, whereas protic solvents, as water, require the inclusion of explicit solvent molecules to account for the effects of hydrogen bonds.

Introduction

The investigation of the interplay between a molecule and its embedding environment is one of the great challenges of today's quantum chemistry and molecular physics.[1-4] Indeed, a wide range of applications in both materials[5-8] and life sciences[9-11] is nowadays rooted in the capability of effectively combining a molecular probe with different substrates and exploiting the specific actions that each constituent plays on the others. One necessary condition to the successful design of such new composites is a deep understanding of the behaviour of each component in different environments, and of the interaction mechanisms that eventually lead to the desired effect.

The simplest example of embedding *medium* is certainly the solvent. Yet, the computational difficulties that arise when handling even simple solvents are manifold: if on the one hand the large dimensions of the whole system and the dynamical nature of the solvent severely limit the use of high level theories, on the other hand the chemical specificity of the local solute-solvent interactions and the subtle effects of the bulk on the solute properties call for reliable procedures. A compromise between computational feasibility and required accuracy can be obtained through the implementation of "focused" models, where the solute is described at a higher level of theory, whereas its embedding medium is treated with more affordable techniques.[1-4, 12-18]

To test the performances of such computational protocols, a valid experimental source of information on solute-solvent interaction certainly comes from spectroscopy.[1, 19-20] Among other observables, solvatochromic shift in different solvents can be used to characterize specific solute-solvent interactions and consequently discern among different embeddings.[14] Furthermore, due to the sensitivity of the solute electronic structure to the detail of the surrounding solvent, the capability of a computational model to adequately reproduce solvatochromic shifts constitutes a good benchmark for evaluating the reliability of the adopted description.

Effective solvent models can be classified into three wide classes, namely *continuum*, discrete and explicit approaches,[2-4, 15] in which the embedding *medium* is treated with an increasing degree of complexity. In the first class, the atomistic nature of the solvent is neglected and the embedding is treated as a macroscopic *continuum*. In particular, among other strategies[4], the polarizable *continuum* method (PCM), in which the solvent is modelled as a macroscopic dielectric *continuum*, is one of the most successfully employed, also thanks to the comprehensive work of the group of Tomasi.[21-24] Turning to the other two classes, the massive increase of computational resources that has characterized the past decades now allows us to take into account some of the microscopic chemical detail of the surrounding solvent.[1-3] In particular, the position of the most relevant solvent

molecules can be included in the calculation either at a quantum mechanical (QM) level (explicit models) or in a more approximate way (discrete models), *e.g.* through point charges or molecular mechanics (MM) representations. The crucial point of such approaches is to achieve a reliable sampling of the solvent relative positions. This is usually done resorting to simulation techniques as Monte Carlo (MC) or Molecular Dynamics (MD) and QM calculations are performed on a series of statistically sound configurations, extracted from previously performed simulation trajectories.

In this paper both a PCM based and an explicit approach will be tested to investigate solvatochromic effects. In both cases the description of the electronic transitions will be achieved through QM calculations performed at two different level of theory. The well-known time extension of density functional theory (TD-DFT) will be first employed to calibrate the method employed for the solvent description, whereas highly accurate multi-reference configuration interaction (MRCI) calculations, complemented with second order perturbation theory (CI-MRPT2), will be exploited to achieve a more realistic representation of the absorption spectra in different solvents. As far as the explicit model is concerned, the statistical sampling of the solvent position will be achieved through a computational protocol recently implemented to accurately reproduce electronic absorption and emission spectra of different chromophores embedded in simple[12, 16-17, 25] and complex[26-28] environments. The sampling will be thus based on MD simulations, performed with accurate force fields (FFs), specifically parameterized with the JOYCE[29-30] protocol on the target molecule.

As a test case we have selected the low-lying n□π* absorption band of pyrimidine, focusing in particular on the shift due to different solvents. The reason underlying this choice is manifold: besides the large amount of experimental[31-36] and computational[12, 37-45] data available for comparison, pyrimidine itself can be considered precursor of single nucleosides, and hence used as a model for larger bio-molecular systems.

Finally, we have recently[12] parameterized a FF for pyrimidine, which was employed in a sequential MD/DFT approach to investigate the solvatochromic shift in water. The main drawback of these previous results was an overestimation of the transition energy of the computed n□π* absorption peak of almost 0.3 eV. To overcome this lack, in the present paper the electronic calculations will be performed with the higher level CI-MRPT2 and a more systematic rationalization of the solvent effect will be tempted by extending the simulation to several kinds of solvents, namely carbon tetrachloride, acetonitrile, and water, as prototypes of apolar, polar aprotic and protic *media*, respectively.

Methods

FF and MD simulations

The total energy in a classical MD simulation can be expressed as a sum of an intra-molecular and an inter-molecular term

$$E^{tot} = E^{intra} + E^{inter} \quad (1)$$

The intra-molecular term is computed for each molecule (either pyrimidine or solvent) as a sum of three different contributions, accounting for stretching, bending and torsional energy, respectively:

$$E^{intra} = \sum_{i=1}^{N_{bonds}} k_i^b (r_i - r_{0_i})^2 + \sum_{i=1}^{N_{angles}} k_i^a (\theta_i - \theta_{0_i})^2 + \sum_{i=1}^{N_{Hdied}} k_i^{hd} (\phi_i - \phi_{0_i})^2 + \sum_{i=1}^{N_{died}} \sum_{j=1}^{M_i} k_{ij}^d (1 + \cos[n_{ij}^d \delta_i - \gamma_{ij}]) \quad (2)$$

The inter-molecular term, which accounts for both solute-solvent and solvent-solvent interactions, is expressed by a standard Lennard-Jones (LJ) term plus a Coulomb charge-charge contribution:

$$E^{inter} = \sum_{i=1}^{N_{atoms}} \sum_{j \neq i}^{N_{atoms}} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \quad (3)$$

A specific FF parameterization procedure was adopted for pyrimidine, following a protocol validated in Ref. [[12]] and illustrated in Figure 1. The bonded parameters entering equation (2) were previously parameterized[12] through the JOYCE procedure[29-30] on DFT Hessian matrix, purposely computed for pyrimidine *in vacuo*. For the intermolecular terms of equation (3) the LJ parameters have been taken from the Optimized Potentials for Liquid Simulations (OPLS) FF[46], whereas the point charges were separately parameterized *in vacuo* and for each investigated solvent through the following protocol.

In a recent computational investigation on the optical and magnetic properties of pyrimidine in water, Canuto and co-workers[38] pointed out the necessity of accurately accounting for solute polarization effects for a quantitative prediction of solvatochromic shifts. The authors compared the molecular dipole moment obtained with the original[46] OPLS charges with that resulting from their sequential MC/QM approach,[38, 47-49] concluding that, for a reliable description of the solvation embedding during the MC sampling, a charge distribution different from the original OPLS should be adopted. Interestingly, they also employed a PCM based procedure, which yielded a dipole value similar to that obtained in their iterative approach. Following this idea, the point charges here adopted were obtained through the Restrained ElectroStatic Potential (RESP) procedure,[50] using the

electrostatic potential arising from the charge density determined by means of DFT/PCM calculations at the equilibrium geometry, separately for each solvent (in the present work for acetonitrile and in Ref. [12] for water and CCl_4).

Finally, as far as the solvent parameters are concerned, they were transferred from TIP3P[51] (water) and AMBER[52] (Ref. [53] and [54] for CCl_4 and acetonitrile, respectively) parameterizations. Further details about FF development can be found in Ref. [12].

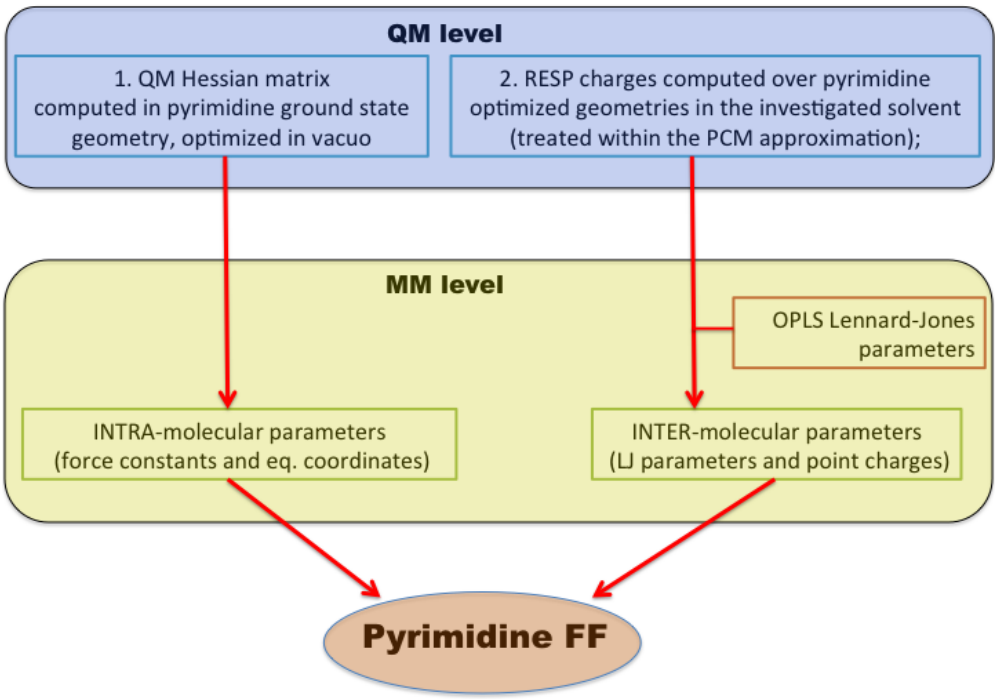

Figure 1. Flow chart of the force-field parameterization protocol for pyrimidine in different solvents.

All MD simulations were performed with the Gromacs4.5 engine.[55] The simulation of pyrimidine *in vacuo* was performed in the NVT ensemble at 298 K, whereas all simulations in solvent were carried out in the NPT ensemble on systems composed of one solute e and ~ 1000 solvent molecules, keeping temperature (298 K) and pressure (1 atm) constant through the Berendsen weak coupling scheme, using coupling constants of 0.1 and 1 ps, respectively. Due to the fast stretching vibrations, a time step of 0.5 fs was employed in all runs. A cut-off of 10 Å was employed for both charge-charge and van der Waals terms, whereas long-range electrostatics were accounted through the particle mesh Ewald procedure. All systems were equilibrated for 2 ns and thereafter production runs were performed for further 5 ns, collecting configurations every 2.5 ps. The solvent structure embedding pyrimidine was investigated for each system by computing both standard center of mass-center of mass (g(r); between pyrimidine and the surrounding solvent molecules) and atom-atom ($g_{kl}(r)$; between atom *k* of the pyrimidine and atom *l* of the solvent) pair correlation functions.

Further information on the structural modifications induced by different environments on the solute was obtained by monitoring the distribution of pyrimidine selected internal coordinates. Finally, three sets of 100, 500 and 1000 snapshots were extracted from each trajectory, and employed for the QM calculations of the electronic excited states.

QM calculations: TD-DFT and CI-MRPT2

QM calculations of the first excited states were performed both at TD-DFT level and through a CI-MRPT2 procedure over the 100 selected snapshots previously sorted by the trajectories of the MD simulations. The interaction of the solute with the surrounding *medium* was described for both methods in different ways:

i) a *continuum* approach, retaining from the snapshots only the solute coordinates and employing the PCM procedure during the QM calculations

ii) an explicit approach was set up for water (labelled $2H_2O$ (QM)) by including in the calculations the two solvent molecules closest to the pyrimidine Nitrogen atoms. The number of waters to be added to the solute in the calculations was determined on the basis of HB analysis performed over the MD trajectories, selecting the two water molecules hydrogen bonded to the N atoms of the solute. For computational convenience, this approach was tested only at TD-DFT level

iii) a hybrid approach ($2H_2O$ (QM)+PCM) in which the super-molecule of previous point ii) interacts with the bulk through the PCM.

The TD-DFT calculations were carried out with the Gaussian09 suite of programs[56] using the standard B3LYP functional coupled with the aug-N07D[57] basis set. The multi-reference CI-MRPT2 calculations for the first 16 eigenstates of pyrimidine follow a route that we have optimized in the past [12, 18, 58-61] and summarized as follows. The MOs basis set for the CI step is obtained by an HF-SCF calculation using GAMESS [62] with the 6-311*G basis set (126 molecular orbitals), followed by the transformation of mono and bi-electronic integrals from atomic to molecular basis. For the calculations in solutions, PCM was included in the HF calculations. For case iii), where two explicit molecules are included, the HF calculations were carried out on the super-molecule with PCM. Then we have applied a localization-fragmentation procedure [13, 63-66] in order to only retain the orbitals of the solute molecule, which are used in the post-HF calculations. This guarantees a well-balanced CI-MRPT2 calculation for the cases with and without the explicit solvent molecules. The 1s orbitals of Carbon and Nitrogen are always neglected in the post HF calculations.

All CI-MRPT2 calculations are performed using the BALOO code,[67] in five subsequent steps. In the first four the initial configurational space is enlarged by a given factor, adding at each step those configurations with the largest second order perturbative correction to the energy of the CI states. In

the present case, starting with an initial space of 236 Configuration Spin Functions (CSFs), the successive steps have dimension ~1100, ~4700 and ~20000 CSFs, respectively. As the well defined increment concerns with the configurations, which originated a different number of CSFs depending on the number of unpaired electrons, the dimension of the CI space (i.e. the number of CSFs) shows a small dependence on the geometry; the changes however are no more than 10%. The final variational space includes CSF up to the excitation level 4 and the subsequent perturbative correction includes CSFs up to excitation level 6. The last MRPT2 step exploits the diagrammatic perturbation theory,[68] recently implemented in the BALOO program,[67] which is much more efficient than traditional many-body perturbation theory. A typical CI-MRPT2 calculation takes a wall time of about 2.2 hours, using 16 CPU units (E5-2670 @ 2.60GHz) and including about 10^{10} CSFs in the final perturbative step. The final spectrum is determined by the perturbed energies and by the oscillator strengths computed with the first-order corrected density matrix.

Spectral lineshapes

Once energy and oscillator strength are determined (either at DFT or CI-MRPT2 level) for a given snapshot, the absorption spectrum is obtained by a convolution with a Gaussian function of FWHM=0.20 eV. A simple average of the spectra for all snapshots is then made to obtain the final absorption profile. It is worth pointing out that in this way the broadening of the spectral lines, related to both solvent effects and solute vibrational dynamics (driven by the specific FF), can be obtained and directly compared to the experimental data.

Results and discussion

FF parameterization

The parameterization of pyrimidine's intra-molecular FF, as well as of the inter-molecular part for water and CCl_4 solvents was performed in Ref. [12] and all the details can be found therein. Conversely, as far as pyrimidine in CH_3CN solvent is concerned, the point charges entering equation (3) were obtained in this work. For the sake of completeness, all the FF parameters obtained for pyrimidine *in vacuo* and in the considered solvents are reported in the Supporting Information (see Tables ST1-ST4).

In Table 1, the molecular dipolar moments computed from the FF point charge sets obtained for each solvent are reported together with those resulting from the reference QM calculations. As already noted in Ref. [38], the dipole moment for H_2O (3.47 D) is remarkably larger that the original[46] OPLS one (2.52 D), and in better agreement with those computed by Canuto and co-workers (3.5 and 4.0 D). Furthermore, by looking at all the dipole values exhibited in the sequence

of the considered solvents, it appears as the PCM derived charges are able to account for the increase of pyridine's polarization when the solvent polarity is augmented.

Method and molecular geometry	*in vacuo*	CCl_4	CH_3CN	H_2O
DFT on optimized structure in PCM	2.38	2.84	3.24	3.47
RESP charges on optimized MM structure	2.34	2.81	3.21	3.41
RESP charges from MD trajectories	2.32 ± 0.02	2.83 ± 0.14	3.23 ± 0.14	3.49 ± 0.15

Table 1. Dipole moments (Debye) obtained for pyrimidine in different solvent through DFT calculations and RESP charges. The latter have been used both on the optimized structure and on MD trajectories.

MD simulations

The structural features of the solvent embedding were investigated by computing several pair correlation functions over the MD trajectories. It should be mentioned that the MD runs performed with carbon tetrachloride and water solvents were carried out employing the same FFs and in the same thermodynamic conditions as in Ref. [12], but extending the simulations to a longer time-scale.

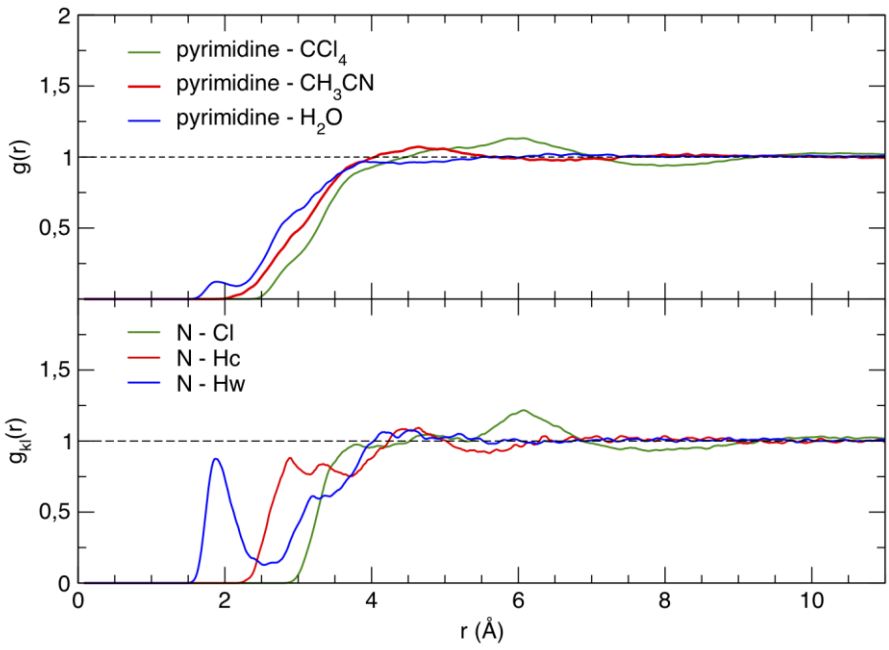

Figure 2. Top: Pair correlation function g(r) between the center of mass of pyrimidine and the different embedding solvents. Bottom: Atomic pair correlation $g_{kl}(r)$ functions between the atom N of pyrimidine and the closest solvent atom. In both panels green, red and blue lines were employed for carbon tetrachloride, acetonitrile and water, respectively.

From inspection of the top panel of Figure 2, where the g(r) functions computed between pyrimidine and solvent centres of mass are displayed, no major differences appear, being the maximum peak position centred around 4 Å for each solvent. On the contrary, the atomic pair

correlation functions are much more effective in distinguishing the three solvents, as the first peak shifts toward larger distances when the embedding medium polarity is decreased. In particular when water is considered, a well-defined N-Hw peak appears around 2 Å: a clear indication that some hydrogen bond (HB) is taking place. This feature is confirmed by an HB analysis over the water solvated pyrimidine trajectories, resulting in an average number of HBs of 2.5, with a residence time of slightly more than 3 ps.

The stronger interaction between the protic water solvent with pyrimidine should reflect in first place in a distortion of pyrimidine's geometry, increased with respect to the one achieved in acetonitrile and CCl_4. To verify this feature a planarity distortion index χ was computed as the average dihedral formed by all heavy atom quadruples forming the ring backbone, and its distribution in the investigated environments is shown in Figure 3. It is apparent that the solvent has the effect of inducing larger geometry distortions with respect to *vacuum*, whereas its specificity is almost irrelevant. The broad curves are consistent with the higher fluctuations of the dipole moment along the MD trajectories in solutions with respect to vacuum (0.14 D *vs.* 0.02 D, see Table 1). This data are also consistent with the general feature that the absorption spectra are broader in solution than *in vacuo*.

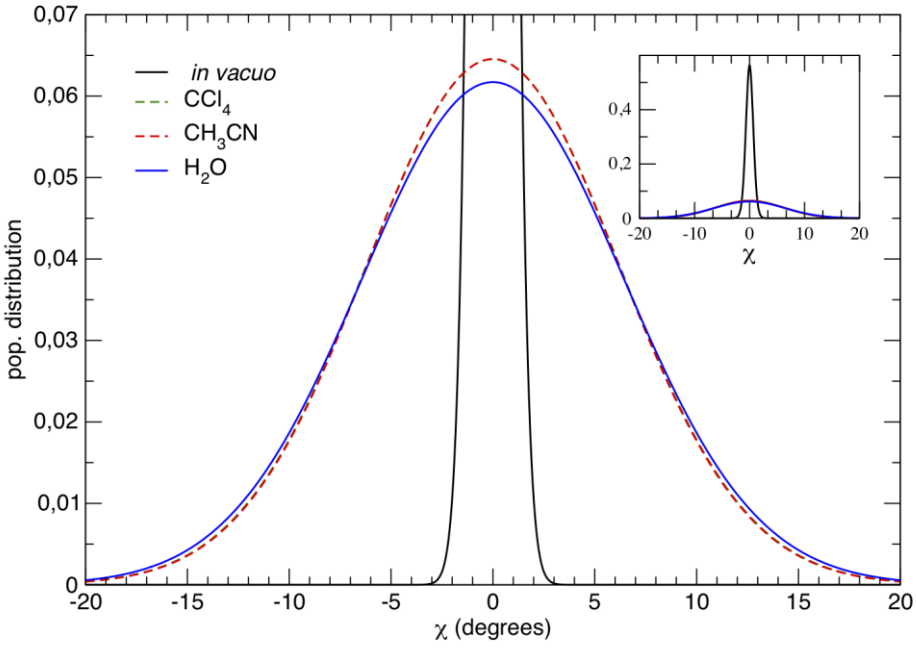

Figure 3. Distribution of pyrimidine planarity distortion index χ in different solvents.

TD-DFT calculations

The TD-DFT calculations of the first 10 electronic excited states were performed over set of 100 snapshots extracted from the different MD trajectories. This number was selected after preliminary

calibrations, computing the pyrimidine absorption spectrum over sets of 100, 500 and 1000 (see Figure S2 in the Supporting Information).

Firstly, a *continuum* model was chosen to account for each solvent, retaining from each snapshot only the pyrimidine geometry. Hence, the subsequent QM calculations were performed representing the solvent entirely by PCM. The final spectra, obtained as described in the Methods section, are reported in Figure 4. To compute solvatochromic shifts, the same procedure was employed over 100 snapshots extracted from the trajectory *in vacuo*, and the resulting absorption band is also reported in Figure 4.

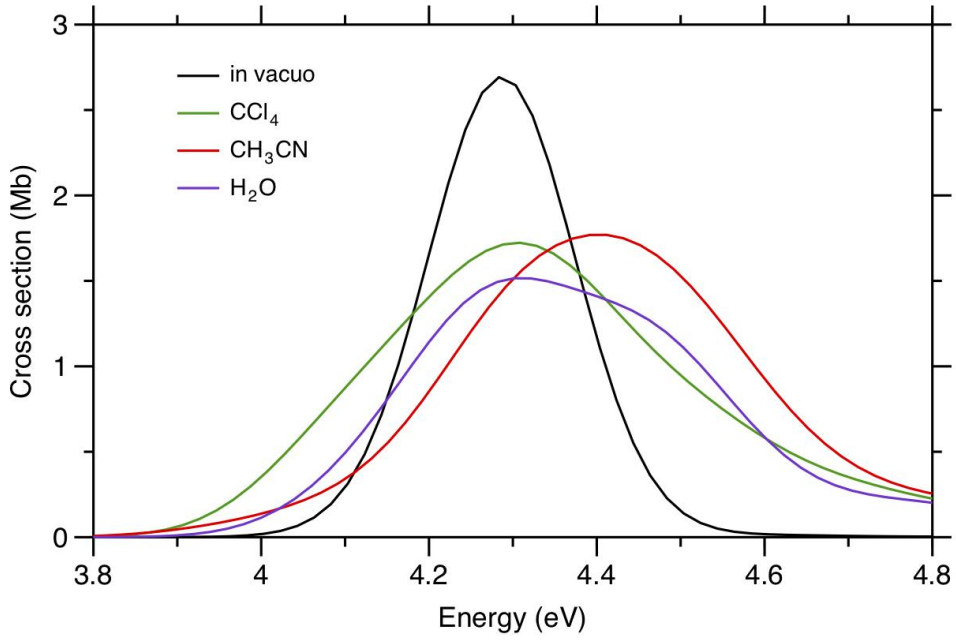

Figure 4. TD-DFT pyrimidine absorption spectra for the n□π* transition *in vacuo* and in different solvents, accounted for at PCM level.

The theoretical peak maximum *in vacuo* at 4.28 eV is in rather good agreement with the experimental[31] value (4.183 eV) suggesting that at least for this system, the B3LYP functional here employed is more effective than the previously used CAM-B3LYP, which gave 4.56 eV.[12] The band is much narrower than those in solution, due to the much minor geometry distortions in gas phase (see Figure 3). The effect of the apolar solvent CCl_4 shifts the peak maximum only slightly, whereas the blue shift is more evident when the polar CH_3CN is considered. The maximum of the band in CH3CN is ~ 4.4 eV is in reasonable agreement with experiment which detects 4.32 eV[35]. Surprisingly, in water a further broadening is observed with no relevant shift with respect acetonitrile. The position of the absorption band in water is only in qualitative agreement with the experimental data,[31, 35] where a rather large (~ 0.39 eV) blue shift was detected with respect to

vacuum. This unpleasant finding accords with previous[12, 38-42, 45] computational studies performed on pyrimidine, where the best results were obtained by the explicit inclusion in the calculation of a number of water molecules.

To get further insight about the behaviour of the solute in water, which appears as the most interacting solvent, we have passed to step ii) including explicitly two water molecules in the TD-DFT calculation. Figure 5 displays all the 100 selected snapshots sorted by MD trajectories, superimposed in order to appreciate the most probable distribution of water around the solute. Driven by the solute symmetry, the sampling accounts for a rather symmetric distribution of the water molecules around the N proton acceptor. However it is worthwhile noticing that, despite the strong HB between the N atoms and water, the thermal motions originate a rather spread distribution, including cases where the N-H distance is much larger than the typical HB distance in N···H-O.

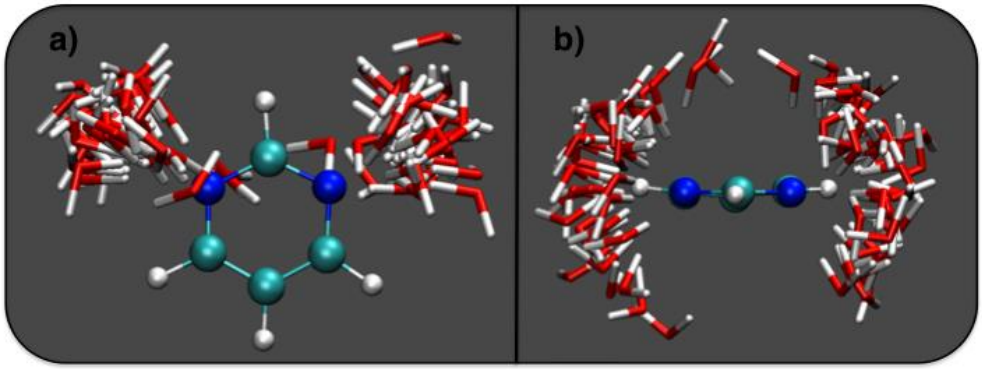

Figure 5. Configurations used for the calculation of the absorption spectra of pyrimidine in water accounting for the solvent through the explicit models ($2H_2O$ (QM) and $2H_2O$ (QM)+PCM). a) Front view b) Side view

The absorption spectra in water, computed with the different solvent models are shown in Figure 6, together with the experimental band.[35] As already noticed, the PCM model severely underestimates the blues shift of the band position. Similarly, the inclusion of the two neighbouring waters, without accounting for the bulk polarizations ($2H_2O$(QM) model), underestimates the maximum peak position of about 0.3 eV. In conclusion it seems that the best results is obtained by the $2H_2O$(QM)+PCM) model, even if the obtained shift of ~ 0.2 eV is still smaller than the value of 0.4 eV, observed in experiment.[35] Moreover the absolute maximum of the band is about 0.1 eV lower than the experimental value of 4.57 eV.

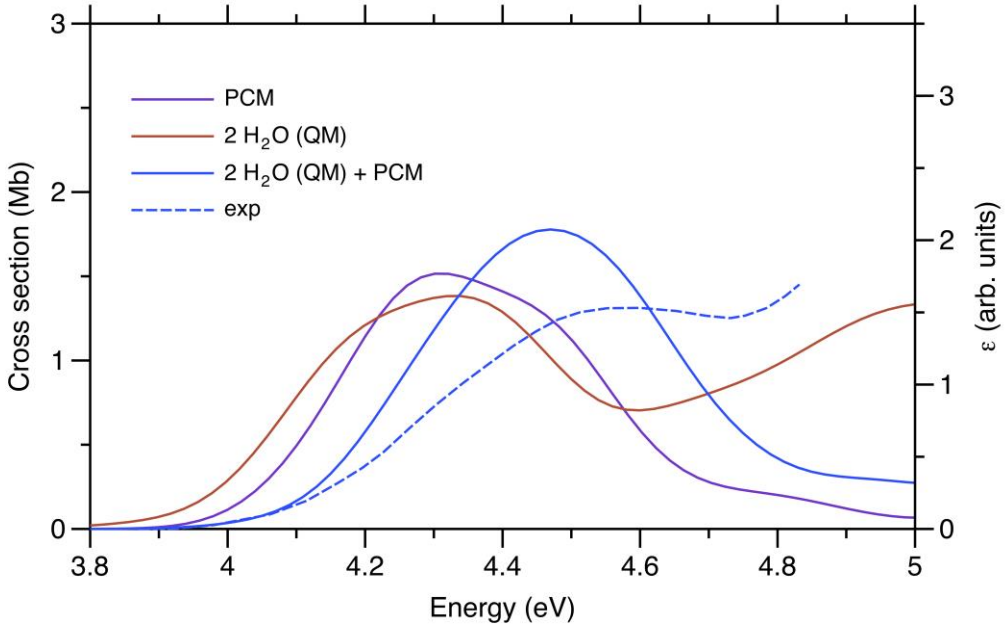

Figure 6. Computed pyrimidine absorption spectra for the n□π* transition in water by TD-DFT, accounting for the solvent at different levels of theory. The experimental molar absorption coefficient in arbitrary units is also reported for comparison (right scale).

CI-MRPT2 calculations

Based on TD-DFT results and in consideration of the computational cost of the post-HF calculations, we decided to account for solvent effect through the PCM for the aprotic solvents, and to resort to the hybrid model for the accurate calculation of the spectra in water.

Despite our CI calculations have been made for the first 16 eigenstates, experimental data in solution are limited to the low energy part of the UV spectrum, that in the range 3.8-4.8 eV (~31000-39000 cm^{-1}). In this region the only allowed, little intense, transition is the n_N□π*, which involves the two orbitals of Figure 7. The n_N orbital, of σ symmetry, is the HOMO in DFT/TDDFT calculations and the HOMO-1 in HF-SCF, where the HOMO is a πorbital.

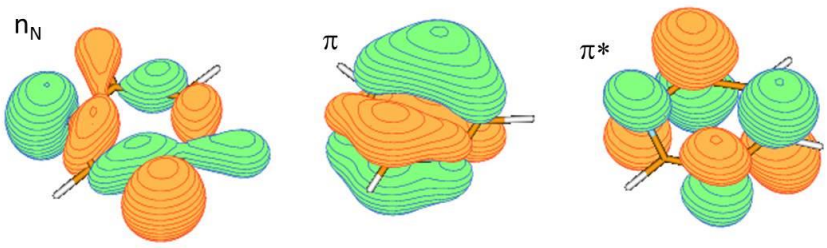

Figure 7. Orbitals involved in the lowest UV transitions of pirymidine.

The absorption spectra in *vacuo* (black), in CCl_4 (green), acetonitrile (red) and water treated with PCM (violet) and also considering in addition two explicit water molecules (blue), computed using

our CI-MRPT2, are reported in Figure 8. The comparison between experimental and computational results is instead shown in Figure 9 for the cases *in vacuo*, in acetonitrile and in water.

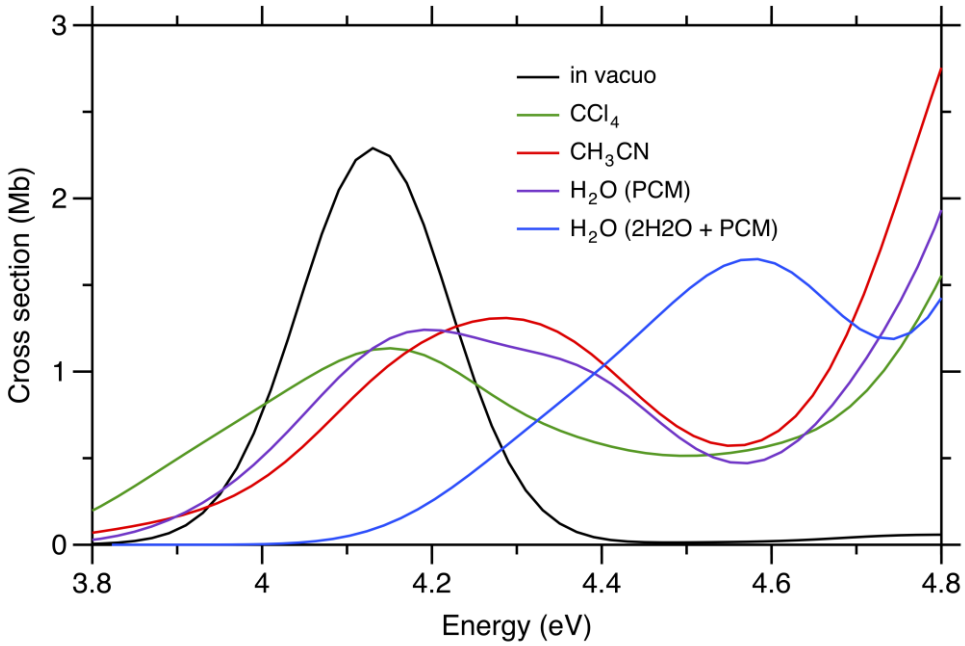

Figure 8. Computed absorption spectra of pyrimidine in vacuum (black), in CCl_4, CH_3CN and water by PCM (green, red and violet, respectively), in water considering two explicit water molecules and PCM (blue).

In Figure 8 it is confirmed how non-polar solvents like CCl_4, whose effect is considered by PCM, do not cause any significant displacement in the position of the band with respect to that observed in vacuum, while polar non-protic solvents like acetonitrile are found to give rise to a blue shift. In all cases a broadening similar to the previous TD-DFT results is observed on going from vacuum to solvent. In the case of water, PCM predicts a small blue shift of the band (less than 0.2 eV), in agreement with TDDFT computations [12]. Such shift is, however, too small in comparison with that observed experimentally of about 0.4 eV. On the contrary the CI-MRPT2 calculations averaged on the 100 snapshots taken from MD, accounting for the solvent through the hybrid $2H_2O$+PCM approach, yields a larger shift (Figure 8) and well reproduce both position and width of the observed band, as shown in Figure 9. The results with PCM alone appear to exhibit a sort of two-component band, not present when including two explicit water molecules. Its origin, according to the tests reported in the Supporting Information (see Figure S2), seems to have no correlation with the variety of geometries of pyrimidine in the two apparent components of the band, but disappears as the statistical sample is enlarged.

Turning to the spectra *in vacuo*, the computed spectrum seems slightly shifted with respect to its experimental counterpart, and shows a clear difference in its width. The first difference is mostly

due to the rather large vibrational distribution and small intensity of the gas phase experimental spectrum of Ref. [31], which makes hard the visual comparison with vertical transitions computed in the present paper. In fact, notwithstanding the experimental band displayed in Figure 9 has been smoothed by a convolution with a Gaussian function of FWHM 0.2 eV, the band maximum seems not coincident with the value of 4.183 eV reported in Ref. [31] (see Figure 2 and Table 1 therein) for the energy of the band maximum. However, in the following discussion, the latter value will be used for the comparison between theoretical predictions and experimental data, as well as for computing the experimental solvent shift of the absorption spectra measured in solvent in Ref. [35] On the contrary, the different broadening exhibited by computed and experimental gas phase spectra can probably be ascribed to the classical treatment of vibrations, which becomes much more critical *in vacuo* than in solution, where instead classical dynamics seems adequate to model vibrational effects (see *e.g.* the CCl_4 case of Figure 8).

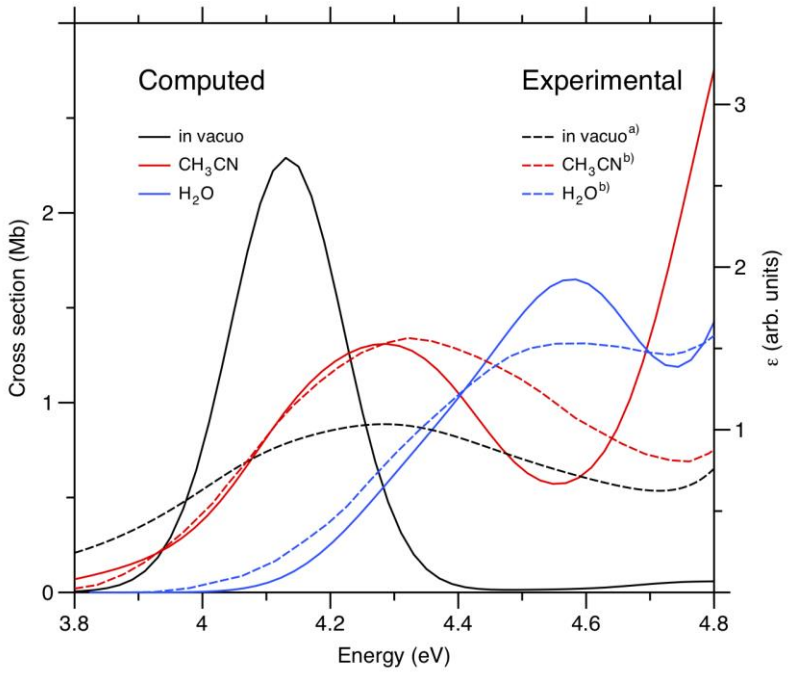

Figure 9. Comparison between computed (solid line, left scale) and experimental (dashed line, right scale) spectra of pyrimidine in vacuum (black), in acetonitrile (red) and in water (blue).

In Tables 2 and 3 the computed absolute peak positions and the solvatochromic shifts in the examined solvents are compared with experimental and literature results. It should be mentioned that in most computational works the solvatochromic shifts and the relative peak position were computed from vertical energy differences and not from the actual maximum of the absorption band.

Ref.	Method	N_w	E_{max} (eV)	ΔE_{max} (eV/cm^{-1})
[45]	ZINDO CIS + SCRF	2	4.68	0.48 (3900)
[44]	MC/QM	1000	-	0.42 (3400)
[43]	CIS QM/MM	510	4.18	0.282 (2275)
[42]	CASPT2/ANO + SCRF	2	4.38	0.19 (990)
[41]	MC/TD-DFT	∞	-	0.275 (2220)
[40]	TD-DFT + PCM	2	4.61	0.32 (2580)
[39]	MD/TD-DFT	4	4.48	0.20 (2580)
[38]	MC/TD-DFT	28	4.58	0.33 (2640)
[12]	MD/TD-DFT + charges	~ 100	4.84[a)]	0.32 (2600) [a)]
this work	MD/CI-MRPT2	2	4.58	0.45
[35] (exp)	UV	-	4.58	0.39 (3160) [b)]

Table 2. Experimental and computed peak maxima and solvent shifts for pyrimidine in water. a) Computed from effective band shapes b) The experimental shift is computed with respect to the value (4.183 eV) reported in gas phase by Ref. [31].

It appears from Table 2 that the present approach is able not only of a very accurate estimate of the blue shift due to hydration, but also quantitatively agrees with the experiment in the absolute position of the band maximum. This accuracy is not lost for the other investigated solvents even when the solvent is not treated explicitly. Indeed, as shown in Table 3, the excellent agreement holds for all the investigated cases, confirming that the adopted protocol is able to account for the different polarization effects at the origin of the solvatochromic series.

	in vacuo		**CCl_4**		**CH_3CN**		**H_2O**	
	E_{max} (eV)	ΔE_{max} (eV/cm^{-1})	E_{max} (eV)	ΔE_{max} (eV/cm^{-1})	E_{max} (eV)	ΔE_{max} (eV/cm^{-1})	E_{max} (eV)	ΔE_{max} (eV/cm^{-1})
this work	4.13	-	4.15	0.02	4.28	0.15	4.58	0.45
Exp. [35]	4.18	-	4.24 [a)]	0.06 [a)]	4.32	0.14	4.58	0.39

Table 3. Experimental and computed peak maxima and solvent shifts for pyrimidine in different solvents. All experimental shifts are computed with respect to the value (4.183 eV) reported in gas phase by Ref. [31] a) The value was measured in iso-octane.

At energies above 4.8 eV, for which, up to our knowledge, no experimental data in water has been reported, we find a more intense $\pi\square\pi^*$ transition (see also Ref. [12]). This band is not expected to show a strong solvent dependence, since involves orbitals not interested in the formation of hydrogen bond with water. This can be seen in Figure 10, where the CI-MRPT2 result *in vacuo* and for the case with explicit water molecules fit very well the experimental spectrum in gas phase of

Ref. [31]. On the other hand PCM alone, i.e. without the explicit solvent molecules seems to fail, as it predicts band positions at lower energies.

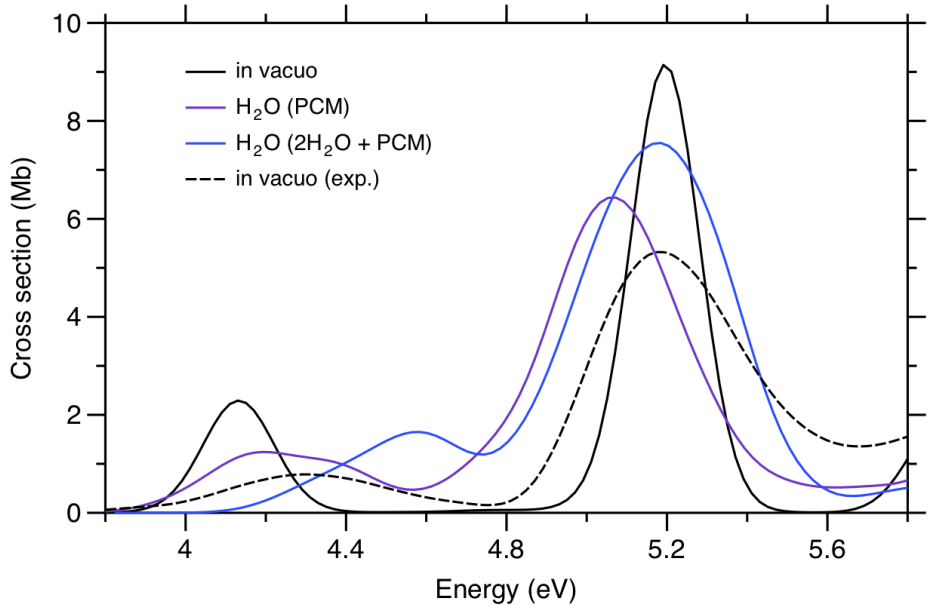

Figure 10. Absorption spectra in the [3.8-5.8] eV range. The experimental absorption *in vacuo* (dashed line) was taken form Ref. [31].

Conclusions

Solvatochromic effects of different solvation environments on the lowest UV bands of pyrimidine were investigated by a sequential MD/QM approach. MD trajectories were exploited both to incorporate vibrational effects at a classical level and to reliably sample the solvent positions, when the local character of the solvent-solute interaction was thought to be relevant. In all cases, one hundred snapshots were selected from the MD trajectories obtained for pyrimidine *in vacuo* or in the investigated solvents. Subsequently, TD-DFT and CI-MRPT2 calculations were performed on each frame to obtain position and oscillator strength of each band. A final average gives rise to the final spectrum that can be directly compared to experimental band shapes.

One of our aims was to investigate by accurate quantum mechanical methods, how a PCM treatment of solvents with different polarity and forming hydrogen bond, is adequate for studies of solvatochromic effects. It turned out that, *in vacuo* and in non-protic solvents the PCM description reliably accounts for the observed solvent effects, for both $n_N \rightarrow \pi^*$ and lowest $\pi \rightarrow \pi^*$ bands. On the contrary, in presence of specific local interactions as hydrogen bonds, it is necessary to consider explicit solvent molecules, at least one molecule for each involved heteroatom. With two water molecules for the present case, we have then been able to reproduce the solvatochromic shift by both TDDFT and CI-MRPT2 methods. The latter, although computationally more expensive, yields

a quantitative agreement not only in the considered shifts but also in the absolute position of the absorption bands. Considering the *ab initio* nature of the CI-MRPT2 methods and their intrinsic predictive abilities, these results strongly suggest that the presented MD/CI-MRPT2 protocol could be a very useful tool in the design of novel chromophore based materials.

Acknowledgements

The authors are grateful to Prof. Vincenzo Barone of the Scuola Normale Superiore for helpful discussions and suggestions.

References

1. Barone, V.; Ed., ***Computational Strategies for Spectroscopy: From Small Molecules to Nano Systems*****. Wiley: Chichester, U. K., 2011.**
2. Canuto, S.; Ed., ***Solvation Effects on Molecules and Biomolecules. Computational Methods and Applications*****. Springer: 2008.**
3. ***Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods, Advances in Quantum Chemistry*****. 2010; Vol. 59.**
4. Mennucci, B.; Cammi R.; Eds, ***Continuum Solvation Models in Chemical Physics*****. Wiley: 2007.**
5. Ciardelli, F.; Ruggeri, G.; Pucci, A., Dye-containing polymers: methods for preparation of mechanochromic materials. ***Chem. Soc. Rev.*** **2013,** ***42*** **(3), 857--70.**
6. Anzenbacher, P.; Li, F.; Palacios, M. A., Toward Wearable Sensors: Fluorescent Attoreactor Mats as Optically Encoded Cross-Reactive Sensor Arrays. ***Angew. Chem. Int. Ed*** **2012,** ***51*****, 2345.**
7. Sagara, Y.; Kato, T., Brightly Tricolored Mechanochromic Luminescence from a Single-Luminophore Liquid Crystal: Reversible Writing and Erasing of Images. ***Angew. Chem. Int. Ed.*** **2011,** ***50*** **(39), 9128--9132.**
8. Bredas, J.-L.; Marder, S. R.; Reichmanis, E., Preface to the Chemistry of Materials Special Issue on \oe\ A -Functional Materials. ***Chem. Mater.*** **2011,** ***23*** **(3), 309.**
9. Campoccia, D.; Montanaro, L.; Arciola, C. R., A review of the biomaterials technologies for infection-resistant surfaces. ***Biomaterials*** **2013,** ***34*** **(34), 8533--54.**
10. Haugland, R. P., ***The Handbook. A Guide to Fluorescent Probes and Labeling Technologies*****. Molecular Probes, Inc.: Eugene, OR, 2005.**
11. Goldys, E., ***Fluorescence Applications in Biotechnology and Life Sciences*****. Wiley-Blackwell: New York, 2005.**
12. Biczysko, M.; Bloino, J.; Brancato, G.; Cacelli, I.; Cappelli, C.; Ferretti, A.; Lami, A.; Monti, S.; Pedone, A.; Prampolini, G.; Puzzarini, C.; Santoro, F.; Trani, F.; Villani, G., Integrated computational approaches for spectroscopic studies of molecular systems in the gas phase and in solution: pyrimidine as a test case. ***Theor. Chem. Acc.*** **2012,** ***131*** **(4).**
13. Barone, V.; Cacelli, I.; Ferretti, A.; Monti, S.; Prampolini, G., An Integrated Protocol for the Accurate Calculation of Magnetic Interactions in Organic Magnets. ***J. Chem. Theory Comput.*** **2011,** ***7*** **(3), 699--706.**
14. Marini, A.; Mu\~oz-Losa, A.; Biancardi, A.; Mennucci, B., What is solvatochromism? ***J. Phys. Chem. B*** **2010,** ***114*** **(51), 17128--35.**

15. Tomasi, J., Thirty years of continuum solvation chemistry: a review, and prospects for the near future. *Theor. Chem. Acc.* 2004, *112* (4), 184-203.
16. De Mitri, N.; Monti, S.; Prampolini, G.; Barone, V., Absorption and Emission Spectra of a Flexible Dye in Solution: A Computational Time-Dependent Approach. *J. Chem. Theory Comput.* 2013, *9* (10), 4507--4516.
17. Barone, V.; Bloino, J.; Monti, S.; Pedone, A.; Prampolini, G., Fluorescence spectra of organic dyes in solution: a time dependent multilevel approach. *Phys. Chem. Chem. Phys.* 2011, *13* (6), 2160--6.
18. Cacelli, I.; Ferretti, A., Solvent effect on the optical properties of (NH3)(5)Ru-pyrazine (+m) (m=2,3) complexes by ab initio calculations. *J. Chem. Phys.* 1998, *109* (19), 8583-8590.
19. Grunenberg, J.; Ed., *Computational spectroscopy*. Wiley- VCH New Jersey 2010.
20. Laane, J., *Frontiers of molecular spectroscopy*. Elevier: Amsterdam, 2009.
21. Tomasi, J.; Mennucci, B.; Cammi, R., Quantum mechanical continuum solvation models. *Chem. Rev.* 2005, *105* (8), 2999--3093.
22. Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J., Ab initio study of ionic solutions by a polarizable continuum dielectric model. *Chem. Phys. Lett.* 1998, *286* (3-4), 253--260.
23. Tomasi, J.; Persico, M., Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chem. Rev.* 1994, *94* (7), 2027--2094.
24. Miertus, S.; Scrocco, E.; Tomasi, J., Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* 1981, *55* (1), 117--129.
25. Barone, V.; Bloino, J.; Monti, S.; Pedone, A.; Prampolini, G., Theoretical multilevel approach for studying the photophysical properties of organic dyes in solution. *Phys. Chem. Chem. Phys.* 2010, *12* (35), 10550--61.
26. Prampolini, G.; Monti, S.; De Mitri, N.; Barone, V., Evidences of Long Lived Cages in Functionalized Polymers: Effects on Chromophore Dynamic and Spectroscopic Properties. *Chem. Phys. Lett.* 2014, *601*, 134.
27. De Mitri, N.; Prampolini, G.; Monti, S.; Barone, V., Structural, dynamic and photophysical properties of a fluorescent dye incorporated in an amorphous hydrophobic polymer bundle. *Phys. Chem. Chem. Phys.* 2014, *16* (31), 16573--87.
28. Pedone, A.; Prampolini, G.; Monti, S.; Barone, V., Realistic Modeling of Fluorescent Dye-Doped Silica Nanoparticles: A Step Toward the Understanding of their Enhanced Photophysical Properties. *Chem. Mater.* 2011, *23* (22), 5016--5023.
29. Barone, V.; Cacelli, I.; De Mitri, N.; Licari, D.; Monti, S.; Prampolini, G., Joyce and Ulysses: integrated and user-friendly tools for the parameterization of intramolecular force fields from quantum mechanical data. *Phys. Chem. Chem. Phys.* 2013, *15* (11), 3736-51.
30. Cacelli, I.; Prampolini, G., Parametrization and Validation of Intramolecular Force Fields Derived from DFT Calculations. *J. Chem. Theory Comput.* 2007, *3* (5), 1803-1817.
31. da Silva, F. F.; Almeida, D.; Martins, G.; Milosavljevic, A. R.; Marinkovic, B. P.; Hoffmann, S. V.; Mason, N. J.; Nunes, Y.; Garcia, G.; Limao-Vieira, P., The electronic states of pyrimidine studied by VUV photoabsorption and electron energy-loss spectroscopy. *Phys. Chem. Chem. Phys.* 2010, *12* (25), 6717-6731.
32. Riese, M.; Grotemeyer, J., Mass-analyzed threshold ionization spectroscopy of pyrimidine: determining the geometry in the first excited and the ionic ground states. *Anal. Bioanal. Chem.* 2006, *386* (1), 59--68.
33. Fischer, G.; Cai, Z.-L.; Reimers, J. R.; Wormell, P., Singlet and Triplet Valence Excited States of Pyrimidine. *J. Phys. Chem. A* 2003, *107* (17), 3093--3106.
34. Palmer, M. H.; Walker, I. C.; Guest, M. F.; Hopkirk, A., The electronic states of the azines. III. Pyrimidine, studied by VUV absorption, near-threshold electron energy-loss spectroscopy and ab initio multi-reference configuration calculations. *Chem. Phys.* 1990, *147* (1), 19--33.

35. Baba, H.; Goodman, L.; Valenti, P. C., Solvent Effects on the Fluorescence Spectra of Diazines. Dipole Moments in the (n,π*) Excited States. ***J. Amer. Chem. Soc.*** **1966,** ***88*** **(23), 5410-5415.**
36. Cavalieri, L. F.; Bendich, A., The Ultraviolet Absorption Spectra of Pyrimidines and Purines 1. ***J. Amer. Chem. Soc.*** **1950,** ***72*** **(6), 2587--2594.**
37. Caricato, M.; Lipparini, F.; Scalmani, G.; Cappelli, C.; Barone, V., Vertical Electronic Excitations in Solution with the EOM-CCSD Method Combined with a Polarizable Explicit/Implicit Solvent Model. ***J. Chem. Theory Comput.*** **2013,** ***9*** **(7), 3035--3042.**
38. Manzoni, V. c.; Lyra, M. L.; Gester, R. M.; Coutinho, K.; Canuto, S., Study of the optical and magnetic properties of pyrimidine in water combining PCM and QM/MM methodologies. ***Phys. Chem. Chem. Phys.*** **2010,** ***12*** **(42), 14023--33.**
39. Kongsted, J.; Mennucci, B., How to model solvent effects on molecular properties using quantum chemistry? Insights from polarizable discrete or continuum solvation models. ***J. Phys. Chem. A*** **2007,** ***111*** **(39), 9890--900.**
40. Mennucci, B., Hydrogen Bond versus Polar Effects: An Ab Initio Analysis on n → π* Absorption Spectra and N Nuclear Shieldings of Diazines in Solution. ***J. Amer. Chem. Soc.*** **2002,** ***124*** **(7), 1506--1515.**
41. de Almeida, K. t. J.; Coutinho, K.; de Almeida, W. B.; Rocha, W. R.; Canuto, S., A Monte Carlo–quantum mechanical study of the solvatochromism of pyrimidine in water and in carbon tetrachloride. ***Phys. Chem. Chem. Phys.*** **2001,** ***3*** **(9), 1583--1587.**
42. Serrano-Andrés, L.; Fulscher, M. P.; Karlstrom, G., Solvent effects on electronic spectra studied by multiconfigurational perturbation theory. ***Int. J. Quantum Chem.*** **1997,** ***65*** **(2), 167--181.**
43. Gao, J.; Byun, K., Solvent effects on the n->pi* transition of pyrimidine in aqueous solution. ***Theor. Chem. Acc.*** **1997,** ***96*** **(3), 151--156.**
44. Zeng, J.; Hush, N. S.; Reimers, J. R., Solvent effects on molecular spectra. III. Absorption to and emission from the lowest singlet (n,π∗) state of dilute pyrimidine in water. ***J. Chem. Phys.*** **1993,** ***99*** **(3), 1508.**
45. Karelson, M. M.; Zerner, M. C., Theoretical treatment of solvent effects on electronic spectroscopy. ***J. Phys. Chem.*** **1992,** ***96*** **(17), 6949--6957.**
46. Jorgensen, W.; McDonald, N., Development of an all-atom force field for heterocycles. Properties of liquid pyridine and diazenes. ***J. Mol. Struct.*** **1998,** ***424,*** **145.**
47. Coutinho, K.; Canuto, S., Solvent Effects form a Sequential Monte Carlo Approach - Quantum Mechanical Approach. ***Adv. Quantum Chem.*** **1997,** ***28,*** **89.**
48. Barreto, R. C.; Coutinho, K.; Georg, H. C.; Canuto, S., Combined Monte Carlo and quantum mechanics study of the solvatochromism of phenol in water. The origin of the blue shift of the lowest p–p* transition. ***Phys. Chem. Chem. Phys.*** **2009,** ***11,*** **1388.**
49. Coutinho, K.; Georg, H. C.; Fonseca, T. L.; Ludwig, V.; Canuto, S., An Efficient Statistically Converged Average Configuration for Solvent Effects. ***Chem. Phys. Lett.*** **2007,** ***437,*** **148.**
50. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A., A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. ***J. Phys. Chem.*** **1993,** ***97*** **(40), 10269-10280.**
51. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. ***J. Chem. Phys.*** **1983,** ***79*** **(2), 926-935.**
52. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, K. M.; Merz, I. R.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. ***J. Am. Chem. Soc.*** **1995,** ***117,*** **5179–5197.**

53. Fox, T.; Kollmann, P. A., Application of the RESP Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B* 1998, *102* 8070–8079.
54. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* 2005, *25*, 1157-1174.
55. van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P.; Tieleman, D.; Sijbers, A.; Feenstra, K.; van Drunen, R.; Berendsen, H. *GROMACS4.5*, 2010.
56. Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. , Gaussian 09, Revision D.01. *Gaussian, Inc., Wallingford CT* 2009.
57. Barone, V.; Cimino, P.; Stendardo, E., Development and Validation of the B3LYP/N07D Computational Model for Structural Parameter and Magnetic Tensors of Large Free Radicals. *J. Chem. Theory Comput.*, 2008, *4*, 751-764
58. Cacelli, I.; Campanile, S.; Denti, G.; Ferretti, A.; Sommovigo, M., (NH3)(5)Ru(1,2,4,5-tetrazine) (2+): Synthesis theoretical study of its solvatochromism and experimental and in the visible spectral region. *Inorg. Chem.* 2004, *43* (4), 1379-1387.
59. Cacelli, I.; Ferretti, A., Theoretical study of the absorption spectrum of (NH3)(5)Ru-(4,4 '-bipyridine) (2+) in solution. *J. Phys. Chem. A* 1999, *103* (23), 4438-4445.
60. Cacelli, I.; Ferretti, A.; Toniolo, A., Absorption and electroabsorption spectra of (NH3)(5)Ru(4,4 '-bipyridine)Ru(NH3)(5) (4+) in water by ab initio calculations. *J. Phys. Chem. A* 2001, *105* (18), 4480-4487.
61. Barone, V.; Ferretti, A.; Pino, I., Absorption spectra of natural pigments as sensitizers in solar cells by TD-DFT and MRPT2: protonated cyanidin. *Phys. Chem. Chem. Phys.* 2012, *14* (46), 16130-16137.
62. Schmidt, M. W.; Baldridge, K. K.; Boats, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., GAMESS. *J. Comput. Chem.* 1993, *14*, 1347.
63. Barone, V.; Cacelli, I.; Ferretti, A., Magnetic coupling in bis-nitronylnitroxide radicals: The role of aromatic bridges. *J. Chem. Phys.* 2009, *130* (9).
64. Barone, V.; Cacelli, I.; Ferretti, A.; Monti, S.; Prampolini, G., Singlet-triplet energy gap of a diarylnitroxide diradical by an accurate many-body perturbative approach. *Phys. Chem. Chem. Phys.* 2011, *13* (10), 4709-4714.
65. Barone, V.; Boilleau, C.; Cacelli, I.; Ferretti, A.; Monti, S.; Prampolini, G., Structure-Properties Relationships in Triplet Ground State Organic Diradicals: A Computational Study. *J. Chem. Theory Comput.* 2013, *9* (1), 300-307.
66. Barone, V.; Boilleau, C.; Cacelli, I.; Ferretti, A.; Prampolini, G., Conformational Effects on the Magnetic Properties of an Organic Diradical: A Computational Study. *J. Chem. Theory Comput.* 2013, *9* (4), 1958--1963.
67. Barone, V.; Cacelli, I.; Ferretti, A.; Prampolini, G., to be submitted. 2014.
68. Cimiraglia, R., Second order perturbation correction to CI energies by use of diagrammatic techniques: an improvement to the CIPSI algorithm. *J. Chem. Phys.* 1984, *83*, 174.