On the Nature of Charge-Transfer Excitations for Molecules in Aqueous Solution: A Polarizable QM/MM Study

Franco Egidi[1], Giulia Lo Gerfo[1], Marina Macchiagodena[1], and Chiara Cappelli*[1]

[1]Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy.

April 27, 2018

Abstract

We illustrate the effect of solvation on the nature of electronic excitations of organic molecules which possess excited states of charge transfer character. The analysis is carried out using both a continuum model and a polarizable QM/MM method that treats the solvent atomistically and embeds each atom in the solvent with a fluctuating charge which responds to the solute quantum-mechanical electrostatic potential in a self-consistent manner. We also show how solvation dynamics can influence the nature of the excited state of molecular systems. The application of the model to aqueous solutions of doxorubicin and a substituted polythiophene derivative, shows that the solvent significantly affects the nature the excited states, which results in an enhanced or reduced charge transfer character as measured using two of the most popular indices for evaluating the distance traveled by the electrons upon excitation.

*chiara.cappelli@sns.it

1 Introduction

The study of electronic excitations for molecules in the condensed phase has long been an interest of computational chemists who wish to interpret and reproduce experimental data, as well as shed light into the mechanistic and physical aspects of the phenomena involved. Given the quantum mechanical (QM) nature of spectroscopic phenomena, particularly those involving the electronic degrees of freedom such as simple one-photon absorption and emission, a QM method must be employed to describe the system. Among the many possible choices the most common is by far time-dependent density functional theory (TDDFT) because of the method's favorable computational cost and many benchmark studies have been reported analyzing the quality of the available functionals for a wide variety of systems.[1,2] In addition to the choice of electronic structure method, a second aspect that should not be overlooked is that solvation effects can drastically affect the spectroscopic response of a system, particularly in the case of highly polar solvents such as water, therefore accurate methods to treat the solute-solvent interaction can be necessary. The most commonly used method to include solvation within QM calculation is the polarizable continuum model (PCM).[3,4] While continuum approaches can successfully model the electrostatics of solvation there are still a number of effects missing and, especially in the case of protic solvents such as water, continuum models may fail to capture the directionality of hydrogen bonds, providing with qualitatively or quantitatively inaccurate results. To address this problem, atomistic solvation models have also been explored. One possibility would be to rely on an electrostatic embedding model, in which the solvent's atoms possess fixed charges which polarize the solute.[5–12] In this case, however, the solvent only affects electronic excitations indirectly through a change in the ground state electrostatic density. In order for the solute-solvent interaction to carry over to response properties a mutual polarization between the two layers must be established, which can be realized through the use of methods that couple a QM description of the solute with a polarizable MM model for the solvent. Of the many different methods that have been described in the literature, we resort to the fluctuating charge model (FQ),[13,14] where each atom in the solvent is endowed with a charge that can change in response to the electrostatic potential of the solute and follows the electronegativity equalization principle.[15,16] This model, coupled with a QM description of the solute, has been successfully employed for the calculation of a variety of molecular properties for systems in aqueous solution,[17] such as Electronic Circular Dichroism,[18] Vibrational Circular Dichroism,[19] Raman and ROA spectroscopy,[20] and Optical Rotation.[21–24]

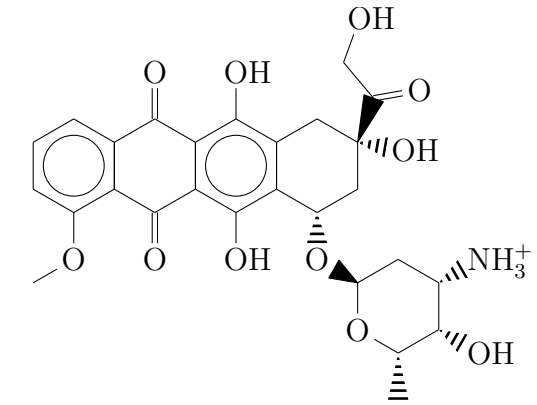
In this contribution we apply the QM/FQ model to a different type of spectroscopic problem, that of the modeling of charge transfer electronic transitions (CT). Molecular systems which present CT transitions are of great interest for their

chemical and technological applications, therefore a number of diagnostic tools for characterizing excited states obtained via TDDFT and readily assess the degree to which CT occurs were developed. An early development was that of Tozer et al.[25] who developed an index based on the absolute overlap of the molecular orbitals (MO) involved in the transition. Following their seminal work, Adamo et al.[26,27] further extended this work to the development of a dimensional index, also based on MOs. Finally, Ciofini et al.[28–30] developed an alternative diagnostic based on the excited state density. A thorough comparison between the aforementioned was also attempted by Adamo[31] showing their strengths and weaknesses.

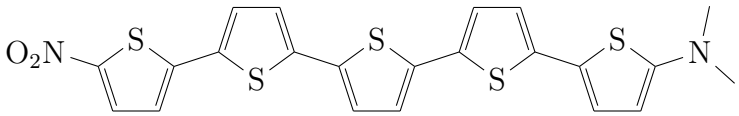
The effect of solvation on the nature of the excited states has been extensively studied using continuum models such as PCM,[32–35] which can be extremely useful in modeling bulk solvent effects, particularly for highly polar solvents, which can affect the CT states quite strongly by stabilizing the charge separation. Continuum models, however, lack the flexibility to include effects such as hydrogen bonding, particularly relevant in the case of water, as well as the effects related to the structural dynamics of flexible solutes. Compared with models that rely on a single-point calculation on fixed molecular structure, a QM/MM method has an additional layer of complexity arising from the underlying molecular dynamics simulation which samples the solute-solvent configurational space providing a large number of structures. The characteristics of the excited states of the solute may vary significantly from one structure to the other, however the picture that arises from the dynamics may provide with additional information. Such information may be in principle different, and complementary, with respect to what is obtained through a mean-field continuum description of the solvation phenomenon. The differences on the CT description arising from the two alternative solvation approaches have only received little attention in the previous literature, and this aspect constitutes the main motivation of the present paper. For our analysis we selected two systems in aqueous solution: doxorubicin, which is a common chemotherapeutic agent, and a substituted polythiophene, henceforth abbreviated as D3, shown in Figure 1. Both are highly conjugated organic molecules which can be easily excited to a charge-transfer state through one-photon absorption. The effect of the solvent modeled with the QM/FQ method upon the excited states of these systems will be analyzed through the use of two of the most common CT indices.

2 Methodology

The solvation method employed in this work relies on a classical molecular dynamics (MD) simulation, which we carry out using periodic boundary conditions, to sample the solute-solvent configurational space. From the simulation, a number



(a) Doxorubicin



(b) D3

Figure 1: Structures of the systems studied in this work.

of snapshots are extracted, and a spherical boundary surface around the solute is cut, retaining the solvent molecules within. A QM/FQ TDDFT calculation is then performed on each snapshot in order to evaluate the CT nature of the first few excited states of the molecule.

Two types of CT indices are compared in this work, one based on the molecular orbitals (MO) involved in the electronic transition by Guido et al.[26,27] henceforth denoted as Δr, and one based on the difference between the unrelaxed excited-state density and ground state density by Ciofini et al.[28–30] denoted as D_{CT}. Both are spacial indices which are designed to roughly indicate the charge displacement associated with the transition. Because the aim of the present work is to study the effect of the explicit solvation environment on these indices rather than provide a thorough analysis of the electronic excited states of the considered systems, we chose to evaluate both indices in their simplest form, i.e. we use Kohn-Sham orbitals rather natural transition orbitals (NTO) for the evaluation of the MO index, and forgo the use of the relaxed excited state density for the density-based index. As previously observed by Guido et al.[31] the unrelaxed density index provides a better correlation between the two indices, simplifying the analysis. Regarding the Δr index, we employ the definition provided in the original work,[26] rather than the slightly altered one that was later presented[31] (though whenever the de-excitations amplitudes are small, both versions yield the same results).

The QM/FQ Approach to Electronic Excited States

The FQ approach[36] is based on the concepts of atomic hardness and electronegativity, which can be rigorously defined within DFT in terms of functional derivatives.[15,37] The FQ model, which couples a polarizable MM environment to a QM description of the inner layer of the system in a multiscale QM/classical fashion, has successfully been exploited to evaluate different kinds of molecular properties and spectra by means of its extension to response theory and analytical derivatives[17–23,38–40], though it can be also exploited in molecular dynamics simulations.[41] In particular, if the QM/FQ is exploited in a time-independent fashion, the standard machinery of computational chemistry can be employed to calculate structural and spectroscopic properties.[18,21,38–40]

In the FQ approach the mutual polarization between the QM and classical portions of the system is represented in terms of a set of classical fluctuating charges whose value depends on the environment[13,14,42] according to the electronegativity equalization principle[15,43] The FQs $\mathbf{q}$, which are placed on the atoms in the classical moiety, are defined through the minimization of the following functional:

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

where α and β run on molecules, i, j on atoms and λ_{α} is a set of Lagrangian multipliers used to impose charge conservation constraints. Q_{α} is the total charge on molecule a. The $\boldsymbol{\chi}$ vector collects the atomic electronegativities, and the $\mathbf{J}$ matrix collects the interaction kernel elements between the FQs and can be expressed in terms of the Ohno functional form[44]:

$$J_{ii} = 2\eta_i, \qquad J_{ij} = \frac{\eta_{ij}}{[1+\eta_{ij}^2 r_{ij}^2]^{1/2}} \tag{2}$$

where η_i is the hardness of the i-th atom,

$$\eta_{ij} = \frac{\eta_i + \eta_j}{2} \tag{3}$$

is the average of the atomic hardnesses of atoms i and j and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between two atoms in the classical, atomistic portion of the system.

The stationarity condition of the $F(\mathbf{q},\boldsymbol{\lambda})$ functional (eq.1) is defined by solving the following equation[36]

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q \tag{4}$$

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

The interaction energy between the FQs and the QM density of charge ρ_{QM} is therefore:

$$E_{\mathrm{QM/MM}} = \sum_{i=1}^{N_q} \Phi[\rho_{QM}](\mathbf{r}_i) q_i \tag{5}$$

where $\Phi[\rho_{QM}](\mathbf{r}_i)$ is the electrostatic potential due to the QM density evaluated at the i-th FQ q_i, which is placed at point $\mathbf{r}_i$.

The Self Consistent Field (SCF) QM/FQ energy functional reads:

$$\mathcal{E}[\mathbf{P}, \mathbf{q}, \boldsymbol{\lambda}] = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{1}{2}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \mathbf{q}^\dagger \boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^\dagger \mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^\dagger \mathbf{q} + \mathbf{q}^\dagger \mathbf{V}(\mathbf{P}) \tag{6}$$

where $\mathbf{h}$ and $\mathbf{G}$ are the usual one- and two-electron matrices, and $\mathbf{P}$ is the QM density matrix. The stationarity conditions under the proper constraints yield a modified set of equations for the QM/FQ Fock matrix:

$$\tilde{F}_{\mu\nu} = \frac{\partial \mathcal{E}}{P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{q}^\dagger \mathbf{V}_{\mu\nu} \tag{7}$$

where μ, ν are atomic basis functions.

The FQs can therefore be obtained by solving the following equation:

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q - \mathbf{V}(\mathbf{P}) \tag{8}$$

Beyond the definition of the theoretical framework and the calculation of the energy, the extension of the computational approach to excited states requires an accurate description of the sample, i.e. a way to determine minimum energy structures. This is achieved through the definition of QM/FQ geometrical energy first (gradients) and second (vibrational frequencies) derivatives. Details on such extension can be found in the relevant literature papers.[17,38,39] In addition, the evaluation of excited state energies can be carried out by applying a linear response approach to the QM/FQ model. The Casida approach to the Time-Dependent Density Functional Theory (TD-DFT)[45] has been extended to the QM/FQ model.[38,46] The TD-DFT equations read

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \Omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{9}$$

where the eigenvalues Ω give vertical excitation energies, while the amplitudes for the single particle excitation and de-excitation are contained in the eigenvectors $\mathbf{X}$ and $\mathbf{Y}$, respectively. The response matrices $\mathbf{A}$ and $\mathbf{B}$ are defined as

$$\begin{aligned} A_{ai\sigma,bj\tau} &= (\epsilon_{a\sigma} - \epsilon_{i\sigma})\delta_{ab}\delta_{ij}\delta_{\sigma\tau} + \frac{\partial F^\sigma_{ai}}{\partial P^\tau_{bj}} \\ B_{ai\sigma,bj\tau} &= \frac{\partial F^\sigma_{ai}}{\partial P^\tau_{jb}} \end{aligned} \tag{10}$$

where i, j are occupied orbitals, a, b are unoccupied ones and σ and τ are spin labels. The usual procedure for the diagonalization of Eq. 9 involves the contraction of the combinations $(\mathbf{A}+\mathbf{B})$ and $(\mathbf{A}-\mathbf{B})$ with the $(\mathbf{X}+\mathbf{Y})$ and $(\mathbf{X}-\mathbf{Y})$ vectors, carried out in the atomic orbital basis.

The FQ contribution to the response matrices reads

$$f_{\mu\nu\sigma,\kappa\lambda\tau} = \frac{\partial v_{\mu\nu\sigma}}{\partial P_{\kappa\lambda\tau}} = \sum_{kl} \langle \mu\nu|k\rangle J^{-1}_{kl} \langle l|\kappa\lambda\rangle \tag{11}$$

where

$$v_{\mu\nu\sigma} = \sum_{k} \langle \mu\nu|k\rangle q_k, \tag{12}$$

Notice that the FQ contributions are purely electrostatic, i.e. they only contribute to the symmetric combination $(\mathbf{A}+\mathbf{B})$. Once the Casida equations are solved, excitation energies and transition amplitudes are obtained for the multiscale QM/FQ system.

3 Computational Details

All calculations were performed with a locally modified version of the Gaussian suite of quantum chemistry programs.[47] We used the CAM-B3LYP[48–50] functional with the 6-31+G* basis set for all DFT calculations. PCM calculations were performed according to Gaussian 16 standards. The MD simulation on doxorubicin was carried out using a previously described protocol.[51]

Classical MD simulations on the D3 system were carried out using GROMACS 4.6.5[52]. Simulations were performed on systems composed of 1 D3 and 5000 TIP3P-FB[53] water molecules. The D3 intramolecular force field is derived by fitting optimized energies, gradients, and Hessian matrices with the procedure implemented in JOYCE[54] as previously used for other flexible molecules.[55,56] Flexible dihedral angles, which allow the rotation of thiophene rings, are described either by a sum of cosine functions or with harmonic terms. In the first case, during the simulation, D3 selectively populates the different torsional angles, with the highest occurrence falling within the minimum-energy configurations; instead using harmonic functions the rotations of thiophene rings are forbidden. Both derived force fields were used in the MD simulations to verify the conformational effect on the nature of the excited states. Atomic partial charges were computed according to class IV CM5 charges[57] at the minimum energy configuration, whereas Lennard-Jones parameters were transferred from the standard OPLS/AA FF.[58] Initially a steepest descent energy minimization was applied. Then the systems were heated up to 298.15 K for 200 ps (using the velocity-rescale thermostat[59] with a coupling

constant of 0.1 ps). Starting from the last configuration obtained from the NVT equilibration, a simulation in the NPT ensemble (using the Parrinello-Rahman barostat[60]) was performed and the system was allowed to converge to uniform density. An additional simulation was run in the NVT ensemble using the average box dimensions obtained from the NPT simulation and increasing the integration time step from 0.2 to 2.0 fs, fixing the fastest degrees of freedom with the LINCS algorithm.[61] The sampling simulations were carried out for 100 ns with $\delta t = 2.0$ fs. Electrostatic interactions were evaluated using the particle-mesh Ewald (PME)[62] method with a grid spacing of 1.2 Å and a spline interpolation of order 4.

For both systems, 200 snapshots were extracted from the MD and a sphere of radius 20 Å was cut around the center of the molecule, and solvent molecules laying outside the sphere were discarded. Subsequent QM/FQ calculations performed on the snapshots employed the SPC FQ parameters given by Rick et al.[13,14]

4 Results

4.1 Doxorubicin

The first system we analyze is doxorubicin, an organic molecule commonly employed in cancer therapy[63] that owes its effectiveness to its ability to intercalate DNA.[64] Because of its broad interest as a chemotherapeutic agent, the spectroscopic properties of this system have been amply studied,[65–68] in particular as they relate to changes in the electronic and structural characteristics underwent by this drug as it leaves its aqueous environment to penetrate the helical frame of the nucleic acid. The understanding of the spectroscopic properties of this system and its interaction with a complex environment must begin with a complete description of its properties as it exists in aqueous solution. In a previous study by some of us,[51] the resonance Raman spectroscopic signature of doxorubicin in water was extensively analyzed using both continuum and discrete models for the solvent. We now move our attention to the electronic structure of this molecule, in particular on the effect of the solvent upon the charge-transfer nature of its excited states.

As seen in figure 1a, doxorubicin consists of a substituted anthraquinone moiety, whose highly conjugated π system allows for extensive mobility of the electron density upon excitation. The gas-phase geometry optimization of this system reveals the existence of three stable isomers, depicted in figure 2, which differ mainly in the orientation the OH groups.[51] Computing the TDDFT excited states on this structures reveals that the first and third excited states both have a strong CT character as evidenced from visual inspection of the molecular orbitals involved in the transitions, which are predominantly HOMO→LUMO and HOMO-1→LUMO (see figure 3), respectively, and are all localized on the anthraquinone moiety.

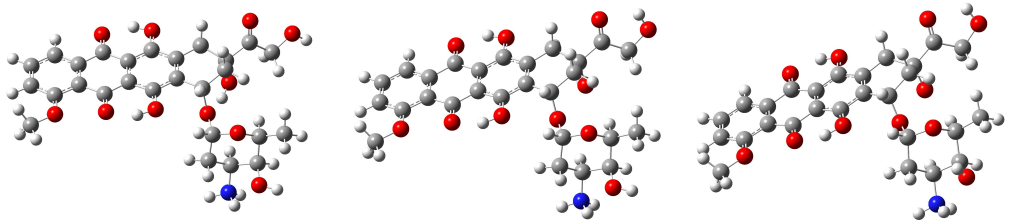
Upon excitations to these states the electron density moves from either side of the molecule toward the central ring. Table 1 reports the values of the two CT indices for the system in vacuo for each conformer and the first three excited states. Though the occupied orbitals involved in the transitions might look very similar for two CT states, being very similar only on either side of the ring structure, the CT indices are actually quite different, with the S_3 state having a significantly higher CT character. The explanation can be visually appreciated by plotting the centroids of the positive and negative density difference for the three states, depicted in figure 4, which are much farther apart for the third excited state, and almost overlap for the second. It can finally be noticed that the three conformers yield very similar CT indices, owing to the fact that the conformational difference is isolated in the orientation of the side chains and only affects the anthraquinone moiety indirectly.

The question is, however, whether the presence of the solvation environment affects the nature of these states and, most importantly, how. To study the effect of an implicit solvation description we have re-optimized the molecular geometries and evaluated the CT indices in the presence of the polarizable continuum. The results are shown in the bottom three rows of table 1. For the S_1 and S_3 states, which present the largest CT character in vacuo, we can see that for all three conformations the effect of the solvent is rather limited; for S_1 we observe a slight increase in CT character whereas for S_3 the trend is different for the two indices, though differences are small. For S_2 we see a much more pronounced increase in CT character for the conformers II and III, particularly when evaluated by means of the Δr index. By looking at the MO contributions to the transition we observe that for the PCM result there are significant contributions from transitions out of occupied orbitals localized on the side chain, which were absent for the molecule in the gas phase. For instance, for the third conformer the S_2 state has a 49% HOMO-4$\rightarrow$LUMO and a 36% HOMO-3$\rightarrow$LUMO contribution where the occupied orbitals are shown in figure 5 (the LUMO orbital is almost unchanged with respect to the gas phase). These orbitals can be described as linear combinations of the non-bonding orbitals of the carbonyl oxygens on both the anthraquinone and side chain, leading to the increase of CT character.

These results suggest that the intrusion from states localized on the side chain can indeed strongly affect the nature of the excitations for this molecule, therefore a more dynamical description that explore the entire conformational space of this system in solution can lead to much more realistic results. In addition, the strong contributions from non-bonding orbitals localized on the carbonyl oxygens suggests that the presence of hydrogen bonds can have a significant effect on the excitations. We have therefore computed the two CT indices using the QM/FQ scheme, which can accurately capture both these effects. In Figure 6 we report how the two

indices vary as the classical MD unfolds for the three lowest-energy excited-states of each of the 200 snapshots. Several considerations can be drawn from the results. We first focus our attention to the lowest-energy state alone. We can see that there is a strong variation for both indices, with Δr varying between 0.6-2.6 Å and D_{CT} between 0.7-2.5. The instantaneous solvation environment therefore appears to strongly affect the nature of the excitation. This variation notwithstanding, the nature of the transition in terms of the orbitals involved remains unchanged, with the HOMO-LUMO contribution varying between 78% and 98%, though of course the exact nature of such orbitals is also affected by the solvent. What this implies is that the first excited state is sufficiently separated from the others to minimize any mixing between excited states that would be brought by the solvent perturbation. This is not true for any of the other states, for which we observe a heavy degree of mixing as evidenced by the orbital contributions to the computed transition amplitudes. This is reflected in the CT indices, which show massive variations across the snapshots. The second excited state can no longer be uniformly classified as a local excitations given that both indices assume values that would place it within the range of a CT excitation, a finding that was also reflected in the PCM results. This behavior is expected as the highly polar environment provided by water can significantly stabilize CT states. Meanwhile, the third excited state shows the greatest variation where for some snapshots it can be classified as a local excitation, while in others the Δr index can be as high as 4 Å. To better appreciate this last point, Figure 7 shows the correlation between the two indices for each excited state. For some of the snapshots the two indices are in complete disagreement as to the nature of the state. Though an in depth comparison of these indices goes beyond the scope of this work it is still interesting to look at an example where the discrepancy is particularly pronounced. We take for instance one snapshot for which the D_{CT} and Δr indices of the S_3 state are 0.38 Å and 3.74 Å, respectively. In this snapshot this state is mainly a HOMO-3→LUMO and HOMO-2→LUMO transition, where these two occupied orbitals are shown in figure 9. The peculiarity of these orbitals is the significant contribution arising from the oxane side chain, which was not observed in either the gas phase or PCM results. A more in-depth analysis of the result can be reached by plotting the actual distribution for the occurrence of each of the two index within a certain interval, which we have reported in Figure 8. It can be seen that for the first two states both indices present a markedly peaked distribution, though D_{CT} is much more skewed towards lower values. For the third excited state, the Δr distributions maintains its symmetric peaked shape, albeit with a larger variance with respect to the other states, meanwhile D_{CT} presents a distribution that is almost flat, and therefore the index has lost its descriptive value. These preliminary results suggest that Δr is a more stable

descriptor when applied to solvated structures originating from a flexible molecular dynamics. Regardless of which index is selected, however, these results show just how much the micro-solvation environment can affect the electronic degrees of freedom of a solute. Remarkably, for this system the chromophore is quite rigid, therefore the observed variation is only due to the instantaneous solute-solvent spacial arrangement, however in more complex systems structural changes in the solute itself can play a large part. This latter effect has been explored in the second molecule present in this study.



(a) I **(b)** II **(c)** III

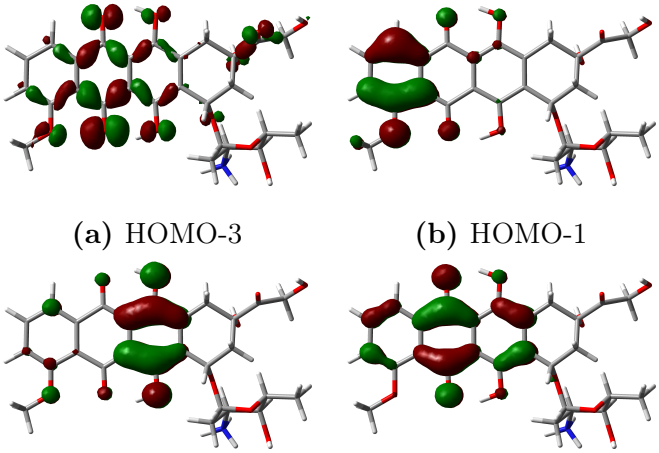
Figure 2: Conformers of doxorubicin.

Table 1: D_{CT} and Δr indices for the three Doxorubicin conformers in vacuo and PCM (in Å).

		I		II		III	
	State	D_{CT}	Δr	D_{CT}	Δr	D_{CT}	Δr
vacuo	S_1	1.72	1.43	1.73	1.44	1.71	1.42
	S_2	0.64	1.41	0.77	1.53	0.73	1.56
	S_3	2.71	2.26	2.69	2.22	2.68	2.21
water	S_1	1.82	1.54	1.85	1.52	1.84	1.51
	S_2	0.90	1.62	0.95	3.10	0.96	3.12
	S_3	2.61	2.41	2.57	2.30	2.59	2.27

4.2 D3

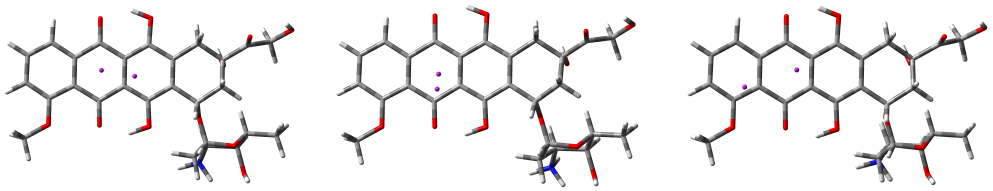
Compared with doxorubicin, for which the considered excitations are localized on the rigid ring structure, D3 presents a much larger conjugated structure which allows for a larger charge displacement upon excitation. This can be observed in Table 2 where the CT indices are reported for the isolated molecule. All reported states can be classified as CT excitations and present values of the CT indices that are significantly higher than those of doxorubicin. The table also shows the



(a) HOMO-3 **(b)** HOMO-1

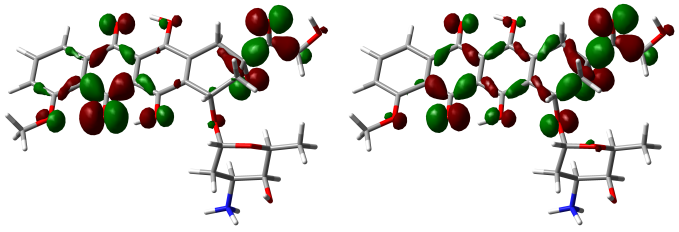
(c) HOMO **(d)** LUMO

Figure 3: Molecular orbitals of doxorubicin III in vacuo.



(a) S_1 **(b)** S_2 **(c)** S_3

Figure 4: Positive and negative difference density centroids for the first three excitations of doxorubicin (purple dots).



(a) HOMO-4 **(b)** HOMO-3

Figure 5: Molecular orbitals of doxorubicin III in water.

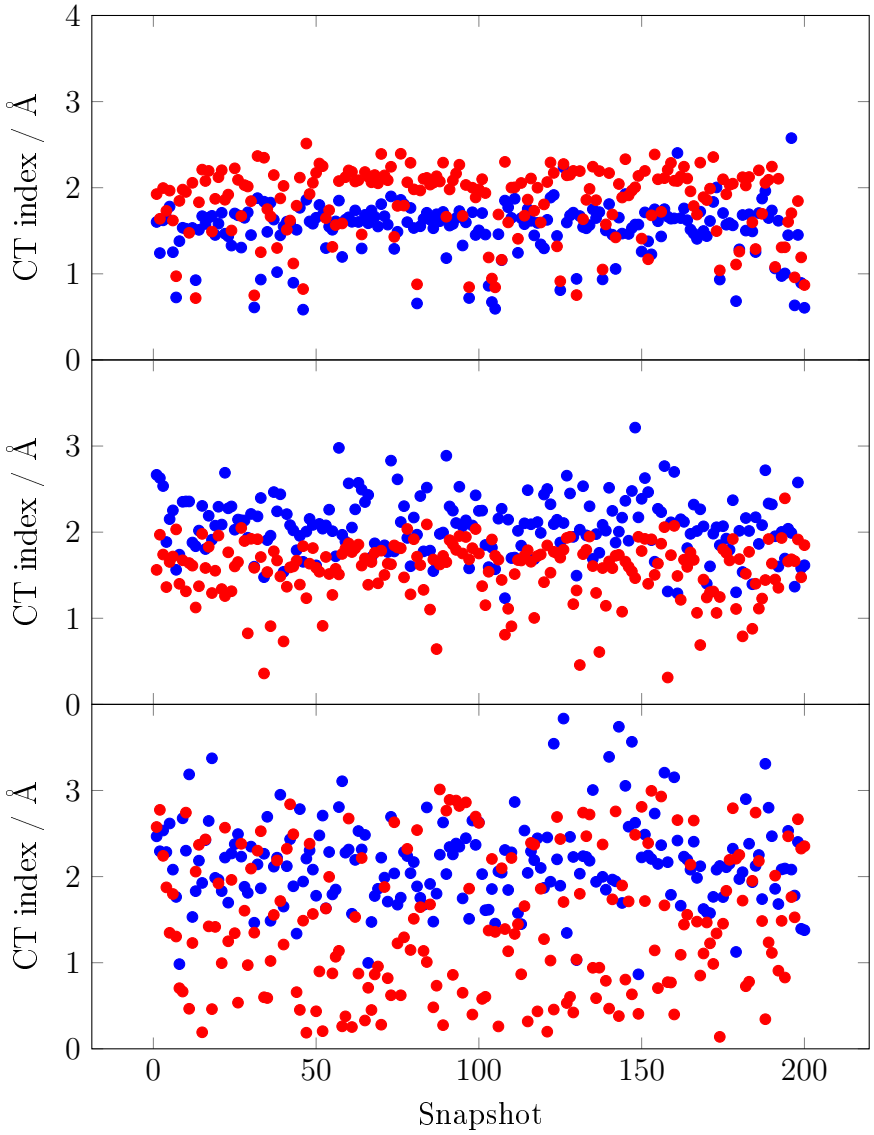
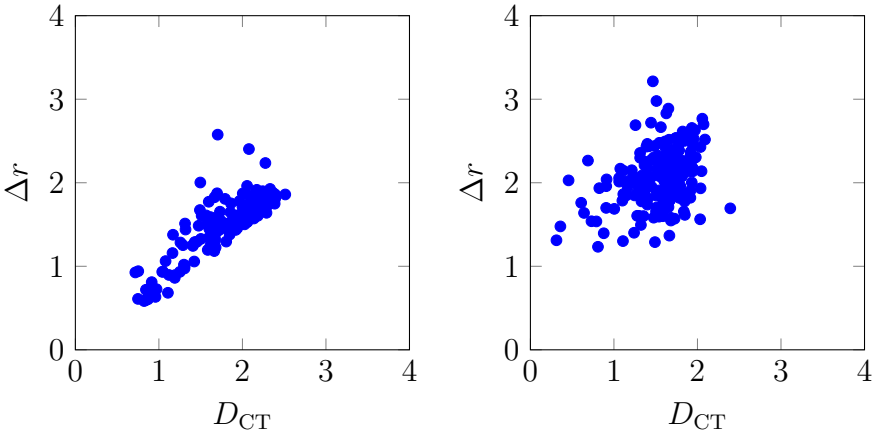
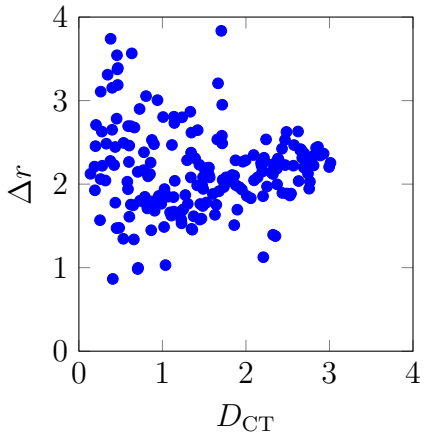


Figure 6: QM/FQ CT indices of doxorubicin for each MD snapshot, for the first (top panel), second (middle panel), and third (bottom panel) excited states. D_{CT} indices are in red, Δr are in blue.



(a) S_1 **(b)** S_2



(c) S_3

Figure 7: Comparison between the two indices for doxorubicin.

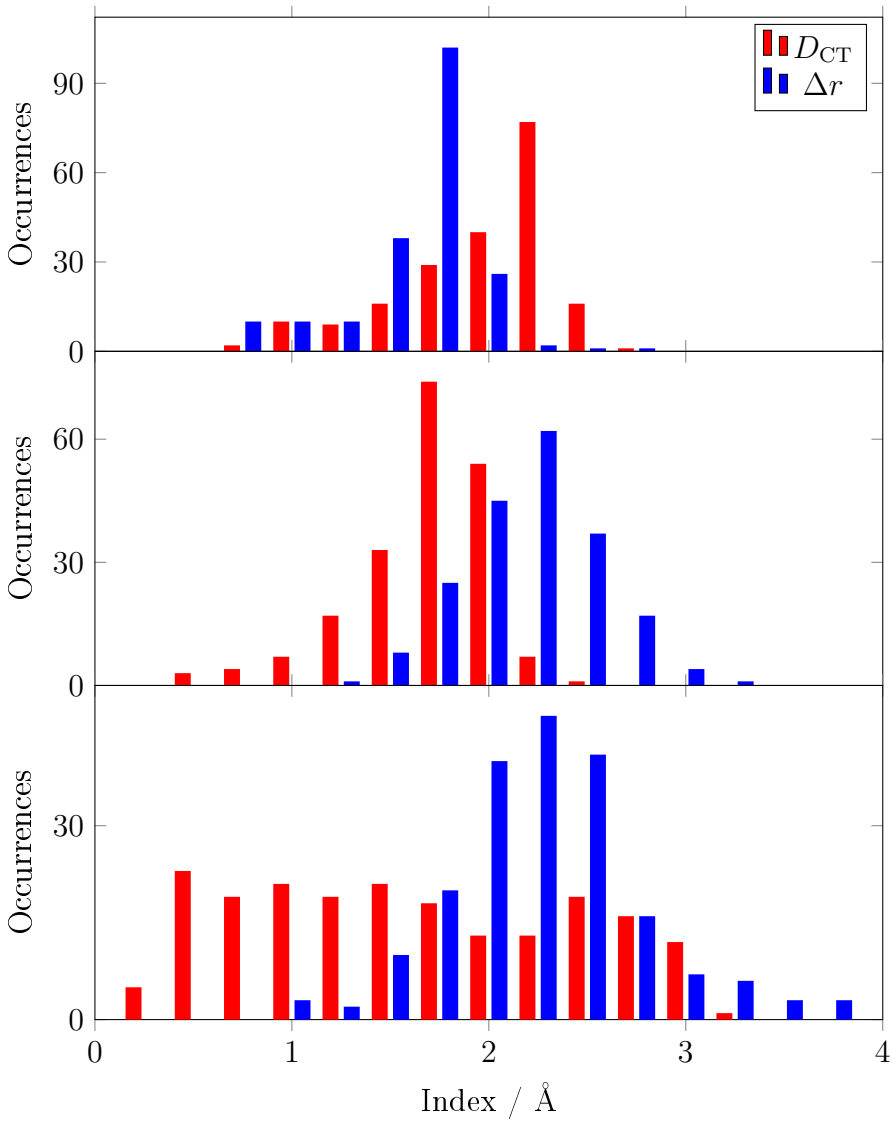
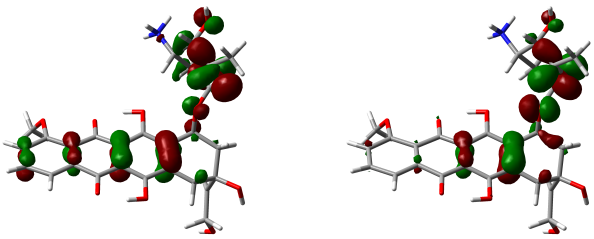


Figure 8: Distribution of the occurrencies of the CT indices computed with the QM/FQ method on the 200 snapshots from the MD, for the S_1 state (top panel), S_2 state (middle panel), and S_3 state (bottom panel).



(a) HOMO-3 **(b)** HOMO-2

Figure 9: Molecular orbitals of one of the snapshots from the doxorubicin/water MD.

orbital contribution to each transition, while the orbitals are shown in Figure 10. Most of the CT character comes from the HOMO-LUMO transition which shows a movement of the electronic density from the amino to the nitro side of the molecule, which explains the higher indices obtained for the first and third excited states.

The introduction of an implicit solvent via the PCM does not significantly affect either the CT indices or the decomposition of the states in terms of orbital transitions. Compared to doxorubicin this molecule shows a smaller variation in its equilibrium geometry and the absence of bulky side chains able to contribute to the transition means that there are no intrusions from electrons outside the conjugated π structure. It should be noted, however, that while both the vacuo and PCM equilibrium geometries are almost perfectly planar, the rotational barrier for two adjacent thiophenes is low enough that the equilibrium structure cannot be representative of the true system as it exists in solution.

The complete exploration of the conformational space of this solute/solvent couple can be achieved through an MD and specific solvent effects introduced via the FQ model. As the thiophene rings reorient themselves during the dynamics, the highly conjugated π structure is broken as the molecule reaches a *cis* configuration. This conformational effect alone may affect the nature of the excited state, in addition to the effect of the solvent itself. In order to disentangle these two contributions we performed three different MD simulations: one where the solute structure is frozen, one employing a harmonic force field for the dihedral angles between the thiophene rings which does not permit a full rotation between the rings, and finally a dynamics where the solute is flexible and therefore allowed to explore the entire conformational space. Figure 11 shows the result for the first excited state. The picture that emerges shows that, overall, there is only a small difference between keeping the structure of the molecule rigid and allowing it to vibrate harmonically, with the average values for both indices varying very little. A significant reduction in CT character is instead observed when the molecule is allowed complete conformational freedom, owing to the more significant breaking

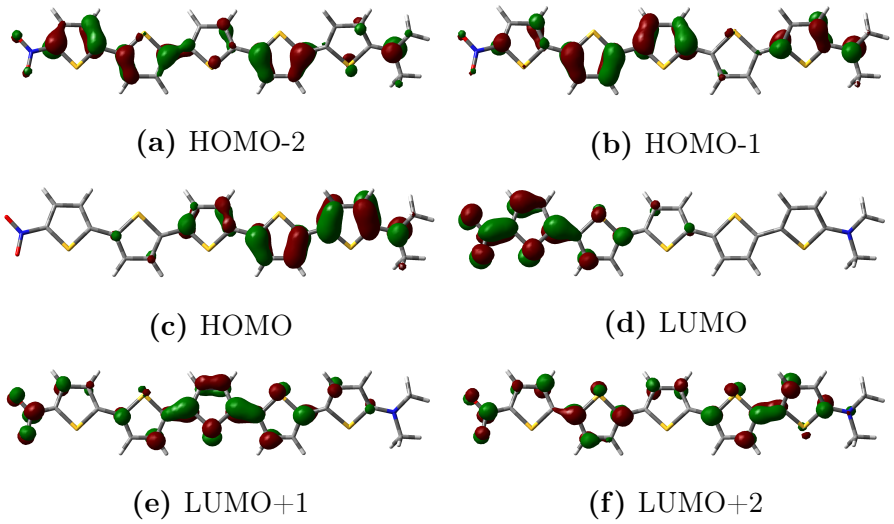
of the aromaticity. Figures 12 and 13 show the same results for the second and third excited state. We note that though the same observations made for the first excited state still apply, the overall picture is much different. In particular, for the second excited state some snapshots present much higher CT character, owing to the fact that as energy increases there is a higher chance of mixing with higher-energy states. Such snapshots can be readily identified from visual inspection of the resulting plots for the rigid and harmonic molecule, however this changes in the case of a flexible molecule, for which the results are much more spread out. Figure 14 reports the distribution of the indices across all simulations for the three excited states. This figure shows immediately shows that freeing all degrees of freedom of the solute in the dynamics significantly increases the spread in the charge-transfer character, particularly for S_2 and S_3. Contrary to what was observed in the case of doxorubicin, the D_{CT} index produces a unimodal distribution in all cases. Given the fact that the two indices often produce different results, it is best to calculate both to obtain a clearer picture regarding the nature of the excitations. Unless the relaxed density for each state is needed, both indices originate from the same set of data, therefore calculating both does not significantly increase the computational cost. It should finally be emphasized that, given the observed high variability of both indices, the picture that emerges is that the solvated system should be regarded as an ensemble of molecules with vastly different electronic properties, particularly when considering the charge-transfer nature of the excited states.

Table 2: D_{CT} and Δr indices for the three D3 in vacuo and PCM (in Å).

	State	D_{CT}	Δr	Character
vacuo	S_1	6.35	9.19	H-1→L(21%) H→L(45%) H→L+1(24%)
	S_2	5.03	6.58	H-1→L(26%) H→L+1(43%) H→L+2(16%)
	S_3	13.80	9.41	H-1→L(26%) H→L(47%) H-2→L(9%)
water	S_1	6.97	9.68	H-1→L(27%) H→L(47%) H→L+1(18%)
	S_2	5.57	6.32	H-1→L(23%) H→L+1(49%) H→L+2(14%)
	S_3	14.41	9.68	H-1→L(24%) H→L(49%) H-2→L(8%)

5 Conclusions

In this work we have shown how solvation dynamics can influence the nature of the excited state of molecular systems. We carried our analysis using both a continuum model as well as the QM/FQ method, a polarizable QM/MM method that treats the solvent atomistically and embeds each atom in the solvent with a fluctuating charge which responds to the solute QM electrostatic potential in a self-consistent manner. Thanks to the self-consistent nature of this model, solvation



(a) HOMO-2 **(b)** HOMO-1

(c) HOMO **(d)** LUMO

(e) LUMO+1 **(f)** LUMO+2

Figure 10: Molecular orbitals of D3 in vacuo.

effects propagate explicitly to the quantum mechanical response equations that determine the excitation energies and electronic transition densities, allowing for explicit solvent effects to be incorporated into the model.

The results show that the solvent can affect significantly the nature the excited states causing strong changes in the solute's electronic structure which result in an enhanced or reduced charge transfer character as measured using two of the most popular indices for evaluating the distance traveled by the electrons upon excitation. Though changes in the electronic structure of the solute can be significant, particularly for high-energy states, large variations are also observed due to the increased conformational freedom of the system in solution. For the chosen systems, in fact, a completely static description that would be implied by single points calculations performed on either the system in vacuo or with PCM would yield an unsatisfactory description of the real molecule in solution. The use of a polarizable QM/MM model such as FQ can greatly improve upon the more common static method afforded by PCM, though it carries the additional computational cost of having to perform a classical dynamics followed by a large number of QM/FQ calculations. Thanks to the increasing availability of powerful computer clusters, however, these types of calculations are nowadays feasible for most systems, therefore the use of QM/MM methods over continuum models should be encouraged, particularly in cases such as those explored in this paper where an implicit solvent cannot capture the chemistry of the system, even qualitatively.

One issue that has not been addressed in this work is the problem of the solvation regime. All calculations presented here, both using the continuum and

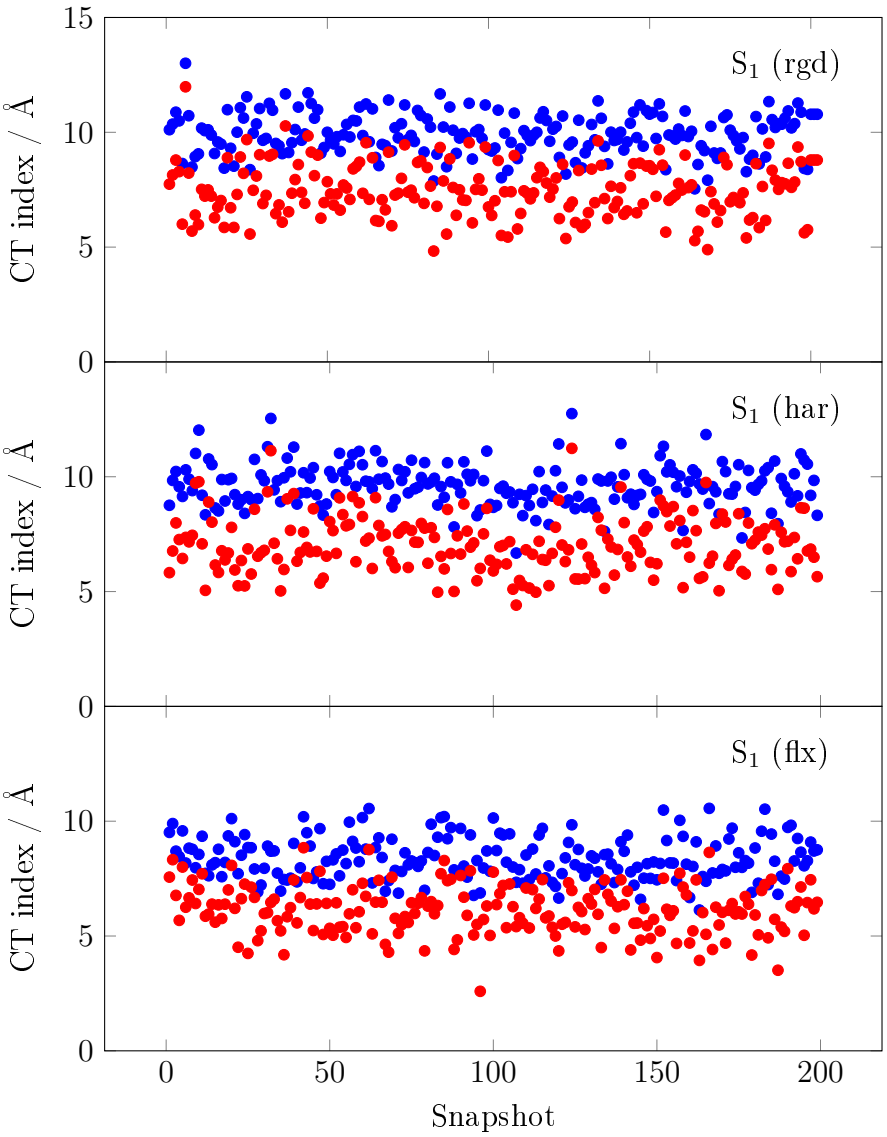


Figure 11: QM/FQ CT indices of D3 S_1 state for each MD snapshot, for the rigid molecule (top panel), harmonic force field (middle panel), and flexible molecule (bottom panel) models. D_{CT} indices are in red, Δr are in blue.

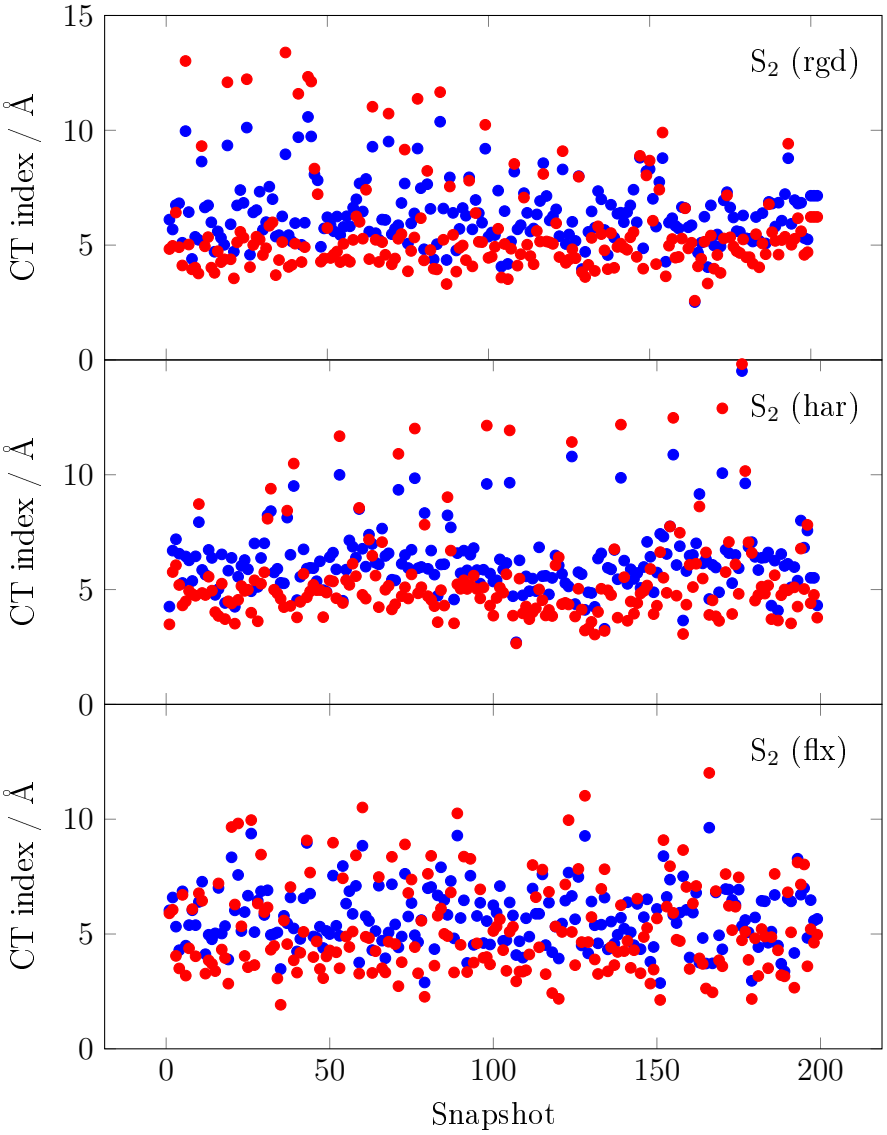


Figure 12: QM/FQ CT indices of D3 S_2 state for each MD snapshot, for the rigid molecule (top panel), harmonic force field (middle panel), and flexible molecule (bottom panel) models. D_{CT} indices are in red, Δr are in blue.

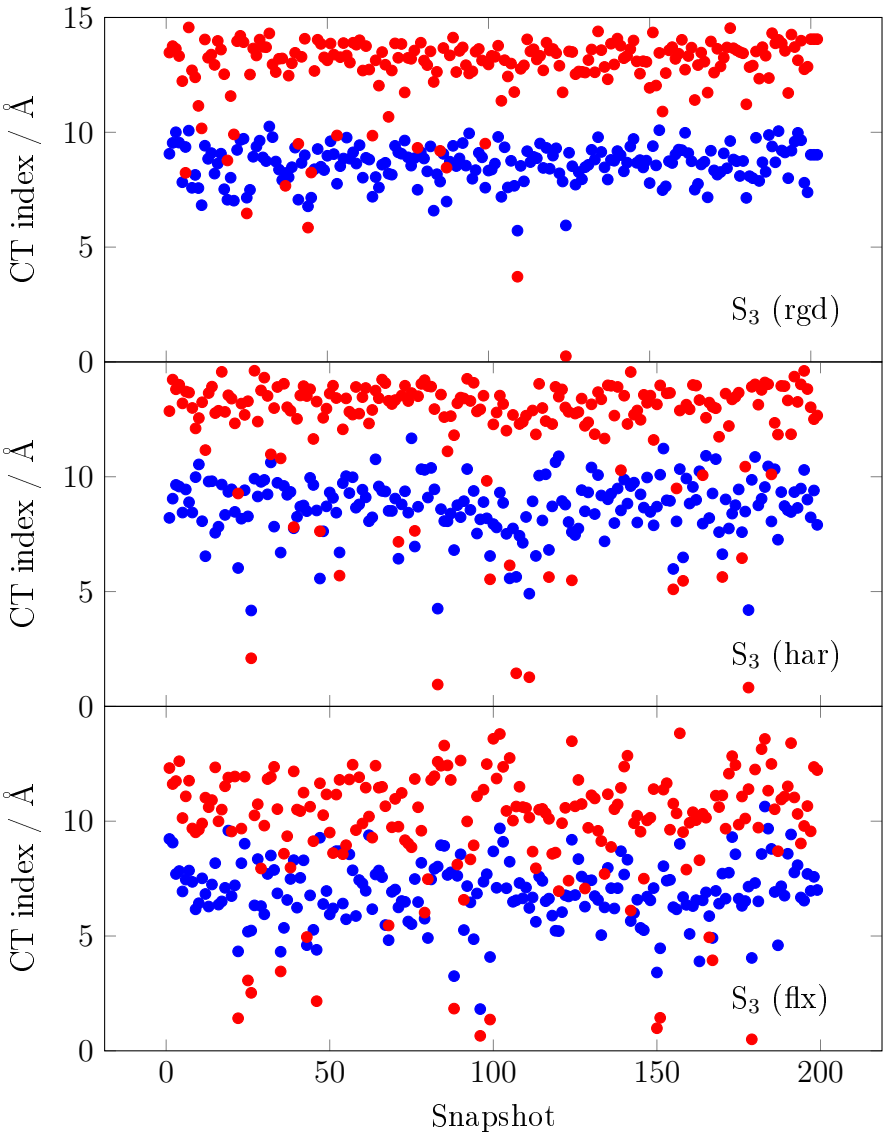


Figure 13: QM/FQ CT indices of D3 S_3 state for each MD snapshot, for the rigid molecule (top panel), harmonic force field (middle panel), and flexible molecule (bottom panel) models. D_{CT} indices are in red, Δr are in blue.

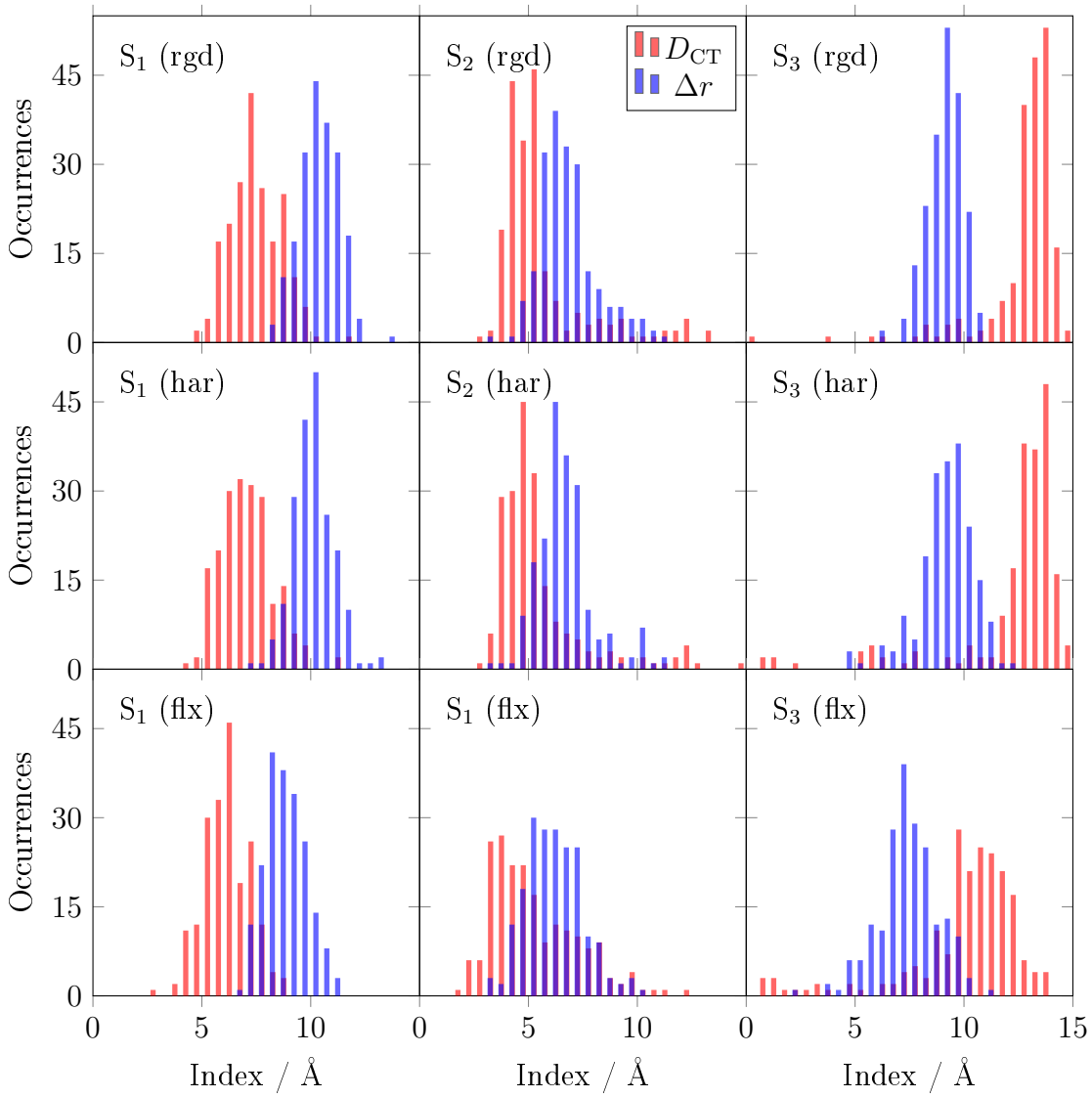


Figure 14: Distribution of the occurrencies of the CT indices computed with the QM/FQ method on the 200 snapshots from the MD of the D3 molecule. The labels S_1-S_3 indicate the first three excited states, and the three types of MDs are denoted as *rgd* (rigid), *har* (harmonic force field), and *flx* (fully flexible molecule). D_{CT} indices are in red, Δr are in blue.

.

discrete model, rely on the linear response formalism which is based on the ground-to-excited state transition density and fails to change in the solvent response due to the relaxation of the excited state density. Methods to properly include relaxation effects into the solvent response have been developed both in the case of continuum[35,69–72] and explicit[73–75] models, however this has never been done for the QM/FQ method. The proper accounting of solvent relaxation through the fluctuating charges has the potential to greatly improve calculated results, particularly for CT transitions. Also, the QM/MM model which has been exploited in this study only focuses on solute-solvent electrostatic effects. Non-electrostatic (repulsion/dispersion) interaction may play a relevant role in describing CT excitations. A model to include them in the calculation of ground-state energies within QM/MM approaches has been recently proposed by some of us[76], and will be extended to the calculation of excitation energies in the future.

Acknowledgments

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

References

[1] Jacquemin, D.; Wathelet, V.; Perpéte, E. A.; Adamo, C. *J. Chem. Theory Comput.* **2009**, *5*, 2420–2435.

[2] Leang, S. S.; Zahariev, F.; Gordon, M. S. *J. Chem. Phys.* **2012**, *136*, 104101.

[3] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3093.

[4] Mennucci, B. *WIREs Comput. Mol. Sci.* **2012**, *2*, 386–404.

[5] Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.

[6] Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.

[7] Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298–305.

[8] Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389–427.

[9] Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185–199.

[10] Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

[11] Egidi, F.; Trani, F.; Ballone, P. A.; Barone, V.; Andreoni, W. *Theor. Chem. Acc.* **2016**, *135*, 264.

[12] Monari, A.; Rivail, J.-L.; Assfeld, X. *Acc. Chem. Res.* **2012**, *46*, 596–603.

[13] Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141–6156.

[14] Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1996**, *118*, 672–679.

[15] Mortier, W. J.; van Genechten, K.; Gasteiger, J. *J. Am. Chem. Soc.* **1985**, *107*, 829–835.

[16] Sanderson, R. T. *Science* **1951**, *114*, 670–672.

[17] Cappelli, C. *Int. J. Quant. Chem.* **2016**, *116*, 1532–1542.

[18] Egidi, F.; Russo, R.; Carnimeo, I.; D'Urso, A.; Mancini, G.; Cappelli, C. *J. Phys. Chem. A* **2015**, *119*, 5396–5404.

[19] Giovannini, T.; Olszówka, M.; Cappelli, C. *J. Chem. Theory Comput.* **2016**, *12*, 5483–5492.

[20] Giovannini, T.; Olszówka, M.; Egidi, F.; Cheeseman, J. R.; Scalmani, G.; Cappelli, C. *J. Chem. Theory Comput.* **2017**, *13*, 4421–4435.

[21] Lipparini, F.; Cappelli, C.; Barone, V. *J. Chem. Phys.* **2013**, *138*, 234108.

[22] Lipparini, F.; Egidi, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2013**, *9*, 1880–1884.

[23] Egidi, F.; Carnimeo, I.; Cappelli, C. *Opt. Mater. Express* **2015**, *5*, 196–209.

[24] Egidi, F.; Bloino, J.; Barone, V.; Cappelli, C. *J. Chem. Theory Comput.* **2012**, *8*, 585–597.

[25] Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.

[26] Guido, C. A.; Cortona, P.; Mennucci, B.; Adamo, C. *J. Chem. Theory Comput.* **2013**, *9*, 3118–3126.

[27] Guido, C. A.; Cortona, P.; Adamo, C. *J. Chem. Phys.* **2014**, *140*, 104101.

[28] Le Bahers, T.; Adamo, C.; Ciofini, I. *J. Chem. Theory Comput.* **2011**, *7*, 2498–2506.

[29] Ciofini, I.; Le Bahers, T.; Adamo, C.; Odobe, F.; Jacquemin, D. *J. Phys. Chem. C* **2012**, *116*, 11946–11955.

[30] Adamo, C.; Le Bahers, T.; Savarese, M.; Wilbraham, L.; García, G.; Fukuda, R.; Ehara, M.; Rega, N.; Ciofini, I. *Coordin. Chem. Rev.* **2015**, *304*, 166–178.

[31] Savarese, M.; Guido, C. A.; Brémond, E.; Ciofini, I.; Adamo, C. *J. Phys. Chem. A* **2017**, *121*, 7543–7549.

[32] Mennucci, B.; Cammi, R.; Tomasi, J. *J. Chem. Phys.* **1998**, *109*, 2798–2807.

[33] Mennucci, B.; Cappelli, C.; Guido, C. A.; Cammi, R.; Tomasi, J. *J. Phys. Chem. A* **2009**, *113*, 3009–3020.

[34] Nguyen, P. D.; Ding, F.; Fischer, S. A.; Liang, W.; Li, X. *J. Phys. Chem. Lett.* **2012**, *3*, 2898–2904.

[35] Guido, C. A.; Jacquemin, D.; Adamo, C.; Mennucci, B. *J. Chem. Theory Comput.* **2015**, *11*, 5782–5790.

[36] Lipparini, F.; Barone, V. *J. Chem. Theory Comput.* **2011**, *7*, 3711–3724.

[37] Chelli, R.; Procacci, P. *J. Chem. Phys.* **2002**, *117*, 9175–9189.

[38] Lipparini, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*, 4153–4165.

[39] Lipparini, F.; Cappelli, C.; Scalmani, G.; Mitri, N. D.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*, 4270–4278.

[40] Caricato, M.; Lipparini, F.; Scalmani, G.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2013**, *9*, 3035–3042.

[41] Mancini, G.; Brancato, G.; Barone, V. *J. Chem. Theory Comput.* **2014**, *10*, 1150–1163.

[42] Rick, S. W.; Bader, S. J. S. J. S.; Berne, B. J. *J. Mol. Liq.* **1995**, *65-66*, 31–40.

[43] Anthony K. Rappe, W. A. G. *J. Phys. Chem.* **1991**, *95*, 3358–3363.

[44] Ohno, K. *Theoret. Chim. Acta* **1964**, *2*, 219–227.

[45] Casida, M. E. In *Recent Advances in Density Functional Methods Part I*; Chong, D. P., Ed.; World Scientific, Singapore, 1995; pp 155–192.

[46] Carnimeo, I.; Cappelli, C.; Barone, V. *J. Comput. Chem.* **2015**, *36*, 2271–2290.

[47] Frisch, M. J. et al. Gaussian 16 Revision A.03. 2016; Gaussian Inc. Wallingford CT.

[48] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[49] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

[50] Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.

[51] Olszówka, M.; Russo, R.; Mancini, G.; Cappelli, C. *Theor. Chem. Acc.* **2016**, *135*, 27.

[52] Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics* **2013**, *29*, 845–854.

[53] Wang, L.-P.; Martinez, T. J.; Pande, V. S. *The Journal of Physical Chemistry Letters* **2014**, *5*, 1885–1891.

[54] Barone, V.; Cacelli, I.; De Mitri, N.; Licari, D.; Monti, S.; Prampolini, G. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3736–3751.

[55] Macchiagodena, M.; Del Frate, G.; Brancato, G.; Chandramouli, B.; Mancini, G.; Barone, V. *Physical Chemistry Chemical Physics* **2017**, *19*, 30590–30602.

[56] Macchiagodena, M.; Mancini, G.; Pagliai, M.; Barone, V. *Phys. Chem. Chem. Phys.* **2016**, *18*, 25342–25354.

[57] Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2012**, *8*, 527–541.

[58] Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.

[59] Bussi, G.; Donadio, D.; Parrinello, M. *The Journal of Chemical Physics* **2007**, *126*, 014101.

[60] Parrinello, M.; Rahman, A. *The Journal of Chemical Physics* **1982**, *76*, 2662–2666.

[61] Hess, B.; Bekker, H.; Berendsen, H.; Fraaije, J. *Journal of Computational Chemistry* **1997**, *18*, 1463–1472.

[62] Darden, T.; York, D.; Pedersen, L. *The Journal of Chemical Physics* **1993**, *98*, 10089–10092.

[63] Arcamone, F.; Cassinelli, G.; Fantini, G.; Grein, A.; Orezzi, P.; Pol, C.; Spalla, C. *Biotechnol. Bioeng.* **1969**, *11*, 1101–1110.

[64] Kersten, H.; Kersten, W. *Inhibitors of Nucleic Acid Synthesis*; Springer: New York, 1974.

[65] Airoldi, M.; Barone, G.; Gennaro, G.; Giuliani, A. M.; Giustini, M. *Biochemistry* **2014**, *53*, 2197–2207.

[66] Fiallo, M. M. L.; Tayeb, H.; Suarato, A.; Garnier-Suillerot, A. *J. Pharm. Sci.* **1998**, *87*, 967–975.

[67] Manfait, M.; Alix, A. J.; Jeannesson, P.; Jardillier, J.-C.; Theophanides, T. *Nucleic Acids Res.* **1982**, *10*, 3803–3816.

[68] Lee, C. J.; Kang, J. S.; Kim, M. S.; Lee, K. P.; Lee, M. S. *Bull. Korean Chem. Soc.* **2004**, *25*, 1211–1216.

[69] Caricato, M.; Mennucci, B.; Tomasi, J.; Ingrosso, F.; Cammi, R.; Corni, S.; Scalmani, G. *J. Chem. Phys.* **2006**, *124*, 124520.

[70] Corni, S.; Cammi, R.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2005**, *123*, 134512.

[71] Improta, R.; Barone, V.; Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2006**, *125*, 054103.

[72] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; Guido, C. A.; Mennucci, B.; Scalmani, G.; Frisch, M. J. *Chem. Sci.* **2011**, *2*, 2143–2161.

[73] Zeng, Q.; Liang, W. *J. Chem. Phys.* **2015**, *143*, 134104.

[74] Guido, C. A.; Scalmani, G.; Mennucci, B.; Jacquemin, D. *J. Chem. Phys.* **2017**, *146*, 204106.

[75] Loco, D.; 'Etienne Polack,; Caprasecca, S.; Lagardère, L.; Lipparini, F.; Piquemal, J.-P.; Mennucci, B. *J. Chem. Theory Comput.* **2016**, *12*, 3654–3661.

[76] Giovannini, T.; Lafiosca, P.; Cappelli, C. *J. Chem. Theory Comput.* **2017**, *13*, 4854–4870.