

# Polarizable Embedding Combined with the Algebraic Diagrammatic Construction: Tackling Excited States in Biomolecular Systems

Maximilian Scheurer,<sup>†,‡</sup> Michael F. Herbst,<sup>†</sup> Peter Reinholdt,<sup>‡</sup> Jógvan Magnus Haugaard Olsen,<sup>¶</sup> Andreas Dreuw,<sup>\*,†</sup> and Jacob Kongsted<sup>\*,‡</sup>

<sup>†</sup>*Interdisciplinary Center for Scientific Computing, Heidelberg University, D-69120 Heidelberg, Germany*

<sup>‡</sup>*Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, DK-5230 Odense M, Denmark*

<sup>¶</sup>*Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, UiT The Arctic University of Norway, N-9037 Tromsø, Norway*

E-mail: dreuw@uni-heidelberg.de; kongsted@sdu.dk

## Abstract

We present a variant of the algebraic diagrammatic construction (ADC) scheme by combining ADC with the polarizable embedding (PE) model. The presented PE-ADC method is implemented through second and third order and is designed with the aim of performing accurate calculations of excited states in large molecular systems. Accuracy and large-scale applicability are demonstrated with three case studies, and we further analyze the importance of both state-specific and linear-response-type corrections to the excitation energies in the presence of the polarizable environment. We demonstrate how our combined method can be readily applied to study photo-induced biochemical processes as we model the charge-transfer (CT) excitation key to the photoprotection mechanism in the dodecin protein with PE-ADC(2). Through direct access to state-of-the-art excited state analysis, we find that the polarizable environment plays a decisive role by significantly increasing the CT character of the electronic excitation in dodecin. PE-ADC is thus suited to decipher photoinduced processes in complex, biomolecular systems at high precision and at reasonable computational cost.

# 1 Introduction

In recent years, high-level *ab-initio* quantum-mechanical (QM) calculations of large and biologically relevant molecules have become increasingly popular.<sup>1-6</sup> This is, on one hand, driven by the steady increase in computational resources and, on the other hand, by improved algorithms. In addition, substantial advances in environmental embedding schemes have enabled a multi-scale treatment of large systems.<sup>1,5-7</sup> A multitude of models has emerged aiming at the description of molecular systems for which a full QM description is not needed but environmental effects are non-negligible. Implicit embedding schemes mimic the medium surrounding a specific molecule by a dielectric continuum. For these continuum solvation models (CSMs), multiple formulations have evolved, e.g., the polarizable continuum model (PCM) and the conductor-like screening model (COSMO).<sup>8-11</sup> Calculations involving CSMs usually aim at the description of a liquid solvent and are straightforward to perform since only a dielectric constant and the cavity need to be chosen. One of the most severe shortcomings of CSMs lies in modeling directionality of explicit solvent interactions, such as hydrogen bonds or  $\pi$ - $\pi$ -stacking.<sup>12,13</sup> These interactions are, however, crucial in anisotropic, heterogeneous environments, such as protic solvents, proteins, nucleic acids and biological membranes. Explicit embedding schemes, on the contrary, capture these heterogeneities by retaining a full atomistic representation of the environment.<sup>1</sup> As a result, explicit quantum-classical embedding schemes are generally more complex and computationally more demanding than CSMs as they require parametrization of the atomistic environment. In the simplest case, the parameters can be obtained by directly transferring charges from commonly available molecular mechanics (MM) force fields. However, such parameters are rarely intended to reproduce the microscopic detail required for accurate embedding calculations, especially with respect to molecular response properties. For the latter, not only a description of permanent electrostatics is of great importance, but also mutual polarization effects between the environment and the core region are desired. It is important to recognize that the increase in complexity and computational effort of explicit embedding arises not only from

the embedding as such, but also from the discrete nature of the environment which requires sampling of a representative number of configurations.

The polarizable embedding (PE) model<sup>14,15</sup> is a fragment-based, hybrid quantum-classical, and explicit embedding scheme, capable of modeling the aforementioned physical effects. Parameters are obtained from first-principles QM calculations performed on small split fragments of the molecular environment. Hence, close agreement with full QM calculations can in general be expected if the property of interest is localized on the quantum part only. In PE, the permanent charge distribution of the environment is modeled by fragment-based multipole expansions through a specific order, usually including monopole (charges), dipole, and quadrupole moments centered at each atom. Polarization effects are taken into account by dipole-dipole polarizabilities placed at the multipole expansion sites. Thus, PE is capable of treating mutual polarization of the environment and the quantum region in a self-consistent manner. In particular, PE has been designed to obtain accurate molecular response properties, e.g., electronic excitation energies. Accordingly, it has already been combined with several methods for calculating excited states in solution, i.e., time-dependent density functional theory (TD-DFT),<sup>14-16</sup> coupled-cluster (CC) methods,<sup>17-19</sup> second-order polarization propagator (SOPPA)<sup>20</sup> as well as a formulation within resonant-convergent response theory<sup>21</sup> and an open-ended response theory framework.<sup>22</sup> Further, these combined approaches have been successfully applied to study biomolecules, e.g., probes for biological membranes,<sup>23,24</sup> nucleic acids,<sup>25</sup> fluorescent proteins<sup>26-28</sup> and light-gated ion channels.<sup>29</sup> For a recent perspective on the PE model see Ref. 30.

The choice of electronic structure method is decisive for the quality of the excited state description. Especially TD-DFT with standard exchange-correlation functionals has well-known limitations, e.g., in the treatment of charge-transfer (CT) excitations, doubly excited states and Rydberg states.<sup>31-34</sup> In recent years, the algebraic diagrammatic construction (ADC) scheme has emerged as an accurate and balanced method for the calculation of excited states.<sup>35</sup> The ADC method builds on a Hermitian representation of the Hamilto-

nian, is size-consistent and systematically improvable. Also, the computational effort of ADC through second order (ADC(2)) is still manageable for medium-sized systems.<sup>35</sup> ADC through third order (ADC(3)) is particularly accurate in treating electronically doubly excited states where ADC(2) has known limitations.<sup>35,36</sup> The ADC method was previously combined with several embedding schemes, such as PCM,<sup>37,38</sup> COSMO,<sup>39,40</sup> and frozen density embedding (FDE).<sup>41,42</sup> There is, to the best of our knowledge, no combination of ADC with a hybrid quantum-classical, explicit embedding scheme including a polarization treatment for modeling electronic excitations in almost arbitrarily complex and large environments. In this paper, the first combined approach of PE with ADC through second and third order, namely PE-ADC(2) and PE-ADC(3), is presented. The PE-ADC schemes are built on top of a PE-HF ground-state calculation and take into account perturbative corrections of the excitation energies in a density-driven manner. To demonstrate the accuracy, applicability and versatility of PE-ADC, three case studies are performed: First, PE-ADC calculations are benchmarked against super-system ADC calculations on para-nitroaniline (PNA) in presence of small water clusters. Second, the excitation energies of the medium-sized chromophore lumiflavin (Lf) in bulk solvent are simulated and compared to PCM-ADC. Third, CT excitations key to the photoprotection mechanism of a flavin-binding protein, dodecin,<sup>43</sup> are modeled, showcasing the suitability of PE-ADC to investigate large biomolecular systems. Thereby, we unravel that the environment is assisting in stabilizing a low-lying CT excitation which is assumed to trigger the photoprotection mechanism.

This article is organized as follows. In the next section, theoretical key aspects leading towards our combination of PE with ADC are discussed. Afterwards, the computational methodology will be outlined. The main results are presented and discussed in detail with respect to the expected physical behavior of the PE-ADC approach in the chosen systems. Finally, a brief conclusion and outlook on possible future applications are given.

## 2 Theoretical Aspects

### 2.1 The Polarizable Embedding model

The general theory of the PE model, including all necessary energy contributions and effective operators, was presented comprehensively in previous work.<sup>14,15</sup> In short, an effective PE operator is constructed as

$$\hat{v}_{\text{PE}} = \hat{v}_{\text{es}} + \hat{v}_{\text{ind}}, \quad (1)$$

with the electrostatics operator  $\hat{v}_{\text{es}}$  and the induction operator  $\hat{v}_{\text{ind}}$ . Using the second-quantization formalism and a multi-index notation,<sup>14,44</sup> the electrostatic operator,  $\hat{v}_{\text{es}}$ , is written as

$$\hat{v}_{\text{es}} = \sum_{s=1}^S \sum_{|k|=0}^{K_s} \frac{(-1)^{|k|}}{k!} M_s^{(k)} \sum_{pq} t_{pq}^{(k)}(\underline{\mathbf{R}}_s) \hat{E}_{pq}, \quad (2)$$

with the one-electron orbital excitation operator  $\hat{E}_{pq}$  and general molecular orbital indices  $p$  and  $q$ . Furthermore,  $M_s^{(k)}$  is the  $k$ -th component of the  $|k|$ -th order Cartesian multipole moment located at  $\underline{\mathbf{R}}_s$  in the environment. The number of total sites in the environment is denoted by  $S$  whereas  $K_s$  is the truncation level of the multipole expansion at a given site  $s$ . The integrals are given by

$$t_{pq}^{(k)}(\underline{\mathbf{R}}_s) = - \int \phi_p^*(\underline{\mathbf{r}}) \left( \frac{\partial^{|k|}}{\partial \underline{\mathbf{r}}^k} \frac{1}{|\underline{\mathbf{r}} - \underline{\mathbf{R}}_s|} \right) \phi_q(\underline{\mathbf{r}}) d\underline{\mathbf{r}}, \quad (3)$$

The induction energy contribution through induced dipole moments amounts to

$$E_{\text{ind}} = -\frac{1}{2} \sum_{s=1}^S \underline{\mu}_s^{\text{ind}}(\underline{\mathbf{F}})^{\text{T}} \underline{\mathbf{F}}(\underline{\mathbf{R}}_s) \quad (4)$$

where  $\underline{\mu}_s^{\text{ind}}(\mathbf{F})$  is the induced dipole moment at site  $s$  in the environment, and  $\mathbf{F}(\mathbf{R}_s)$  is the electric field vector acting on site  $s$ , comprising the field from nuclei and electrons, as well as the fields caused by the permanent multipole moments, i.e.,

$$\mathbf{F} = \mathbf{F}_{\text{nuc}} + \mathbf{F}_{\text{el}}[\mathbf{P}] + \mathbf{F}_{\text{mul}} \quad (5)$$

Note that the electric field from the electrons, and in turn the total field vector,  $\mathbf{F}$ , *explicitly* depend on the electronic density matrix  $\mathbf{P}$ . The induced dipole moments can be solved for with classical response equations.<sup>45</sup> Finally, the induction operator is defined as

$$\hat{v}_{\text{ind}} = - \sum_{s=1}^S \sum_{a=x,y,z} \mu_{a,s}^{\text{ind}}(\mathbf{F}) \hat{F}_a^e(\mathbf{R}_s), \quad (6)$$

using  $a$  for the respective Cartesian component  $x, y$ , or  $z$  and the electric field operator  $\hat{F}_a^e(\mathbf{R}_s)$ . Since the embedding operator depends on the wave function itself, namely through the electric fields, the overall embedding operator is non-linear. As for a usual SCF procedure, where the exchange and Coulomb terms depend on the density, the embedding operator is updated in every iteration using the current SCF density matrix. Thus, polarization effects are treated in a fully self-consistent way for the Hartree-Fock electronic ground state.

## 2.2 Algebraic Diagrammatic Construction of the Polarization Propagator

A comprehensive derivation of the algebraic diagrammatic construction of the polarization propagator (ADC) can be found in literature.<sup>35,46–48</sup> In brief, the intermediate state (IS) formalism is used to obtain the ADC working equations. A set of creation and annihilation operators is applied to the exact ground state wave function, and the set of resulting IS wave functions are subject to a Gram-Schmidt orthogonalization. Subsequently, the exact excited

wave function  $|\Psi_n\rangle$  is expanded in the basis of intermediate states,  $\{|\tilde{\Psi}_J\rangle\}$ , as

$$|\Psi_n\rangle = \sum_J X_{nJ} |\tilde{\Psi}_J\rangle, \quad (7)$$

with the expansion coefficients  $X_{nJ}$  given through the eigenvectors of the ADC matrix. The ADC matrix  $M_{IJ}$  is a matrix representation of the Hamiltonian  $\hat{H}$  shifted by the exact ground state energy  $E_0$  in the basis of intermediate states, i.e.,

$$M_{IJ} = \langle \tilde{\Psi}_I | \hat{H} - E_0 | \tilde{\Psi}_J \rangle, \quad (8)$$

The corresponding Hermitian eigenvalue problem reads

$$\mathbf{M}\mathbf{X} = \mathbf{\Omega}\mathbf{X}, \quad \mathbf{X}^\dagger\mathbf{X} = \mathbf{1}, \quad (9)$$

with the eigenvector matrix  $\mathbf{X}$  and the diagonal matrix  $\mathbf{\Omega}$  containing excitation energies, i.e.,  $\Omega_n = E_n - E_0$ , where  $n$  denotes the  $n$ -th excited state. In practice, the IS basis is constructed from a Møller-Plesset (MP) perturbation expansion of the Hartree-Fock (HF) ground state. This leads to a perturbation expansion of the ADC matrix, In this manner, MP(2) and MP(3) ground states yield ADC(2) and ADC(3) working equations, respectively. Explicit equations and implementations can be found in previous work.<sup>36,46–49</sup> Since our approach to combine PE with ADC is completely density-driven, we will briefly comment on how ADC transition density matrices and excited state density matrices are obtained.<sup>47,48</sup> The one-particle reduced transition density matrix is given by

$$P_{pq}^{\text{tr},0 \rightarrow n} = \langle \Psi_n | c_p^\dagger c_q | \Psi_0 \rangle = \sum_I X_{n,I}^\dagger \langle \tilde{\Psi}_I | c_p^\dagger c_q | \Psi_0 \rangle \quad (10)$$

Further, one constructs the one-particle reduced density matrices and the one-particle reduced transition density matrices between different excited states by

$$P_{pq}^{n \rightarrow m} = \langle \Psi_m | c_p^\dagger c_q | \Psi_n \rangle = \sum_{I,J} X_{m,I}^\dagger \langle \tilde{\Psi}_I | c_p^\dagger c_q | \tilde{\Psi}_J \rangle X_{n,J}. \quad (11)$$

By setting  $n = m$  in the above equation, one obtains the  $n$ -th excited state density matrix. The difference density matrix between the ground state and the  $n$ -th excited state, which will become important in our density-driven PE-ADC approach, is simply

$$\Delta \mathbf{P}^{0 \rightarrow n} = \mathbf{P}^n - \mathbf{P}^0. \quad (12)$$

with the electronic ground state density  $\mathbf{P}^0$  and the density of the  $n$ -th excited state  $\mathbf{P}^n$ . Employing the IS representation (ISR), it is straightforward to compute the density matrices from the ADC eigenvectors.

### 2.3 Combination of PE and ADC

Our approach to include contributions from PE in ADC calculations of excitation energies is fully density-driven and based on perturbative, *a posteriori* corrections. As a consequence, the expressions for the ADC secular matrix (eq (8)) elements do not change. A similar approach was pursued for the combination of PCM with ADC<sup>37,38</sup> or the effective fragment potential (EFP)<sup>50-54</sup> method with EOM-CCSD,<sup>55,56</sup> for example. In recent publications,<sup>16,39,57</sup> it has been argued that both state-specific and linear-response-type contributions should be accounted for in order to obtain accurate excitation energies. We will follow these considerations in our approach combining PE with ADC, and in the following assume the environment to remain in the electronic ground state, thus, the electronic excitation to be localized on the core molecule, i.e., the chromophore. Generally, the shift in excitation energy of the core molecule in presence of an environment, also called *perichromatic* shift, can be written up

to second order<sup>57</sup> as

$$\Delta E_{\text{shift}}^{0 \rightarrow n} = \underbrace{\Delta E_{\text{es}}^{0 \rightarrow n}}_{\text{(i)}} + \underbrace{\Delta E_{\text{ind}}^{0 \rightarrow n}}_{\text{(ii)}} + \underbrace{\Delta E_{\text{disp}}^{0 \rightarrow n}}_{\text{(iii)}} + \underbrace{\Delta E_{\text{excoup}}^{0 \rightarrow n}}_{\text{(iv)}}. \quad (13)$$

The individual terms denote (i) the difference in Coulomb interaction energy with the electrostatic environment, (ii) the change in mutual induction energy upon excitation, (iii) energy difference in London dispersion energy, and (iv) the non-resonant excitonic coupling. In the PE model, London dispersion energy is not introduced by construction, so term (iii) is neglected. Moreover, these contributions have been argued to be rather small.<sup>16,57</sup> Since the PE model includes electrostatics, term (i) is already accounted for in the reference HF ground state entering the ADC calculation. This leaves us with the change in induction energy (ii) and non-resonant excitonic coupling (iv). The change in induction energy depends on the density matrix of the excited state and the resulting induced moments in the environment. The induced moments enter the ADC problem as a static dipole field through the Fock operator since the electric field of the electrons on environmental sites is not being updated during the ADC procedure. Hence, the change in induction energy included by construction amounts to

$$\Delta E(0)_{\text{el,ind}}^{0 \rightarrow n} = -\Delta \mathbf{F}_{\text{ind}}^{0 \rightarrow n} \underline{\mu}_{\text{ind}}^0 \quad (14)$$

$$= -\mathbf{F}_{\text{el}}^i \underline{\mu}_{\text{el,ind}}^0 + \mathbf{F}_{\text{el}}^0 \underline{\mu}_{\text{el,ind}}^0, \quad (15)$$

with  $\Delta E(0)_{\text{el,ind}}^{0 \rightarrow n}$  emphasizing the *frozen* induced moments of the ground state. The perichromatic shift with just these terms, i.e., a frozen PE ground state, is expressed by

$$\Delta E(0)_{\text{shift}}^{0 \rightarrow n} = \Delta E_{\text{es}}^{0 \rightarrow n} + \Delta E(0)_{\text{el,ind}}^{0 \rightarrow n} \quad (16)$$

with only zeroth-order contributions to the excitation energies. To improve upon this description, induced moments created through the density of the excited state need to be taken

into account, that is, a correction term will be added to eq (15) using a state-specific approach.<sup>57</sup> The total induction energy through the electric fields  $\underline{\mathbf{F}}_{\text{el}}^n$  in a given state  $n$  due to electrons (nuclear and multipole fields do not change upon excitation) is given by

$$E_{\text{ind}}^n = -\frac{1}{2}\underline{\mathbf{F}}_{\text{el}}^n\mu_{\text{el,ind}}^n, \quad (17)$$

and the difference in induction energy between the ground state and an excited state is simply

$$\Delta E_{\text{ind}}^{0\rightarrow n} = -\frac{1}{2}\left(\underline{\mathbf{F}}_{\text{el}}^n\mu_{\text{el,ind}}^n - \underline{\mathbf{F}}_{\text{el}}^0\mu_{\text{el,ind}}^0\right). \quad (18)$$

Taking the difference of eq (15) and the above equation, one finds the perturbative state-specific (ptSS) correction term for the excitation energy as

$$\Delta E_{\text{ptSS}}^{0\rightarrow n} = -\frac{1}{2}\underline{\mathbf{F}}_{\text{el}}[\Delta\mathbf{P}]^\dagger\mu_{\text{ind}}[\Delta\mathbf{P}]. \quad (19)$$

For a comprehensive derivation, see Ref. 16. The last term in eq (13) missing in our description is the non-resonant excitonic coupling energy contribution (iv). As has been shown in great detail using long-range perturbation theory by Schwabe,<sup>57</sup> and described in a similar manner for PCM,<sup>58</sup> the non-resonant excitonic coupling describes the interaction of the transition dipole moment with the induced moments in the environment, consistently taken into account through response theory approaches.<sup>14-16,57</sup> Since we do not include the embedding operator explicitly in the ADC problem, we correct for non-resonant excitonic coupling with a perturbative linear-response-type (ptLR) expression,

$$\Delta E_{\text{ptLR}}^{0\rightarrow n} = -\underline{\mathbf{F}}_{\text{el}}[\mathbf{P}^{\text{tr}}]^\dagger\mu_{\text{ind}}[\mathbf{P}^{\text{tr}}]. \quad (20)$$

Here, couplings between different excited states are neglected, as done in preceding work as well.<sup>16,57</sup> In total, the correction for the  $n$ -th excitation energy amounts to

$$\Delta E_{\text{corr}}^{0 \rightarrow n} = \Delta E_{\text{ptSS}}^{0 \rightarrow n} + \Delta E_{\text{ptLR}}^{0 \rightarrow n}. \quad (21)$$

To summarize, contributions from the environment using the PE approach are included indirectly through the frozen HF ground state and excitation energies are corrected *a posteriori* with perturbative treatments (eq (21)) using the difference and transition densities of the obtained excited states. All excited state properties are, however, calculated without further corrections in presence of the frozen PE-HF ground state.

### 3 Computational Methodology

In order to perform PE calculations, the already existing library `PElib`<sup>14,15,59</sup> was interfaced with *Q-Chem*.<sup>60</sup> Further, a stand-alone C++ code was implemented to execute PE in a shared-memory parallel way for fast performance on large systems. All ADC calculations were performed using a local development version of the *Q-Chem* program package.<sup>60</sup> Data analysis of all calculations was carried out using the Python packages `cclib`,<sup>61</sup> `orbkit`,<sup>62</sup> `pandas`,<sup>63,64</sup> `numpy`,<sup>65</sup> `scipy`.<sup>66</sup> `Matplotlib`<sup>67</sup> and `seaborn`<sup>68</sup> were used to create plots.

#### 3.1 para-Nitroaniline with small water clusters

Para-Nitroaniline (PNA) structures with two, four or six water molecules were taken from previous work by Slipchenko,<sup>55</sup> in which state-specific perturbative corrections were calculated with the EFP method at the EOM-CCSD level of theory. For these systems, the three lowest singlet excitation energies were calculated at the ADC(2) and ADC(3) levels using the cc-pVDZ basis set<sup>69-71</sup> for the full super-system including solvent, PNA in vacuum, and PE with ptSS and ptLR corrections for the excitation energies. Embedding potentials for the solvent molecules were obtained using `PyFraME`<sup>72</sup> and are based on `LoProp`<sup>73</sup> multipole

moments through second order and anisotropic dipole-dipole polarizabilities calculated at the PBE0/cc-pVDZ level of theory<sup>74</sup> using the Dalton program<sup>75</sup> and LoProp for Dalton script.<sup>76</sup> Detachment and attachment densities of the lowest  $n\pi^*$  ( $1^1A_2$ ) and  $\pi\pi^*$  ( $1^1A_1$ ) excitations were obtained using the `libwfa` library<sup>77,78</sup> and visualized with VMD.<sup>79</sup>

### 3.2 Lumiflavin in bulk solvent

To compare the performance of a continuum solvation model in ADC(2) calculations with our PE-ADC approach, a benchmark study has been performed using the flavin derivative lumiflavin (Lf). Excitation energies at the ADC(2)/cc-pVDZ level of theory were computed using three different setups, summarized in Table 1. We have chosen water ( $H_2O$ ) as a polar, protic solvent, and cyclohexane (cyHex) as an apolar, aprotic solvent. Geometries

Table 1: Benchmark systems for Lf in bulk solvent.

#	geometries	excited states	solvent
1	HF-3c	ADC(2)	none
2	HF-3c/PCM	PCM-ADC(2)	$H_2O$ , cyHex
3	QM/MM MD	PE-ADC(2)	$H_2O$ , cyHex

were obtained as indicated in Table 1 using the Orca 4 program package,<sup>80,81</sup> with a HF-3c<sup>82</sup> geometry optimization for the first two systems, and QM/MM MD sampling with the same method for the PE-ADC approach. A general workflow of how to simulate UV/Vis spectra with PCM-ADC or PE-ADC is depicted in Figure 1. Note that explicit solvent modeling, i.e., PE-ADC, requires configurational sampling of the solvent molecules through MD and/or QM/MM MD simulations. By contrast, implicit solvent modeling with PCM-ADC usually uses quantum mechanically optimized structures since the sampling is implicitly contained in the model by virtue of the macroscopic dielectric constants. Of note, the quantum-chemical method used for sampling in QM/MM MD simulations should be carefully chosen to provide suitable input geometries for high-level *ab-initio* excited states calculations. Note that HF-3c is being used here for fast sampling in the QM/MM MD simulations and that excitation

energies compared to MP2-optimized structures are blue-shifted by approximately 0.3 eV (Tab. S1). The detailed protocol for MD and QM/MM MD configurational sampling is thoroughly explained in the following.

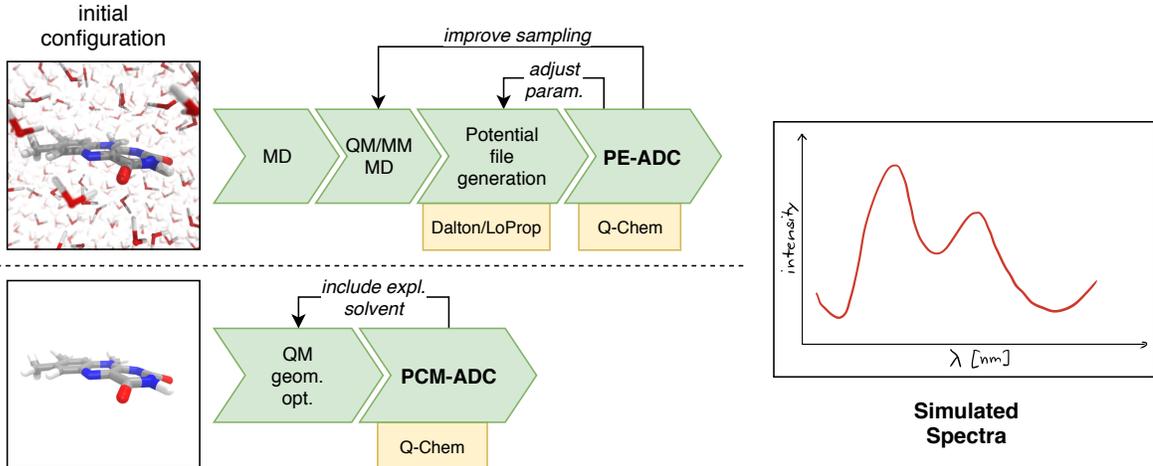


Figure 1: Workflows for PE-ADC and PCM-ADC calculations. From an initial chromophore configuration with explicit solvent, the workflow towards PE-ADC is described in the upper diagram, whereas implicit solvent modeling is described in the lower diagram. The PE-ADC method requires configurational sampling through MD and QM/MM MD simulations and subsequent parametrization of the environment. Results can be improved through longer sampling and adjusting the level of parametrization. Initial structures ought to be optimized first before entering a PCM-ADC calculation, and spectra can be refined by adding explicit solvent molecules to the QM system.

**QM/MM MD setup:** Lf was placed in a cube of 50 Å edge length filled with water molecules or cyHex using Packmol.<sup>83</sup> For configurational sampling, MD simulations of these systems were carried out using NAMD 2.12.<sup>84</sup> The CHARMM36 force field<sup>85</sup> was used for solvents and Lf was parametrized using SwissParam.<sup>86</sup> System setup was performed with the *psfgen* plugin in VMD.<sup>79</sup> The systems were first minimized for 10000 steps and then simulated in the NpT ensemble with periodic boundary conditions for 5 ns using an integration time step of 2 fs. Temperature was maintained at 310 K through the Langevin thermostat and pressure was kept at 1 atm with the Langevin Nosé-Hoover barostat<sup>87,88</sup> as implemented in NAMD.<sup>84</sup> To constrain water molecule geometries and bond vibrations, SETTLE and RATTLE algorithms were applied, respectively.<sup>89,90</sup> Electrostatics at long range were

treated using PME,<sup>91</sup> using a smooth 12 Å cutoff for short-range electrostatics. From the resulting trajectory, ten random snapshots were taken per system and refined in a QM/MM MD simulation with the NAMD QM/MM interface.<sup>92</sup> The HF-3c method<sup>82</sup> as implemented in Orca 4<sup>80,81</sup> was used for the quantum region which consisted of Lf. Point charges of the MM region were included via electrostatic embedding up to 12 Å distance from the quantum region. The QM/MM systems were first minimized for 100 steps and subsequently propagated for 12 ps at a 0.5 fs time step. Temperature and pressure were maintained as explained above. No constraints were applied to the simulation systems. This setup resulted in ten individual QM/MM trajectories per system from which ten random snapshots were extracted each with MDAnalysis.<sup>93,94</sup> Embedding potentials for every extracted snapshot, i.e., 100 per solvent system, were generated with PyFraME<sup>72</sup> including solvent molecules within 15 Å proximity. It has been shown in previous work that this distance cutoff is sufficient for convergence of polarization effects in PE approaches.<sup>95</sup> Multipoles through second order and polarizabilities were obtained at the PBE0/cc-pVDZ level of theory as implemented in the Dalton program<sup>75</sup> using the LoProp approach.<sup>73,76</sup> For each final snapshot, the five lowest singlet excited states were calculated using PE-ADC(2)/cc-pVDZ with perturbative excitation energy corrections. An ordinary least-squares (OLS) fit of the ptSS and ptLR corrections versus change in dipole moment squared,  $\Delta\mu^2$ , and oscillator strength (length gauge),  $f$ , respectively, was performed using the equations

$$\Delta E_{\text{ptLR}} = -\alpha f \tag{22}$$

for the ptLR correction energy and

$$\Delta E_{\text{ptSS}} = -\alpha \Delta\mu^2 \tag{23}$$

for the ptSS correction energy. The fitting parameters,  $\alpha$ , are related to the effective polarizability of the environment, neglecting pre-factors and dimensions of the fit. However, this is

not important since in the analysis, only ratios of fitting parameters are taken into account. The OLS fits were calculated including all five excited states of all snapshots.

### 3.3 Lumiflavin embedded in dodecin environment

From an MD trajectory obtained in previous work,<sup>43</sup> we extracted ten snapshots for subsequent refinement with QM/MM MD simulations. As quantum region, we chose one Lf ligand and its adjacent tryptophan residue (W36). The QM region is depicted in Figure 2b. Bonds between the quantum and classical region in the W36 backbone were treated as explained in Ref. 43. To prevent over-polarization at the QM/MM boundary, the charge shifting (CS) scheme<sup>96</sup> was applied to shift charges adjacent to QM/MM bonds. Point charges of the MM region were included via electrostatic embedding within a 12 Å distance cutoff. As explained before, NAMD 2.12 was used for the MM region, whereas the QM region was treated with HF-3c.<sup>82</sup> Electrostatics, temperature, and pressure were handled as explained above. The QM/MM MD simulation was run for 12 ps at a 0.5 fs integration time step after 100 minimization steps.

From the ten QM/MM MD simulations, ten snapshots were extracted per trajectory, and embedding potentials were generated for the extracted snapshots using PyFraME.<sup>72</sup> The dodecin environment setup is graphically illustrated in Figure 2a. The core quantum region in the PE-ADC calculation (Fig. 2b) consisted of Lf and the adjacent W36 residue, which will be referred to as Lf-W36 in the following. The protein was parametrized using PBE0/cc-pVDZ for multipole moments up to second order and polarizabilities using the LoProp approach. All ions within a distance of 50 Å from the core were taken into account through the formal charge of the ion and isotropic polarizabilities calculated with Dalton<sup>75</sup> at the B3LYP/aug-cc-pVTZ level of theory.<sup>69,97</sup> Water molecules within a 10 Å radius were parametrized using LoProp including multipole moments through quadrupoles and dipole-dipole polarizabilities. From 10 Å to 20 Å, water molecules were modeled with averaged charges and isotropic polarizabilities.<sup>98</sup> TIP3P water was further used to model the solvent between 20 Å and

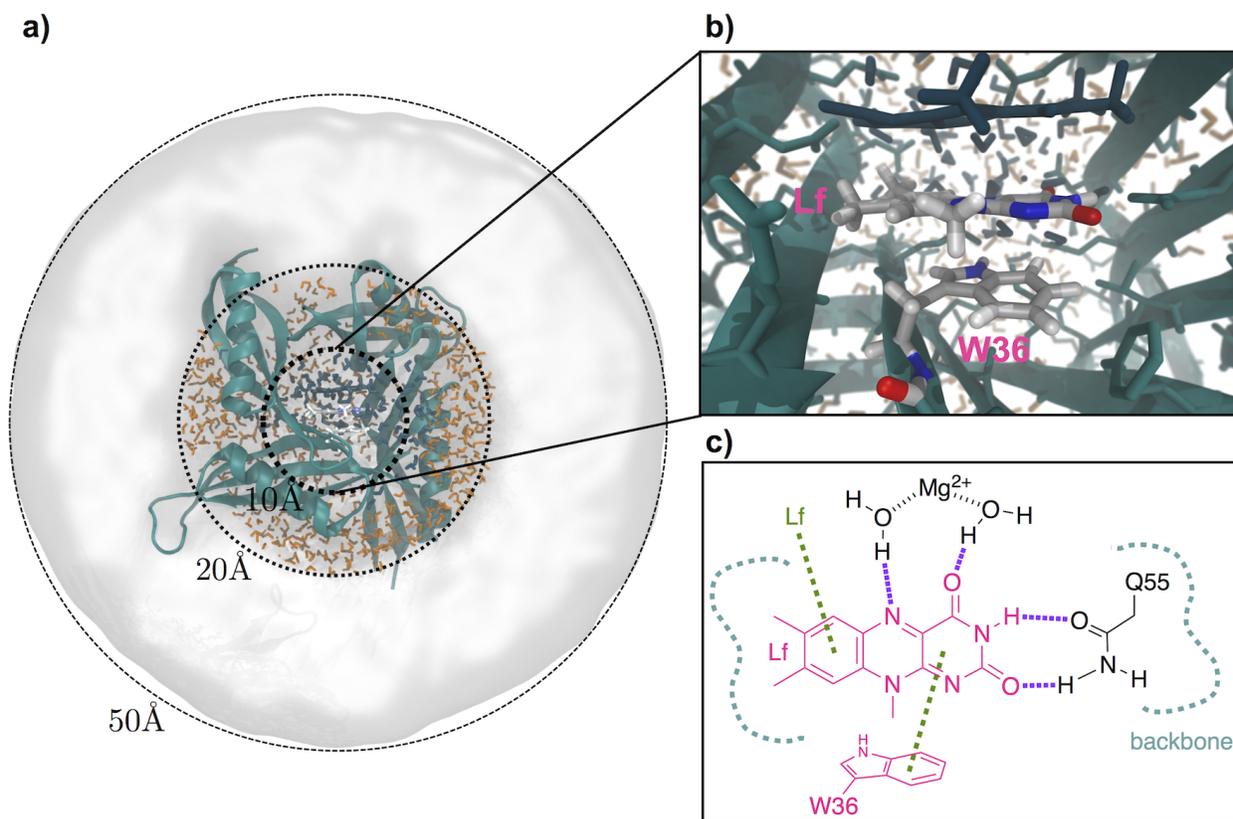


Figure 2: Dodecin environment setup for PE-ADC(2). **a)** Solvent shells in parametrization for dodecin PE-ADC calculation. Inner solvent shell ( $r \leq 10\text{\AA}$ , dark blue) was parametrized with LoProp. Middle solvent shell ( $10\text{\AA} < r \leq 20\text{\AA}$ , orange) uses pre-defined charges and isotropic polarizabilities. Outer solvent shell ( $20\text{\AA} < r \leq 50\text{\AA}$ , white surface) uses TIP3P water. Protein (cyan) is fully parametrized with LoProp. Ions are included up to  $50\text{\AA}$  distance and have pre-defined charges and isotropic polarizabilities. Ions are not shown for simplicity. **b)** Zoom-in of the dodecin active site. **c)** Interaction diagram. Purple dashed lines: hydrogen bonds, green dashed lines:  $\pi$ -stacking. The quantum region for PE-ADC calculation is drawn in magenta.

$50\text{\AA}$ . To avoid over-polarization, site polarizabilities were removed in  $1.2\text{\AA}$  distance of the quantum region, and charges of these sites were redistributed to the nearest neighbor site.<sup>99</sup> Higher-order multipole moments of the redistributed sites were removed. The energetically lowest five singlet excited states were obtained using PE-ADC(2)/cc-pVDZ while keeping the core and the highest 70 virtual orbitals frozen, following the restricted virtual space (RVS) ADC scheme.<sup>100</sup> To further investigate the contribution of the environment to the possible CT, ADC(2)/cc-pVDZ calculations of the isolated Lf-W36 system, i.e., without PE, were

further performed on snapshots with a ptSS correction smaller than -0.05 eV for the first excitation.

## 4 Results and Discussion

PE-ADC has been employed to perform a set of case studies, ranging from the small water-coordinated chromophore, PNA, over the medium-sized Lf molecule in bulk solvent to a protein system that embeds Lf. The benchmark studies probe different physical aspects of the PE-ADC method, which we are going to present and discuss in the following, also with respect to the type of electronic transition. In that sense, we cover  $n\pi^*$ ,  $\pi\pi^*$ , as well as inter- and intramolecular CT excitations.

### 4.1 para-Nitroaniline with small water clusters

Two different types of electronic excitations have been benchmarked in the case of PNA (Fig. 3) using PE-ADC(2) and PE-ADC(3), i.e., the energetically lowest singlet  $n\pi^*$  and  $\pi\pi^*$  excitations, in the presence of small water clusters. PNA forms hydrogen bonds with the surrounding water molecules *via* the nitro ( $\text{NO}_2$ ) group in the subset of structures studied here. The  $\text{NO}_2$  group acts as a hydrogen bond acceptor via the oxygen lone pairs. All results for PNA excitations using ADC(2) and ADC(3) calculations in vacuum, super-system, and PE are summarized in Table 2. Even though hydrogen bonds between the solvent and solute exist, neither over-polarization nor electron spill-out<sup>101</sup> was observed. The ground and excited state dipole moments are displayed in Figure 3a, and the detachment-attachment densities for both states of the structure including six water molecules obtained within an ADC(2) super-system calculation are shown in Figure 3b. For the lowest state, which has  $n\pi^*$  character, a blue-shift in the presence of water is expected because the dipole moment of PNA is reduced upon excitation (Fig. 3, Tab. 2). This should result in a destabilization of the excited state.<sup>102,103</sup> The blue-shift can also be rationalized by the detachment of electron

density at the lone pair of the nitro group, destabilizing hydrogen bonds with coordinated water molecules (Fig. 3b). Since the change in density upon excitation as well as the oscillator strength of  $S_0 \rightarrow S_1$  are rather small, minor contributions through perturbative corrections are to be expected.

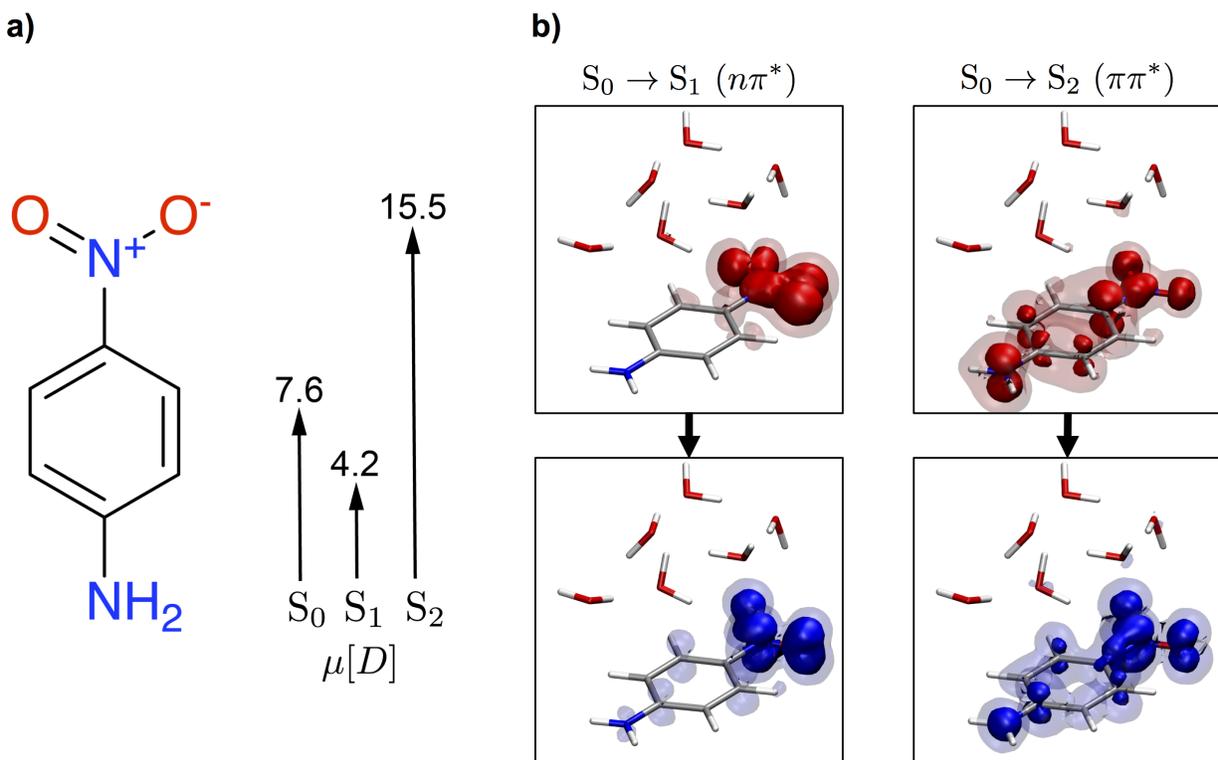


Figure 3: **a)** Lewis structure and dipole moments (arrows) of PNA for different electronic states in vacuum. **b)** Detachment (red) and attachment (blue) densities of the two lowest excitations of PNA in the presence of six water molecules. The  $n\pi^*$  excitation is local in nature, where electron density is transferred from the lone pair to the  $\pi^*$  orbital of the nitro group. Hence, the proximal hydrogen bond is weakened upon photoexcitation. An intramolecular charge-transfer takes place upon  $\pi\pi^*$  excitation, dramatically increasing the dipole moment of PNA.

Table 2: PE-ADC(2) and PE-ADC(3) results for PNA singlet excitations.<sup>a</sup>

		<b>ADC(2)</b>					
		$n\pi^*$			$\pi\pi^*$		
		vacuum	super-system	PE	vacuum	super-system	PE
PNA + 2 H <sub>2</sub> O	E [eV]	3.827	3.923	3.930	4.599	4.299	4.328
	f	0	0	0.000	0.453	0.436	0.447
	$\Delta\mu$ [D]	-3.38	-3.182	-3.214	7.882	8.611	8.519
	ptSS [eV]	-	-	-0.001	-	-	-0.010
	ptLR [eV]	-	-	-0.000	-	-	-0.005
PNA + 4 H <sub>2</sub> O	E [eV]	3.828	3.884	3.894	4.597	4.386	4.402
	f	0	0	0.000	0.454	0.433	0.451
	$\Delta\mu$ [D]	-3.383	-3.327	-3.282	7.862	8.264	8.332
	ptSS [eV]	-	-	-0.002	-	-	-0.010
	ptLR [eV]	-	-	-0.000	-	-	-0.005
PNA + 6 H <sub>2</sub> O	E [eV]	3.829	3.939	3.966	4.6	4.295	4.367
	f	0	0	0.000	0.454	0.427	0.445
	$\Delta\mu$ [D]	-3.383	-3.117	-3.155	7.862	8.638	8.435
	ptSS [eV]	-	-	-0.002	-	-	-0.018
	ptLR [eV]	-	-	-0.000	-	-	-0.011
		<b>ADC(3)</b>					
		$n\pi^*$			$\pi\pi^*$		
		vacuum	super-system	PE	vacuum	super-system	PE
PNA + 2 H <sub>2</sub> O	E [eV]	4.082	4.175	4.177	4.555	4.308	4.335
	f	0	0	0.000	0.400	0.398	0.407
	$\Delta\mu$ [D]	-2.464	-2.237	-2.298	5.944	6.541	6.449
	ptSS [eV]	-	-	-0.000	-	-	-0.007
	ptLR [eV]	-	-	-0.000	-	-	-0.005
PNA + 4 H <sub>2</sub> O	E [eV]	4.083	4.146	4.148	4.551	4.379	4.392
	f	0	0	0.000	0.401	0.391	0.407
	$\Delta\mu$ [D]	-2.466	-2.43	-2.385	5.935	6.229	6.307
	ptSS [eV]	-	-	-0.000	-	-	-0.006
	ptLR [eV]	-	-	-0.000	-	-	-0.005
PNA + 6 H <sub>2</sub> O	E [eV]	4.084	4.203	4.218	4.554	4.318	4.376
	f	0	0	0.000	0.401	0.390	0.404
	$\Delta\mu$ [D]	-2.466	-2.168	-2.208	5.933	6.579	6.405
	ptSS [eV]	-	-	-0.001	-	-	-0.012
	ptLR [eV]	-	-	-0.000	-	-	-0.010

<sup>a</sup>  $n\pi^* = 1^1A_2$ ,  $\pi\pi^* = 1^1A_1$  for gas-phase optimized geometries with point group symmetry.  $\Delta\mu$  refers to the difference dipole moment between ground state and excited state. Perturbative corrections have already been incorporated into the PE excitation energies.

In our calculations, we consistently observed a blue-shift in presence of water molecules compared to the *in vacuo* case. This blue-shift was also reproduced by PE-ADC, albeit only through contributions of the polarizable environment to the ground state wave function. As expected, both ptSS and ptLR corrections are almost equal to zero, and the solvent blue-shift is solely modeled through the PE-HF ground state. On the contrary, the  $S_0 \rightarrow S_2$  transition exhibits quite different properties. This  $\pi\pi^*$  excitation corresponds to an intramolecular charge-transfer, giving rise to a significant increase in dipole moment. Electron density is detached mostly from the amino group of PNA and attached to the nitro group on the opposite site of the molecule (Fig. 3b). In this case, a red-shift of the excitation energy is expected, since the excited state is stabilized by a polar solvent.<sup>104,105</sup> As can be seen in Table 2, this holds true for the super-system calculation and is accurately captured by PE-ADC. Even though the change in dipole moment is dramatic and the oscillator strength is much larger for the CT excitation, the perturbative corrections are small compared to the zeroth-order contribution through the environment, but correct the overestimated excitation energy in the right direction. Hence, the magnitude of the perturbative corrections can be seen as an error estimate for the fact that a) the excited state wave function and the induced moments are not self-consistently obtained (ptSS) and b) no response treatment of the PE operator is directly included (ptLR). These perturbative corrections always alter overestimated excitation energies in the right direction, since the correction terms are quadratic in the induced moments and carry a negative pre-factor (eqs (19) and (20)). The ptSS and ptLR corrections for the  $\pi\pi^*$  excitation are largest in the presence of six water molecules and would be even higher if the number of solvent molecules were further increased.<sup>16</sup>

The maximum absolute error observed when comparing PE-ADC with the super-system ADC calculations are as small as 0.07 eV for ADC(2) and 0.06 eV for ADC(3) due to the localization of the electronic excitations processes to the chromophore only (Fig. 3b). These errors are well below the intrinsic errors of the used ADC methods.<sup>35,36,106</sup> All in all, the expected solvent effects are accurately captured for both types of excitations in water-

coordinated PNA. Also, the zeroth-order contribution just through the PE-HF ground state has been found to have the largest effect on the solvent shift, as has been shown using other approaches as well.<sup>16,28,55,56</sup> Over-estimated excitation energies were narrowed by ptSS corrections which are larger for significant changes in dipole moment upon excitation, whereas ptLR corrections contribute in case of a large oscillator strength.

## 4.2 Lumiflavin in bulk solvent

Table 3: Comparison of ADC(2), PE-ADC(2) and PCM-ADC(2) excitations of Lf in water and cyHex solvent.<sup>a,b</sup>

$S_0 \rightarrow S_i^c$	properties	vacuum	H <sub>2</sub> O		cyHex	
			PE	PCM	PE	PCM
$S_1 (\pi\pi^*)$	E [eV]	3.230	3.073	3.070	3.129	3.122
	f	0.257	0.243	0.262	0.238	0.257
	$\Delta\mu$ [D]	1.901	1.950	2.312	2.009	2.074
$S_2 (n\pi^*)$	E [eV]	3.610	3.579	3.773	3.565	3.663
	f	0.001	0.005	0.001	0.003	0.001
	$\Delta\mu$ [D]	-4.841	-3.490	-2.844	-3.637	-3.59
$S_3 (n\pi^*)$	E [eV]	3.804	3.842	3.971	3.842	3.787
	f	0.000	0.002	0.000	0.001	0.000
	$\Delta\mu$ [D]	-6.224	-6.272	-8.521	-6.706	-8.021
$S_4 (\pi\pi^*, CT)$	E [eV]	4.498	4.245	4.066	4.301	4.247
	f	0.121	0.139	0.161	0.129	0.138
	$\Delta\mu$ [D]	6.581	6.315	8.341	6.574	7.546
$S_5 (n\pi^*)$	E [eV]	4.599	4.568	n.c.	4.546	n.c.
	f	0.000	0.030	n.c.	0.013	n.c.
	$\Delta\mu$ [D]	-4.700	-5.588	n.c.	-5.707	n.c.

<sup>a</sup> Median values of corrected excitation energies and properties are shown for PE-ADC(2).

<sup>b</sup> n.c. = not converged.

<sup>c</sup> Character of the transition in parentheses.

The following case study aims to model bulk solvent effects on the electronic excitations of Lf through PE-ADC(2). In this context, a super-system ADC calculation is not possible due to the system size. For this reason, we will compare bulk solvent effects from water and cyHex to PCM-ADC(2) calculations. The latter is well-established and has proven a reliable

technique to obtain excitation energies in bulk solvent.<sup>37,38</sup>

For PE-ADC(2), solute-solvent configurations were sampled by means of QM/MM MD simulations, whereas gas-phase ADC(2) and PCM-ADC(2) excitation energies were obtained from optimized structures. The results including excitation energies, oscillator strengths, and changes in dipole moment upon excitation are shown in Table 3. Characterization of the excitations was achieved using natural transition orbitals (NTOs) depicted in Figure S1 in the Supporting Information. The solvent shift will be referred to as the difference between vacuum and solvent excitation energies. For the lowest, bright  $\pi\pi^*$  excitation, PE and PCM results give almost identical results. The expected red-shift is present for both solvents, where water shows a slightly larger red-shift than cyHex for both embedding schemes. Predicted solvent shifts for the  $S_2$   $n\pi^*$  dark state differ between PCM-ADC(2) and PE-ADC(2): PE predicts a red-shift of the excitation energy, whereas PCM shows a blue-shift. However, the red-shift is small and could simply result from the sampled geometries underlying the PE-ADC(2) calculations, whereas an optimized structure is used in the PCM case. This hypothesis would, nevertheless, have to be further investigated. Differing solvent shifts are also observed for the third singlet state, which is an  $n\pi^*$  excitation as well. A  $\pi\pi^*$  excitation with CT character corresponds to the fourth excited state of Lf. In this case, a red-shift is again correctly predicted by PE and PCM. This red-shift is more pronounced in water solvent, where PCM predicts a shift approximately 0.2 eV larger than PE. The comparison of PE and PCM excitation energies and excited state properties shows a consistent behavior for spectroscopically relevant bright states, i.e., the  $S_0 \rightarrow S_1$  and  $S_0 \rightarrow S_4$  transitions in this case. The trends for solvent shifts through bulk solvation modeled by PE-ADC(2) are thus comparable to PCM-ADC(2) which is a thoroughly benchmarked method for these applications.<sup>37</sup>

To gain further insight into the excitation energies and perturbative corrections, a statistical analysis of the ADC calculations is needed. Accordingly, the PE-ADC(2) results for corrected excitation energies, ptSS, and ptLR corrections are depicted as box plots in Figure

4. The box plots show the ranges of corrected excitation energies resulting from the config-

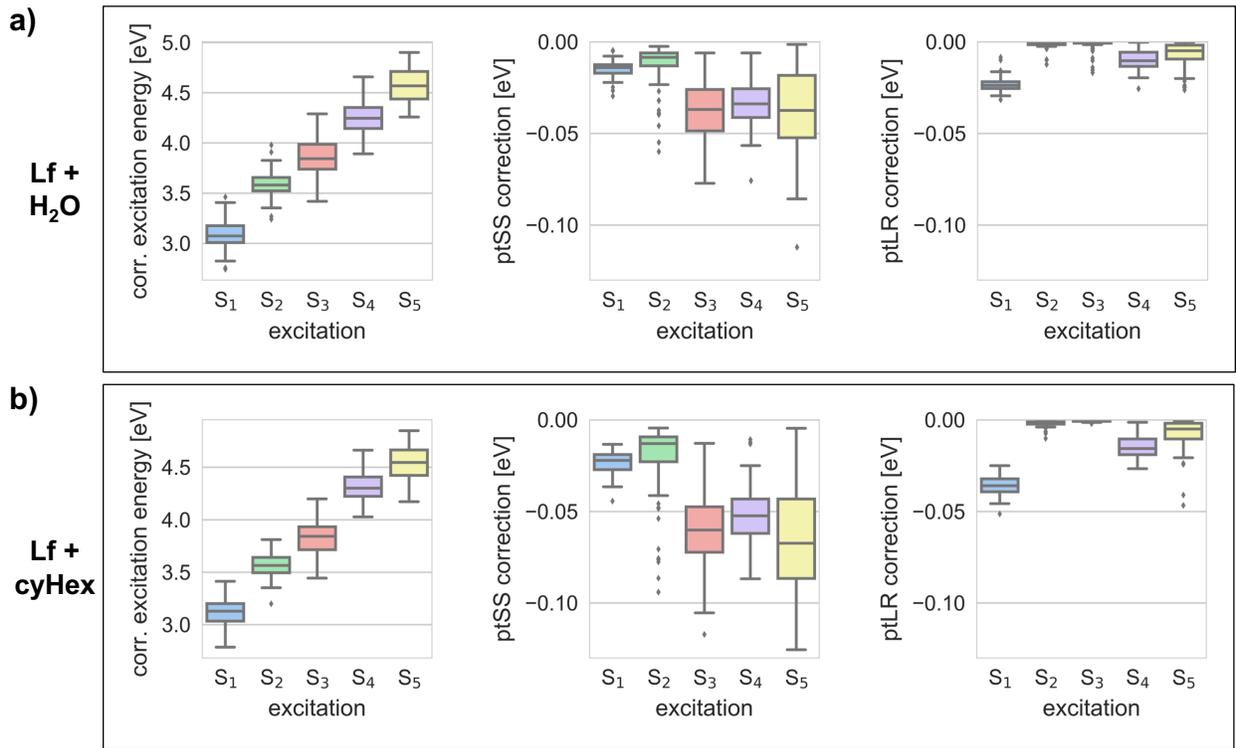


Figure 4: PE-ADC(2) excitation energies and corrections for Lf in **a)** water and **b)** cyHex solvent. Corrected excitation energies and ptSS and ptLR contributions are depicted as box plots for the five lowest excited singlet states of Lf. The boxes comprise 50% of the observed data, where the antennae mark lower and upper boundaries of 1.5 times inter-quartile range (IQR). Data points outside 1.5 IQR are defined as outliers (diamonds).

urational sampling. A minor amount of outliers is observed here for all states except for  $S_2$ . However, the energy range spanned by the boxes (50% of the observations) is rather narrow. Higher-lying states cover a broader energy range for ptSS correction terms, whereas ptLR corrections are very localized. Thus, the ptSS corrections term seems to be rather sensitive to the underlying geometry. Generally, the box plots show a rather condensed representation of the manifold of PE-ADC excited state calculations, and are thus the ideal tool to assess the geometrical sampling for the final excited state calculations. In case of highly fluctuating excitation energies, one would have to improve the sampling, as explained in the general work flow (Fig. 1). As discussed before, the total solvent red-shift on  $\pi\pi^*$  excitations is larger for water than for cyHex. On the contrary, both the ptSS and ptLR corrections display a

larger magnitude in cyHex solvent compared to water. The interpretation of the perturbative

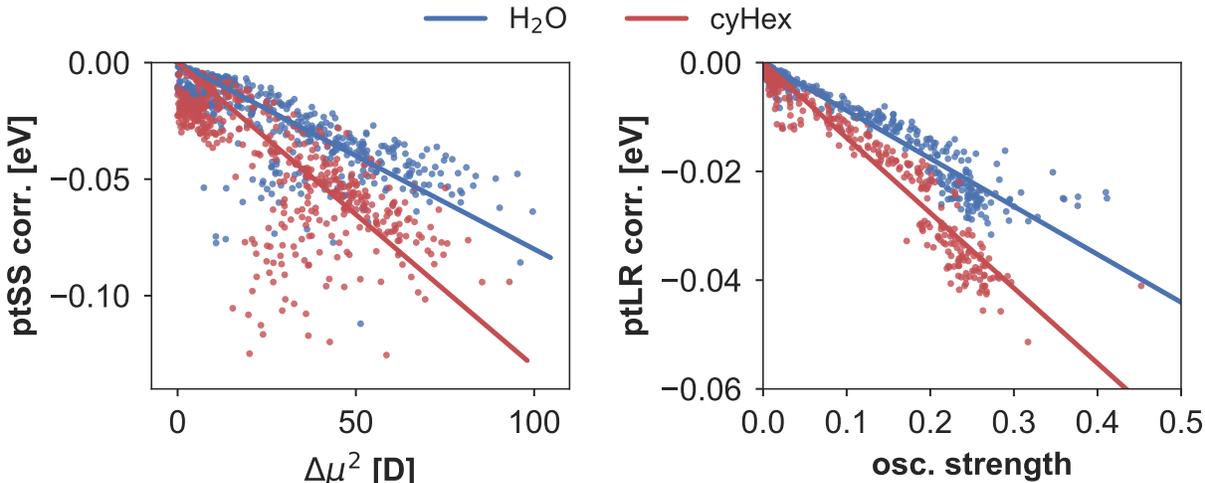


Figure 5: Correlation between perturbative corrections and excited state/transition properties. For both water and cyHex solvent, ptSS and ptLR corrections are shown vs. squared difference dipole moment ( $\Delta\mu^2$ ) and oscillator strength, respectively, for all states.

corrections in this context will be elaborated by correlating the ptSS and ptLR terms with the squared change in dipole moment and the oscillator strength, respectively, because these properties are inherently linearly related to the correction terms through equations (19) and (20). Corresponding linear fits, yielding what is here called the ‘effective’ solvent polarizability as a fitted parameter, are plotted in Figure 5. Detailed fitting results are shown in Table S2. The ratios between the effective polarizabilities of cyHex and water solvent are 1.63 for ptSS and 1.57 for ptLR. That is, the cyHex environment is effectively more polarizable than the water environment, leading to larger perturbative energy correction terms. Using the configurational sampling through QM/MM MD, the ‘effective’ solvent polarizabilities are easily obtained directly from fitting the perturbative energy corrections. All in all, this analysis nicely illustrates the general relationship between excited state properties and the expected magnitude of ptSS and ptLR corrections.

### 4.3 Protein environment effects on the charge-transfer excitation in dodecin

To showcase the capabilities of PE-ADC for addressing excited states of chromophores embedded in large, complex environments, we investigate the CT excitation involved in the photoprotection mechanism of the flavin-binding protein dodecin.<sup>107–109</sup> Dodecin stores flavin derivatives and efficiently photoprotects them by a multi-step quenching mechanism.<sup>43</sup> The initial key step in the mechanism most likely comprises a CT excitation, where an electron is transferred from a tryptophan residue (W36) to Lf, forming a positively charged tryptophan radical and a negatively charged flavosemiquinone radical. In previous work,<sup>43</sup> the CT excitation of Lf-W36 was modeled using ADC(2) calculations *in vacuo*. Since such a CT excitation creates two oppositely charged fragments, it exhibits a large change in electron density. We aim at extending the previous study by showcasing the excitation process using PE-ADC(2), thereby including explicit environmental effects. Not only is the system challenging due to the nature of the excited state, but also with respect to the heterogeneous interactions of the core region with the environment: Lf forms hydrogen bonds with solvent water molecules and a bifurcated hydrogen bond with Q55 in the protein. Further, it interacts with another Lf molecule by  $\pi$ -stacking. The heterogeneity of the core-environment interactions modeled through PE is depicted in Figure 2c. Obviously, these interactions are impossible to capture using a CSM. We also found PE-ADC(2) to be computationally affordable on the Lf-W36 system since the calculations only require additional time during the SCF procedure compared to ADC(2) calculations without PE (data not illustrated).

The results for PE-ADC(2) excitation energies and perturbative corrections for the three energetically lowest singlet states are presented in Figure 6. The two lowest states are almost degenerate, as observed in the *in vacuo* case,<sup>43</sup> while the third excited state covers a spectrum of almost 2 eV among all snapshots. The observed similarity between the lowest two states can further be examined with the dominant amplitudes involved in the transition (Fig. S2). In fact,  $S_1$  and  $S_2$  show an almost identical pattern, including HOMO-2 $\rightarrow$ LUMO, HOMO-

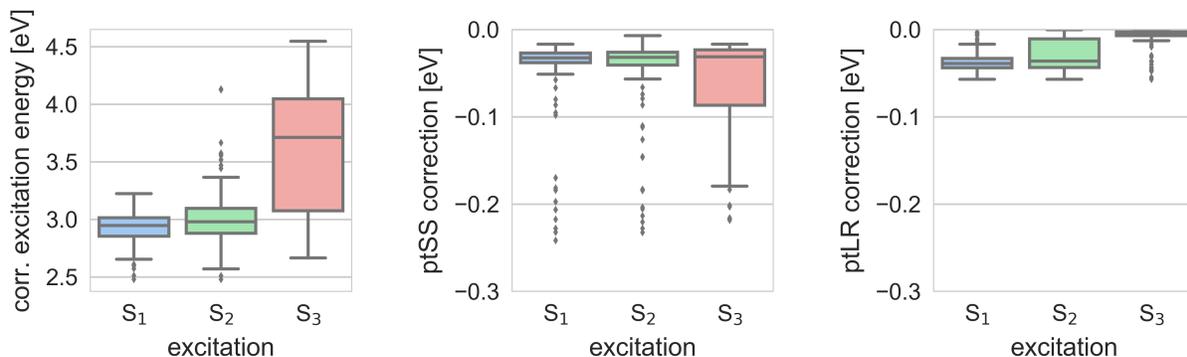


Figure 6: PE-ADC(2) excitation energies and corrections for Lf-W36 in dodecin environment. The presented box plots were prepared as for Fig. 4.

$1 \rightarrow \text{LUMO}$ , and  $\text{HOMO} \rightarrow \text{LUMO}$  as dominant transitions. However,  $S_3$  only contains these amplitudes to some extent in addition to other major orbital contributions. The HOMO and LUMO of a representative snapshot are depicted in Figure 7. Clearly, the HOMO and LUMO are localized on the W36 and Lf, respectively, giving rise to a CT from W36 to Lf. The other dominant amplitudes show a mixture of CT and local excitation (LE)<sup>43</sup> (data not illustrated), where  $\text{HOMO} \rightarrow \text{LUMO}$  exhibits the most prominent fragment separation. Since

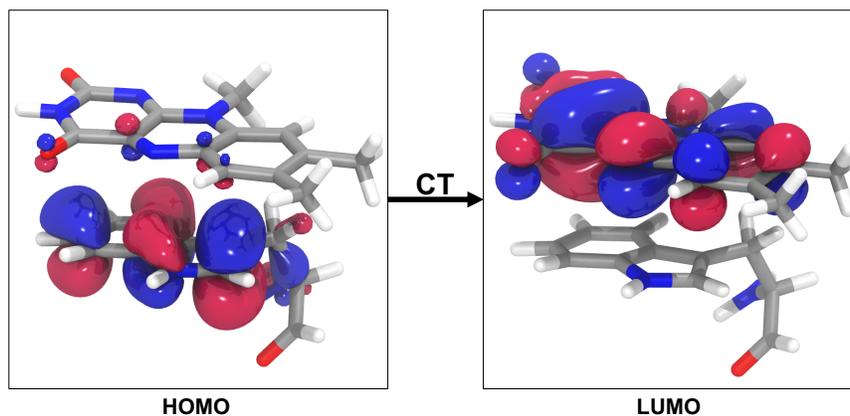


Figure 7: HOMO to LUMO transition comprising a possible Lf-W36 charge-transfer.

the energy range of the third singlet excitation also covers  $S_1$  and  $S_2$  excitation energies, it is possible that in a few snapshots,  $S_0 \rightarrow S_3$  corresponds to a CT excitation. This is also corroborated by the aforementioned orbital contributions.

The ptSS correction energies are localized around -0.03 eV and -0.2 eV for the first two

states, and a broader range of values is covered by the data for the third state. The data for ptLR correction energies contain very few outliers for  $S_1$  and  $S_2$  with a median of -0.039 eV and -0.036 eV, respectively. The median ptLR correction for  $S_3$  is -0.003 eV, thus, almost equal to zero. Accordingly, ptSS corrections for  $S_3$  and the outliers clustered around -0.2 eV for  $S_1$  and  $S_2$  are dominating over ptLR corrections in the Lf-W36 system.

For a CT state, the ptSS correction is expected to be large due to the charge separation between the two fragments and the resulting change in dipole moment upon excitation. To further unravel the effect of the polarizable environment on the CT character, snapshots that showed a ptSS correction smaller than -0.05 eV for the first excited state were re-calculated without PE, i.e., on the isolated Lf-W36 system only. These snapshots should supposedly exhibit a strong CT character due to the large magnitude of the ptSS correction. To avoid visual inspection of all snapshots, statistical analysis of the CT character was performed using the electron-hole distance  $d_{h \rightarrow e}$ <sup>34,77,78</sup> which is obtained from the excitonic wave function. It quantifies a CT through the distance of the hole and electron centroid and is thus ideal to characterize a large amount of ADC calculations. A CT results in a large shift in electron density from one fragment to another. This should result in a large increase in dipole moment upon excitation and consequently a large ptSS energy correction. Figure 8a shows the expected relationship between ptSS correction and electron-hole distance: The larger the electron-hole distance, the larger the magnitude of the ptSS correction. Interestingly, both quantities are related to different properties of the electronic excitations: electron-hole distances are obtained from the excitonic wave function, i.e., from the transition density matrix, whereas the ptSS term originates from the difference density matrix. Nevertheless, the electron-hole distance quantifies a charge transfer through electronic excitation which is also reflected by the ptSS correction term.

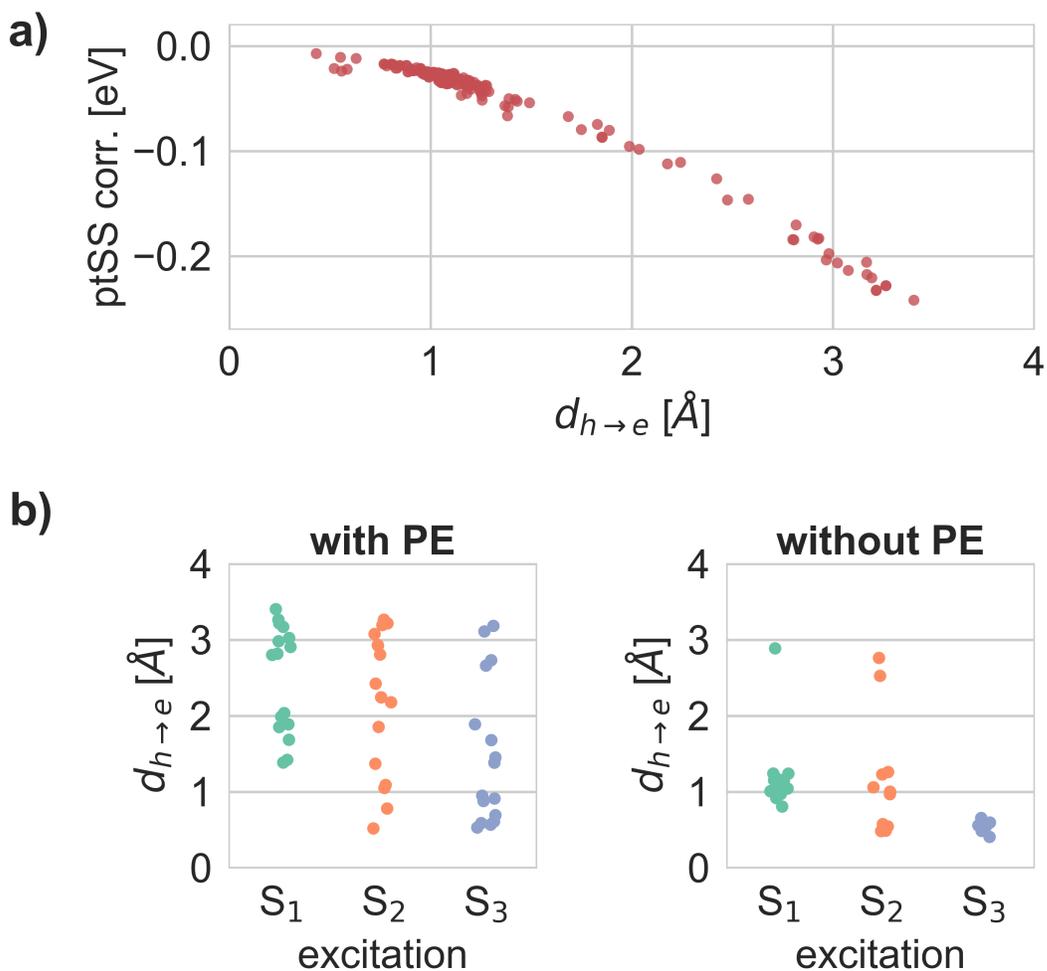


Figure 8: Analysis of Lf-W36 CT states using electron-hole distances. **a)** Relationship between ptSS correction and electron-hole distance  $d_{h \rightarrow e}$  for S<sub>1</sub> and S<sub>2</sub>. The electron-hole distance quantifies a permanent CT, which is also reflected by a large shift in electron density, leading to a larger ptSS correction term. **b)** Effect of PE on electron-hole distances. In isolation, i.e., without the protein environment included through PE, the character of the excitation is shifted from CT to LE.

To quantify the environmental effect on the CT character, electron-hole distances for the extracted snapshots with and without PE are compared in Figure 8b. With PE, S<sub>1</sub> electron-hole distances are clustered above 1 Å, whereas the values for S<sub>2</sub> and S<sub>3</sub> are almost evenly distributed between 1 Å and 3 Å. The corresponding median values are 2.81 Å, 2.33 Å, and 1.17 Å, respectively. This picture is dramatically changed if no environment through PE is

included for the same snapshots: For all states, electron-hole distances are clustered around 1 Å with the median values 1.07 Å, 0.97 Å, and 0.56 Å, respectively. As a consequence, the electron-hole separation is diminished without the polarizable environment, increasing the LE character of the excited states. Clearly, not all observed electronic excitations are pure CT states, but rather a mixture of CT and LE. The presence of the environment, however, increases the CT character of the low-lying excited states. In that way, the protein environment promotes CT excitations to trigger the photoprotection mechanism. A local excitation on the Lf fragment would not result in subsequent proton transfer and follow-up excited state quenching.<sup>43</sup> From a biochemical point of view, the protein environment seems to be essential in mediating the CT excitation in order to efficiently protect flavin derivatives from photodegradation.

Our calculations with PE-ADC(2) clearly show that the protein and solvent environment is capable of stabilizing and thus promoting the CT excitation on Lf-W36 which is essential to the current understanding of the dodecin photocycle. In experimental studies, the photocycle was manipulated through W36 mutants with different redox potentials.<sup>109</sup> Similarly, one could computationally probe and design future active site mutants guided by PE-ADC.

## 5 Conclusion

In the presented work, the PE model was combined with ADC to investigate electronic excitations of chromophores in complex, polarizable environments at reasonable computational cost but high precision. Excitation energies are perturbatively corrected in our approach both by a state-specific (ptSS) and linear-response-type (ptLR) term. PE-ADC(2) and PE-ADC(3) calculations on water-coordinated PNA were performed, showing that the presented approach captures the most important effects of the solvent shift through the PE-HF ground state. We have further compared PE-ADC with PCM-ADC and presented a case study where the dodecin photoprotection mechanism was thoroughly investigated in presence of

the protein and solvent environment through PE-ADC(2). Direct access to excited state analysis and visualization in *Q-Chem* enabled us to show that a possible CT state is stabilized through the polarizable protein and solvent environment. This CT excitation is crucial to the functionality of the dodecin photoprotection mechanism. The inclusion of environment interactions with the core system does not significantly increase the computation time compared to canonical ADC, in particular the cost of the ADC calculation is virtually the same as without PE. As an outlook on possible future applications, our PE-ADC approach works independently of the ADC variant, i.e., resolution-of-the-identity (RI) ADC<sup>110,111</sup> and core-valence-separation (CVS) ADC<sup>112,113</sup> could just be invoked in combination with PE without additional derivation of equations or programming. The procedure to perform PE-ADC calculations was outlined, where we took advantage of automated environment parametrization with PyFraME.<sup>72</sup> Recently, a tutorial review on performing PE calculations was published,<sup>99</sup> explaining the setup and caveats of these calculations in general. We believe that the PE-ADC approach can be employed to study a plethora of biologically relevant systems and will thus be of interest for a broad user base seeking highly accurate excited state descriptions in complex, polarizable environments.

## Acknowledgement

The authors thank Adrian L. Dempwolff, Erik Rosendahl Kjellgren, and Dr. Tim Stauch for fruitful discussions. Simulations and computations for the work described in this paper were supported by the DeIC National HPC Centre, SDU. Support from H2020-MSCA-ITN-2017 training network “COSINE – Computational Spectroscopy In Natural sciences and Engineering” is also acknowledged.

## Supporting Information Available

Supporting Figures and Tables.

This material is available free of charge via the Internet at <http://pubs.acs.org/>.

## References

- (1) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chemie Int. Ed.* **2009**, *48*, 1198–1229.
- (2) Wallrapp, F. H.; Guallar, V. Mixed quantum mechanics and molecular mechanics methods: Looking inside proteins. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 315–322.
- (3) Van Der Kamp, M. W.; Mulholland, A. J. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* **2013**, *52*, 2708–2728.
- (4) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013**, *139*.
- (5) Brunk, E.; Rothlisberger, U. Mixed quantum mechanical/molecular mechanical molecular dynamics simulations of biological systems in ground and electronically excited states. *Chem. Rev.* **2015**, *115*, 6217–6263.
- (6) Morzan, U. N.; Alonso de Armiño, D. J.; Foglia, N. O.; Ramírez, F.; González Lebrero, M. C.; Scherlis, D. A.; Estrin, D. A. Spectroscopy in Complex Environments from QM–MM Simulations. *Chem. Rev.* **2018**, *118*, 4071–4113.
- (7) Gomes, A. S. P.; Jacob, C. R. Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. *Annu. Reports Sect. "C" (Physical Chem.)* **2012**, *108*, 222–277.

- (8) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.
- (9) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (10) Mennucci, B. Polarizable continuum model. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 386–404.
- (11) Stamm, B.; Lagardère, L.; Scalmani, G.; Gatto, P.; Cancès, E.; Piquemal, J.-P.; Maday, Y.; Mennucci, B.; Lipparini, F. How to make continuum solvation incredibly fast in a few simple steps: a practical guide to the domain decomposition paradigm for the Conductor-like Screening Model Continuum Solvation, Linear Scaling, Domain Decomposition. *Int. J. Quantum Chem.* **2018**, in press.
- (12) Mennucci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Gabriel, S.; Stephens, P. J. Polarizable Continuum Model (PCM) Calculations of Solvent Effects on Optical Rotations of Chiral Molecules. *J. Phys. Chem. A* **2002**, *106*, 6102–6113.
- (13) Klein, R. A.; Mennucci, B.; Tomasi, J. Ab Initio Calculations of  $^{17}\text{O}$  NMR-Chemical Shifts for Water. The Limits of PCM Theory and the Role of Hydrogen-Bond Geometry and Cooperativity. *J. Phys. Chem. A* **2004**, *108*, 5851–5863.
- (14) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited States in Solution through Polarizable Embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721–3734.
- (15) Olsen, J. M. H.; Kongsted, J. Molecular properties through polarizable embedding. In *Adv. Quantum Chem.*; Elsevier, 2011; Vol. 61; pp 107–143.

- (16) Schröder, H.; Schwabe, T. Corrected Polarizable Embedding: Improving the Induction Contribution to Perichromism for Linear Response Theory. *J. Chem. Theory Comput.* **2018**, *14*, 833–842.
- (17) Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. Scrutinizing the effects of polarization in QM/MM excited state calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551.
- (18) Sneskov, K.; Schwabe, T.; Kongsted, J.; Christiansen, O. The polarizable embedding coupled cluster method. *J. Chem. Phys.* **2011**, *134*.
- (19) Schwabe, T.; Sneskov, K.; Hugaard Olsen, J. M.; Kongsted, J.; Christiansen, O.; Hättig, C. PERI-CC2: A polarizable embedded RI-CC2 method. *J. Chem. Theory Comput.* **2012**, *8*, 3274–3283.
- (20) Eriksen, J. J.; Sauer, S. P.; Mikkelsen, K. V.; Jensen, H. J.; Kongsted, J. On the importance of excited state dynamic response electron correlation in polarizable embedding methods. *J. Comput. Chem.* **2012**, *33*, 2012–2022.
- (21) Pedersen, M. N.; Hedegård, E. D.; Olsen, J. M. H.; Kauczor, J.; Norman, P.; Kongsted, J. Damped response theory in combination with polarizable environments: The polarizable embedding complex polarization propagator method. *J. Chem. Theory Comput.* **2014**, *10*, 1164–1171.
- (22) Steindal, A. H.; Beerepoot, M. T. P.; Ringholm, M.; List, N. H.; Ruud, K.; Kongsted, J.; Olsen, J. M. H. Open-ended response theory with polarizable embedding: multiphoton absorption in biomolecular systems. *Phys. Chem. Chem. Phys.* **2016**, *18*, 28339–28352.
- (23) Nåbo, L. J.; List, N. H.; Steinmann, C.; Kongsted, J. Computational Approach to Evaluation of Optical Properties of Membrane Probes. *J. Chem. Theory Comput.* **2017**, *13*, 719–726.

- (24) Nåbo, L. J.; Modzel, M.; Krishnan, K.; Covey, D. F.; Fujiwara, H.; Ory, D. S.; Szomek, M.; Khandelia, H.; Wüstner, D.; Kongsted, J. Structural design of intrinsically fluorescent oxysterols. *Chem. Phys. Lipids* **2018**, *212*, 26–34.
- (25) Nørby, M. S.; Steinmann, C.; Olsen, J. M. H.; Li, H.; Kongsted, J. Computational Approach for Studying Optical Properties of DNA Systems in Solution. *J. Chem. Theory Comput.* **2016**, *12*, 5050–5057.
- (26) List, N. H.; Olsen, J. M. H.; Jensen, H. J. A.; Steindal, A. H.; Kongsted, J. Molecular-level insight into the spectral tuning mechanism of the dsred chromophore. *J. Phys. Chem. Lett.* **2012**, *3*, 3513–3521.
- (27) Beerepoot, M. T. P.; Steindal, A. H.; Kongsted, J.; Brandsdal, B. O.; Frediani, L.; Ruud, K.; Olsen, J. M. H. A polarizable embedding DFT study of one-photon absorption in fluorescent proteins. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4735.
- (28) Schwabe, T.; Beerepoot, M. T. P.; Gvan, J.; Haugaard, M.; Cd, O.; Kongsted, J. Analysis of computational models for an accurate study of electronic excitations in GFP. *Phys. Chem. Chem. Phys.* **2015**, *17*, 2582–2588.
- (29) Sneskov, K.; Olsen, J. M. H.; Schwabe, T.; Hättig, C.; Christiansen, O.; Kongsted, J. Computational screening of one- and two-photon spectrally tuned channelrhodopsin mutants. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7567.
- (30) List, N. H.; Olsen, J. M. H.; Kongsted, J. Excited states in large molecular systems through polarizable embedding. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20234–20250.
- (31) Dreuw, A.; Head-Gordon, M. Failure of Time-Dependent Density Functional Theory for Long-Range Charge-Transfer Excited States: The Zincbacteriochlorin-Bacteriochlorin and Bacteriochlorophyll-Spheroidene Complexes. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.

- (32) Dreuw, A.; Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* **2005**, *105*, 4009–4037.
- (33) Mewes, S. A.; Plasser, F.; Dreuw, A. Communication: Exciton analysis in time-dependent density functional theory: How functionals shape excited-state characters. *J. Chem. Phys.* **2015**, *143*, 171101.
- (34) Mewes, S. A.; Plasser, F.; Krylov, A.; Dreuw, A. Benchmarking Excited-State Calculations Using Exciton Properties. *J. Chem. Theory Comput.* **2018**, *14*, 710–725.
- (35) Dreuw, A.; Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *5*, 82–95.
- (36) Harbach, P. H.; Wormit, M.; Dreuw, A. The third-order algebraic diagrammatic construction method (ADC(3)) for the polarization propagator for closed-shell molecules: Efficient implementation and benchmarking. *J. Chem. Phys.* **2014**, *141*, 064113.
- (37) Mewes, J. M.; You, Z. Q.; Wormit, M.; Kriesche, T.; Herbert, J. M.; Dreuw, A. Experimental benchmark data and systematic evaluation of two a posteriori, polarizable-continuum corrections for vertical excitation energies in solution. *J. Phys. Chem. A* **2015**, *119*, 5446–5464.
- (38) Mewes, J.-M.; Herbert, J. M.; Dreuw, A. On the accuracy of the general, state-specific polarizable-continuum model for the description of correlated ground- and excited states in solution. *Phys. Chem. Chem. Phys.* **2017**, *19*, 1644–1654.
- (39) Lunkenheimer, B.; Köhn, A. Solvent Effects on Electronically Excited States Using the Conductor- Like Screening Model and the Second-Order Correlated Method. *J. Chem. Theory Comput.* **2013**, *9*, 977–994.

- (40) Karbalaeei Khani, S.; Marefat Khah, A.; Hättig, C. COSMO-RI-ADC(2) excitation energies and excited state gradients. *Phys. Chem. Chem. Phys.* **2018**, *20*, 16354–16363.
- (41) Prager, S.; Zech, A.; Aquilante, F.; Dreuw, A.; Wesolowski, T. A. First time combination of frozen density embedding theory with the algebraic diagrammatic construction scheme for the polarization propagator of second order. *J. Chem. Phys.* **2016**, *144*, 204103.
- (42) Prager, S.; Zech, A.; Wesolowski, T. A.; Dreuw, A. Implementation and Application of the Frozen Density Embedding Theory with the Algebraic Diagrammatic Construction Scheme for the Polarization Propagator up to Third Order. *J. Chem. Theory Comput.* **2017**, *13*, 4711–4725.
- (43) Scheurer, M.; Brisker-Klaiman, D.; Dreuw, A. Molecular Mechanism of Flavin Photoprotection by Archaeal Dodecin: Photo-Induced Electron Transfer and Mg<sup>2+</sup>-Promoted Proton Transfer. *J. Phys. Chem. B* **2017**, *121*, 10457–10466.
- (44) Saint Raymond, X. *Elementary Introduction to the Theory of Pseudodifferential Operators*; CRC Press: Boca Raton, 1991.
- (45) Applequist, J.; Carl, J. R.; Fung, K.-K. Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities. *J. Am. Chem. Soc.* **1972**, *94*, 2952–2960.
- (46) Schirmer, J. Beyond the random-phase approximation: A new approximation scheme for the polarization propagator. *Phys. Rev. A* **1982**, *26*, 2395–2416.
- (47) Schirmer, J.; Trofimov, A. B. Intermediate state representation approach to physical properties of electronically excited molecules. *J. Chem. Phys.* **2004**, *120*, 11449–11464.

- (48) Wormit, M.; Rehn, D. R.; Harbach, P. H.; Wenzel, J.; Krauter, C. M.; Epifanovsky, E.; Dreuw, A. Investigating excited electronic states using the algebraic diagrammatic construction (ADC) approach of the polarisation propagator. *Mol. Phys.* **2014**, *112*, 774–784.
- (49) Trofimov, A. B.; Schirmer, J. An efficient polarization propagator approach to valence electron excitation spectra. *J. Phys. B At. Mol. Opt. Phys.* **1995**, *28*, 2299–2324.
- (50) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. An effective fragment method for modeling solvent effects in quantum mechanical calculations. *J. Chem. Phys.* **1996**, *105*, 1968–1986.
- (51) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. The Effective Fragment Potential Method: A QM-Based MM Approach to Modeling Environmental Effects in Chemistry. *J. Phys. Chem. A* **2001**, *105*, 293–307.
- (52) Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. The Effective Fragment Potential: A General Method for Predicting Intermolecular Interactions. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177–193.
- (53) Gordon, M. S.; Mullin, J. M.; Pruitt, S. R.; Roskop, L. B.; Slipchenko, L. V.; Boatz, J. A. Accurate methods for large molecular systems. *J. Phys. Chem. B* **2009**, *113*, 9646–9663.
- (54) Ghosh, D.; Kosenkov, D.; Vanovschi, V.; Williams, C. F.; Herbert, J. M.; Gordon, M. S.; Schmidt, M. W.; Slipchenko, L. V.; Krylov, A. I. Noncovalent interactions in extended systems described by the effective fragment potential method: Theory and application to nucleobase oligomers. *J. Phys. Chem. A* **2010**, *114*, 12739–12754.
- (55) Slipchenko, L. V. Solvation of the excited states of chromophores in polarizable envi-

- ronment: Orbital relaxation versus polarization. *J. Phys. Chem. A* **2010**, *114*, 8824–8830.
- (56) Kosenkov, D.; Slipchenko, L. V. Solvent Effects on the Electronic Transitions of p-Nitroaniline : A QM / EFP Study. *J. Phys. Chem. A* **2010**, *115*, 392–401.
- (57) Schwabe, T. General theory for environmental effects on (vertical) electronic excitation energies. *J. Chem. Phys.* **2016**, *145*, 154105.
- (58) Caricato, M.; Mennucci, B.; Tomasi, J.; Ingrosso, F.; Cammi, R.; Corni, S.; Caricato, M.; Mennucci, B.; Tomasi, J. 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*, 124520.
- (59) Olsen, J. M. H.; List, N. H.; Steinmann, C.; Steindal, A. H.; Nørby, M. S. PELib: The Polarizable Embedding library. 2018; <https://doi.org/10.5281/zenodo.1209196>.
- (60) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W.; Harbach, P. H.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyayev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.;

Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscammann, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M.; Head-Gordon, M. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

- (61) Trofimov, A. B.; Stelter, G.; Schirmer, J. A consistent third-order propagator method for electronic excitation. *J. Chem. Phys.* **1999**, *111*, 9982–9999.
- (62) Hermann, G.; Pohl, V.; Tremblay, J. C.; Paulus, B.; Hege, H. C.; Schild, A. ORBKIT: A modular python toolbox for cross-platform postprocessing of quantum chemical wavefunction data. *J. Comput. Chem.* **2016**, *37*, 1511–1520.
- (63) McKinney, W. Data Structures for Statistical Computing in Python. Proc. 9th Python Sci. Conf. 2010; pp 51–56.
- (64) McKinney, W. pandas: a Foundational Python Library for Data Analysis and Statistics. 2011.

- (65) Van Der Walt, S.; Colbert, S. C.; Varoquaux, G. The NumPy array: A structure for efficient numerical computation. *Comput. Sci. Eng.* **2011**, *13*, 22–30.
- (66) Oliphant, T. E. SciPy: Open source scientific tools for Python. *Comput. Sci. Eng.* **2007**, *9*, 10–20.
- (67) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 99–104.
- (68) Waskom, M.; Botvinnik, O.; O’Kane, D.; Hobson, P.; Lukauskas, S.; Gempertline, D. C.; Augspurger, T.; Halchenko, Y.; Cole, J. B.; Warmenhoven, J.; de Ruiter, J.; Pye, C.; Hoyer, S.; Vanderplas, J.; Villalba, S.; Kunter, G.; Quintero, E.; Bachant, P.; Martin, M.; Meyer, K.; Miles, A.; Ram, Y.; Yarkoni, T.; Williams, M. L.; Evans, C.; Fitzgerald, C.; Brian,; Fonnesbeck, C.; Lee, A.; Qalieh, A. Mwaskom/Seaborn: V0.8.1. 2017; <https://doi.org/10.5281/zenodo.883859>.
- (69) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (70) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (71) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- (72) Olsen, J. M. H. PyFraME: Python tools for Fragment-based Multiscale Embedding (version 0.1.1). 2018; <https://doi.org/10.5281/zenodo.1168860>.
- (73) Gagliardi, L.; Lindh, R.; Karlström, G. Local properties of quantum chemical systems: The LoProp approach. *J. Chem. Phys.* **2004**, *121*, 4494–4500.

- (74) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (75) Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Ekström, U.; Enevoldsen, T.; Eriksen, J. J.; Ettenhuber, P.; Fernández, B.; Ferrighi, L.; Fliegl, H.; Frediani, L.; Hald, K.; Halkier, A.; Hättig, C.; Heiberg, H.; Helgaker, T.; Hennum, A. C.; Hettema, H.; Hjertenæs, E.; Høst, S.; Høyvik, I.-M.; Iozzi, M. F.; Jansík, B.; Jensen, H. J. Aa.; Jonsson, D.; Jørgensen, P.; Kauczor, J.; Kirpekar, S.; Kjærgaard, T.; Klopper, W.; Knecht, S.; Kobayashi, R.; Koch, H.; Kongsted, J.; Krapp, A.; Kristensen, K.; Ligabue, A.; Lutnæs, O. B.; Melo, J. I.; Mikkelsen, K. V.; Myhre, R. H.; Neiss, C.; Nielsen, C. B.; Norman, P.; Olsen, J.; Olsen, J. M. H.; Osted, A.; Packer, M. J.; Pawłowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybkin, V. V.; Salek, P.; Samson, C. C. M.; de Merás, A. S.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sneskov, K.; Steindal, A. H.; Sylvester-Hvid, K. O.; Taylor, P. R.; Teale, A. M.; Tellgren, E. I.; Tew, D. P.; Thorvaldsen, A. J.; Thøgersen, L.; Vahtras, O.; Watson, M. A.; Wilson, D. J. D.; Ziolkowski, M.; Ågren, H. The Dalton quantum chemistry program system. *WIREs Comput. Mol. Sci.* **2014**, *4*, 269–284.
- (76) Vahtras, O. LoProp for Dalton. 2014; <https://doi.org/10.5281/zenodo.13276>.
- (77) Plasser, F.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations. I. Formalism. *J. Chem. Phys.* **2014**, *141*, 024106.
- (78) Plasser, F.; Bäppler, S. A.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations. II. Applications. *J. Chem. Phys.* **2014**, *141*, 024107.

- (79) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.
- (80) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (81) Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*.
- (82) Sure, R.; Grimme, S. Corrected small basis set Hartree-Fock method for large systems. *J. Comput. Chem.* **2013**, *34*, 1672–1685.
- (83) Martinez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **2009**, *30*, 2157–2164.
- (84) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (85) Huang, J.; Mackerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34*, 2135–2145.
- (86) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: A fast force field generation tool for small organic molecules. *J. Comput. Chem.* **2011**, *32*, 2359–2368.
- (87) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **1994**, *101*, 4177–4189.
- (88) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *The Journal of chemical physics* **1995**, *103*, 4613–4621.

- (89) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (90) Andersen, H. C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983**, *52*, 24–34.
- (91) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (92) Melo, M. C. R.; Bernardi, R. C.; Rudack, T.; Scheurer, M.; Riplinger, C.; Phillips, J. C.; Maia Julio D C, R. G. B.; Ribeiro, J. V.; Stone, J. E.; Neese, F.; Schulten, K.; Luthey-Schulten, Z. NAMD goes quantum: An integrative suite for QM/MM simulations. *Nat. Methods* **2018**, *15*, 351–354.
- (93) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.
- (94) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domański, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python Sci. Conf.* **2016**, 98–105.
- (95) Beerepoot, M. T.; Steindal, A. H.; Ruud, K.; Olsen, J. M. H.; Kongsted, J. Convergence of environment polarization effects in multiscale modeling of excitation energies. *Comput. Theor. Chem.* **2014**, *1040-1041*, 304–311.
- (96) Sherwood, P.; H. de Vries, A.; J. Collins, S.; P. Greatbanks, S.; A. Burton, N.; A. Vincent, M.; H. Hillier, I. Computer simulation of zeolite structure and reactivity using embedded cluster methods. *Faraday Discuss.* **1997**, *106*, 79–92.

- (97) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of chemical physics* **1993**, *98*, 5648–5652.
- (98) Beerepoot, M. T.; Steindal, A. H.; List, N. H.; Kongsted, J.; Olsen, J. M. H. Averaged Solvent Embedding Potential Parameters for Multiscale Modeling of Molecular Properties. *Journal of chemical theory and computation* **2016**, *12*, 1684–1695.
- (99) Steinmann, C.; Reinholdt, P.; Nørby, M. S.; Kongsted, J.; Olsen, J. M. H. Response properties of embedded molecules through the polarizable embedding model. **2018**, <https://arxiv.org/abs/1804.03598>.
- (100) Yang, C.; Dreuw, A. Evaluation of the restricted virtual space approximation in the algebraic-diagrammatic construction scheme for the polarization propagator to speed-up excited-state calculations. *J. Comput. Chem.* **2017**, *38*, 1528–1537.
- (101) Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. Polarizable Density Embedding: A Solution to the Electron Spill-Out Problem in Multiscale Modeling. *J. Phys. Chem. Lett.*
- (102) Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Åstrand, P. O.; Christiansen, O. Solvent effects on the  $n \pi^*$  electronic transition in formaldehyde: A combined coupled cluster/molecular dynamics study. *J. Chem. Phys.* **2004**, *121*, 8435–8445.
- (103) Kongsted, J.; Osted, A.; Pedersen, T. B.; Mikkelsen, K. V.; Christiansen, O. The  $n \pi^*$  electronic transition in microsolvated formaldehyde. A coupled cluster and combined coupled cluster/molecular mechanics study. *J. Phys. Chem. A* **2004**, *108*, 8624–8632.
- (104) Thomsen, C. L.; Thøgersen, J.; Keiding, S. R. Ultrafast Charge-Transfer Dynamics: Studies of *p*-Nitroaniline in Water and Dioxane. *J. Phys. Chem. A* **1998**, *102*, 1062–1067.
- (105) Kovalenko, S.; Schanz, R.; Farztdinov, V.; Hennig, H.; Ernsting, N. Femtosecond relax-

- ation of photoexcited para-nitroaniline: solvation, charge transfer, internal conversion and cooling. *Chem. Phys. Lett.* **2000**, *323*, 312–322.
- (106) Hättig, C. Structure optimizations for excited states with correlated second-order methods: CC2 and ADC (2). *Adv. Quantum Chem.* **2005**, *50*, 37–60.
- (107) Grininger, M.; Staudt, H.; Johansson, P.; Wachtveitl, J.; Oesterhelt, D. Dodecin is the key player in flavin homeostasis of archaea. *J. Biol. Chem.* **2009**, *284*, 13068–13076.
- (108) Staudt, H.; Oesterhelt, D.; Grininger, M.; Wachtveitl, J. Ultrafast excited-state deactivation of flavins bound to dodecin. *J. Biol. Chem.* **2012**, *287*, 17637–17644.
- (109) Staudt, H.; Hoesl, M. G.; Dreuw, A.; Serdjukow, S.; Oesterhelt, D.; Budisa, N.; Wachtveitl, J.; Grininger, M. Directed manipulation of a flavoprotein photocycle. *Angew. Chemie - Int. Ed.* **2013**, *52*, 8463–8466.
- (110) Hättig, C. Structure optimizations for excited states with correlated second-order methods: CC2 and ADC (2). *Adv. Quantum Chem.* **2005**, *50*, 37–60.
- (111) Wormit, M. Development and application of reliable methods for the calculation of excited states : from light-harvesting complexes to medium-sized molecules. Ph.D. thesis, 2009.
- (112) Wenzel, J.; Wormit, M.; Dreuw, A. Calculating X-ray absorption spectra of open-shell molecules with the unrestricted algebraic-diagrammatic construction scheme for the polarization propagator. *J. Chem. Theory Comput.* **2014**, *10*, 4583–4598.
- (113) Wenzel, J.; Wormit, M.; Dreuw, A. Calculating core-level excitations and x-ray absorption spectra of medium-sized closed-shell molecules with the algebraic-diagrammatic construction scheme for the polarization propagator. *J. Comput. Chem.* **2014**, *35*, 1900–1915.

# Graphical TOC Entry

