ORIGINAL ARTICLE

Journal Section

Simulating Vertical Excitation Energies of Solvated Dyes: from Continuum to Polarizable Discrete Modeling

Tommaso Giovannini [1] | Marina Macchiagodena [1] | Matteo Ambrosetti [1] | Alessandra Puglisi [1] | Piero Lafiosca [1] | Giulia Lo Gerfo [1] | Franco Egidi [1] | Chiara Cappelli [1]

[1]Scuola Normale Superiore, Classe di Scienze, Piazza dei Cavalieri 7, 56126, Pisa

Correspondence
Chiara Cappelli and Tommaso Giovannini, Classe di Scienze, Scuola Normale Superiore, Pisa, 56126, Italy
Email: chiara.cappelli@sns.it, tommaso.giovannini@sns.it

Funding information

We present a computational study on the spectroscopic properties of UV-Vis absorbing dyes in water solution. We model the solvation environment by using both continuum and discrete models, with and without polarization, in order to establish how the physical and chemical properties of the solute-solvent interaction may affect the spectroscopic response of aqueous systems. Seven different compounds were chosen, representing different classes of organic molecules. The classical atomistic description of the solvent molecules was enriched with polarization effects treated by means of the Fluctuating Charges (FQ) model, propagated to the first-order response function of the quantum-mechanical (QM) solute to include its effects withing the modeling of the electronic excitations of the systems. Results obtained with the QM/FQ model were compared with those from continuum solvation models as well as non-polarizable atomistic models, and then confronted with the experimental values in order to determine the accuracy that can be expected with each level of theory. Moreover, a thorough structural analysis using Molecular Dynamics simulations is provided for each system.

KEYWORDS
Solvent Effects, QM/PCM, QM/MM, QM/FQ, Excitation Energies, TD-DFT

1 | INTRODUCTION

One-photon absorption spectroscopy within the UV-Visible range is often the most direct and inexpensive analytical tool that can be used to study the electronic properties of a system. Most commonly, such measurements are carried out on solvated samples, with water being a ubiquitous choice.

With the gradual increase in the complexity of the systems under investigation, the correct interpretation of experimental data is increasingly reliant upon their calculated *ab-initio* counterparts. Many theoretical models based on quantum mechanics (QM), accompanied by their computational implementations, have been presented over the years offering different levels of compromise between the computational cost and the accuracy of the results [1, 2, 3]. At present, methods based on density functional theory (DFT) and its time-dependent counterpart (TD-DFT) have become the most popular choice for the simulation of absorption spectra of medium-large organic molecular systems thanks to their versatility stemming from the freedom of choice of density functional and basis set, as well as the favorable scaling with system size which allows their application to increasingly large systems [4, 5, 6, 1].

Many benchmarks studies have been presented elaborating upon the merits and limitations of TD-DFT for the simulation of UV-Vis spectroscopy, as well as on the most appropriate choice of functional and basis set combination for different types of system [7, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16]. And though many computational studies are carried out on isolated systems, solvent effects should not be neglected for the presence of the solvation environment can significantly alter the electronic absorption properties of a system, both qualitatively and quantitatively [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32]. For this reason, theoretical models have been developed to tackle this problem and then combined with DFT and TD-DFT to include solvent effects within the theoretical model.

The standard protocol for such cases requires the usage of Polarizable Continuum Model (PCM) to describe solvation [33, 34, 29, 28]. This approach falls into the category of implicit solvent models where the environment is represented in a continuous way, while the solute molecule sits within a cavity and the surrounding continuum possesses dielectric properties that mimic the given solvent. Implicit methods prove to simulate correctly the properties of non-aqueous solutions lacking specific interactions between the solute and solvent molecules. Meanwhile, in water as well as in numerous other media the directional interactions, such as hydrogen bonds (HB), can play a crucial role. Hydrogen bonding can be introduced within continuum models by including explicit solvent molecules treated quantum mechanically in the system. This quantum treatment assures that both the directional nature of hydrogen and its covalent contribution are treated, however this still relies upon a static description of the system, whereas in reality the solvent moves about around the solute and a physically correct picture should not neglect the fact that the system is but an ensemble of many different configurations that may have varying spectroscopic properties. In fact, even though the positions of the explicit solvent molecules may be optimized to obtain a minimum-energy-structure, many such structures may be obtained in principle because of the high flexibility of the supramolecular system, but none of them taken singularly may be representative of the whole. In pure PCM, the converged quantum mechanical density is considered an *implicit* average over the configurational space of the solvent, as the latter is viewed as smeared out, which justifies the continuum picture. A super-molecule approach which re-introduces some solvent molecules as explicit static entities therefore helps to include some crucial interaction into the picture [32] but carries its own problems.

In order to overcome the limitations of implicit solvent models, explicit solvation models have been developed in the past years. The widest used explicit approach is based on the Quantum Mechanics/Molecular Mechanics (QM/MM) multi-scale scheme. The system is divided into the portion directly responsible for a given property (e.g. a chromophore interacting with light) that is described at the QM level and the surrounding molecules described at the MM (classical) level through *ad-hoc* constructed Force Field (FF). In the most commonly used QM/MM approaches, only the electrostatic interaction between the two portion is considered. In particular, each atom is endowed with a fixed

pre-parametrized charges, giving rise to the so called non-polarizable QM/MM models. Hovewer, to recover a better and more physical description of the electrostatic interaction between the QM and MM portions, several polarizable QM/MM models, in which the MM atoms can be polarized by the QM density, have been developed. Such models can be based on distributed multipoles [35, 36, 37, 38], induced dipoles [39, 40, 41, 42], Drude oscillators [43] or Fluctuating Charges (FQ) [44].

The FQ model was firstly developed into a 3-layer fully polarizable approach with non-periodic boundary conditions (QM/FQ/PCM). The method has subsequently been extended to allow calculations of numerous molecular properties by including features like analytical first and second derivatives [45], response equations [46], magnetic perturbations with Gauge Including Atomic Orbitals (GIAOs) [47], excitation energies (at the TD-DFT and equation-of-motion coupled cluster model with single and double substitutions levels of theory) [48, 49], Vibrational Optical Activity [50, 51], excitation energy gradients [52].

In the present work several approaches to include solvation effects, varying from the implicit QM/PCM to QM/MM approaches both including or not mutual polarization are challenged to reproduce the absorption spectra of a series of organic dyes by resorting to TD-DFT. The results are compared with experimental data in order to assess the accuracy of the different employed approaches.

The article is organized as follows: in the next section we briefly recall the fundamentals of the QM/FQ model, by also focusing on its extension to calculate TD-DFT/FQ excitation energies. After a section dedicated to the details of the computations, the results for selected organic dyes in aqueous solution, previously studied with different approaches [18, 53], are discussed. In particular, the discussion focuses on Molecular Dynamics (MD) and TD-DFT results. Eventually, we come to conclusions and considerations for further perspectives.

2 | THEORETICAL MODEL

The FQ model provides a computationally efficient and chemically consistent way of introducing polarization effects within both classical dynamics simulations and QM/MM calculations [54, 44]. In the FQ model, each atom is endowed with a charge that is allowed to fluctuate. Such fluctuation is ruled by the difference in atomic electronegativities. Thus, two set of parameters are needed to describe the FQ energy, namely atomic hardnesses and electronegativities. It is worth noticing that such terms can be rigorously defined in the "conceptual DFT" framework [55, 56]. Through them, the atomic charges can be calculated by solving a linear system [57, 58, 59]. More in detail, the value of the fluctuating charge on each MM atom is related to the electronegativity [57, 58, 59] according to the Electronegativity Equalization Principle (EEP)[55, 60], which states that, at equilibrium, the instantaneous electronegativity χ of each atom has the same value [60, 55]. The FQs ($\mathbf{q}$) can be defined as those minimizing the following functional [61]:

$$\begin{aligned} F(\mathbf{q},\lambda) &= \sum_{\alpha,i} q_{\alpha i}\chi_{\alpha i} + \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j} q_{\alpha i}J_{\alpha i,\beta j}q_{\beta j} + \sum_{\alpha}\lambda_{\alpha}\Big(\sum_{i} q_{\alpha i} - Q_{\alpha}\Big) \\ &= \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} \end{aligned} \tag{1}$$

where $\mathbf{q}$ is a vector containing the FQs, the Greek indices α run over molecules and the Latin ones i over the atoms of each molecule. $\boldsymbol{\lambda}$ is a set of Lagrangian multipliers used to impose charge conservation constraints on each molecule. In this work, the charge interaction kernel $\mathbf{J}$ is the Ohno kernel [62]. Atomic units are used throughout the manuscript. The

stationarity conditions of the functional in eq.1 are defined through the following equation [61]:

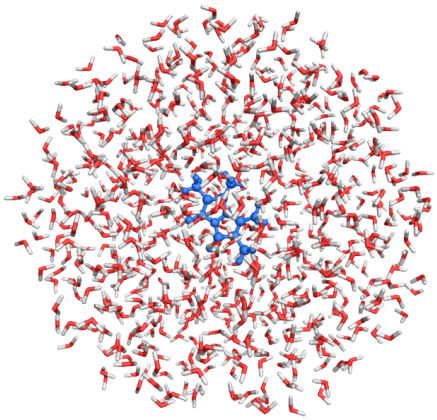
$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q \tag{2}$$

where $\mathbf{C}_Q$ collects atomic electronegativities and total charge constraints, whereas charges and Lagrangian multipliers are collected in $\mathbf{q}_{\boldsymbol{\lambda}}$, and $\mathbf{D}$ includes the $\mathbf{J}$ matrix and the Lagrangian blocks.

The FQ force field (FF) can be effectively coupled to QM methods. The resulting QM/FQ approach [44] has been shown to be especially suited to the modeling of response and spectral properties because, as it is shown below, its energy expression can be easily differentiated up to high orders. The QM/FQ describes also polarization effects: in this contest, the charges equilibrate to both the electrostatic potential generated by the QM moiety and their electronegativies, while the QM core feels the presence of the FQs through specific additional terms in the QM Hamiltonian, in a mutual polarization fashion.

The QM/FQ model system is usually partitioned in a QM core region placed at the center of a spherical region defining the environment (see Figure 1), which is described classically by exploiting the FQ FF. The size of this region is chosen to guarantee the convergence of the desired property/spectrum. Notice that the position of QM and MM atoms is obtained by a previous performed classical Molecular Dynamics (MD) allowing the exploration of the configuration space. This gives rise to a dynamic approach to the solvation phenomenon, which is instead neglected in purely continuum approaches.

FIGURE 1 Representation of the QM/FQ scheme. The QM portion is in blue.



The coupling between the QM and the MM portions is defined as the classical electrostatic interaction [46]:

$$E_{\text{QM/FQ}} = \sum_{i=1}^{N_q} V_{\text{QM}}[\rho](\mathbf{r}_i) q_i \tag{3}$$

where $V_{\text{QM}}[\rho](\mathbf{r}_i)$ is the electrostatic potential due to the QM density of charge at the i-th FQ q_i placed at $\mathbf{r}_i$. Notice that non-electrostatic interaction terms, which have been recently proposed by some of us [63], will not be considered in this work. By exploiting a Self Consistent Field (SCF) description of the QM moiety, the global QM/MM energy functional

reads[64, 46, 45]:

$$\mathcal{E}[\mathbf{P},\mathbf{q},\boldsymbol{\lambda}] = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{1}{2}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} + \mathbf{q}^{\dagger}\mathbf{V}(\mathbf{P}) \tag{4}$$

where **h** and **G** are the one and two electron contributions to the energy and Fock operator, respectively, and **P** is the density matrix. Finally, the FQs are obtained by solving the following equation

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_{Q} - \mathbf{V}(\mathbf{P}) \tag{5}$$

which includes the coupling term $\mathbf{V}(\mathbf{P})$ between the QM and MM moieties.

In case of the calculation of response/spectroscopic properties, such terms propagate to the solute's response equations, so that polarization effects are fully considered also in the computed final spectral data [46, 45, 47, 50, 51, 65].

2.1 | Linear Response Theory in QM/FQ

In order to calculate excitation UV-VIS spectra, we briefly recall how linear response equations have to be changed to account for the presence of the FQ portion. For a more detailed discussion on this topic, we refer the reader to ref. [46]. The following matrices, depending on the FQ charges, are defined:

$$\tilde{A}_{ai,bj} = (\varepsilon_a - \varepsilon_i)\delta_{ab}\delta_{ij} + \langle aj||ib\rangle - \sum_{kl}^{N_q} V_{ia}^{\dagger} D^{-1} V_{jb} \tag{6}$$

$$\tilde{B}_{ai,bj} = \langle ab||ij\rangle - \sum_{kl}^{N_q} V_{ia}^{\dagger} D^{-1} V_{bj} \tag{7}$$

where i, j are occupied orbitals whereas a, b are virtual orbitals. ε are orbital energies. The sum runs over the molecules in the classical portion and V is the electrostatic potential. D is the FQ matrix introduced in the previous section (see Eq. 5). Then, excitation energies and transition amplitudes are obtained by solving the so-called Casida's equations for the QM/FQ linear response theory [46]:

$$\begin{pmatrix} \tilde{A} & \tilde{B} \\ \tilde{A}^{*} & \tilde{B}^{*} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{8}$$

3 | COMPUTATIONAL DETAILS

For this work we have selected seven molecules, depicted in Figure 2. These systems are all organic molecules for which experimental measurements of their UV-Vis absorption properties in aqueous solution exist. Furthermore, several of these system are capable of forming intermolecular hydrogen bonds, and their absorption spectra exhibit bright excitations with varying degree of charge-transfer character, and can therefore be affected by the presence of the highly polar solvation environment to a different extent. All QM calculations were performed using the Gaussian16 program [66], and employed CAM-B3LYP density functional [67] and 6-311++G(d,p) basis set. Continuum solvation effects were included using the polarizable continuum model (PCM). The ground-state geometry for each molecule

was optimized at the QM/PCM taking into account the possible presence of multiple conformers, then the first five excited states for each were converged at the TD-DFT/PCM level. Non-equilibrium solvation effects [29, 33, 68, 69] were modeled by resorting to the Linear Response (LR) formalism. In order to estimate the effect of the covalent and directional components of hydrogen bond, both the geometry optimizations and spectra calculations were repeated after saturating every hydrogen bonding site with a water molecule (QM/QM$_w$/PCM results). The QM/MM calculations of excitation energies and intensities were performed by resorting to the following computational steps:

1. *Definition of the systems and calculation of atomic charges.* The solute molecules were surrounded by a number of water molecules large enough to represent all the solute-solvent interactions. The atomic charges of the solute were computed by using the Charge Model 5 (CM5) [70].
2. *Classical MD simulations in aqueous solution.* The MD simulations were performed in a cubic box reproduced periodically in every direction, satisfying the Periodic Boundary Conditions (PBC). A minimization step ensures that the several simulations were started from a minimum of the classical PES. From the MD runs, a set of snapshots was extracted to be used in the QM/MM and QM/FQ calculations.
3. *Definition of the different regions of the two-layer scheme and their boundaries.* Each snapshot extracted from the MD runs was cut into a sphere centered on the solute. The radius of the sphere was chosen in order to include all specific water-solute interactions.
4. *QM/MM or QM/FQ calculations and comparison with experimental data.* QM/MM or QM/FQ excitation energies calculations were performed on the set of structures obtained for the seven molecules in the previous step of the protocol. The results obtained for each spherical snapshot were extracted and averaged to produce the final value.

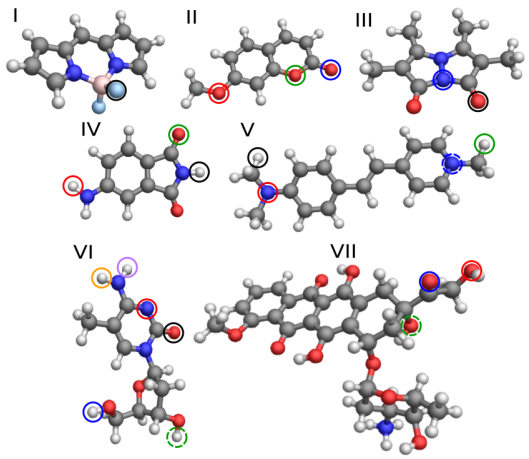


FIGURE 2 Representation of the studied molecules: **I** (bodipy); **II** (7-methoxycoumarin); **III** (bimane); **IV** (5-aminophtalimide) ; **V** (pyridinium dye); **VI** (5-methylcytidine); **VII** (doxorubucin).

The systems studied in this manuscript are depicted in Figure 2 and their simplified name are reported in Table 1. Notice that IUPAC names are reported in Table S1, given as Supporting Information (SI).

In step 1, the systems were optimized and CM5 charges were calculated at the B3LYP/6-31+G* level of theory including solvent effects by means of the PCM [34, 33]. For sake of completeness, the vertical energies were also

computed using the PCM model at the CAM-B3LYP/6-311++G** level of theory. Depending on the case of study, explicit solvent molecules were included (QM/QM_w/PCM).

The MD simulations were performed by using GROMACS[71], with the GROMOS [72], GAFF [73] and Amber11 [74] force fields to describe intra-/inter-molecular interactions. The force field used depend on the considered solute **I**-**VII** and are summarized in Table 1. CM5 charges were used to account for electrostatic interactions. The TIP3P-FB FF was used to describe the water molecules [75]. A single molecule was dissolved in a cubic box containing at least 3000 water molecules. The number of water molecules varies depending on the dimension of the considered molecules **I**-**VII** (see Table 1 for the exact number for each structure). For molecule **V**, a chloride ion has been included in the box to neutralize the system. The molecular systems were initially brought to 0 K with the steepest descent minimization procedure and then heated to 298.15 K in an NVT ensemble using the velocity-rescaling [76] method with an integration time step of 0.2 fs and a coupling constant of 0.1 ps for 200 ps. The time step and temperature coupling constant were then increased to 2.0 fs and 0.2 ps, respectively, and an NPT simulation (using the Parrinello-Rahman barostat and a coupling constant of 1.0 ps) for 1 ns was performed to obtain a uniform distribution of molecules in the box. 100 ns production runs in the NVT ensemble were then carried out, fixing the fastest internal degrees of freedom by means of the LINCS algorithm (δt=2.0 fs) [77]. Electrostatic interactions are treated by using particle-mesh Ewald (PME) [78] method with a grid spacing of 1.2 Å and a spline interpolation of order 4. The cross interactions for Lennard-Jones terms are calculated using the Lorentz-Berthelot [79, 80] mixing rules and we have excluded intramolecular interactions between atom pairs separated up to three bonds. A snapshot every 500 ps was extracted in order to obtain a total of 200 uncorrelated snapshots for each system. For each snapshot a solute-centered sphere with radius of at least 17 Å of explicit waters was cut. The radii used for each molecule **I**-**VII** are summarized in Table 1. Notice however that for molecule **V**, the chloride ion was not present in any of the extracted spherical snapshots. For each snapshot, the excitation energies were then calculated with two QM/MM models, treating the QM portion at the CAM-B3LYP/6-311++G** level. The water molecules were modeled by means of the non-polarizable TIP3P FF [81], and the FQ SPC parametrization proposed by Rick *et al.* [57]. Further calculations were performed by adding the closest water molecules in the QM portion, in a QM/QM_w/FQ framework. The water molecules to be included in the QM/QM_w/FQ calculations were chosen case by case by looking to the maximum/a in the Radial Distribution Function (RDF or $g(r)$) calculated from the MDs. The average number of water molecules included in the QM portion are reported in Table 1. All the QM(/QM_w)/MM(FQ) calculations were performed by using a locally modified version of Gaussian 16 package [66].

TABLE 1 Assigned number and simplified names for the studied molecules. Number of water molecules (N_{H_2O}) and Force Field used for molecules **I**-**VII** for MD simulations. The sphere radius for the QM/MM calculations and the average number of water molecules included in QM/QM_w/FQ calculations (N_{QM_w}) are also reported.

Chromophore	Simplified Name	N_{H_2O}	FF	Sphere radius (Å)	N_{QM_w}
I	bodipy	3000	GROMOS	17	2
II	7-methoxycoumarin	3000	GAFF	17	2
III	bimane	3000	GAFF	17	3
IV	5-aminophtalimide	3000	GAFF	17	6
V	pyridinium dye	5000	GAFF	20	2
VI	5-methylcytidine	3500	Amber11	17	6
VII	doxorubicin	5000	GAFF	20	10

4 | RESULTS

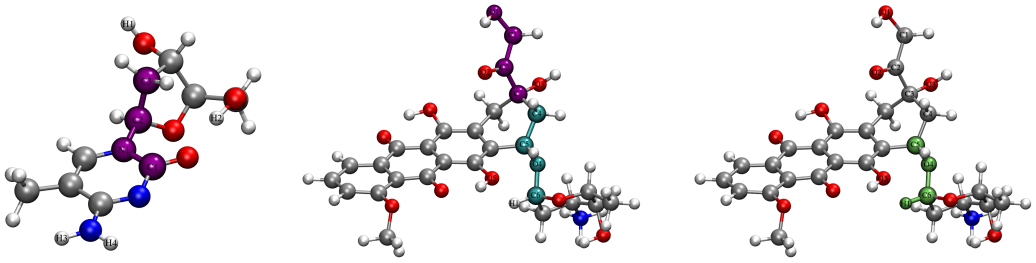
In the first part of this section, the results from the MD simulations are presented. In particular, we focus on the sampling of the conformational space of each molecule observed in the dynamics, as well as the emergence of hydrogen bonding patterns. Following the analysis of the results of the classical dynamics, the excitation energy and absorption spectra obtained with the QM/PCM, QM/QM$_w$/PCM, non-polarizable QM/MM, polarizable QM/FQ and QM/QM$_w$/FQ methods are presented. An analysis of how the different levels of theory employed in the modeling of the solute-solvent interaction can affect the predicted spectroscopic signature is provided for each system.

4.1 | MD Analysis

The MD trajectories of molecules **I**-**VII** were analyzed to provide information about how the solvation environment affects the conformational space explored by the systems and about the intermolecular interaction through Hydrogen Bond (HB) with water molecules. This analysis was performed by using TRAVIS package [82].

4.1.1 | Conformational analysis based on MD simulations

Among the seven considered molecules, only **VI** and **VII** are flexible and present different minima in the potential energy landscape that can be optimized using QM techniques. In Figure 3 the molecular structures of these molecules are reported highlighting the dihedral angles that define the different conformations.



(a) **VI.** D0 (purple). (b) **VII.** D1 (purple) and D2 (cyan) (c) **VII.** D3 (green)

FIGURE 3 (a) **VI** (5-methylcytidine, *syn* conformer),(b) and (c) **VII** (doxorubicin) structures. The dihedral angles studied in the conformation analysis are colored. Relevant atoms are also labeled.

In molecule **VI**, the flexibility is due to the rotation around the D0 dihedral angle (see Figure 3(a)), which defines the relative orientation of the sugar and the pyrimidine ring. Two main conformers can be present, namely *syn* and *anti* (see section S2 in the Supporting Information). Notice that in Figure 3(a), only the *syn* conformer is depicted. Figure 4 reports the time evolution of the D0 dihedral angle highlighted in Figure 3(a) as the classical dynamics unfolds, together with its resulting distribution. The *syn* conformer (D0 ≈ -70 degrees) results to be by far the most abundant, however the *anti* conformer (D0 ≈ 130 degrees) is also present, albeit with a much lower and less sharp population.

The plot in figure 4 also points out to one of the problems that may be encountered when resorting to a continuum solvation model, as a QM/PCM geometry optimization yields one structure which may not be representative of the

whole conformational space spanned by the solute around the relative minimum which is, in this case, quite wide.

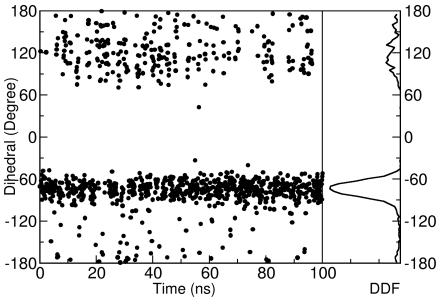


FIGURE 4 Time development and dihedral distribution of the dihedral angle D0 of molecule **VI**.

Concerning molecule **VII**, the main free rotation is given by the amino-sugar moiety with respect to the rest of the molecule. Three different dihedrals can be identified, namely D1, D2, D3 (see Figures 3(b) and 3(c)). The distributions of such dihedrals during the MD are depicted in Figure 5. The D1 maximum value is placed at about -135 degree indicating that no intramolecular interaction between O3 and the hydrogen bounded to O1 can occur. In figure 5, the D2 distribution shows a three peaks profile (81, 110, 157 degrees), whereas two peaks are identified in the D3 profile (-34, 44 degrees). It is worth noticing that D2 and D3 profiles show a correlation. In fact, such dihedrals are responsible of intramolecular interaction O5$\cdots$H1, which occurs when D2 and D3 are at about 44 and 110/157 degrees, respectively, which is not a highly populated area of the conformational landscape. This is confirmed also by the Radial Distribution Function (RDF, or $g(r)$) reported in the inset of Figure 5. It shows that the least intense peak at about 2.5 Å is associated with the intramolecular interaction just discussed.

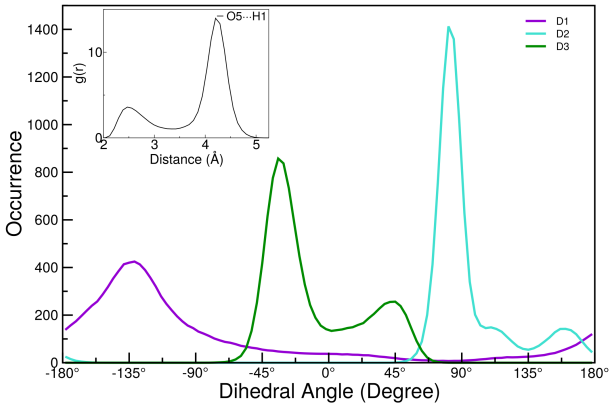


FIGURE 5 Population analysis of D1, D2 and D3 depicted in Figure 3(b) for molecule **VII**. In the inset, the radial distribution function between O5 and H1 is also reported (see Figure 3(b) for atom labeling).

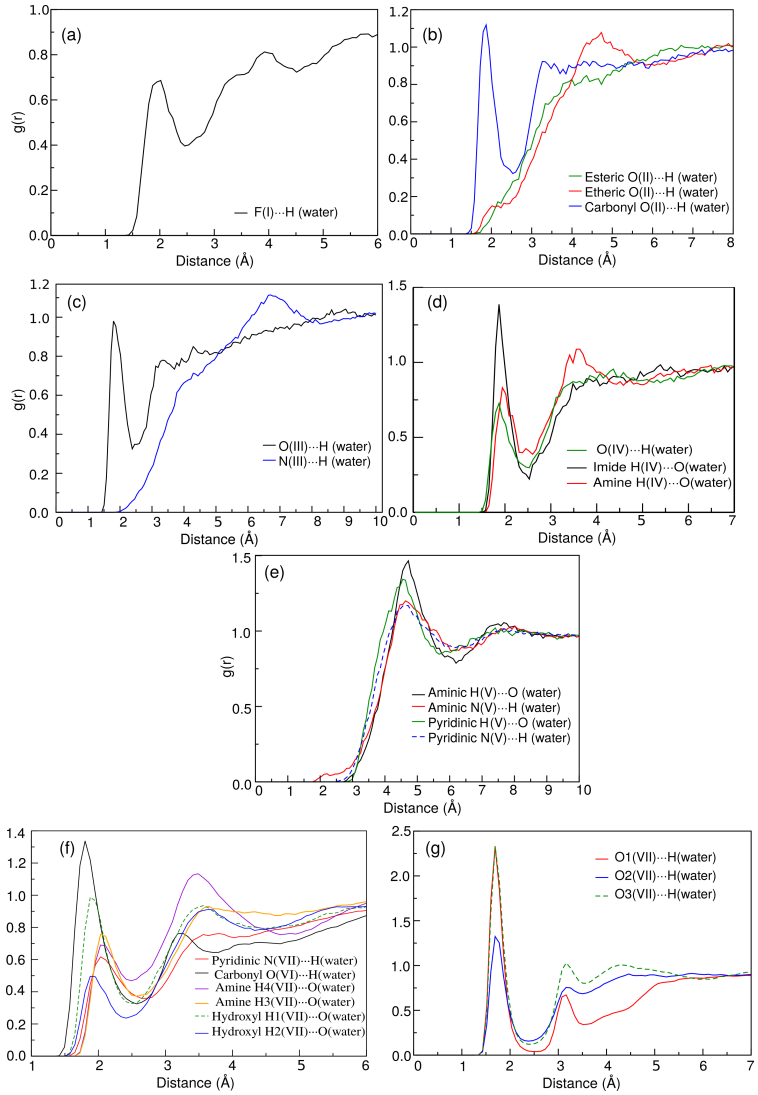


FIGURE 6 Radial distribution function between selected sites of all the analyzed chromophores and water molecules: (a) **I**, (b) **II**, (c) **III**, (d) **IV**, (e) **V**, (f) **VI**, (g) **VII**. Sites are highlighted in Figure 2.

4.1.2 | Hydration Patterns

Molecule **I**, whose solvation properties were previously investigated, albeit with a lack of polarization effects in the classical portion [83], is characterized by two fluorine atoms that can be involved in Hydrogen Bonds (HB) with water hydrogen atoms. In Figure 6(a), the $g(r)$ between F(**I**) and H(water) atoms is reported. The first peak at about 1.95-2.00 Å is intense and broad and it clearly represents an HB pattern.

Concerning molecule **II**, the three chromophore oxygen atoms are potentially responsible for HB interactions with water molecules. The $g(r)$ of molecule **II** is depicted in Figure 6(b). The carbonyl oxygen atom is the only one that presents values of O(**II**)$\cdots$H(water) distances characteristic of hydrogen bond. To further analyze this interaction, the Combined Distribution Function (CDF), i.e. $g(r)$ as function of the O(**II**)$\cdots$H(water) distance and the angle O(chromophore)$\cdots$HO(water), was calculated. The plot is depicted in Figure S3 given as SI. This analysis confirms that the carbonyl oxygen atom is involved in an HB with the surrounding water molecules.

Molecule **III** is characterized by two carbonyl oxygen atoms which are potentially bonded to water molecules through an HB interaction. The $g(r)$ is plotted in Figure 6(c). The O(**III**)-H(Water) rdfs are equal due to the molecular symmetry and they present a peak at about 1.8 Å which is an indicator of HB interaction. Notice that the integral value of the $g(r)$ shows that molecule **III** interacts on average with two water molecules at the same time.

Molecule **IV** presents two carbonyl oxygen atoms which may act as HB acceptors as well as an amine and an imide group which instead can act as HB donors. Molecule **IV** $g(r)$ is depicted in Figure 6(d). The HB donor character is the predominant one and the imide hydrogen atom is the most involved in the HB interaction. This is also confirmed by the Spatial Distribution Function (SDF) reported in Figure S4, given as SI.

Molecule **V** is characterized by the presence of two nitrogen atoms (aminic and pyridinic) whose ibridization character does not allow for the formation of hydrogen bonds with the solvent. The $g(r)$ presented in Figure 6(e) confirms this, with distances around 4.5-5.0 Å.

Molecule **VI** is characterized by several potential HB sites. The carbonyl and ether oxygen atoms, together with the iminic atom are potentially HB acceptors, whereas the oxygen and nitrogen atoms which are bound to an hydrogen atoms are potential HB donors. The $g(r)$ is depicted in Figure 6(f). The most intense peak of the several $g(r)$ is shown by the carbonyl oxygen atom, which means that molecule **VI** behaves mostly as HB acceptor. It is worth noticing that $g(r)$ of the hydrogen atoms (H3 and H4 in Figure 3(a)) of the amine group are not equivalent. In fact, at about 3.6 Å, the $g(r)$(H4$\cdots$O$_{H_2O}$) presents a broad peak which is instead absent in case of H3. This result suggests that the rotation around the C-N bond is blocked, and also that this bond has a partially double character. This can be chemically explained through the resonance of the lone electron pair on nitrogen with the π electrons of the aromatic ring. This conclusion is also supported by the CDF between the intermolecular distance of the pyridinic nitrogen and the water hydrogen atoms and the $g(r)$ of the amine H4$\cdots$O$_{H_2O}$ (see Figure S5 given as SI). The CDF shows that when a water molecule interacts with the pyridinic nitrogen atom the relative distance H4$\cdots$O(water) is exactly around 3.6 Å explaining the second broad peak in the g(r)H4$\cdots$O(water).

Molecule **VII** hydration pattern has already been studied extensively in a previous paper by some of the present authors and we direct an interested reader to this publication for more details. In this manuscript, the main intermolecular interactions between the chromophore and the water molecules are reported (see Figure 6(g)). Such interactions involve the oxygen atoms O1, O2 and O3 (see Figure 3(b), for atom labeling) and the hydrogen atoms of water. The three $g(r)$ are characterized by a peak placed at about 1.8 Å , and the most intense ones are related to the hydroxyl oxygen O1 and O3.

4.2 | Excitation Energies

We now move on to present the results obtained by exploiting the continuum and explicit (with or without polarization effects) approach to the calculation of the absorption spectra of each of the seven molecules. In general, five models of increasing complexity will be considered: (1) a purely continuum QM/PCM model, (2) a QM/QM$_w$/PCM model where explicit solvent molecules are included in the QM part to saturate hydrogen bond sites, (3) a non-polarizable QM/TIP3P hybrid quantum-classical model where the solvent is non-polarizable but rather treated using fixed charges, (4) the polarizable QM/FQ model to expose the role of solvent polarization in generating the spectroscopic response, and finally (5) a QM/QM$_w$/FQ model which treats some water molecules close to the solute hydrogen bonding site(s) at the DFT level to model any covalent effects that might be of importance. As aforementioned, to calculate the QM/FQ spectra 200 uncorrelated snapshots were extracted from the MD simulations; such a number is enough to yield a converged spectrum as already pointed out by some of the present authors [51, 50, 84, 85].

The QM/FQ convoluted and averaged spectra for molecules **I-VII** are reported in Figure 7. Therein the QM/PCM, QM/QM$_w$/PCM, QM/TIP3P, QM/QM$_w$/FQ and experimental data are also plotted. For QM/QM$_w$/PCM (i.e. with some water molecules explicitly introduced in the QM portion) further details are given in Figure S6 and in Table S2, given as SI. Notice that for molecule **V** and **VII** no explicit water are included in QM/PCM calculations: for molecule **V** this is due to the fact that no specific interactions are present in aqueous solution, as confirmed by Figure 6. Molecule **VII** presents three distinct conformers and several hydrogen bonding sites. Because of this, converging the geometries for all structures proved difficult, therefore we omitted this molecule from the analysis, though the inclusion of explicit QM solvent molecules was still done for the QM/QM$_w$/FQ model, which showed no significant change compared to the QM/FQ results (vide infra). For all molecules we considered the first bright excitation. Before comparing the different solvation models, the character of each excitation was investigated by looking at the molecular orbitals (MO) involved in the transition , which are depicted in Figure S7, given as SI. The MOs show that most of the excitations can be classified as charge-transfer states, however the exact degree to which the electron density is displaced during the excitations should be evaluated in order to provide more quantitative results. To this end we employ a simple and intuitive numerical index recently developed [86] which considers the baricenters of the positive and negative difference density. To analyze the Charge Transfer (CT) nature of the first electronic transition, the extension of the length of the electron transfer, we used a simple and intuitive index, denoted as D_{CT}, that was recently developed [86]. The barycenters of the positive and negative density distributions are calculated by the difference of the Ground State (GS) and Excited State (ES) densities. The CT length (D_{CT}) is defined as the distance between the two barycenters. In table 3, the D_{CT} for molecule **I-VII** are reported. Notice that also another quantitative index (Δr) proposed by Guido *et al.* based on MOs was employed in the analysis of the CT nature [87]. The values obtained by exploiting this alternative index are reported in Table S3, given as SI.

We start the discussion on excitation energies by focusing on molecule **I**. The experimental spectrum in Figure 7(a) is characterized by a band at about 500 nm. To guarantee a direct comparison with the experimental spectrum, for this molecule all the data have been convoluted with a Gaussian function with a Full Width Half Maximum (FWHM) of 0.13 eV. The calculated D_{CT} reported in Table 3 clearly show that the first excitation has no CT character. This is also confirmed by the MO involved in the transition depicted in Figure S7, given as SI. From Figure 7(a) and Table 2, it is clear that the several solvation models used in this work result in very similar excitation energies which differ from the experiment of about 20%, except for the non polarizable QM/TIP3P which is the worst method employed. Because the inclusion of polarization effects into the solvation model, while moving the results in right direction does not lead to a quantitative agreement, and considering the rigidity of the molecular structure, the observed difference between theory and experiment is most likely due to poor description afforded by chosen DFT functional. In fact, it has been shown that

bodipy dyes do require higher level QM theory models that adequately describe differential electron correlation to produce accurate results [88, 89]. On the other hand, all the polarizable models predict the same energy, meaning that polarization effects play indeed a role in capturing solvation effects for this molecule. Also, the inclusion of some water molecules in the QM portion, both in the case of QM/PCM and QM/FQ, does not give a relevant improvement in the comparison with experimental data.

Concerning molecule **II**, the experimental spectrum depicted in Figure 7(b) presents one main broad band with a maximum placed at about 325 nm. The calculated D_{CT} reported in Table 3 show that the first excitation has a low CT character, which is enhanced by the explicit solvation models. This is also confirmed by the MO involved in the transition depicted in Figure S7, given as SI. The general trend discussed for molecule **I** is valid also in this case. In fact, QM/PCM and QM/QM_w/PCM give very similar results, with a discrepancy of about 0.5 eV from the experimental value. The QM/TIP3P model is again the most inaccurate one, resulting in a shift of 0.05 eV with respect all the other methods. Notice however that the inclusion of water molecule in the QM portion in QM/QM_w/FQ calculations results in a shift of the excitation energy towards the experiment. This can be rationalized by considering that the a strong and directional HB interaction is present in the MD (see previous section). Also in this case, it is worth noticing that the several methods employed give similar excitation energies, with an error that is almost constant (12% on average). Thus, such a discrepancy is again probably due to the computational level of theory.

The experimental spectrum of molecule **III** presents a double peak profile in the region 350-400 nm, which is probably due to vibronic effects (see Figure 7(c)). The calculated D_{CT} data reported in Table 3 and the associated MO depicted in Figure S7 in the SI show that the first excitation has no CT character. The most intense is the peak at about 400 nm. The QM/PCM approach predicts an excitation energy which is 0.62 eV blue shifted with respect to the experimental value. The inclusion of three explicit water molecules redshift the energy of about 0.2 eV, however resulting in a discrepancy with respect the experiment of about 0.4 eV. The explicit solvation models coupled with the dynamic approach of the MD results are in fair agreement with the experimental data, also at the QM/TIP3P level of theory. This is particularly interesting and it can be due to the fact the in this case polarization effects are not crucial in the description of the excitation energy.

Molecule **IV** experimental UV-VIS spectrum has a maximum at about 375 nm (see Figure 7(d)). In this case, MO (see Figure S7, given as SI) and D_{CT} values show for the first transition a CT character, although it is generally small. Similarly to the previous cases, QM/PCM predicts an excitation energy lower of about 0.4 eV with respect the experiment. The inclusion of explicit water molecules, however, gives an almost perfect agreement with only a 1% error. The errors obtained by using the explicit solvation models are lower with respect the QM/PCM model, with an error of 2% in the case of QM/QM_w/FQ approach. Notice that in this case the purely QM/FQ model gives a discrepancy of about 0.25 eV with respect the experimental value. Considering that the inclusion of explicit water molecules both in QM/QM_w/PCM and QM/QM_W/FQ models, are crucial in the reproduction of the excitation energies, some non-electrostatic effects can play a relevant role in this case.

Concerning molecule **V**, the experimental spectrum is characterized by a broad band placed at about 450 nm (see Figure 7(e)). As resulting from the D_{CT} calculations (see Table 3 and from the MO involved (see Figure S7 in the SI), the first transition has a CT character. As pointed out before, in this case no explicit water molecules were included in the QM/PCM calculations, thus no results for the QM/QM_w/PCM are discussed. This was due to the fact that no specific solute-solvent interactions were identified from the RDF depicted in Figure 6, panel (e). This explains also why the results obtained by using an implicit or an explicit model are very similar and in general in fair agreement. Notice that the QM/TIP3P error is the highest, and again this can be explained by the fact that polarization effects in the solvation model may be crucial in this case.

The experimental excitation spectrum of molecule **VI** presents a main band at about 280 nm (see Figure 7(f)). The

calculated D_{CT} reported in Table 3 clearly show that the first excitation has not a CT character. This is also confirmed by the MO involved in the transition depicted in Figure S7, given as SI. QM/PCM and QM/QM_w/PCM were calculated by weighting the spectra of the two conformers by their Boltzmann population (see Section S2 and Figure S6 given as SI). The excitation energies predicted by exploiting such methods differ of an average 14% with respect the experimental value. In particular, the discrepancy is of 0.6 eV on average. A shifting towards the experiment is recovered by using the explicit solvation models: the QM/FQ gives the best agreement (0.4 eV, 8%), however no great differences between the three exploited models is reported.

The last molecule (**VII**)'s experimental spectrum presents a huge vibronic band with a maximum at about 500 nm (see Figure 7(g)). Similarly to molecule **V**, also in this case no explicit water molecules were added to the QM region in QM/QM_w/PCM calculations. Again, also in this case, the first excitation has no CT character (see Table 3 and Figure S7 in the SI). The four solvation approaches give very similar results, with the maximum shifting giving passing through the QM/PCM to the QM/QM_w/FQ model (0.11 eV). However, as pointed out before for molecule **I**, the similarity in the results may be due to the computational level adopted in this work. Probably, vibronic contributions influence the position and the shape of the absorption band, thus resulting in a shifting and a better agreement with the experiment.

To conclude the discussion on calculated excitation energies, in Table 2 the Mean Relative Deviation (MRD), Root Mean Square Deviation (RMSD) and Maximum Absolute Deviation (MAD) obtained for all the considered models are reported. As it was expected the less MRD is shown by the QM/QM_w/FQ, however the inclusion of explicit water molecules is not crucial to reproduce the excitation energy of the variety of studied molecules. In fact, QM/FQ model results in a MRD of about 10%. RMSD confirms this trend. In fact, QM(/QM_w)/PCM reports an RMSD 0.13 eV greater than the QM/FQ model. The same conclusions can be also extracted by the MAD values, which are reported for molecule **I** in every adopted model.

TABLE 2 QM/PCM, QM/QM_w/PCM, QM/TIP3P, QM/QM_w/FQ excitation energies. Experimental data taken from the indicated references are reported in the last column. Relative deviations with respect the experiments are given in parentheses. Mean Relative Deviation (MRD), Root Mean Square Deviation (RMSD) and Maximum Absolute Deviation (MAD) are also reported. All data are in eV.

Chromophore	QM/PCM	QM/QM_w/PCM	QM/TIP3P	QM/FQ	QM/QM_w/FQ	Exp
I	2.99 (21%)	2.98 (20%)	3.04 (23%)	2.97 (20%)	2.97 (20%)	2.48 [90]
II	4.28 (12%)	4.28 (12%)	4.32 (13%)	4.28 (12%)	4.23 (11%)	3.81 [18]
III	3.82 (20%)	3.63 (14%)	3.11 (3%)	3.14 (2%)	3.04 (5%)	3.20 [91]
IV	3.72 (11%)	3.32 (1%)	3.52 (5%)	3.60 (7%)	3.42 (2%)	3.36 [92]
V	2.96 (6%)	-	3.04 (9%)	2.91 (5%)	2.94 (6%)	2.78 [18]
VI	5.05 (13%)	5.09 (14%)	4.87 (9%)	4.83 (8%)	4.88 (9%)	4.46 [53]
VII	2.99 (16%)	-	2.94 (14%)	2.93 (13%)	2.88 (12%)	2.58 [93]
MRD	14%	12%	11%	10%	9%	
RMSD	0.47	0.46	0.37	0.34	0.32	
MAD	0.63	0.63	0.56	0.49	0.49	

To end the discussion, we show how the spectra are obtained in the case of the QM/MM methods from raw data.

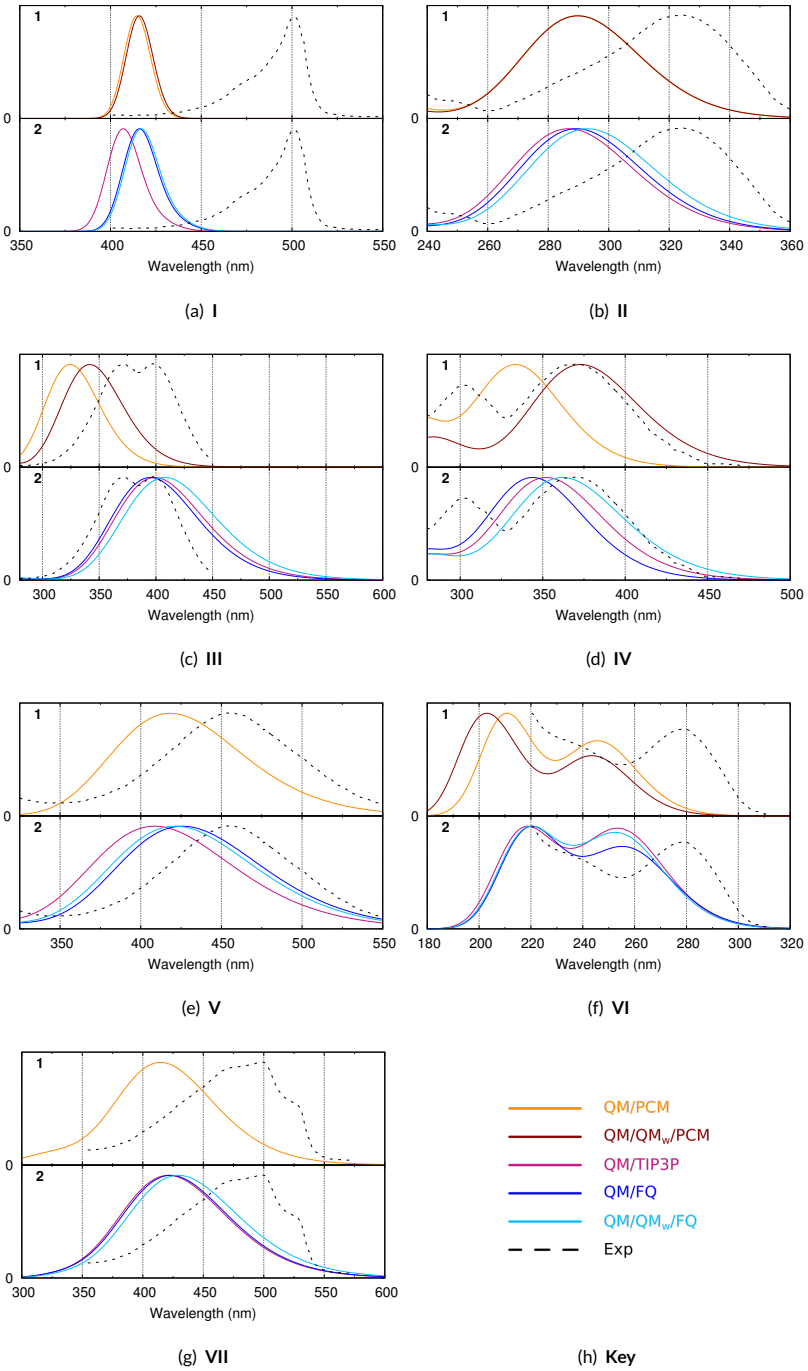


FIGURE 7 QM/PCM, QM/QM$_w$/PCM, QM/TIP3P, QM/QM$_w$/FQ and experimental UV-VIS spectra for molecules **I-VII**. The key is shown in panel h.

TABLE 3 QM/PCM, QM/QM_w/PCM, QM/TIP3P, QM/QM_w/FQ calculated D_{CT} (Å) index for the first excitation.

Chromophore	QM/PCM	QM/QM_w/PCM	QM/TIP3P	QM/FQ	QM/QM_w/FQ
I	0.55	0.56	0.59	0.59	0.59
II	1.95	1.96	2.16	2.07	2.19
III	0.46	0.50	0.73	0.69	0.73
IV	2.62	2.64	2.87	2.82	2.88
V	4.55	-	5.89	5.90	5.86
VI	1.00	0.88	1.13	1.21	1.05
VII	1.85	-	1.80	1.87	1.75

The data extracted from the single QM/FQ calculations on each snapshot for molecule **VI** in aqueous solution are reported as a stick spectrum in Figure 8. The same spectra for all of the investigated molecules **I-VII** are reported in Figure S8, given as SI.

Clearly, the overall shape of the final, averaged spectrum is already visible from the data reported in Figure 8, which also gives insight into the spreading of the transition bands, both in wavelengths and intensities. This is due to the fact that in the different snapshots the spatial distribution of water molecules around the molecule varies and also the conformational freedom of the molecule is sampled. The same also applies to the other investigated systems (see SI). In order to obtain the final, averaged spectrum, each transition in Fig. 8 was convoluted with a Gaussian function and averaged. It should be clear that the final spectrum emerges as an ensemble average of many different snapshots which may have widely varying spectroscopic response properties. This is in stark contrast with results obtained using methods like standard QM/PCM which most commonly only takes minimum-energy-structures as representative of the whole, and whose results are then usually convoluted with wide empirical lineshape functions meant to represent the spread of conformational and solvation degrees of freedom, and whose true distribution may be far from being represented by a simple Gaussian or Lorentzian lineshape function. Notice however that other more sophisticated approaches have been proposed in the literature to overcome this problem [32].

5 | SUMMARY AND CONCLUSIONS

We have presented a computational study examining the merits and shortcomings of five different solvation models in the reproduction of UV-Vis absorption spectra of organic molecules in aqueous solution. The picture that has emerged shows that the performance of each model is highly dependent on the specific properties of each system, particularly the extent of the charge transfer character of the excitations. However, a general conclusion that can be drawn is that the inclusion of solute-solvent polarization effects, whether using continuum or discrete models, can often be crucial and lead to a significant improvement in the results. The inclusions of such effects through the fluctuating charge model (FQ) does not lead to any significant increase in the computational effort because the cost of solving the QM/FQ equations is negligible compared to the cost of optimizing the QM wavefunction or solving the linear response equations from which excited state properties are extracted, and can therefore be safely applied in all cases. The use of polarizable QM/MM methods, in addition, offer the advantage of being able to sample the solute-solvent conformational space completely, without having to rely on a minimum-energy-structure picture. This is particularly useful in the case of flexible systems, for which larger differences between QM/PCM and QM/FQ results can be observed. Finally, while the PCM model

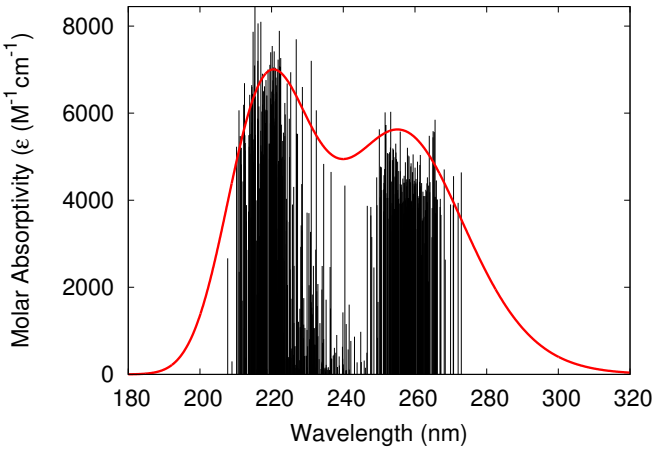


FIGURE 8 Molecule **VI** QM/FQ calculated data reported as stick spectrum and convoluted with a Gaussian band shape (FWHM=0.5 eV)

is often used on its own to model the effect of solvation, it is unable to correctly model the directional component of hydrogen bond interactions. This has also transpired through the analysis of our results in several cases where a large difference could be observed between QM/FQ and QM/PCM spectra. The inclusion of some explicit water molecule treated at the QM level did improve the QM/PCM results in some cases, bringing them closer to those obtained using the QM/FQ method. However this procedure rests on the assumption that such solvent molecules rest at fixed positions with respect to the solute. As evidenced from the classical dynamics, this is not the case as solvent molecules move about spanning a large space of configurations which collectively serve to produce the specific interaction that is peculiar to hydrogen bonding. This interaction is fully recovered in the QM/FQ picture which offers both the advantage of including polarization effects as when using PCM, and building upon a dynamical solute-solvent picture as is commonly done in hybrid non-polarizable QM/MM methods. If necessary, some of the solvent molecules closer to the solute can still be treated quantum mechanically in order to include any covalent contributions to hydrogen bonding that may be present, however the difference between the results obtained this way compared to the difference between the QM/PCM and the QM/QM$_w$/PCM results is not as large because the QM/FQ method already includes directional contributions in the picture. For the systems studied in this work, the effect of the eventual covalent character of the solute-solvent interaction alone was not crucial and only contributed in a minor capacity to the final results.

These results should still be seen as preliminary as still much more work remains to be done in benchmarking polarizable QM/MM methods for the purpose of calculating spectra of systems in solution. UV-Vis absorption spectroscopy is but the simplest type of observable that may be studied, and more complex spectroscopic observables might be considered for future benchmarks, including mixed electric-magnetic properties such as circular dichroism intensities or higher order spectroscopies such as Raman or Raman optical activity. In this work we have only considered a set of medium-sized organic molecules, however more complex systems, with larger CT effects, may be of interest. Biological molecules such as peptides or nucleic acids, for which aqueous solution is the natural environment, are often both highly flexible and able to form multiple hydrogen bonds with the solvent, and are therefore the ideal systems for the application of the method. The study of the spectroscopic properties of solvated inorganic systems such as transition metal complexes through polarizable QM/MM methods is also a largely unexplored field and may present

its own peculiar challenges. Finally, this method can be extended to non-aqueous solvents both polar and non-polar. For apolar solvents the effect of polarization may be much less important compared to water, in that case significant contributions to the solute-solvent interactions may instead come from non-electrostatic forces such as dispersion and repulsion. Methods for the inclusion of such effects in a QM/MM picture have been recently presented [63] however their extension to the computation of molecular spectra is still lacking and will be the topic of future investigations.

6 | ACKNOWLEDGEMENTS

We are thankful for the computer resources provided by the high performance computer facilities of the SMART Laboratory (http://smart.sns.it/).

REFERENCES

[1] Barone V, Baiardi A, Biczysko M, Bloino J, Cappelli C, Lipparini F. Implementation and validation of a multi-purpose virtual spectrometer for large systems in complex environments. Phys Chem Chem Phys 2012;14(36):12404–12422.

[2] Hättig C, Weigend F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. J Chem Phys 2000;113(13):5154–5161.

[3] Egidi F, Segado M, Koch H, Cappelli C, Barone V. A benchmark study of electronic excitation energies, transition moments, and excited-state energy gradients on the nicotine molecule. J Chem Phys 2014;141(22):224114.

[4] Improta R, Barone V. Absorption and fluorescence spectra of uracil in the gas phase and in aqueous solution: A TD-DFT quantum mechanical study. J Am Chem Soc 2004;126(44):14320–14321.

[5] Barone V. Computational strategies for spectroscopy: from small molecules to nano systems. John Wiley & Sons; 2011.

[6] Jacquemin D, Wathelet V, Perpete EA, Adamo C. Extensive TD-DFT benchmark: singlet-excited states of organic molecules. J Chem Theory Comput 2009;5(9):2420–2435.

[7] Jacquemin D, Perpete EA, Scuseria GE, Ciofini I, Adamo C. TD-DFT performance for the visible absorption spectra of organic dyes: conventional versus long-range hybrids. J Chem Theory Comput 2008;4(1):123–135.

[8] Laurent AD, Jacquemin D. TD-DFT benchmarks: A review. Int J Quantum Chem 2013;113(17):2019–2039.

[9] Goerigk L, Grimme S. Assessment of TD-DFT methods and of various spin scaled CIS (D) and CC2 versions for the treatment of low-lying valence excitations of large organic dyes. J Chem Phys 2010;132(18):184103.

[10] Li J, Cramer CJ, Truhlar DG. Two-response-time model based on CM2/INDO/S2 electrostatic potentials for the dielectric polarization component of solvatochromic shifts on vertical excitation energies. Int J Quantum Chem 2000;77(1):264–280.

[11] Silva-Junior MR, Schreiber M, Sauer SP, Thiel W. Benchmarks for electronically excited states: Time-dependent density functional theory and density functional theory based multireference configuration interaction. J Chem Phys 2008;129(10):104103.

[12] Caricato M, Trucks GW, Frisch MJ, Wiberg KB. Electronic transition energies: a study of the performance of a large range of single reference density functional and wave function methods on valence and Rydberg states compared to experiment. J Chem Theory Comput 2010;6(2):370–383.

[13] Isegawa M, Peverati R, Truhlar DG. Performance of recent and high-performance approximate density functionals for time-dependent density functional theory calculations of valence and Rydberg electronic transition energies. J Chem Phys 2012;137(24):244104.

[14] Isegawa M, Truhlar DG. Valence excitation energies of alkenes, carbonyl compounds, and azabenzenes by time-dependent density functional theory: Linear response of the ground state compared to collinear and noncollinear spin-flip TDDFT with the Tamm-Dancoff approximation. J Chem Phys 2013;138(13):134111.

[15] Alipour M. On the performance of time-dependent double-hybrid density functionals for description of absorption and emission spectra of heteroaromatic compounds. Theor Chem Acc 2016;135(3):67.

[16] Maier TM, Bahmann H, Arbuznikov AV, Kaupp M. Validation of local hybrid functionals for TDDFT calculations of electronic excitation energies. J Chem Phys 2016;144(7):074106.

[17] Jacquemin D, Mennucci B, Adamo C. Excited-state calculations with TD-DFT: from benchmarks to simulations in complex environments. Phys Chem Chem Phys 2011;13(38):16987–16998.

[18] Jacquemin D, Planchat A, Adamo C, Mennucci B. TD-DFT Assessment of Functionals for Optical 0–0 Transitions in Solvated Dyes. J Chem Theory and Comput 2012;8(7):2359–2372. http://dx.doi.org/10.1021/ct300326f.

[19] Reichardt C. Solvatochromism, thermochromism, piezochromism, halochromism, and chiro-solvatochromism of pyridinium N-phenoxide betaine dyes. Chem Soc Rev 1992;21(3):147–153.

[20] Buncel E, Rajagopal S. Solvatochromism and solvent polarity scales. Acc Chem Res 1990;23(7):226–231.

[21] Reichardt C. Solvatochromic dyes as solvent polarity indicators. Chem Rev 1994;94(8):2319–2358.

[22] Cannelli O, Giovannini T, Baiardi A, Carlotti B, Elisei F, Cappelli C. Understanding the interplay between the solvent and nuclear rearrangements in the negative solvatochromism of a push–pull flexible quinolinium cation. Phys Chem Chem Phys 2017;19(48):32544–32555.

[23] Prampolini G, Bellina F, Biczysko M, Cappelli C, Carta L, Lessi M, et al. Computational Design, Synthesis, and Mechanochromic Properties of New Thiophene-Based π-Conjugated Chromophores. Chemistry Eur J 2013;19(6):1996–2004.

[24] Tosi I, Segado Centellas M, Campioli E, Iagatti A, Lapini A, Sissa C, et al. Excitation Dynamics in Hetero-bichromophoric Calixarene Systems. ChemPhysChem 2016;17(11):1686–1706.

[25] Carlotti B, Cesaretti A, Cannelli O, Giovannini T, Cappelli C, Bonaccorso C, et al. Evaluation of Hyperpolarizability from the Solvatochromic Method: Thiophene Containing Push- Pull Cationic Dyes as a Case Study. J Phys Chem C 2018;122:2285–2296.

[26] Lapini A, Fabbrizzi P, Piccardo M, di Donato M, Lascialfari L, Foggi P, et al. Ultrafast resonance energy transfer in the umbelliferone–alizarin bichromophore. Phys Chem Chem Phys 2014;16(21):10059–10074.

[27] Budzák S, Laurent AD, Laurence C, Medved' M, Jacquemin D. Solvatochromic Shifts in UV–Vis Absorption Spectra: The Challenging Case of 4-Nitropyridine N-Oxide. J Chem Theory Comput 2016;12(4):1919–1929.

[28] Mennucci B, Cappelli C, Guido CA, Cammi R, Tomasi J. Structures and properties of electronically excited chromophores in solution from the polarizable continuum model coupled to the time-dependent density functional theory. The Journal of Physical Chemistry A 2009;113(13):3009–3020.

[29] Marenich AV, Cramer CJ, Truhlar DG, Guido CA, Mennucci B, Scalmani G, et al. Practical computation of electronic excitation in solution: vertical excitation model. Chem Sci 2011;2(11):2143–2161.

[30] Labat F, Le Bahers T, Ciofini I, Adamo C. First-principles modeling of dye-sensitized solar cells: challenges and perspectives. Acc Chem Res 2012;45(8):1268–1277.

[31] Marenich AV, Cramer CJ, Truhlar DG. Uniform Treatment of Solute–Solvent Dispersion in the Ground and Excited Electronic States of the Solute Based on a Solvation Model with State-Specific Polarizability. J Chem Theory Comput 2013;9(8):3649–3659.

[32] Marenich AV, Cramer CJ, Truhlar DG. Electronic absorption spectra and solvatochromic shifts by the vertical excitation model: solvated clusters and molecular dynamics sampling. J Phys Chem B 2014;119(3):958–967.

[33] Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. Chem Rev 2005;105(8):2999–3094.

[34] Barone V, Cossi M. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. J Phys Chem A 1998;102(11):1995–2001.

[35] Day PN, Jensen JH, Gordon MS, Webb SP, Stevens WJ, Krauss M, et al. An effective fragment method for modeling solvent effects in quantum mechanical calculations. J Chem Phys 1996;105(5):1968–1986.

[36] Kairys V, Jensen JH. QM/MM boundaries across covalent bonds: a frozen localized molecular orbital-based approach for the effective fragment potential method. J Phys Chem A 2000;104(28):6656–6665.

[37] Mao Y, Demerdash O, Head-Gordon M, Head-Gordon T. Assessing Ion–Water Interactions in the AMOEBA Force Field Using Energy Decomposition Analysis of Electronic Structure Calculations. J Chem Theory Comput 2016;12(11):5422–5437.

[38] Loco D, Polack É, Caprasecca S, Lagardere L, Lipparini F, Piquemal JP, et al. A QM/MM approach using the AMOEBA polarizable embedding: from ground state energies to electronic excitations. J Chem Theory Comput 2016;12(8):3654–3661.

[39] Thole BT. Molecular polarizabilities calculated with a modified dipole interaction. Chem Phys 1981;59(3):341–350.

[40] Steindal AH, Ruud K, Frediani L, Aidas K, Kongsted J. Excitation energies in solution: the fully polarizable QM/MM/PCM method. J Phys Chem B 2011;115(12):3027–3037.

[41] Jurinovich S, Curutchet C, Mennucci B. The Fenna–Matthews–Olson Protein Revisited: A Fully Polarizable (TD) DFT/MM Description. ChemPhysChem 2014;15(15):3194–3204.

[42] Donati G, Wildman A, Caprasecca S, Lingerfelt DB, Lipparini F, Mennucci B, et al. Coupling Real-Time Time-Dependent Density Functional Theory with Polarizable Force Field. J Phys Chem Lett 2017;8(21):5283–5289.

[43] Boulanger E, Thiel W. Solvent boundary potentials for hybrid QM/MM computations using classical drude oscillators: a fully polarizable model. J Chem Theory Comput 2012;8(11):4527–4538.

[44] Cappelli C. Integrated QM/Polarizable MM/Continuum Approaches to Model Chiroptical Properties of Strongly Interacting Solute-Solvent Systems. Int J Quantum Chem 2016;116(21):1532–1542.

[45] Lipparini F, Cappelli C, Scalmani G, De Mitri N, Barone V. Analytical first and second derivatives for a fully polarizable QM/classical hamiltonian. J Chem Theory Comput 2012;8(11):4270–4278.

[46] Lipparini F, Cappelli C, Barone V. Linear response theory and electronic transition energies for a fully polarizable QM/classical hamiltonian. J Chem Theory Comput 2012;8(11):4153–4165.

[47] Lipparini F, Cappelli C, Barone V. A gauge invariant multiscale approach to magnetic spectroscopies in condensed phase: General three-layer model, computational implementation and pilot applications. J Chem Phys 2013;138(23):234108.

[48] 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.

[49] Egidi F, Russo R, Carnimeo I, D'Urso A, Mancini G, Cappelli C. The Electronic Circular Dichroism of Nicotine in Aqueous Solution: A Test Case for Continuum and Mixed Explicit-Continuum Solvation Approaches. J Phys Chem A 2015;119(21):5396–5404.

[50] Giovannini T, Olszowka M, Cappelli C. Effective Fully Polarizable QM/MM Approach To Model Vibrational Circular Dichroism Spectra of Systems in Aqueous Solution. J Chem Theory Comput 2016;12(11):5483–5492.

[51] Giovannini T, Olszòwka M, Egidi F, Cheeseman JR, Scalmani G, Cappelli C. Polarizable Embedding Approach for the Analytical Calculation of Raman and Raman Optical Activity Spectra of Solvated Systems. J Chem Theory Comput 2017;13(9):4421–4435.

[52] Carnimeo I, Cappelli C, Barone V. Analytical gradients for MP2, double hybrid functionals, and TD-DFT with polarizable embedding described by fluctuating charges. J Comput Chem 2015;36(31):2271–2290.

[53] Martínez-Fernández L, Pepino AJ, Segarra-Martí J, Banyasz A, Garavelli M, Improta R. Computing the Absorption and Emission Spectra of 5-Methylcytidine in Different Solvents: A Test-Case for Different Solvation Models. J Chem Theory Comput 2016;12(9):4430–4439. http://dx.doi.org/10.1021/acs.jctc.6b00518.

[54] Mancini G, Brancato G, Barone V. Combining the Fluctuating Charge Method, Non-Periodic Boundary Conditions and Meta-Dynamics: Aqua Ions as Case Studies. J Chem Theory Comput 2014;10(3):1150–1163.

[55] Mortier WJ, Van Genechten K, Gasteiger J. Electronegativity equalization: application and parametrization. J Am Chem Soc 1985;107(4):829–835. http://pubs.acs.org/doi/abs/10.1021/ja00290a017.

[56] Chelli R, Procacci P. A transferable polarizable electrostatic force field for molecular mechanics based on the chemical potential equalization principle. J Chem Phys 2002;117(20):9175–9189.

[57] Rick SW, Stuart SJ, Berne BJ. Dynamical fluctuating charge force fields: Application to liquid water. J Chem Phys 1994;101(7):6141–6156.

[58] Rick SW, Stuart SJ, Bader JS, Berne B. Fluctuating charge force fields for aqueous solutions. J Mol Liq 1995;65:31–40.

[59] Rick SW, Berne BJ. Dynamical Fluctuating Charge Force Fields: The Aqueous Solvation of Amides. J Am Chem Soc 1996;118(3):672–679. http://pubs.acs.org/doi/abs/10.1021/ja952535b.

[60] Sanderson R. An interpretation of bond lengths and a classification of bonds. Science 1951;114(2973):670–672.

[61] Lipparini F, Barone V. Polarizable force fields and polarizable continuum model: a fluctuating charges/PCM approach. 1. theory and implementation. J Chem Theory Comput 2011;7(11):3711–3724.

[62] Ohno K. Some remarks on the Pariser-Parr-Pople method. Theor Chim Acta 1964;2(3):219–227.

[63] Giovannini T, Lafiosca P, Cappelli C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. J Chem Theory Comput 2017;13(10):4854–4870.

[64] McWeeny R. Methods of molecular quantum mechanics. Academic press; 1992.

[65] Egidi F, Carnimeo I, Cappelli C. Optical rotatory dispersion of methyloxirane in aqueous solution: assessing the performance of density functional theory in combination with a fully polarizable QM/MM/PCMapproach. Opt Mater Express 2015;5(1):196–209.

[66] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al., Gaussian 16 Revision A.03; 2016. Gaussian Inc. Wallingford CT.

[67] Yanai T, Tew DP, Handy NC. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). Chem Phys Lett 2004;393(1):51–57.

[68] Egidi F, Barone V, Bloino J, Cappelli C. Toward an accurate modeling of optical rotation for solvated systems: Anharmonic vibrational contributions coupled to the polarizable continuum model. J Chem Theory Comput 2012;8(2):585–597.

[69] Mennucci B, Cammi R, Tomasi J. Excited states and solvatochromic shifts within a nonequilibrium solvation approach: A new formulation of the integral equation formalism method at the self-consistent field, configuration interaction, and multiconfiguration self-consistent field level. J Chem Phys 1998;109(7):2798–2807.

[70] Dodda LS, Vilseck JZ, Cutrona KJ, Jorgensen WL. Evaluation of cm5 charges for nonaqueous condensed-phase modeling. J Chem Theory Comput 2015;11(9):4273–4282.

[71] Abrahama MJ, Murtola T, Schulz R, Pálla S, Smith JC, Hess B, et al. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. SoftwareX 2015;1-2:19–25.

[72] Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 2004;25(13):1656–1676. http://dx.doi.org/10.1002/jcc.20090.

[73] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25(9):1157–1174. http://dx.doi.org/10.1002/jcc.20035.

[74] Sorin EJ, Pande VS. Exploring the Helix-Coil Transition via All-Atom Equilibrium Ensemble Simulations. Biophys J 2005;88(4):2472 – 2493. http://www.sciencedirect.com/science/article/pii/S000634950573304X.

[75] Wang LP, Martinez TJ, Pande VS. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. J Phys Chem Lett 2014;5(11):1885–1891. http://dx.doi.org/10.1021/jz500737m.

[76] Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys 2007;126(1). http://scitation.aip.org/content/aip/journal/jcp/126/1/10.1063/1.2408420.

[77] Hess B, Bekker H, Berendsen HJ, Fraaije JG. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997;18(12):1463–1472.

[78] Darden T, York D, Pedersen L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. J Chem Phys 1993;98(12):10089–10092.

[79] Antoon LH. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. Annals Physics 1881;248:127–136.

[80] Marcellin B. Sur Le Mélange Des Gaz. Comptes rendus hebdomadaires des séances de l'Académie des sciences 1898;126:1703–1855.

[81] Mark P, Nilsson L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. J Phys Chem A 2001;105(43):9954–9960.

[82] Brehm M, Kirchner B. TRAVIS - A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories. J Chem Inf Model 2011;51(8):2007–2023. http://dx.doi.org/10.1021/ci200217w, pMID: 21761915.

[83] Egidi F, Trani F, Ballone PA, Barone V, Andreoni W. Low-Lying Electronic Excitations of a Water-Soluble BODIPY: from the Gas Phase to the Solvated Molecule 2016;135:264.

[84] Giovannini T, Del Frate G, Lafiosca P, Cappelli C. Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution. Phys Chem Chem Phys 2018;20:9181–9197.

[85] Giovannini T, Ambrosetti M, Cappelli C. A polarizable embedding approach to second harmonic generation (SHG) of molecular systems in aqueous solutions. Theor Chem Acc 2018 May;137(6):74.

[86] Le Bahers T, Adamo C, Ciofini I. A qualitative index of spatial extent in charge-transfer excitations. J Chem Theory Comput 2011;7(8):2498–2506.

[87] Guido CA, Cortona P, Mennucci B, Adamo C. On the metric of charge transfer molecular excitations: a simple chemical descriptor. J Chem Theory Comput 2013;9(7):3118–3126.

[88] Momeni MR, Brown A. Why do TD-DFT excitation energies of BODIPY/aza-BODIPY families largely deviate from experiment? Answers from electron correlated and multireference methods. J Chem Theory Comput 2015;11(6):2619–2632.

[89] Le Guennic B, Jacquemin D. Taking up the cyanine challenge with quantum tools. Acc Chem Res 2015;48(3):530–537.

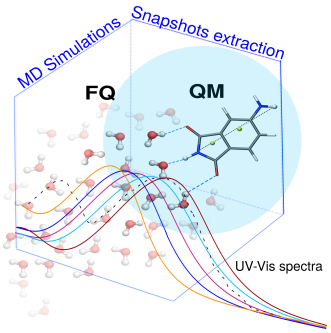
[90] Arroyo IJ, Hu R, Merino G, Tang BZ, Pena-Cabrera E. The smallest and one of the brightest. Efficient preparation and optical description of the parent borondipyrromethene system. J Org Chem 2009;74(15):5719–5722.

[91] Politzer IR, Crago KT, Kiel DL, Hampton T. The Effects of β-Cyclodextrin On the Fluorescence, Uv Absorption and Solubility of Selected Bimanes In Aqueous Solutions. Anal Lett 1989;22(6):1567–1580. http://dx.doi.org/10.1080/00032718908051621.

[92] Bucsiová L, Hrdlovič P. Medium Effect of Polymer Matrices on Spectral Properties of 4-Aminophthalimide and 4-Dimethylaminophthalimide. J Macromol Sci A 2007;44(9):1047–1053. http://dx.doi.org/10.1080/10601320701424495.

[93] Hillig KW, Morris MD. Pre-resonance Raman spectra of adriamycin. Biochem Bioph Res Comm 1976;71(4):1228 – 1233. http://www.sciencedirect.com/science/article/pii/0006291X76907853.

GRAPHICAL ABSTRACT



Continuum PCM and discrete QM/MM solvation methodologies (also including polarization effects modeled by means of a Fluctuating Charge approach) are challenged to reproduce excitation energies of seven organic dyes in aqueous solution.