

A Δ SCF model for excited states within a polarizable embedding

Michele Nottoli,^a Patrizia Mazzeo,^{a,b} Filippo Lipparini,^a Lorenzo Cupellini,^a Benedetta Mennucci^{a*}

^a Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via G. Moruzzi 13, 56124 Pisa, Italy, ^b Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

ABSTRACT

Hybrid TDDFT/MM approaches are very popular methods for describing electronic transitions of molecules in solution or embedded in more complex (bio)matrices. However, when combined with a polarizable force field some problems can appear depending on the type of environment response scheme that is used. In particular, specific a posteriori corrections are generally needed to accurately describe charge-transfer states implying a large reorganization of the electron density in the excited state. Here, we present a possible strategy to solve this issue introducing a Δ SCF formulation. As the Δ SCF strategy has the advantage of being intrinsically state specific, its coupling to a polarizable model is expected to be particularly suited to describe all cases where the standard, linear response, formulation of a polarizable TDDFT/MM approach is not sufficient.

KEYWORDS

QM/MM, polarizable embedding, excited states, solvatochromism, electron-transfer

1. Introduction

Linear-response time-dependent DFT (from now on TDDFT) is a very widely employed approach for obtaining electronic transition energies and densities in a computationally efficient way. TDDFT obtains transition energies and properties via the linear response (LR) of a ground state DFT solution to time-dependent electric fields. TDDFT however has some well-known shortcomings. First of all, the method is limited to describing single excitations, making it impossible to model doubly (or higher) excited states. Moreover, the LR formulation becomes less and less accurate with increasing difference between excited and ground state densities, which occurs for example with charge transfer (CT) transitions. In these cases, the results of TDDFT become extremely dependent on the selected DFT functional.

A further problematic aspect of TDDFT arises when the method is coupled to a polarizable classical model to account for environment effects. Since the very first formulations using continuum solvation models, it appeared clearly that the non-linear character of the QM/classical Hamiltonian generated by the density-dependent term due to the polarizable embedding introduces a theoretical problem not present in the standard TDDFT formulation [1, 2]. A solution that is consistent with the linear-response (LR) nature of the TDDFT approach was proposed and is generally known as LR scheme. Within such a framework, the response of the polarizable environment to the excitation follows the field generated by the transition density that characterizes the excitation, rather than the excited state density. Because of this characteristic, the LR response has been also classified as an excited-state “dispersion” or “resonance” contribution [2, 3]. Successively the LR scheme was extended to polarizable TDDFT/MM formulations [4, 5].

The LR scheme is extremely effective, as it naturally merges within the TDDFT equations, but it has the limit of accounting only for one component of the environment effect. In fact, in the excitation process, the electron density will change, and this change should be reflected in the

polarization of the environment. However, this effect is neglected in the LR scheme, which can lead to a non-negligible error in the estimation of the transition energies especially for excitations involving a large change in the densities, such as a CT. Different approximate models have been proposed to introduce this additional state-specific (SS) response within a TDDFT framework. Once more, the first proposals came from continuum solvation models: among them, one of the most successful schemes thanks of its simplicity and computational efficiency is the so called corrected LR (cLR) proposed by Caricato et al. [6]. The approach is based on a first-order perturbative scheme and introduces an a posteriori correction to the TDDFT energies using the polarization of the environment due to the difference between excited and ground state densities. The cLR approach was later extended to polarizable QM/MM descriptions [5, 7]. The cLR correction represents a physically separate environment contribution to the transition energy, and can be added to the LR energy, in the so-called cLR² approach [8].

An alternative strategy for introducing SS effects is to abandon the TDDFT formulation and to use state-specific orbital optimized DFT approaches [9]. The idea beyond this class of methods is that one can converge the self-consistent field (SCF) procedure to an excited state, and obtain the excitation energies as the difference between the excited-state and ground-state SCF energies. For this reason, this approach is also known as Δ SCF. An excited-state SCF solution can be obtained by using a non-Aufbau ad-hoc criterion, such as choosing the orbitals that maximize the overlap with the ones that were occupied at the previous step, as done in the Maximum Overlap Method (MOM) [10, 11], or with those of the guess determinant, as done in the Initial Maximum Overlap Method (IMOM) [12]. These methods have been shown to give accurate excitation energies not only for low-lying valence excitations [13] but also for some doubly excited states [14, 15], low-lying charge transfer states [9, 16], and Rydberg states [17]. Recently, Δ SCF methods have been applied to the calculation of electronic couplings [18], to the calculations of singlet-triplet energy gaps in solution [16], and to nonadiabatic excited-state dynamics in condensed phase [19, 20].

As the Δ SCF strategy has the advantage of being intrinsically state specific, its coupling to a polarizable model automatically includes the effects of a SS response of the environment.

In this paper we present a comparative analysis of this approach with respect to the different TDDFT response schemes (LR and cLR²).

Here we employ the AMOEBA polarizable force field [21, 22] which uses an induced-dipole formulation to account for the polarization of the MM atoms. However, our conclusions can be extended to other polarizable QM/MM methods [4, 7, 23–25].

We consider three distinct scenarios for the description of excited states. We first focus on the valence excitation in three green fluorescent protein (GFP) mutants. Then, we focus our attention to the dual emission from a locally excited (LE) state and an intramolecular CT (ICT) state of the dimethylaminobenzonitrile (DMABN) molecule in solution, and finally we consider an intermolecular CT state in the AppA photosensory protein. By analyzing these scenarios, we highlight the strengths and weaknesses of the Δ SCF and TDDFT approaches in describing excitations within a polarizable environment.

2. Methods

In this section, we briefly review the main aspects of the DFT/AMOEBA and its extension to excited states both within a TDDFT and Δ SCF formulation. For TDDFT, both LR and cLR² schemes will be presented.

In the AMOEBA FF, each MM atom bears static multipoles up to the quadrupole and an isotropic polarizability, which in the presence of an electric field gives rise to an induced dipole. A peculiarity of AMOEBA is that the MM contribution to the electric field that induces the dipoles, E_d , is different from the one that appears in the definition of the AMOEBA polarization energy, E_p , which makes the force field non variational. In particular, the two fields E_d, E_p differ for the exclusion rules used in their definition, being them based on polarization group for E_d and on the connectivity for E_p , respectively. The coupling of the AMOEBA FF to a SCF-based

description can therefore be achieved through a polarization Lagrangian [5, 26–28]

$$\begin{aligned} \mathcal{L}(P, \mu_d, \mu_p) = & E^{\text{QM}}(P) + E^{\text{self}}(M) + \langle q, V(P) \rangle - \langle \mu_s, E(P) \rangle + \langle \Theta, G(P) \rangle \\ & - \frac{1}{2} \langle \mu_d, E_p + E(P) \rangle + \frac{1}{2} \langle \mu_p, T\mu_d - E_d - E(P) \rangle. \end{aligned} \quad (1)$$

In eq. 1, P is the SCF density matrix, q , μ_s , and Θ are the static charges, dipoles and quadrupoles that make the multipolar distribution, μ_d are the induced dipoles, μ_p a set of Lagrange multipliers that enforce the polarization equations $T\mu_d = E_d + E(P)$, $V(P)$, $E(P)$, and $G(P)$ are the potential, field and field gradient produced by the quantum mechanical density of charge, and T is the polarization matrix, that has the inverse polarizabilities on the diagonal block and (damped) dipole-dipole interaction tensors on the off-diagonal ones. The brackets are shorthand notation for the scalar product. All the detailed expressions for the aforementioned quantities can be found in ref. 5. A thorough, complete discussion of the Lagrangian formalism can be found in ref. [28].

The coupled QM/AMOEBA equations are easily obtained by imposing the stationarity of the Lagrangian in eq. 1 with respect to the induced dipoles μ_d , to the associated Lagrange multiplier μ_p , and, under the usual orthonormality constraints, to the density matrix. The resulting equations are

$$\begin{aligned} FC &= SCE, \quad P = CC^\dagger, \\ T\mu_d &= E_d + E(P), \\ T\mu_p &= E_p + E(P), \end{aligned} \quad (2)$$

where the elements of the effective Fock (or Kohn-Sham) matrix F are given by

$$F_{\mu\nu} = F_{\mu\nu}^{\text{vac}} + \langle q, V_{\mu\nu} \rangle - \langle \mu_s, E_{\mu\nu} \rangle + \langle \Theta, G_{\mu\nu} \rangle - \frac{1}{2} \langle \mu_p + \mu_d, E_{\mu\nu} \rangle. \quad (3)$$

In eq. 3, F^{vac} is the usual Fock or KS matrix, while $V_{\mu\nu}$, $E_{\mu\nu}$, and $G_{\mu\nu}$ are one-electron electrostatic potential, field and field gradient integrals, respectively. As the effective Fock matrix depends on the induced dipoles while at the same time the induced dipoles depend on the electronic density, mutual polarization is included when the coupled equations are solved self-consistently. This in turn requires one to solve, at each SCF iteration, the polarization equations and the corresponding linear equations for the Lagrange multipliers, which, for large systems, can be an overwhelming computational burden. In our implementation, thanks to the use of an efficient preconditioned conjugate gradient iterative solver [29, 30] and the Fast Multipole Method to compute matrix-vector products [31–33], all the operations introduced by the polarizable embedding are performed at a linear-scaling computational cost, which makes it possible to treat very large systems with a modest computational overhead.

The DFT/AMOEBA formulation has been extended to obtain transition energies and transition densities using the TDDFT formulation [5, 34].

Using Casida’s formalism [35] and assuming that real orbitals are used, the TDDFT/AMOEBA equations are given by

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}, \quad (4)$$

where the A and B matrices are given by

$$\begin{aligned} A_{ia,jb} &= \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + (ia|jb) - c_x(ab|ij) + c_l f_{aibj}^{xc} + \mathcal{V}_{iajb}^{\text{env}}, \\ B_{ia,jb} &= (ia|jb) - c_x(aj|ib) + c_l f_{aibj}^{xc} + \mathcal{V}_{iajb}^{\text{env}}. \end{aligned} \quad (5)$$

In eq. 5, we use the indices i, j to refer to occupied molecular orbitals (MO), and a, b to refer to virtual one. ϵ are orbital energies, $(pq|rs)$ two-electron integrals in the MO basis in the Mulliken notation, and f_{aibj}^{xc} is the exchange–correlation kernel. The coefficients c_x and c_l control the amount of exact exchange for hybrid functionals by interpolating between HF ($c_x = 1, c_l = 0$) and pure functionals ($c_x = 0, c_l = 1$).

The use of a polarizable embedding such as AMOEBA affects the TDDFT equations in two ways. The first effect is visible in the molecular orbitals and their energies which are solutions of the effective Fock operator reported in eq.3. Moreover, a new term appears in A and B matrices, namely [4, 5]:

$$\mathcal{V}_{iajb}^{\text{env}} = -\langle \mu_{jb}, E_{ia} \rangle, \quad (6)$$

where μ_{jb} are the AMOEBA dipoles induced by the field arising from orbitals $\phi_j\phi_b$, and E_{ia} is the field arising from $\phi_i\phi_a$ orbitals. At convergence of Casida’s equations, this LR environment term gives rise to an excitation energy contribution of the type [36]

$$\Delta\omega^{LR} = -\langle \mu^{X+Y}, E^{X+Y} \rangle, \quad (7)$$

where E_k^{X+Y} is the electric field generated by the *transition density* ($X + Y$) at the polarizable site k , and μ_k^{X+Y} is the k -th dipole induced by the same field. Such LR term can be interpreted as the instantaneous response of the AMOEBA polarizable sites to the transition density of the electronic excitation. In the literature, this term has also been described as a dispersion-like interaction [2]. Being proportional to the square of the transition density, for dark CT excitations the LR term is negligible.

A polarizable environment presents an additional type of response, namely the one due to the change in the electronic charge densities of the ground and excited states. This further effect is commonly described as a state-specific response [2, 37]. Within TDDFT, the SS response can be added as a posteriori correction to the transition energies following an approach based on first-order perturbation theory. The resulting corrected LR approach was originally proposed for continuum solvation models [6] but later it has been extended to polarizable QM/MM descriptions [5, 7]. Within AMOEBA, the cLR correction to the excitation energy reads:

$$\Delta\omega^{\text{cLR}} = -\frac{1}{2}\langle \mu^\Delta, E^\Delta \rangle, \quad (8)$$

where E_k^Δ is the electric field generated by the change in the electronic charge densities of the ground and excited states at the polarizable site k and μ^Δ are the induced dipoles generated by the same field.

As the LR and the SS responses describe two different physical interactions, they should be simultaneously taken into account in order to get a complete description of the environment effect on an electronic transition. The most straightforward way of doing that is to sum the cLR correction 8 correction to the TDDFT/AMOEBA excitation energy obtained by solving eq.4. This computational protocol has been proposed for continuum solvation models and called cLR² [8], but it can equivalently be applied to a QM/AMOEBA description.

In order to quantify the effect of the LR response on excitation energies, it is useful to compute excitation energies after removing the $\mathcal{V}_{iajb}^{\text{env}}$ terms from the TDDFT equations (4). The resulting excitation energy contains the effect of induced dipoles in the orbital energies only, and we denote it as ω_0 .

Computing excitation energies using a Δ SCF approach in combination with AMOEBA has the advantage of being intrinsically state specific. In fact, each electronic state is variationally obtained, together with the proper polarization. This means that the Δ SCF excitation energies automatically include the effects of a SS response of the environment. In contrast to the cLR, which is a first-order correction, here the SS response is fully self consistent. On the other hand, the Δ SCF excitation energies completely neglect the LR effect.

3. Results

In this section we report three applications of the Δ SCF/AMOEBAs approach and its comparison with three different TDDFT/AMOEBAs schemes. These schemes correspond to the LR and cLR² models described in the Methods section along with the “ ω_0 ” excitation energy obtained by solving TDDFT equations without the explicit AMOEBA term, V_{iajb}^{env} reported in eq.6.

The three selected applications refer to (a) vertical excitations in three different mutants of the green fluorescent protein (GFP), (b) the emission process of DMABN in acetonitrile solution and (c) a photo-induced intermolecular electron transfer between tyrosine and flavine in the AppA protein (see fig. 1).

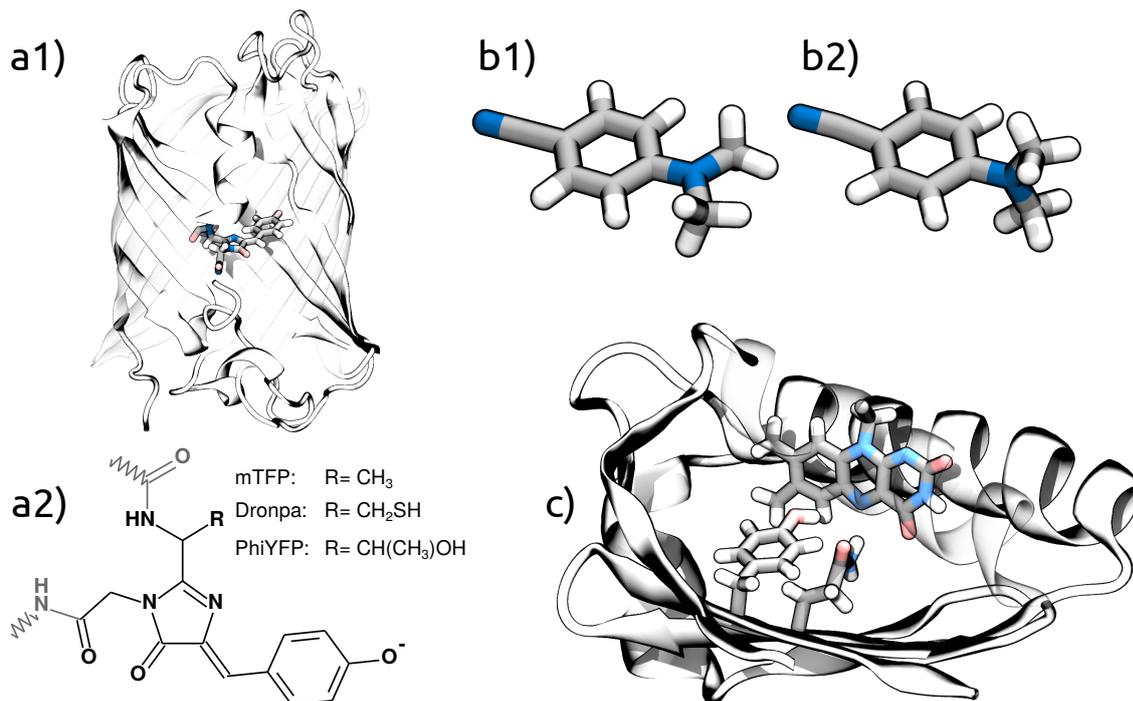


Figure 1. General structure of GFP (a1) and molecular structures of the chromophores (a2) of the three investigated mutants — mTFP, Dronpa and PhiYFP. The QM part includes also the atoms reported in gray. Structures of DMABN in the locally excited (b1) and intramolecular charge transfer (b2) excited states, these respectively have the dimethylamino moiety coplanar and twisted with respect to the phenyl ring. Illustration of the AppA protein (c). The flavine chromophore and the tyrosine and glutamine residues which are included in the QM part are highlighted.

3.1. Computational details

To characterize the excited states of the selected systems, we performed a series of calculations using two hybrid (B3LYP and PBE0) and two long-range corrected (CAM-B3LYP, ω B97XD) functionals combined with the 6-311+G(d,p) basis set and with an environment described using the AMOEBA force field [21, 22].

The Δ SCF is computed as the difference between the regular SCF energy and the IMOM SCF energy [12]. In both cases, the AMOEBA induced dipoles are calculated self-consistently with the specific state. IMOM calculations are performed on top of a guess density in which one α electron is promoted to the LUMO. For what concerns the IMOM calculations, we used a decreased convergence threshold for the SCF density of 10^{-6} in the root-mean square norm of the density difference between two subsequent iterations, which is justified by the fact that in these cases we only use the value of the SCF energy, without performing further post-SCF calculations. A more conservative 10^{-8} SCF threshold has been used for the TDDFT calculations.

To characterize the solvatochromic shifts, we computed reference values in vacuum. Since

ω_0 , LR and cLR2 on the one side and Δ SCF on the other side correspond to two different QM levels of theory, two different reference values are needed: a TDDFT calculation in vacuum for the first set of models and a Δ SCF calculation in vacuum.

Finally, to further investigate the origin of the solvatochromic shifts, we calculated the electrostatic potential due to the induced dipoles at the QM atoms for the different molecules. This analysis has been performed only with the CAM-B3LYP functional.

All the calculations have been performed with a locally modified version of the Gaussian code [38].

3.2. Absorption in Mutants of the Green Fluorescent Protein

In this section, we present a study of the light absorption in three different mutants of the green fluorescent protein (GFP).

The structures of the three GFP mutants, namely mTFP, Dronpa and PhiYFP, were taken from previous works [39, 40]. Here, we included an AMOEBA shell of around ~ 20 Å; the exact radius of the shell for each mutant was chosen to have a global neutral system (charge of the QM part -1, charge of the MM part +1). The resulting values are 20.9 Å, 7376 MM atoms for mtFP, 25.5 Å, 11574 MM atoms for Dronpa and 20.0 Å, 6855 MM atoms for PhiYFP [40]. The protein tertiary structure is similar among the three mutants but the molecular structure of the chromophore presents some differences (see figure 1a).

The investigated excitations of the three GFPs present a strong HOMO–LUMO π – π^* component, therefore the guesses used for the IMOM calculations were obtained by promoting one α electron from the HOMO to the LUMO.

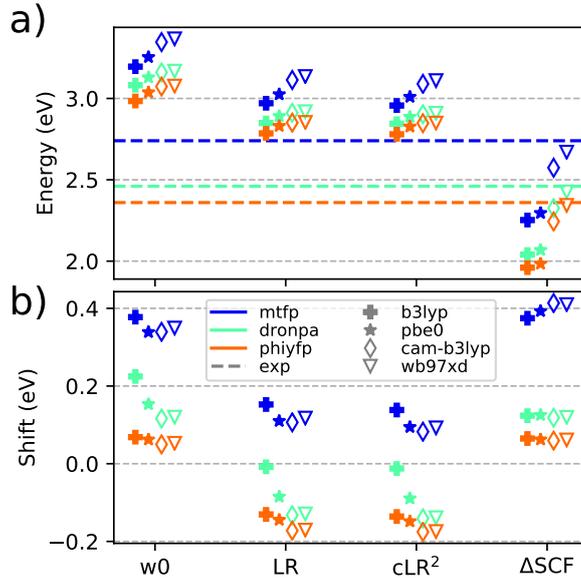


Figure 2. a): GFP excitation energies (in eV) for the three investigated systems — mTFP, Dronpa and PhiYFP — reported using different colors. The horizontal dotted lines indicate the measured excitation energies[41–43]. b): GFP solvatochromic shifts for the three investigated systems. The reference values are a vacuum TDDFT calculation for ω_0 , LR and cLR values, and an in vacuum Δ SCF calculation for the Δ SCF values. In both figures the four investigated functionals are reported using different markers: long-range corrected functionals are indicated with empty markers.

Figure 2 reports the excitation energies (top panel) computed for the three different mutants with the four selected DFT functionals and the different models. We also analyze the solvatochromic shifts (bottom panel) in order to remove the dependence of the absolute transition energy on the functional or method.

The three TDDFT sets of energies show a relatively weak sensitivity to the selected functional for all the three mutants. Conversely, the Δ SCF description shows significant higher

energies when long-range corrected functionals are used. This is in contrast to most findings, which suggest a weaker dependence of Δ SCF transition energies on the functional [9, 15, 44].

As expected on the basis of the π - π^* nature of the excitation, which has a strong transition dipole, the TDDFT-LR energies are significantly lower than the ω_0 ones. On the contrary, adding a SS correction does not further affect the energies: cLR² and LR values are almost indistinguishable. For the same reason, we do not expect Δ SCF to be an accurate approach for this excitation as it includes only a SS response. This is confirmed in the solvathocromic shift, where the Δ SCF results are very similar to the ones obtained using a ω_0 description. However, by cancellation of errors, the absolute Δ SCF transition energies obtained with long-range corrected functionals are very close to the experiment.

Figure 3 reports plots of the values of electrostatic potential generated by the AMOEBA induced dipoles at the chromophore atoms of the three GFP mutants. The potential from induced dipoles depends on the inducing QM density, so we distinguish the dipoles induced by the ground state density of each mutant (first column) and the ones induced by the excited state density, calculated either at TDDFT or IMOM level (second and third column, respectively).

As it can be observed, the three systems present a qualitatively similar picture but with some not negligible differences which can be related to the different solvatochromic shifts. We also note that moving from the GS induced dipoles to the ES ones virtually no differences are observed, meaning that the ground and excited state densities are very similar. This explains why the SS contributions to the transition energies are negligible.

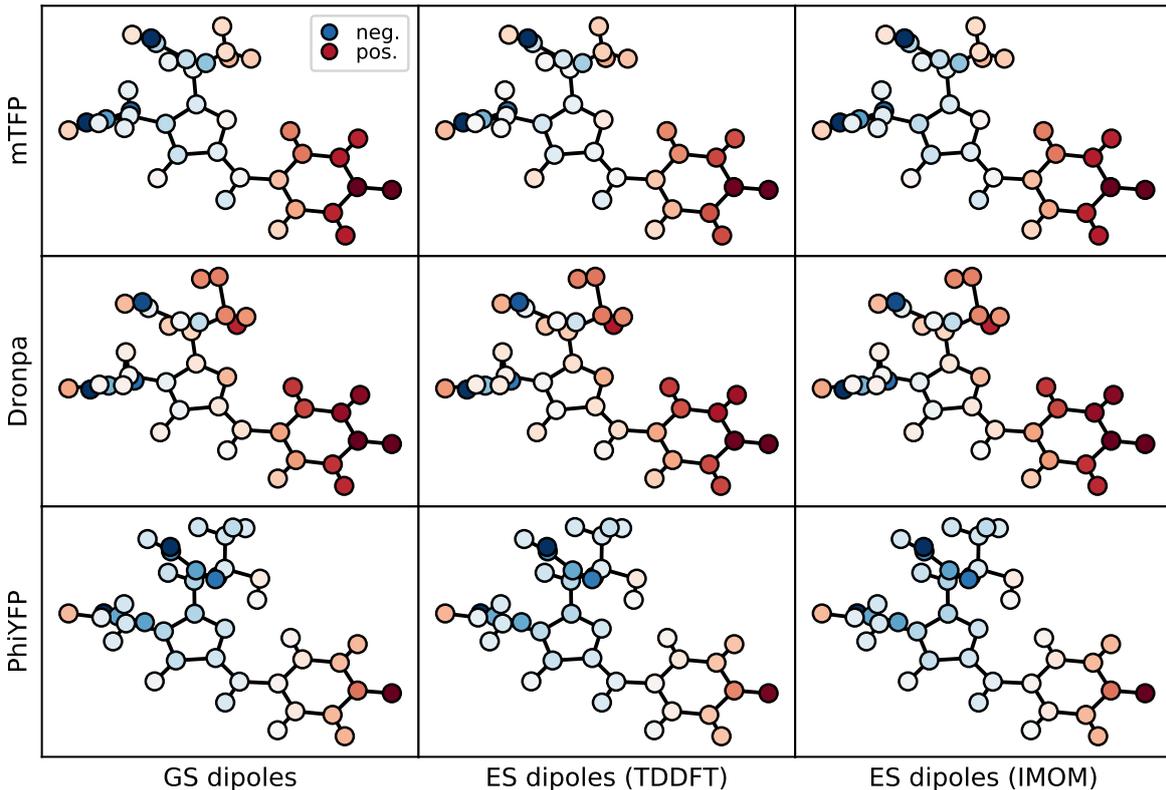


Figure 3. Electrostatic potential from the environment at the chromophore atoms for the three mutants. The potential reported is from the dipoles induced by the GS density (first column), by the ES density computed using TDDFT (second column) and by the ES density computed using IMOM (third column). We present only results obtained with the CAM-B3LYP functional. The color map is in arbitrary units, red values are positive, blue values are negative. All the plots are normalized to the same value.

3.3. Emission of DMABN

In this section, we focus on the emission from two excited states — a locally excited (LE) state and an intermolecular charge transfer (ICT) state — for DMABN in acetonitrile solution. The molecular structures of the two excited states are reported in figure 1b. We note that the two structures present a dimethylamino moiety which is twisted in the ICT state, and coplanar with the phenyl ring in the LE state.

The solute-solvent configurations were taken from 6 excited-state QM/AMOEBA Born-Oppenheimer MD trajectories presented in a previous work [34]. Those trajectories describe the adiabatic interconversion from the LE to the ICT states of DMABN. From each trajectory, we selected 5 frames in the 50–75 fs time window after the initial excitation and 5 frames in the 5000-5025 fs window. In the first set of frames, the solute is still in the LE state vertically reached from the ground state, with little vibrational relaxation. In the second set of frames, instead the solute has converted into the ICT state, and it features a twisted geometry. In addition, the acetonitrile solvent has reoriented to solvate the ICT state. All the reported values are the average over 30 frames for each state, and in the plots, the error bars report the standard error.

In both the LE and ICT cases, the investigated transition is from the first excited state dominated by a HOMO–LUMO component. For this reason, the initial excited-state guess for each of these IMOM calculations was obtained by promoting one α electron from the HOMO to the LUMO.

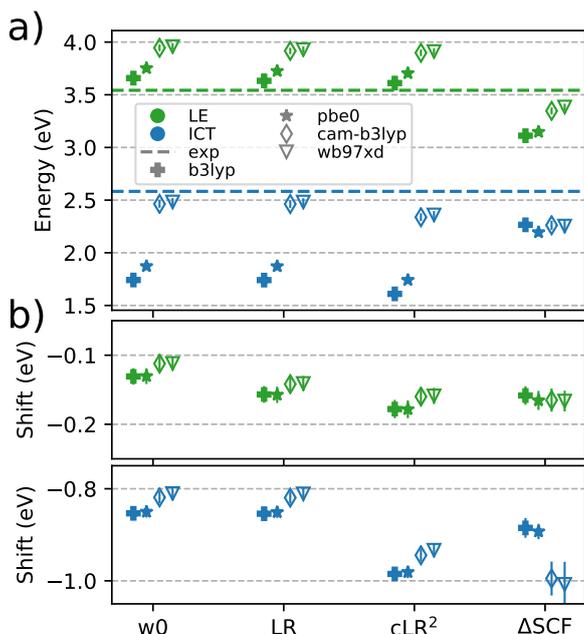


Figure 4. DMABN average emission energies from the LE and ICT states (a) and corresponding average solvatochromic shifts (b). In both panels, LE results are reported in green and ICT results in blue. The four investigated functionals are reported using different markers: long-range corrected functionals are reported using empty markers. The standard error is reported through error bars. For the emission energies, available experimental values are also reported as dashed line[45]. For the solvatochromic shifts, the reference values are an in vacuum TDDFT calculation for w_0 , LR and cLR values, and an in vacuum ΔSCF calculation for the ΔSCF values.

Figure 4 reports the average LE and ICT emission energies and the corresponding average solvatochromic shifts computed with the four selected functionals and the different models.

We first compare TDDFT and ΔSCF behaviors with respect to the choice of the DFT functional. Contrarily to the GFP case, here the ΔSCF results depend very weakly on the functional, whereas the TDDFT results for ICT strongly depend on the amount of long-range exact exchange. The latter finding is due to the well known limits of hybrid functionals in describing CT states. On the contrary, the ΔSCF approximation seems to make hybrid and

long-range corrected functionals very similar for the ICT. This result is more in line with the literature findings [9].

Moving to the different environment models, the TDDFT sets of data show what is expected on the basis of the very different nature of LE and ICT states. The former presents similar LR and cLR² results while for the latter the LR component of the solvatochromic shift is negligible (LR and ω_0 are in fact almost identical) and the SS effect is very strong which leads to very different LR and cLR² solvatochromic shifts.

The Δ SCF description behaves differently for the two states. For LE, we observe a similar behavior with respect to TDDFT descriptions even if the absolute excitation energies are significantly smaller with all the DFT functionals. For ICT, instead, the excitation energies calculated with Δ SCF are very similar for all DFT functionals and similar to the TDDFT-cLR² ones obtained with long-range corrected functionals.

When we move to solvatochromic shifts, Δ SCF and all TDDFT schemes behave similarly for the LE state. The solvatochromic shifts are all quite similar, with the Δ SCF and cLR² both predicting a shift of ~ -0.2 eV. However, in the case of cLR² this shift also contains the LR contribution, which is not included in Δ SCF.

The picture changes for the ICT state, where Δ SCF becomes more sensitive to the DFT functional than TDDFT. For this dark state, the LR component of the solvatochromic shift is negligible, therefore one would expect Δ SCF and cLR² to give the same results. This is not exactly true. We attribute this discrepancy to the fact that in Δ SCF the state-specific contribution is calculated self-consistently, whereas in cLR² it is estimated to first order on the basis of the TDDFT difference density.

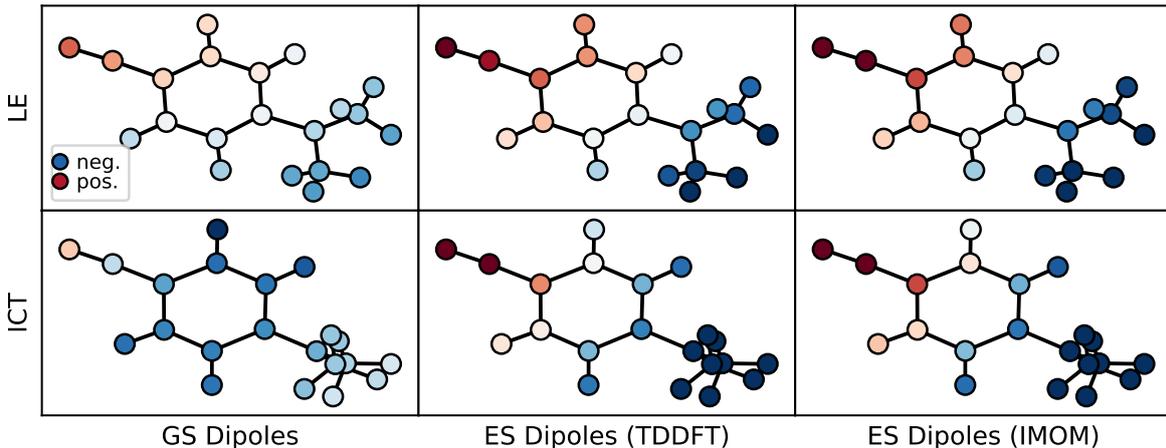


Figure 5. Electrostatic potential from the environment at the DMABN atoms for the LE and ICT structures/states. The potential reported is from the dipoles induced by the GS density (first column), by the ES density computed using TDDFT (second column) and by the ES density computed using IMOM (third column). We present only results obtained with the CAM-B3LYP functional. The color map is in arbitrary units, red values are positive, blue values are negative. All the plots are normalized to the same value.

Figure 5 reports a representation of the electrostatic potential due to the AMOEBA induced dipoles at the QM atoms. The electrostatic potentials are reported for two sample structures, corresponding respectively to a LE and an ICT state.

The induced dipoles in the GS behave differently for the two structures: in the LE one, they show a clear dipolar distribution, whereas in the ICT, they show a much less dipolar distribution. The two structures/states instead generate very similar induced dipoles when the excited state density is used. Moreover, TDDFT and IMOM give very similar pictures. In the LE case, the dipolar character observed in the GS becomes more accentuated and the induced dipoles stabilize this effect, whereas in the ICT case, the neutral character found in the GS is completely replaced by a strong dipolar character, and the induced dipoles strongly stabilize this state.

3.4. Charge Transfer in AppA

The last investigated process is an intermolecular charge transfer state in the AppA protein. The state corresponds to an electron transfer from a tyrosine residue to the flavine mononucleotide (FMN) chromophore. Figure 1c shows the system and the QM part which includes also a glutamine residue hydrogen-bonded to both tyrosine and flavine.

To study this process, we selected 10 frames from a classical NVT MD trajectory, then, for each of them, we performed a QM/AMOEBA minimization until a root-mean square norm on the forces of 4 kcal/mol/Å and finally a 2 ps QM/AMOEBA NPT equilibration to obtain the structures which are used in this work. In this case, the AMOEBA shell has a radius of 30 Å, of which 22 Å are polarizable. In the QM/AMOEBA MD simulations, the non polarizable shell is kept frozen to provide non periodic boundary conditions.

Here, the intermolecular CT is not guaranteed to be the first excited state, therefore we performed a series of preliminary LR calculations, solving for 6 excited states, using both a QM/AMOEBA and in vacuum descriptions. From the results of the preliminary calculations, we identified the states with the largest electron density displacement ($\Delta r > 5$ Å) [46], which correspond to the charge transfer state. This information was used to set up the state-specific cLR and IMOM calculations.

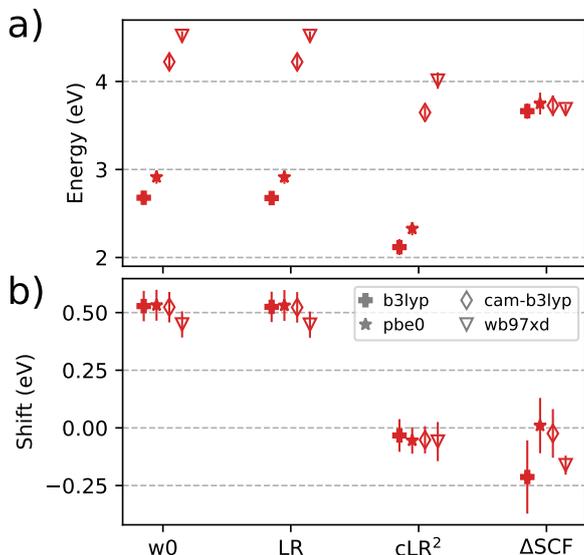


Figure 6. Average absorption energies (a) and solvatochromic shifts for the charge transfer state of AppA (b), the standard error is reported through error bars. The four investigated functionals are reported using different marks; long-range corrected functionals are reported using empty markers. The reference values are a in vacuum TDDFT calculation for ω_0 , LR and cLR values, and a in vacuum Δ SCF calculation for the Δ SCF values.

Figure 6 reports the excitation energies to the CT state and the corresponding solvatochromic shifts for the various models.

The excitation energies behave very similarly to what previously found for the ICT state of DMABN: the three TDDFT schemes show large differences between hybrid and long-range corrected functionals. Moreover, ω_0 and LR energies are almost indistinguishable, whereas introducing the SS shift in cLR² leads to substantially smaller excitation energies. As for ICT in DMABN, Δ SCF gives excitation energies which are insensitive to the functional and similar to the TDDFT-cLR² ones obtained with long-range corrected functionals.

The results obtained for the solvatochromic shifts qualitatively follow the ones found for ICT in DMABN, as well, but this time the differences between the three different TDDFT schemes are largely amplified. Here in fact, ω_0 and LR (which are almost indistinguishable) give a large and positive shift which becomes slightly negative when SS effects are introduced with cLR². This larger SS effect compared to the DMABN case arises from the intermolecular CT nature

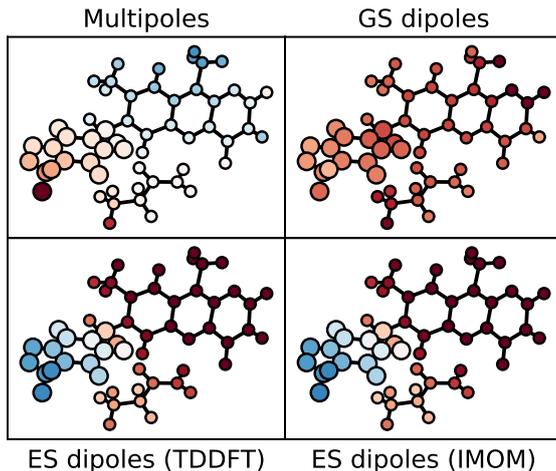


Figure 7. Electrostatic potential from the environment at the QM atoms. The potential reported is from the static multipoles and from the induced dipoles. The latter are respectively induced by the GS density, by the ES density computed using TDDFT and by the ES density computed using IMOM. We present only results obtained with the CAM-B3LYP functional. The color map is in arbitrary units, red values are positive, blue values are negative. All the plots are normalized to the same value, the tyrosine is plotted with larger markers to highlight its position in space.

of the state considered here, where negative and positive charges are well separated.

Figure 7 shows the electrostatic potential due to the AMOEBA atoms at the QM atoms. To explain the positive shift (~ 0.5 eV) seen at the ω_0 and LR levels, in this case, we also investigated the field generated by the fixed AMOEBA multipoles. It is evident that the electron transfer from the tyrosine to the FMN is greatly hampered by the static field of the protein pocket. In fact, the potential from the fixed multipoles goes against the electron transfer.

Including a SS polarizable description greatly stabilizes the CT state, as the dipoles reorient to counter the resulting displacement of charge. This can be seen from the electrostatic potential from the ES and IMOM dipoles, and by the solvatochromic shift computed using the cLR² and Δ SCF strategies.

4. Conclusions

In this work we have used the combination of Δ SCF and a polarizable MM force field for describing electronic transitions of embedded molecules and compared it with the same hybrid approach but within a TDDFT formulation. For TDDFT we have considered two different schemes for the environment response, accounting for the linear response only (LR scheme) or for both the linear and the state specific responses (cLR² scheme), respectively. We focused on three distinct systems characterized by bright valence excitations, intra- and intermolecular charge transfer states.

Our results highlight very different behavior for bright valence excitations and CT states. For valence excitations, excitation energies are more strongly dependent on the functional for Δ SCF than for TDDFT. In addition, Δ SCF misses the most important contribution to the solvatochromic shift, i.e. the LR one. For both intra- and inter-molecular CT states, Δ SCF proves very robust with respect to the choice of functional. Moreover, the Δ SCF solvatochromic effect on CT energies is very similar to those obtained with the most complete TDDFT cLR² scheme. Although we have considered a limited set of electronic transitions, our sample spans excitations with different characters and allows us to demonstrate clear trends. A larger sample of chromophores in diverse environments will allow confirming these trends in the general case.

This analysis shows that the polarizable Δ SCF/MM can represent a very effective approach to describe electronic transitions of molecules in solution or embedded in more complex environments when excitation showing a significant CT character are investigated. When bright valence

states are considered, instead, a further improvement seems necessary, namely the inclusion of the LR component of the environment response. This can be achieved with a correction based on LR TDDFT calculations, or by explicitly computing it from the transition density at the Δ SCF level [13].

References

- [1] Cammi, R.; Corni, S.; Mennucci, B.; Tomasi, J. Electronic excitation energies of molecules in solution: State specific and linear response methods for nonequilibrium continuum solvation models, *J. Chem. Phys.* **2005**, *122* (10), 104513.
- [2] Corni, S.; Cammi, R.; Mennucci, B.; Tomasi, J. Electronic excitation energies of molecules in solution within continuum solvation models: Investigating the discrepancy between state-specific and linear-response methods, *J. Chem. Phys.* **2005**, *123* (13), 134512.
- [3] Schwabe, T. General theory for environmental effects on (vertical) electronic excitation energies, *J. Chem. Phys.* **2016**, *145* (15), 154105.
- [4] Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G.D.; Mennucci, B. Electronic Energy Transfer in Condensed Phase Studied by a Polarizable QM/MM Model, *J. Chem. Theory Comput.* **2009**, *5* (7), 1838–1848.
- [5] Loco, D.; Polack, É.; Caprasecca, S.; Lagardère, L.; Lipparini, F.; Piquemal, J.P.; Mennucci, B. 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.
- [6] Caricato, M.; Mennucci, B.; Tomasi, J.; Ingrosso, F.; Cammi, R.; Corni, S.; Scalmani, G. Formation and relaxation of excited states in solution: A new time dependent polarizable continuum model based on time dependent density functional theory, *J. Chem. Phys.* **2006**, *124* (12), 124520.
- [7] Schröder, H.; Schwabe, T. Corrected Polarizable Embedding: Improving the Induction Contribution to Perichromism for Linear Response Theory, *J. Chem. Theory Comput.* **2018**, *14* (2), 833–842.
- [8] Guido, C.A.; Chrayteh, A.; Scalmani, G.; Mennucci, B.; Jacquemin, D. Simple Protocol for Capturing Both Linear-Response and State-Specific Effects in Excited-State Calculations with Continuum Solvation Models, *J. Chem. Theory Comput.* **2021**, *17* (8), 5155–5164.
- [9] Hait, D.; Head-Gordon, M. Orbital Optimized Density Functional Theory for Electronic Excited States, *J. Phys. Chem. Lett.* **2021**, *12* (19), 4517–4529.
- [10] Gilbert, A.T.B.; Besley, N.A.; Gill, P.M.W. Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method (MOM), *J. Phys. Chem. A* **2008**, *112* (50), 13164–13171.
- [11] Besley, N.A.; Gilbert, A.T.B.; Gill, P.M.W. Self-consistent-field calculations of core excited states, *J. Chem. Phys.* **2009**, *130* (12), 124308.
- [12] Barca, G.M.J.; Gilbert, A.T.B.; Gill, P.M.W. Simple Models for Difficult Electronic Excitations, *J. Chem. Theory Comput.* **2018**, *14* (3), 1501–1509.
- [13] Worster, S.B.; Feighan, O.; Manby, F.R. Reliable transition properties from excited-state mean-field calculations, *J. Chem. Phys.* **2021**, *154* (12), 124106.
- [14] Barca, G.M.J.; Gilbert, A.T.B.; Gill, P.M.W. Excitation Number: Characterizing Multiply Excited States, *J. Chem. Theory Comput.* **2017**, *14* (1), 9–13.
- [15] Hait, D.; Head-Gordon, M. Excited State Orbital Optimization via Minimizing the Square of the Gradient: General Approach and Application to Singly and Doubly Excited States via Density Functional Theory, *J. Chem. Theory Comput.* **2020**, *16* (3), 1699–1710.
- [16] Kunze, L.; Hansen, A.; Grimme, S.; Mewes, J.M. PCM-ROKS for the Description of Charge-Transfer States in Solution: Singlet–Triplet Gaps with Chemical Accuracy from Open-Shell Kohn–Sham Reaction-Field Calculations, *J. Phys. Chem. Lett.* **2021**, *12* (35), 8470–8480.
- [17] Cheng, C.L.; Wu, Q.; Voorhis, T.V. Rydberg energies using excited state density functional theory, *J. Chem. Phys.* **2008**, *129* (12), 124112.
- [18] Liu, J.; Zhang, Y.; Bao, P.; Yi, Y. Evaluating Electronic Couplings for Excited State Charge Transfer Based on Maximum Occupation Method Δ SCF Quasi-Adiabatic States, *J. Chem. Theory Comput.* **2017**, *13* (2), 843–851.
- [19] Vandaele, E.; Mališ, M.; Lubner, S. The photodissociation of solvated cyclopropanone and its hydrate explored via non-adiabatic molecular dynamics using Δ SCF, *Phys. Chem. Chem. Phys.* **2022**, *24* (9), 5669–5679.
- [20] Vandaele, E.; Mališ, M.; Lubner, S. The Δ SCF method for non-adiabatic dynamics of systems in the liquid phase, *J. Chem. Phys.* **2022**, *156* (13), 130901.

- [21] Ponder, J.W.; Wu, C.; Ren, P.; Pande, V.S.; Chodera, J.D.; Schnieders, M.J.; Haque, I.; Mobley, D.L.; Lambrecht, D.S.; DiStasio, R.A.; et al. Current Status of the AMOEBA Polarizable Force Field, *J. Phys. Chem. B* **2010**, *114* (8), 2549–2564.
- [22] Ren, P.; Wu, C.; Ponder, J.W. Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules, *J. Chem. Theory Comput.* **2011**, *7* (10), 3143–3161.
- [23] Bondanza, M.; Nottoli, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. Polarizable embedding QM/MM: the future gold standard for complex (bio)systems?, *Phys. Chem. Chem. Phys.* **2020**, *22* (26), 14433–14448.
- [24] Olsen, J.M.; Aidas, K.; Kongsted, J. Excited States in Solution through Polarizable Embedding, *J. Chem. Theory Comput.* **2010**, *6* (12), 3721–3734.
- [25] Giovannini, T.; Riso, R.R.; Ambrosetti, M.; Puglisi, A.; Cappelli, C. Electronic transitions for a fully polarizable QM/MM approach based on fluctuating charges and fluctuating dipoles: Linear and corrected linear response regimes, *J. Chem. Phys.* **2019**, *151* (17), 174104.
- [26] Loco, D.; Lagardère, L.; Caprasecca, S.; Lipparini, F.; Mennucci, B.; Piquemal, J.P. Hybrid QM/MM Molecular Dynamics with AMOEBA Polarizable Embedding, *J. Chem. Theory Comput.* **2017**, *13* (9), 4025–4033.
- [27] Loco, D.; Lagardère, L.; Cisneros, G.A.; Scalmani, G.; Frisch, M.; Lipparini, F.; Mennucci, B.; Piquemal, J.P. Towards large scale hybrid QM/MM dynamics of complex systems with advanced point dipole polarizable embeddings, *Chem. Sci.* **2019**, *10* (30), 7200–7211.
- [28] Nottoli, M.; Lipparini, F. General formulation of polarizable embedding models and of their coupling, *J. Chem. Phys.* **2020**, *153* (22), 224108.
- [29] Wang, W.; Skeel, R.D. Fast evaluation of polarizable forces, *J. Chem. Phys.* **2005**, *123* (16), 164107.
- [30] Lipparini, F.; Lagardère, L.; Stamm, B.; Cancès, E.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J.P. Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular Dynamics: I. Toward Massively Parallel Direct Space Computations, *J. Chem. Theory Comput.* **2014**, *10* (4), 1638–1651.
- [31] Greengard, L.; Rokhlin, V. A fast algorithm for particle simulations, *J. Comput. Phys.* **1987**, *73* (2), 325–348.
- [32] Caprasecca, S.; Jurinovich, S.; Lagardère, L.; Stamm, B.; Lipparini, F. Achieving Linear Scaling in Computational Cost for a Fully Polarizable MM/Continuum Embedding, *J. Chem. Theory Comput.* **2015**, *11* (2), 694–704.
- [33] Lipparini, F. General Linear Scaling Implementation of Polarizable Embedding Schemes, *J. Chem. Theory Comput.* **2019**, *15* (8), 4312–4317.
- [34] Nottoli, M.; Mennucci, B.; Lipparini, F. Excited state Born–Oppenheimer molecular dynamics through coupling between time dependent DFT and AMOEBA, *Phys. Chem. Chem. Phys.* **2020**, *22* (35), 19532–19541.
- [35] Casida, M.E.; Jamorski, C.; Casida, K.C.; Salahub, D.R. Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: Characterization and correction of the time-dependent local density approximation ionization threshold, *J. Chem. Phys.* **1998**, *108* (11), 4439–4449.
- [36] Nottoli, M.; Cupellini, L.; Lipparini, F.; Granucci, G.; Mennucci, B. Multiscale Models for Light-Driven Processes, *Annu. Rev. Phys. Chem.* **2021**, *72* (1), 489–513.
- [37] Lunkenheimer, B.; Köhn, A. Solvent Effects on Electronically Excited States Using the Conductor-Like Screening Model and the Second-Order Correlated Method ADC(2), *J. Chem. Theory Comput.* **2013**, *9* (2), 977–994.
- [38] Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. Gaussian Development Version Revision J.06+, 2020. Gaussian Inc. Wallingford CT.
- [39] Nifosi, R.; Mennucci, B.; Filippi, C. The key to the yellow-to-cyan tuning in the green fluorescent protein family is polarisation, *Phys. Chem. Chem. Phys.* **2019**, *21* (35), 18988–18998.
- [40] Nottoli, M.; Nifosi, R.; Mennucci, B.; Lipparini, F. Energy, Structures, and Response Properties with a Fully Coupled QM/AMOEBA/ddCOSMO Implementation, *J. Chem. Theory Comput.* **2021**, *17* (9), 5661–5672.
- [41] Henderson, J.N.; wang Ai, H.; Campbell, R.E.; Remington, S.J. Structural basis for reversible photobleaching of a green fluorescent protein homologue, *Proc. Natl. Acad. Sci.* **2007**, *104* (16), 6672–6677.
- [42] Fron, E.; Flors, C.; Schweitzer, G.; Habuchi, S.; Mizuno, H.; Ando, R.; Schryver, F.C.D.; Miyawaki, A.; Hofkens, J. Ultrafast Excited-State Dynamics of the Photoswitchable Protein Dronpa, *J. Am.*

- Chem. Soc.* **2007**, *129* (16), 4870–4871.
- [43] Pletneva, N.V.; Pletnev, V.Z.; Souslova, E.; Chudakov, D.M.; Lukyanov, S.; Martynov, V.I.; Arhipova, S.; Artemyev, I.; Wlodawer, A.; Dauter, Z.; et al. Yellow fluorescent protein phiYFPv (Phialidium): structure and structure-based mutagenesis, *Acta Crystallogr. D* **2013**, *69* (6), 1005–1012.
- [44] Peng, B.; Kuiken, B.E.V.; Ding, F.; Li, X. A Guided Self-Consistent-Field Method for Excited-State Wave Function Optimization: Applications to Ligand-Field Transitions in Transition-Metal Complexes, *J. Chem. Theory Comput.* **2013**, *9* (9), 3933–3938.
- [45] Coto, P.B.; Serrano-Andrés, L.; Gustavsson, T.; Fujiwara, T.; Lim, E.C. Intramolecular charge transfer and dual fluorescence of 4-(dimethylamino)benzonitrile: ultrafast branching followed by a two-fold decay mechanism, *Phys. Chem. Chem. Phys.* **2011**, *13* (33), 15182.
- [46] Guido, C.A.; Cortona, P.; Adamo, C. Effective electron displacements: A tool for time-dependent density functional theory computational spectroscopy, *J. Chem. Phys.* **2014**, *140* (10), 104101.