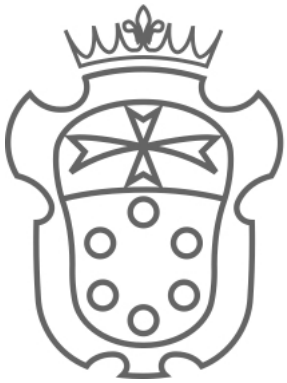


SCUOLA NORMALE SUPERIORE

Classe di Scienze

Corso di perfezionamento in Methods and Models for Molecular Science

XXXII Ciclo

Polarizable atomistic approaches to model electronic properties of complex molecular systems

Settore Scientifico Disciplinare Chim/02

Candidato
Dr. Matteo Ambrosetti

Relatore
Prof. Chiara Cappelli

Anno accademico 2021/2022

Contents

List of publications

1. *Fully polarizable QM/fluctuating charge approach to two-photon absorption of aqueous solutions* - 2019 J. Chem. Theory Comput. (**Chapter 1**)

2. *A polarizable embedding approach to second harmonic generation (SHG) of molecular systems in aqueous solutions* - 2018 Theoretical Chemistry Accounts (**Chapter 1**)

3. *QM/Fluctuating Charge Protocol to Compute Solvatochromic Shifts* - 2021 J. Chem. Theory Comput. (**Chapter 2**)

4. *Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQFμ model* - 2019 J. Chem. Theory Comput. (**Chapter 3**)

5. *Electronic transitions for a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles: linear and corrected linear response regimes* - 2019 J. Chem. Phys. (**Chapter 3**)

6. *Calculation of ir spectra with a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles* - 2019 J. Chem. Theory Comput. (**Chapter 3**)

7. *Quantum confinement effects on solvatochromic shifts of molecular solutes* - 2020 J. Phys. Chem. Lett. (**Chapter 4**)

8. *A 4,4 0 -bis(2-benzoxazolyl)stilbene luminescent probe: assessment of aggregate formation through photophysics experiments and quantum-chemical calculations Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling* - 2018 Phys. Chem. Chem. Phys.

9. *Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling* - 2019 Int. J. Quantum Quantum Chem.

10. *Investigation of electronic energy transfer in a BODIPY-decorated calix[4]arene* - 2019 Dyes and Pigments

11. *Simulating Absorption Spectra of Flavonoids in Aqueous Solution: A Polarizable QM/MM Study* - 2020 Molecules

Al ragazzo che ero,
all'uomo che sono

Introduction

In 2013 the Nobel prize for chemistry was awarded to Martin Karplus, Michael Levitt and Arieh Warshel "*for the development of multiscale models for complex chemical systems*". [1]

The meaning of *multiscale models* is twofold. It can describe a model which is able to treat at least two different computational approaches, and use them to describe different portions of the same system at the same time. As an alternative, it may refer to a model where the output obtained from a highly accurate calculation is used to parameterize another model, which is less accurate but able to deal with bigger systems on a longer time scale at the same computational cost.

The term *complex chemical systems* has instead a more vague definition. The complexity of the system could be related to its dimensions (i.e., the large number of atoms/molecules that needs to be considered in order to have a reliable representation of the physico-chemical phenomenon) or to the interactions occurring at the molecular level, which may not be included in the adopted model or even to the description of the effect of an external perturbation, such as an external electric field, and its effect on the system's properties.

The main goal of this thesis work is the development of a theoretical framework for the description of both energy and response properties of complex molecular systems by means of multiscale approaches.

In multiscale models the chemical system is generally divided into a variable number of portions, which are modeled using different levels of theory, namely, quantum mechanics (QM) and molecular mechanics (MM). The principal benefit of such a partitioning is the possibility of reducing the computational cost without loosing the information related to the properties of interest.

A typical application of multiscale modeling is the study of electronic properties of solvated systems. In this case, it is common practice to describe the solute at the QM level, whereas the solvent is described classically. In this way all the electronic degrees of freedom related to the solute are considered explicitly, whereas the presence of the environment, i.e., the solvent, is modeled as an external "perturbation" acting on the system under study. Solvent effects can be described in different ways. The most adopted approaches describe the solvent either as a dielectric medium, e.g., by means of the Polarizable Continuum Model, PCM, giving rise to the QM/PCM method, or atomistically, thus defining a QM/MM approach. [2–15] The MM atomistic layer can be described in terms of fixed point charges (Electrostatic Embedding, EE) or by exploiting more sophisticated methods belonging to the family of the so-called Polarizable Embedding (PE) approaches. [16–31]

Although the distinctive inability of PCM to take into account solute-solvent specific interactions is recovered by exploiting EE approaches, there is still a strong simplification in adopting such a model within a QM/MM framework, i.e., the assumption that the QM portion is perturbed by the presence of the MM environment, whereas the MM portion is not perturbed by the presence of the QM electronic density. In other words, solute-solvent mutual polarization effects are not considered, thus yielding a physically inconsistent description of the system.

For this reason, this thesis focuses on PE approaches to solvation, and in particular on the Fluctuating Charge (FQ) model. [27, 28] The polarization is in this case introduced by means of two fundamental parameters, which control the charge flow among the atoms of the MM portion, namely, electronegativity and chemical hardness.

The starting point is the definition and partitioning of the energy of the QM/MM system into three contributions:

$$E = E_{QM} + E_{MM} + E_{QM/MM} \tag{1}$$

where E_{QM} and E_{MM} are the energies of the QM and MM portions, respectively, whereas $E_{QM/MM}$ represents the interaction energy between the two moieties. $E_{QM/MM}$ is generally partitioned as:

$$E_{QM/MM} = E^{Ele}_{QM/MM} + E^{Pol}_{QM/MM} + E^{Rep}_{QM/MM} + E^{Dis}_{QM/MM} \tag{2}$$

where $E^{Ele}_{QM/MM}$ is the electrostatic term and $E^{Pol}_{QM/MM}$ is the polarization energy. The last two terms $E^{Rep}_{QM/MM}$ and $E^{Dis}_{QM/MM}$ represent the so-called Pauli repulsion and dispersion energies, respectively; they can be grouped together and refereed to as "van der Waals" contributions.

The different QM/MM approaches can be hierarchically classified based on how the two (QM and MM) portions are individually modelled, but also based on the specific formulation of $E^{Ele}_{QM/MM}$ and $E^{Pol}_{QM/MM}$. Three principal schemes are exploited, i.e. the mechanical embedding (ME), electrostatic embedding (EE) and polarizable embedding (PE).
In the mechanical embedding scheme QM computations for the QM portion are performed in the absence of the MM portion, and the interaction between the QM and MM portions are treated at the MM level. [32, 33] These interactions usually include both bonded (stretching, bending, and torsional) and non-bonded (electrostatic and van der Waals) interactions. ME handles the interaction at the MM level, however such a treatment has drawbacks. First, it requires an accurate set of MM parameters such as atom-centered point charges for both the QM and MM portions. Second, ME ignores the perturbation of the electronic structure of the QM portion due to the electrostatic interaction with the MM portion (the atom-centred charges in the MM portion do not polarize the QM portion).
In the EE scheme the QM computation for the QM portion is carried out in the presence of the MM portion, by including the electrostatic interaction between the QM electronic density and MM point charges in termso of one-electron operators that enter the QM Hamiltonian. [34, 35] In this way the polarization of the QM portion by the charge distribution of the MM portion is taken into account automatically. Essentially, a QM code must be able to perform the self-consistent-field (SCF) treatment in the presence of the external point-charge field, that represents the MM charge model. EE is limited to the first term of Equation 2, which can be expressed as:

$$E^{EE}_{QM/MM} = E^{Ele}_{QM/MM}(\rho_{QM}) = \sum_{i}^{N_q} q_i V_i(\rho_{QM}) \tag{3}$$

In particular, the MM portion, and the resulting QM/MM electrostatic interaction energy, are represented by means of a set of fixed charges, placed on the MM atoms, which interact with the QM density, ρ_{QM}. The index i runs over the number of charges N_q, and $V_i(\rho_{QM})$ is the QM electric potential calculated at the i-th charge q_i placed in the MM portion. $E^{Ele}_{QM/MM}(\rho_{QM})$ is expressed in terms of Coulomb's law,

and it is explicitly dependent on the QM density. The use of MM atom-centered partial charges is very efficient, and it is the most popular way to construct the effective QM Hamiltonian. MM point charges actually include the contributions due to higher-order multipoles implicitly, i.e., the higher-order contributions are folded into the zero-order parameters. It is also possible to include higher-order multipole contributions explicitly, but this results in harder implementations and increase in the computational cost. Furthermore, it is important to stress that higher-order terms are generally much more sensitive to geometry or conformational changes, thus limiting the transferability of the model. The bonded (stretching, bending, and torsional) interactions and non-bonded van der Waals interactions between the QM and MM portions are retained at the MM level. As in the case of ME, MM parameters are generally taken from the most popular force fields.
From a more physically oriented point of view, QM and MM portions will polarize each other until their charge distributions are self-consistent.

In PE, the coupling between the two portions is expressed as:

$$E^{PE}_{QM/MM} = E^{Ele}_{QM/MM}(\rho_{QM}) + E^{Pol}_{QM/MM}(\rho_{QM}) = \sum_{i}^{N_q} q_i V_i(\rho_{QM}) + \sum_{i} x_i(\rho_{QM}) s_i(\rho_{QM}) \tag{4}$$

The various PE approaches which have been proposed in the literature differ in the way they define the polarization energy $E^{Pol}_{QM/MM}$ in terms of the QM density, ρ_{QM}. In particular, they prescribe different specifications for the electrostatic quantities x and the QM electric sources s, both employed to represent the MM portion.
In the present work, we will use the FQ model, which is based on the electronegativity equalization (EE) principle. [36, 37] The atomic charges are redistributed to equalize the electronegativity/chemical potential at each site.
The energy required to create a charge q on an atom can be expressed by a Taylor expansion truncated at the second order as:

$$E(q) = E^0 + q\left(\frac{\partial E}{\partial q}\right) + \frac{1}{2}q^2\left(\frac{\partial^2 E}{\partial^2 q}\right) = E^0 + q\chi^0 + \frac{1}{2}q^2 J \tag{5}$$

where χ^0 represents the Mulliken's definition of electronegativity, [38] whereas J depends on the hardness ($\eta = J/2$). [39]
The electrostatic energy of an atomic system composed by a set of M molecules each containing N_α atoms can be computed as follows:

$$E(\boldsymbol{q}, \boldsymbol{r}) = \sum_{\alpha}^{M}\sum_{i}^{N_\alpha}\left(\chi^0_{\alpha,i} q_{\alpha,i} + \frac{1}{2}J_{\alpha,ii} q^2_{\alpha,i}\right) + \sum_{\alpha}^{M}\sum_{\beta}^{M}\sum_{i}^{N_\alpha}\sum_{j>i}^{N_\beta} J_{\alpha\beta,ij}(\boldsymbol{r}_{\alpha\beta,ij}) q_{\alpha,i} q_{\beta,j} \tag{6}$$

where $\mathbf{q}$ is a vector representing a set of charges, whereas $J_{\alpha\beta,ij}(\mathbf{r}_{\alpha\beta,ij})$ takes the name of interaction kernel and can assume different functional forms. [40–42] Optimum charge distribution is achieved by minimizing the energy expression reported in Equation 6 with respect to the charges on each atom, i.e.:

$$\frac{\partial E(\boldsymbol{q}, \boldsymbol{r})}{\partial q_{\alpha,i}} = 0 \tag{7}$$

that is another way of equalizing electronegativities (see Equation 5). This kind of minimization problem can be solved by matrix inversion, iteratively or exploiting the extended Lagrangian methods.

In order to capture solute-solvent interactions by means of computational approaches, a sensible description of the solvated systems need to be achieved. In fact, in most cases, a systems composed of a single solute molecule surrounded by few nearby solvent molecules is not a good approximation of the interactions taking place in the actual sample. On the other hand, accurate calculations on a very large number of molecules cannot be computationally afforded. It should be additionally emphasized that, contrary to bulk properties which are characteristics of the whole solution, the origin of a spectral response is in most cases the solute molecule. One of the principal consequences of this solute-localized spectral response is the possibility of reducing the degrees of freedom needed to accurately describe the system, i.e. by using a QM description for the solute and a classical approach for the solvent. This approach identifies to the so-called "focused model" where the focus is always the solute molecule and its interactions with the solvent.
Before moving to a more detailed discussion of the techniques and computational tools adopted and developed in this thesis, it is important to summarize the computational workflow applied in all the studies presented in the following chapters.

1. Definition of the QM and MM portions: the part of the system to be treated as the QM portion, and that defined as the MM portion, need to be accurately chosen. The definition may change depending on the system's chemical properties, its potential interactions with the surrounding environment, and the final spectral property to be modelled. As a general approach, the solute molecule is always included in the QM portion. In some special cases, the QM portion can be extended to the nearest solvent molecules, e.g. in the case of an aqueous environment the water molecules forming H-bonds with the solute may strongly influence its electron density distribution and consequently the final spectra. The boundary between the two moieties (i.e., the QM/MM boundary) is defined accordingly.

2. Conformational sampling: the solute molecule may generally be described in a rigid, semi-rigid, or flexible fashion. A highly rigid structure may reasonably be modelled by freezing its geometry into a minimum energy structure (possibly optimized at a given QM level). In this way it is possible to directly observe the contribution to the final spectra arising from solvation. On the other hand, since only one solute structure is considered, this approach should be applied with care to flexible systems. For the latter, the configurational space needs to be reliably sampled. To avoid spurious effects coming from solute-solute interactions, the system is generally composed of a single solute molecule. This assumption corresponds to the case of extremely diluted systems, where each solute molecule is surrounded by an almost infinite number of bulk solvent molecules. In addition to intramolecular conformational changes, a reliable sampling of solvent configurations around the solute is of out-most importance. An effective way of sampling the phase space of a given solution consists of resorting to classical Molecular Dynamics, where temperature and pressure are chosen to mirror the desired experimental setup.

For many common systems, such as organic and biological molecules, suitable parametrizations are available in the literature, whereas for less common systems it is possible to either use a general-purpose force field or, in the worst case, parametrize a new force field. Once the force field and boundary conditions (e.g., temperature, pressure, number of solvent molecules etc.) are defined, the simulation may be run. It is imperative for such simulations to be long enough to take into account the principal conformational changes of the investigated molecules and that the simulation parameters correctly reproduce all possible system configurations and their relative energy. Note that the force field used in the MD sampling may differ from the one employed in subsequent QM/MM spectroscopy calculations. The force field used for the MD sampling and the force field used in the QM/MM calculations have different purposes, and generally a different functional form. The former is employed to obtain an accurate conformational sampling, while the latter has to accurately model the effect of the solvent upon the spectroscopic observables.

3. Extraction of representative structures: a number of structures (snapshots) are extracted from MD runs and employed for the subsequent QM/MM calculations. Two slightly different approaches can be adopted to generate these structures: I) solute-centered sphere and II) skin-shell shape. In the first case, which is generally applied to small or globular molecules, the snapshots have the shape of spherical "droplets" obtained by cutting spheres of given radii centered on the solute. In the second case, which is more appropriate for molecules extending along a preferential direction (e.g., linear molecules), all the solvent molecules within a specific distance from each atom of the solute are considered. Although the first approach benefits from the spherical symmetry of the solvent, which better mimic an homogeneous environment, the second approach has the advantage of reducing the number of solvent molecules included in the MM portion without loosing accuracy in the description of the environment. The radius of the droplet is chosen to be large enough to retain solute-solvent interactions in a physically consistent way, and is usually of the order of tens of Ångstrom. The total number of snapshots to be extracted, which constitutes the computational sample, is chosen so to reach the convergence of the desired property/spectroscopy (steps 4 and 5 below). Notice that such a number is highly variable (from hundreds to thousands of snapshots), and it strongly depends on the property/spectroscopy to be computed.

4. QM/MM calculations: once the snapshots have been extracted from the MD trajectory, a QM/MM calculation of the target property is performed on each of them. This step requires to choose the QM level of theory, which can depend on the system and the desired spectroscopic response, and should follow state-of-the-art of QM calculations of the same property for isolated systems. Then, a model for the coupling of the QM and MM portions need to be defined.

5. Extraction of spectra: energies, structures, properties, and spectra obtained for each snapshot are extracted and averaged to produce final spectra. The results can then be analyzed and finally compared with experimental data. At this stage, any shortcomings of the procedure may emerge, e.g. insufficient number of snapshots, insufficiently long MD, poor choice of classical force field,

or inadequacies in the QM/MM or electronic structure methods. Then, the procedure may be restarted from the step(s) that need refinement.

At the time when this work started, there were three main issues affecting the QM/FQ model, namely: 1) absence of extension to non-linear optical properties, such as second harmonic generation and two photo absorption spectroscopy; 2) lack of reliable parameterization for non-aqueous solutions and 3) limitation in the treatment of out-of-plane polarization of the MM portion, which was based on point charges only, thus discarding higher orders (e.g. dipoles). These three limitations are discussed in details in the next chapters, and methods to treat such phenomena/effects are proposed.

The first original contribution that is presented in this Thesis (Chapter 1) reports on the extension of QM/FQ to the computation of one, two photon absorption (OPA and TPA) and second harmonic generation (SHG) spectra of molecular solutes in aqueous solution. The aim of this study is twofold; from one side it shows the ability of QM/FQ to accurately reproduce several electronic properties and spectroscopic observables. At the same time, QM/FQ is compared with other embedding approaches, such as QM/PCM or non-polarizable QM/MM approaches.
Chapter 2 solves a limitation of QM/FQ, i.e., the lack of reliable parametrization for non-aqueous environments. In particular, a machine learning-based parameterization procedure is proposed and applied to parametrize dioxane, tetrahydrofurane, acetonitrile, ethanol and methanol. Then QM/FQ and the new parametrization are challenged to reproduce solvatochromic shifts of organic dyes of increasing physico-chemical complexity.
QM/FQ models the MM classical environment through a set of charges. This description is improved in Chapter 3, where a novel fully polarizable approach based on fluctuating charges and fluctuating dipoles (FQFμ) is proposed and coupled to a quantum mechanical SCF Hamiltonian (QM/FQFμ). The peculiarity of the resulting QM/FQFμ approach stands in the fact that polarization effects are modeled in terms of both fluctuating charges and dipoles, which vary as a response to the external electric field/potential. The resulting QM/FQFμ model is then extended to the calculation of vertical excitation energies of solvated dyes under two different solvation regimes, i.e., linear response (LR), and corrected-Linear Response (cLR). In the same chapter, QM/FQFμ is further extended to the analytical evaluation of nuclear gradients and infrared (IR) spectra of molecular systems in aqueous solution. Finally, in Chapter 4 the purely electrostatic description of the solute-solvent interaction is questioned and the pivotal role of quantum confinement effects on the description of solvatochromic shifts of two prototypical organic molecules in aqueous environment is demonstrated. In this case particular attention is paid to dissect the effects of the single forces acting on the solute-solvent couple, thus allowing for a rationalization of experimental findings in terms of physico-chemical quantities.

Chapter 1

Non-linear spectroscopies

Fully Polarizable QM/Fluctuating Charge Approach to Two-Photon Absorption of Aqueous Solutions

Multiphoton absorption is the synchronous absorption of multiple photons leading to an excitation of a molecule from one electronic state to another. [43] Such an effect was originally predicted by Göppert-Mayer in 1931, [44] but only first measured in 1961 because its intensity was too weak to be detected before the advent of laser sources. [45] Among multiphoton processes, the most common is two-photon absorption (TPA), in which the simultaneous absorption of two photons takes place. [43] Nowadays, TPA measurements are not as common as compared to one-photon absorption (OPA), however the study of TPA processes is growing and rapidly becoming a well-established research field. [46, 47] TPA is governed by different symmetry selection rules, states that are dark in OPA experiments might thus be accessible in TPA. Furthermore, TPA is a non-linear process whose intensity depends on the square of the incoming light. This affords a greater spatial resolution than in OPA experiments. TPA has a number of technological applications, in particular in molecular devices. [43, 48–51] The design of molecular systems with large TPA cross sections is thus a challenge both from an experimental and computational point of view. [43, 52–66]

High TPA cross sections are generally measured for large chromophores, [61, 67] for which the computational description at high level of accuracy is difficult and sometimes not affordable. For this reason, most of the computational studies on this kind of spectroscopy have been performed by resorting to density functional theory (DFT), due to the good compromise between accuracy and computational cost. [52, 68, 69] In addition to the quantum-mechanical issue in the description of the target molecule, it is worth remarking that most of the experimental TPA cross sections are measured in the condensed phase. [63–65, 70, 71] For instance, a 50% increase in TPA cross sections has been reported by changing the solvent. [46] In order to successfully reproduce experimental data, such effects need to be taken into consideration.

The problem of treating solvent effects on observable properties is one of the pillars in quantum chemistry. The most successful approaches make use of multiscale and focused models, where the environment is treated at a lower level of accuracy with respect to the target molecule: [2, 6, 72–80] the latter is generally treated at the quantum-mechanical (QM) level, whereas the former is treated classically.

In the resulting QM/classical approaches, the classical portion can range from an atomistic description (giving rise to quantum mechanics/molecular mechanics (QM/MM) models[22, 72, 73, 81–85]) to a dielectric continuum (DC) description. Among the latter, the polarizable continuum model (PCM), [2–4, 86–89] in which the environment is depicted as a homogeneous continuum dielectric with given dielectric properties, has been particularly successful. In such an approach, the QM described target molecule is accommodated into a molecular shaped cavity. The QM electron density and the dielectric mutually polarize. The QM/PCM approach has been extended to the description of TPA spectra by some of the present authors, [56, 90–92] and an open-ended response formulation was put forth in a recent communication. [93] One of the main problems related to a continuum description of the environment is that all information about the atomistic structure of the environment is neglected. Thus, the specific molecule-environment interactions (*e. g.*

hydrogen bonding), cannot be described.

In order to recover the atomistic description of the environment, QM/MM is exploited, where the target molecule is still described at the QM level, whereas the environment is described by resorting to molecular mechanics (MM) force fields. [72, 74, 77, 94–99] In electrostatic QM/MM embedding approaches, a set of fixed charges is placed on the MM portion and the interaction between QM density and MM charges is introduced in the QM Hamiltonian. Mutual polarization, *i. e.* the polarization of the MM portion arising from the interaction with the QM density and viceversa, can be introduced by employing polarizable force fields. These can be based on distributed multipoles, [16–20] induced dipoles, [21, 24, 25] Drude oscillators, [26] capacitances and polarizabilities, [100–102] fluctuating charges (FQs) [27–29] or fluctuating charges and fluctuating dipoles (FQFμ). [30]

Thanks to the availability of a variational formulation of the quantum/classical polarizable coupling, the QM/FQ approach has been extended to the analytical calculation of a large variety of properties and spectroscopies: molecular gradients and Hessians, [103] linear response properties, [104, 105] including optical rotation, [106–108] and electronic circular dichroism, [109] vibrational circular dichroism, [110] third order mixed electric/magnetic/geometric properties [111, 112] and second harmonic generation. [113] Remarkably, QM/FQ has already been shown to accurately model some of the systems where PCM and other continuum models completely fail.

Non-polarizable QM/MM approaches and polarizable QM/MM based on induced dipoles have been extended to the calculation of TPA spectra of molecules in solution. [24, 85, 114] In this paper, we extend the QM/FQ model to the computation of TPA spectra. In particular, we have selected the challenging case of rhodamine 6G (R6G) in aqueous solution, which has been studied extensively both theoretically and experimentally. [58, 115–121] The large interest in such a molecule – in particular to its TPA spectrum – is due to the transition between the ground and second excited state, which is dark in OPA due to symmetry selection rules. From a theoretical point of view, this is the first time that solvent effects on TPA of R6G in aqueous solution are considered. This is achieved using a variety of models: a continuum approach (PCM), an electrostatic embedding (QM/TIP3P [122]) and a polarizable embedding (QM/FQ).

The QM/FQ Approach

In the FQ approach, each MM atom is endowed with a charge which can vary according to the electronegativity equalization principle (EEP) [123, 124] which states that, at equilibrium, the instantaneous electronegativity (χ) of each atom has the same value. [123, 124] The model is based on a set of two parameters, *i. e.* atomic electronegativies and chemical hardnesses (η), which can be rigorously defined within *conceptual DFT* [123, 125] as the first and second derivatives of the energy with respect to the charges, respectively. Through these parameters, FQs (q) can be defined as those minimizing the functional: [27, 126]

$$\begin{aligned} F[\boldsymbol{q},\boldsymbol{\lambda}] &= \sum_{\alpha,i} q_{\alpha i}\chi_{\alpha i} + \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j} q_{\alpha i}J_{\alpha i,\beta j}q_{\beta j} + \sum_{\alpha}\lambda_{\alpha}\left(\sum_{i} q_{\alpha i} - Q_{\alpha}\right) \\ &= \boldsymbol{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\boldsymbol{q}^{\dagger}\boldsymbol{J}\boldsymbol{q} + \boldsymbol{q}^{\dagger}\boldsymbol{\lambda} \end{aligned} \tag{1.1}$$

where $\boldsymbol{q}$ is a vector containing the FQs, the Greek indices α run over molecules and the Latin ones i over the atoms of each molecule. $\boldsymbol{\lambda}$ is a set of Lagrangian multipliers used to impose charge conservation constraints on each molecule. The charge interaction kernel $\boldsymbol{J}$ is, in our implementation, the Ohno kernel and the diagonal terms of $\boldsymbol{J}$ kernel are the chemical hardnesses η. The stationarity conditions of the functional in Equation (1.1) are defined through a linear system: [113, 126]

$$\begin{pmatrix} \boldsymbol{J} & \mathbf{1}_{\boldsymbol{\lambda}} \\ \mathbf{1}_{\boldsymbol{\lambda}}^{\dagger} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{\lambda} \end{pmatrix} = - \begin{pmatrix} \boldsymbol{\chi} \\ \boldsymbol{Q} \end{pmatrix} \tag{1.2}$$

We note in passing that the capacitance MM model employed by Rinkevicius, Li, Sandberg, Mikkelsen, and Ågren includes induced charges in its modeling of the metallic portions of the MM environment. [101] Hence, despite the largely different physical setting of the models, their polarization equations bear a significant resemblance.

The QM/FQ model system is constituted by a QM core region placed at the center of a spherical region defining the environment (see Figure 1.1), *i. e.* containing a number of solvent molecules, which are described in terms of FQ force field (FF).

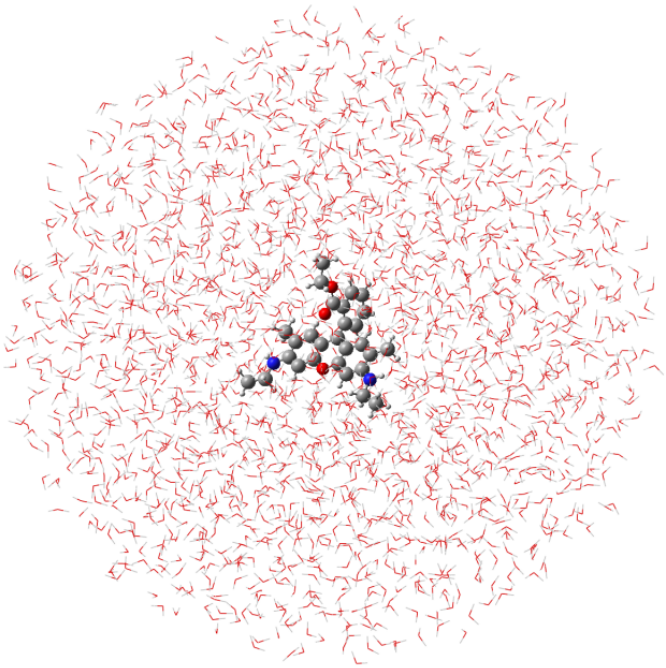


Figure 1.1: Schematic representation of the QM/FQ model.

The FQ FF can be effectively coupled to *any* QM method. It suffices to augment the QM energy functional with the classical functional in Equation (1.1) and the quantum/classical interaction. For the FQ FF the latter is defined as the classical electrostatic interaction between the FQs and the QM density. [104] Explicitly, the interaction takes the form:

$$E_{QM/FQ} = \sum_{i=1}^{N_{FQ}} V[\rho](\mathbf{r}_i) q_i = \sum_{i=1}^{N_{FQ}} \sum_{\mu,\nu=1}^{N_{AO}} \left\langle \chi_\mu \left| \frac{-q_i}{|\mathbf{r} - \mathbf{r}_i|} \right| \chi_\nu \right\rangle D_{\nu\mu}, \tag{1.3}$$

with $V[\rho](\mathbf{r}_i)$ the electrostatic potential due to the QM density of charge at the i-th FQ q_i placed at $\mathbf{r}_i$, $\{\chi_\mu\}$ a Gaussian atomic orbital (AO) basis set and $D_{\mu\nu}$ the AO density matrix.

For a hybrid Kohn–Sham (KS) DFT description of the QM moiety, the global QM/MM energy functional reads: [103, 104]

$$\mathcal{E}[\boldsymbol{D}, \boldsymbol{q}, \boldsymbol{\lambda}] = E_{SCF}[\boldsymbol{D}] + F[\boldsymbol{q}, \boldsymbol{\lambda}] + \boldsymbol{q}^\dagger \boldsymbol{V}(\boldsymbol{D}). \tag{1.4}$$

Stationarity of Equation (1.4) with respect to the density matrix yields the self-consistent field (SCF) equations:

$$\mathbf{F}\boldsymbol{C} = \boldsymbol{S}\boldsymbol{C}\boldsymbol{\epsilon}, \quad \boldsymbol{F} = \boldsymbol{h} + \boldsymbol{G}^\gamma(\boldsymbol{D}) + \boldsymbol{F}_{xc} + \boldsymbol{F}_{FQ} \tag{1.5}$$

where the various terms in the KS matrix $\boldsymbol{F}$ are:

$$h_{\mu\nu} = \left\langle \chi_\mu \middle| -\frac{1}{2}\nabla^2 - \sum_K \frac{Z_K}{|\boldsymbol{R}_K - \boldsymbol{r}|} \middle| \chi_\nu \right\rangle, \tag{1.6a}$$

$$G^\gamma_{\mu\nu}(\boldsymbol{D}) = \sum_{\alpha\beta} D_{\beta\alpha}(g_{\mu\nu\alpha\beta} - \gamma g_{\mu\beta\alpha\nu}), \tag{1.6b}$$

$$F_{xc,\mu\nu} = \int \mathrm{d}\boldsymbol{r}\chi_\mu(\boldsymbol{r})\chi_\nu(\boldsymbol{r}) \left.\frac{\partial E_{xc}}{\partial \rho(\boldsymbol{r})}\right|_{\rho(\boldsymbol{r})} = \int \mathrm{d}\boldsymbol{r}\chi_\mu(\boldsymbol{r})\chi_\nu(\boldsymbol{r}) v_{xc}(\boldsymbol{r}), \tag{1.6c}$$

$$F_{FQ,\mu\nu} = \boldsymbol{q}^\dagger \boldsymbol{V}_{\mu\nu} = \sum_{i=1}^{N_{FQ}} \left\langle \chi_\mu \middle| \frac{-q_i}{|\mathbf{r} - \mathbf{r}_i|} \middle| \chi_\nu \right\rangle. \tag{1.6d}$$

The FQs consistent with the QM density are obtained by solving the stationarity conditions with respect to the polarization variational degrees of freedom $\boldsymbol{q}$, *i. e.* by solving Equation (1.2) with a modified right-hand side (RHS), including the QM potential as additional source term, effectively coupling the QM and MM moieties and ensuring mutual polarization:

$$\begin{pmatrix} \boldsymbol{J} & \mathbf{1}_{\boldsymbol{\lambda}} \\ \mathbf{1}^\dagger_{\boldsymbol{\lambda}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{\lambda} \end{pmatrix} = - \begin{pmatrix} \boldsymbol{\chi} + \boldsymbol{V}(\boldsymbol{D}) \\ \boldsymbol{Q} \end{pmatrix} \tag{1.7}$$

.

Linear and Quadratic Response Functions in a QM/FQ framework

Thanks to its variational formalism, the QM/FQ approach [29] is especially suited to the modeling of response and spectral properties because its energy expression can be easily differentiated up to high orders. The quantum/classical coupling terms needed for the calculation of response properties, can be easily derived and implemented so that polarization effects are fully considered also in the computed final spectral data. [103–105, 110, 111, 127, 128]

For a QM/FQ system subject to a Hermitian, time-periodic, one-electron perturbation V^t, response functions and response equations can be formulated in the atomic orbital, density matrix-based quasienergy formalism [129] of [130]. To the best of our knowledge, this is the first time the quasienergy formalism is employed in the QM/FQ framework. The starting point is the time-averaged quasienergy

Lagrangian, $\{\tilde{L}(\tilde{\boldsymbol{C}}, \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{q}}, t)\}_T$, parametrized in terms of the desired perturbed coefficient matrix $\tilde{\boldsymbol{C}}$, the Lagrange multipliers ensuring orthonormality of the one-electron basis $\tilde{\boldsymbol{\mu}}$ and the perturbed FQs $\tilde{\boldsymbol{q}}$. The tilde is here used for quantities evaluated at general perturbation strengths. We can obtain this Lagrangian by augmenting the quasienergy in Equation (52) of Ref. [130] with the perturbed FQ functional of Equation (1.1). The time-averaged quasienergy Lagrangian is not suitable for an atomic orbital-based theory, since it features the molecular orbital (MO) coefficient matrix. However, its *perturbation-strength-differentiated* counterpart $\{\tilde{L}^a(\tilde{\boldsymbol{D}}, \tilde{\boldsymbol{q}}, t)\}_T$ can be expressed in terms of the desired variational degrees of freedom $\tilde{\boldsymbol{D}}$ and $\tilde{\boldsymbol{q}}$:

$$\tilde{L}^a(\tilde{\boldsymbol{D}}, \tilde{\boldsymbol{q}}, t) \overset{\{Tr\}_T}{=} \tilde{\mathcal{E}}^{00,a} - \tilde{\boldsymbol{S}}^a \tilde{\boldsymbol{W}}, \tag{1.8}$$

where we have borrowed notation from Ref. [93]. Indeed, the close similarity between the PCM and FQ models allows us to leverage the same arguments in Ref. [93] to formulate linear and quadratic response functions in a QM/FQ framework. The generalized KS energy $\tilde{\mathcal{E}}$ is the time-dependent equivalent of Equation (1.4):

$$\tilde{\mathcal{E}}[\tilde{\boldsymbol{D}}, \tilde{\boldsymbol{q}}] \overset{Tr}{=} \left[\tilde{\boldsymbol{h}} + \tilde{\boldsymbol{V}}^t + \frac{1}{2}\tilde{\boldsymbol{G}}^\gamma(\tilde{\boldsymbol{D}}) - \frac{i}{2}\tilde{\boldsymbol{T}}\right] \tilde{\boldsymbol{D}} + \tilde{E}_{xc}[\tilde{\rho}(\tilde{\boldsymbol{D}})] + h_{nuc} + F[\tilde{\boldsymbol{q}}, \tilde{\boldsymbol{\lambda}}] + \tilde{\boldsymbol{q}}^\dagger \tilde{\boldsymbol{V}}(\tilde{\boldsymbol{D}}). \tag{1.9}$$

where $\tilde{\boldsymbol{V}}^t$ is the AO basis representation of the perturbation operator and $\tilde{\boldsymbol{T}} = \left\langle \tilde{\chi}_\mu \middle| \dot{\tilde{\chi}}_\nu \right\rangle - \left\langle \dot{\tilde{\chi}}_\mu \middle| \tilde{\chi}_\nu \right\rangle$ is the time-differentiation AO overlap matrix. Evaluation of Equation (1.8) at zero perturbation strength yields the the first-order property formula. For an electric field perturbation this corresponds to the electric dipole moment and, thanks to the Hellmann–Feynman theorem, only requires the unperturbed density matrix. Further differentiation of Equation (1.8) and evaluation at zero perturbation strength yields higher order response functions. Detailed expressions can be obtained from eqs. (22a)-(22c) and Appendix A in Ref. [93] by replacing perturbed and unperturbed apparent surface charges σ with perturbed and unperturbed FQs, respectively and the generalized free energy $\mathcal{G}$ with $\mathcal{E}$:

$$L^{ab} \overset{\{Tr\}_T}{=} \mathcal{E}^{00,ab} + \mathcal{E}^{10,a}\boldsymbol{D}^b + \mathcal{E}^{01,a}\boldsymbol{q}^b - \boldsymbol{S}^{ab}\boldsymbol{W} - \boldsymbol{S}^a\boldsymbol{W}^b \tag{1.10a}$$

$$\begin{aligned} L^{abc} \overset{\{Tr\}_T}{=} & \mathcal{E}^{00,abc} + \mathcal{E}^{10,ac}\boldsymbol{D}^b + \mathcal{E}^{10,ab}\boldsymbol{D}^c + \mathcal{E}^{20,a}\boldsymbol{D}^b\boldsymbol{D}^c + \mathcal{E}^{10,a}\boldsymbol{D}^{bc} + \mathcal{E}^{11,a}\boldsymbol{D}^b\boldsymbol{q}^c \\ & + \mathcal{E}^{01,ac}\boldsymbol{q}^b + \mathcal{E}^{01,ab}\boldsymbol{q}^c + \mathcal{E}^{01,a}\boldsymbol{q}^{bc} + \mathcal{E}^{11,a}\boldsymbol{q}^b\boldsymbol{D}^c \\ & - \boldsymbol{S}^{abc}\boldsymbol{W} - \boldsymbol{S}^{ab}\boldsymbol{W}^c - \boldsymbol{S}^{ac}\boldsymbol{W}^b - \boldsymbol{S}^a\boldsymbol{W}^{bc} \end{aligned} \tag{1.10b}$$

Response parameters need to be determined in order to assemble the property expressions from the perturbed variational parameters $\boldsymbol{D}^a$, $\boldsymbol{q}^a$ and so forth appearing in the expressions given above. Zero-field perturbation-strength differentiation of the orthonormality, TDSCF, and FQ equations yields the desired response equations. [93, 130, 131] Solution of the N-th order response equations for the b_N perturbation tuple yields the desired response parameters $\boldsymbol{X}^{b_N}$. The perturbed variational parameters are further partitioned into a sum of *homogeneous* and *particular* contributions. [130] Whereas the former depend on the N-th order response parameters, the latter depend only on lower order response parameters. With this partition, the response equations to any order can be compactly rearranged as: [132, 133]

$$\left[\boldsymbol{E}^{[2]} - \omega_{b_N}\boldsymbol{S}^{[2]}\right]\boldsymbol{X}^{b_N} = \boldsymbol{M}^{b_N}_{\mathrm{RHS}}. \tag{1.11}$$

The left-hand side (LHS) includes the generalized Hessian and metric matrices, $\boldsymbol{E}^{[2]}$ and $\boldsymbol{S}^{[2]}$, respectively and $\omega_{b_N} = \sum_{i=1}^{N}\omega_{b_i}$. The RHS $\boldsymbol{M}^{b_N}_{\mathrm{RHS}}$ collects contributions from lower-order perturbed density matrices and N-th order particular contributions, see Refs. [93, 130, 131].

For a single, electric-field type perturbation, the matrix-vector products $\boldsymbol{E}^{[2]}\boldsymbol{X}^{b}$ and $\boldsymbol{S}^{[2]}\boldsymbol{X}^{b}$ assume the form:

$$\begin{aligned}\boldsymbol{E}^{[2]}\boldsymbol{X}^{b} &= \boldsymbol{G}^{KS}([\boldsymbol{X}^{b},\boldsymbol{D}]_{\boldsymbol{S}})\boldsymbol{D}\boldsymbol{S} - \boldsymbol{S}\boldsymbol{D}\boldsymbol{G}^{KS}([\boldsymbol{X}^{b},\boldsymbol{D}]_{\boldsymbol{S}}) + \boldsymbol{F}[\boldsymbol{X}^{b},\boldsymbol{D}]_{\boldsymbol{S}}\boldsymbol{S} \\ &\quad - \boldsymbol{S}[\boldsymbol{X}^{b},\boldsymbol{D}]_{\boldsymbol{S}}\boldsymbol{F} + \boldsymbol{q}^{b}_{H}\boldsymbol{V}\boldsymbol{D}\boldsymbol{S} - \boldsymbol{S}\boldsymbol{D}\boldsymbol{V}\boldsymbol{q}^{b}_{H} \qquad (1.12)\\ \boldsymbol{S}^{[2]}\boldsymbol{X}^{b} &= \boldsymbol{S}[\boldsymbol{X}^{b},\boldsymbol{D}]_{\boldsymbol{S}}\boldsymbol{S}. \qquad (1.13)\end{aligned}$$

where $\boldsymbol{G}^{KS}$ now collects the two-electron and exchange-correlation contributions. The general expression for the RHS (see Equation (46) in Ref. [93]) simplifies to the matrix elements of the electric dipole perturbation operator, with no contributions from the classical polarizable model. These equations are equivalent, upon transformation to the MO basis, to their more familiar formulation as Casida's equations. [104, 134] Perturbed FQs are obtained by solving:

$$\boldsymbol{J}\boldsymbol{q}^{b}_{H} = -\boldsymbol{V}(\boldsymbol{D}^{b}_{H}), \tag{1.14}$$

once again highlighting the introduction of the mutual QM/MM polarization.

Response equations for the second-order response parameters $\boldsymbol{X}^{bc}$ can be derived in a similar fashion. Restricting ourselves to electric-field type perturbations only, the linear transformations are expressed as:

$$\begin{aligned}\boldsymbol{E}^{[2]}\boldsymbol{X}^{bc} &= \boldsymbol{G}^{KS}([\boldsymbol{X}^{bc},\boldsymbol{D}]_{\boldsymbol{S}})\boldsymbol{D}\boldsymbol{S} - \boldsymbol{S}\boldsymbol{D}\boldsymbol{G}^{KS}([\boldsymbol{X}^{bc},\boldsymbol{D}]_{\boldsymbol{S}}) + \boldsymbol{F}[\boldsymbol{X}^{bc},\boldsymbol{D}]_{\boldsymbol{S}}\boldsymbol{S} \\ &\quad - \boldsymbol{S}[\boldsymbol{X}^{bc},\boldsymbol{D}]_{\boldsymbol{S}}\boldsymbol{F} + \boldsymbol{q}^{bc}_{H}\boldsymbol{V}\boldsymbol{D}\boldsymbol{S} - \boldsymbol{S}\boldsymbol{D}\boldsymbol{V}\boldsymbol{q}^{bc}_{H} \qquad (1.15)\\ \boldsymbol{S}^{[2]}\boldsymbol{X}^{bc} &= \boldsymbol{S}[\boldsymbol{X}^{bc},\boldsymbol{D}]_{\boldsymbol{S}}\boldsymbol{S}. \qquad (1.16)\end{aligned}$$

In contrast to the linear response equations, the RHS will contain FQ contributions:

$$\boldsymbol{M}^{bc}_{RHS,FQ} = \left[(\boldsymbol{q}^{bc}_{P})^{\dagger}\boldsymbol{V}\boldsymbol{D}\boldsymbol{S} + \boldsymbol{q}^{\dagger}\boldsymbol{V}\boldsymbol{D}^{bc}_{\mathrm{P}}\boldsymbol{S} + (\boldsymbol{q}^{b}_{H})^{\dagger}\boldsymbol{V}\boldsymbol{D}^{c}_{\omega}\boldsymbol{S} + (\boldsymbol{q}^{c}_{H})^{\dagger}\boldsymbol{V}\boldsymbol{D}^{b}_{\omega}\boldsymbol{S}\right]^{\ominus}, \tag{1.17}$$

with the second-order particular FQs calculated as the solution to the linear equation:

$$\boldsymbol{J}\boldsymbol{q}^{bc}_{P} = -\boldsymbol{V}(\boldsymbol{D}^{bc}_{P}), \tag{1.18}$$

and the perturbed density matrix $\boldsymbol{D}^{bc}_{P}$ is in turn assembled from first-order perturbed density matrices.

As for the PCM, there are two classes of contributions from the classical polarizable region: implicit, through the unperturbed Fock matrix, and explicit, through the N-th order perturbed homogeneous FQs. This is indeed a trait shared by any quantum/classical polarizable model.

One- and Two-Photon Absorption

We can formulate one- and two-photon absorption parameters in terms of single residues of the linear and quadratic response functions, respectively. [135] have presented a density matrix-based, open-ended formulation of single residues that can be coupled to classical polarizable models. [93, 114]

TPA cross sections of randomly oriented systems can be calculated from the imaginary part of the third susceptibility. Alternatively, they can be obtained as the individual two-photon transition matrix elements S_{ab} between the initial state $|i\rangle$ and final state $|f\rangle$, with the sum-over-states expression: [136]

$$S_{ab} = \sum_s \left(\frac{\langle i|\mu_a|s\rangle \langle s|\mu_b|f\rangle}{\omega_{si} - \omega} + \frac{\langle i|\mu_b|s\rangle \langle s|\mu_a|f\rangle}{\omega_{si} - \omega} \right) \tag{1.19}$$

where $a, b \in x, y, z$ and ω is the frequency of the external radiation, which is half of the excitation energy ω_f to the final state $|f\rangle$ ($2\omega = \omega_f$). The summation runs over all s states, including initial and final state. $\omega_{si} = \omega_s - \omega_i$ is the transition energy between s and i states.

For linearly polarized light with parallel polarization, rotationally averaged microscopic OPA and TPA cross sections can be written in terms of the transition matrix tensors $\mathbf{S}$ and their complex conjugates $\overline{\mathbf{S}}$ as:

$$\langle \delta^{\mathrm{1PA}} \rangle = \frac{1}{3} \sum_a S_a \bar{S}_a \tag{1.20}$$

$$\langle \delta^{\mathrm{2PA}} \rangle = \frac{1}{15} \sum_{ab} \left(2 S_{ab} \bar{S}_{ab} + S_{aa} \bar{S}_{bb} \right). \tag{1.21}$$

Finally, the macroscopic TPA cross section in cgs units can be obtained from the rotationally averaged TPA strengths ($\langle \delta^{\mathrm{TPA}} \rangle$) expressed in atomic units as:

$$\sigma^{\mathrm{TPA}} = \frac{N \pi^3 \alpha a_0^5 \omega^2}{c} \langle \delta^{\mathrm{TPA}} \rangle \, g(2\omega, \omega_0, \Gamma) \tag{1.22}$$

where $N = 4$ in case of single beam experiments, α is the fine structure constant, a_0 is the Bohr radius, ω is the photon energy in atomic units, c is the speed of light and $g(2\omega, \omega_0, \Gamma)$ the lineshape function describing spectral broadening effects. The common unit for TPA cross sections is GM in honour of the work of Maria Goppert-Mayer (1 GM = 10^{-50} cm^4 s photon^{-1}). We refer the reader to Ref. [52] for further details on the computational approach to TPA cross sections.

Computational Details

The close similarity of the PCM and FQ models renders the implementation of the latter a rather straightforward extension of the former. Our implementation relies on the PCMSolver library, [137] which already provides the infrastructure for the PCM and lends itself to the extensions needed for a FQ implementation. The calculation of TPA cross sections with the FQ discrete model was implemented in the LSDALTON program package. [138] LSDALTON provides single residues of the quadratic response function *in vacuo* and we exploited the existing interface with PCMSolver [139] to extend this functionality to include the PCM and FQ

quantum/classical polarizable models. The TPA implementation in LSDALTON leverages the $2n + 1$ rule. Ensuring correctness of the linear response calculations is enough to guarantee correctness up to cubic response properties. We performed extensive testing of the PCM and FQ linear and quadratic response functions, including their residues, against DALTON. [56, 90–92]

The use of a modular programming paradigm, based on open-source libraries and programs, is particularly beneficial for the work here presented. The most significant programming investment was the extension of the PCMSolver library.

Rhodamine 6G (R6G) (see Figure 1.2) in aqueous solution has been amply studied experimentally using both OPA and TPA techniques. [58, 115–121] Such a system is capable of forming solute-solvent hydrogen bonds, and is thus a good test case for our atomistic polarizable QM/FQ approach.

We adopted the following computational protocol for our QM/MM calculations of excitation energies, OPA and TPA intensities:

Definition of the systems and calculation of atomic charges The solute molecule was surrounded by a number of water molecules large enough to represent all the solute-solvent interactions (at least 8500). The atomic charges of the solute were computed by using the Charge Model 5 (CM5). [140]

Classical MD simulations in aqueous solution The MD simulations were performed in a cubic box reproduced periodically in every direction, satisfying periodic boundary conditions (PBC). A minimization step ensures that simulations were started from a minimum of the classical PES. From the MD run, a set of snapshots was extracted to be used in the following QM/TIP3P [122] and QM/FQ calculations.

Definition of the different regions of the two-layer scheme and their boundaries Each snapshot extracted from the MD run was cut into a sphere centered on the solute. A radius of 25 Å was chosen in order to include all specific water-solute interactions.

QM/classical calculations QM/TIP3P, and QM/FQ OPA and TPA calculations were performed on 100 structures obtained in the previous step. The results obtained for each spherical snapshot were extracted and averaged to produce the final value.

In step 1, R6G was optimized and CM5 charges were calculated at the B3LYP/6-31+G* level of theory including solvent effects by means of the PCM. [2]

The MD simulation was performed using GROMACS, [141] with the OPLS-AA [142] force fields to describe intra- and inter-molecular interactions. CM5 charges were used to account for electrostatic interactions. The TIP3P FF was used to describe the water molecules. [122] A single molecule was dissolved in a cubic box containing at least 8500 water molecules. A chloride ion has been included in the box to neutralize the system. The chromophore was kept fixed during all the steps of the MD run. This neglects geometric relaxation due to solvation and is an approximation. However, our primary goal is to compare *direct* (electronic) rather than *indirect* (geometrical) solvent effects on the properties. Furthermore, this approximation affords a fairer comparison with implicit solvent models. R6G in aqueous solution was initially brought to 0 K with the steepest descent minimization procedure

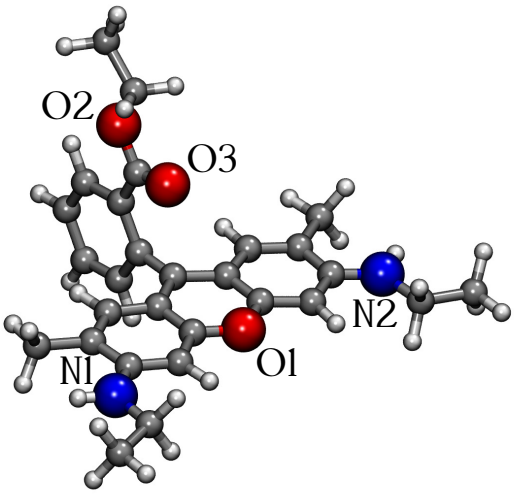


Figure 1.2: Rhodamine 6G structure and atom labeling.

and then heated to 298.15 K in an NVT ensemble using the velocity-rescaling [143] method with an integration time step of 0.2 fs and a coupling constant of 0.1 ps for 200 ps. The time step and temperature coupling constant were then increased to 2.0 fs and 0.2 ps, respectively, and an NPT simulation (using the Berendsen barostat and a coupling constant of 1.0 ps) for 1 ns was performed to obtain a uniform distribution of molecules in the box. A 10 ns production run in the NVT ensemble was then carried out, fixing the fastest internal degrees of freedom by means of the LINCS algorithm (δt=2.0 fs), [144] and freezing the R6G at the center of the simulation box. Electrostatic interactions are treated by using particle-mesh Ewald (PME) [145] method with a grid spacing of 1.2 Å and a spline interpolation of order 4. We have excluded intramolecular interactions between atom pairs separated up to three bonds. A snapshot every 100 ps was extracted in order to obtain a total of 100 uncorrelated snapshots.

For each snapshot a solute-centered sphere with a radius of 25 Å was cut. Notice that the chloride ion was not present in any of the extracted spherical snapshots. For each snapshot, OPA and TPA spectra were then calculated with two QM/MM approaches: the water molecules were modeled by means of the non-polarizable TIP3P FF, [122] and the FQ SPC parametrization proposed by. [27] For comparison, we also ran QM calculations on the isolated chromophore and embedded in a PCM continuum modelling the water solution.

We performed all OPA calculations using a locally modified version of the Gaussian 16 package, [146] whereas we used a locally modified version of the LSDALTON program, [138] interfaced to the PCMSolver library, [137] for the TPA calculations. The CAM-B3LYP/6-31+G* model chemistry was used in all calculations. For the LSDALTON calculations we leveraged the implementation of density fitting, with the df-def2 auxiliary fitting basis, to accelerate the evaluation of the Coulomb matrix.

For the PCM calculations, the cavity was generated from a set of atom-centered, interlocking spheres. PCMSolver implements the GePol algorithm for cavity generation and uses the Bondi–Mantina set of van der Waals radii [147, 148] 1.20 Å

for hydrogen, 1.70 Å for carbon, 1.55 Å for nitrogen and 1.52 Å for oxygen. All radii were scaled by a factor of 1.2. Values of the static and optical permittivities of $\varepsilon_s = 78.39$ and $\varepsilon_\infty = 1.776$, respectively, were used for the PCM calculations in water in LSDALTON.

For the PCM calculations in Gaussian, the following atomic radii were used and scaled by a factor of 1.1: 1.443 Å for hydrogen, 1.9255 Å for carbon, 1.83 Å for nitrogen and 1.75 Å for oxygen. Values of the static and optical permittivities of $\varepsilon_s = 78.3553$ and $\varepsilon_\infty = 1.77785$, respectively, were instead used to model water solvent effects.

The differences in PCM parametrization between the two codes used in this study did not lead to significant differences (less than 1%) between computed ground state energies, excitation frequencies and oscillator strengths.

MD analysis: Hydration Pattern

First, we analyze the MD runs in terms of the hydration patterns with a particular focus on hydrogen bonds (HBs) formed between R6G and solvent water molecules. Second, OPA and TPA spectra are presented and compared with their experimental counterparts.

R6G is characterized by a keto oxygen and by amino groups, which can act as HB donors, whereas the ether oxygen atoms (O1 and O2) together with the amino nitrogen atoms (N1 and N2) can act as HB acceptors (see Figure 1.2 for atom labeling). HB patterns were analyzed by extracting the radial distribution functions $g(r)$ from the MD trajectories. For this analysis the TRAVIS package was used. [149] The radial distribution functions were computed taking as reference oxygen atoms (O1, O2 and O3), nitrogen atoms (N1 and N2) and amino hydrogen atoms (H1 and H2) of the solute; they are plotted in Figure 1.3. The most intense peak of the $g(r)$ refers to the carboxylic oxygen (O3), whereas the other atoms are not involved in HBs with the solvent water molecules. The average number of HBs between water molecules and the carboxylic oxygen (O3) is 1.7, thus confirming strong HB patterns.

OPA Spectra of R6G in aqueous solution

100 uncorrelated snapshots were extracted from the MD run. The convergence of the studied properties (OPA and TPA spectra) was checked by considering an increasing number of snapshots. These results are reported in Figures S1 and S2 in the supporting information (SI) [150].

QM/FQ OPA results are depicted in Figure 1.4. Raw data (sticks) and their Gaussian convolution, with full width at half maximum (FWHM) of 0.1 eV, are shown.

The OPA spectrum is characterized by an intense transition (S1) at about 2.8 eV (440 nm) which is related to a pure HOMO→LUMO transition. The second transition (S2) is a pure HOMO-1→LUMO transition and is located at about 3.2 eV (380 nm). This transition is dark due to the symmetry of the involved orbitals, see Figure 1.5. Our findings confirm what already reported in previous theoretical studies, [116] and is in contrast with some experimental works, [58] where S2 is instead assigned to a visible transition. The third transition predicted by the QM/FQ

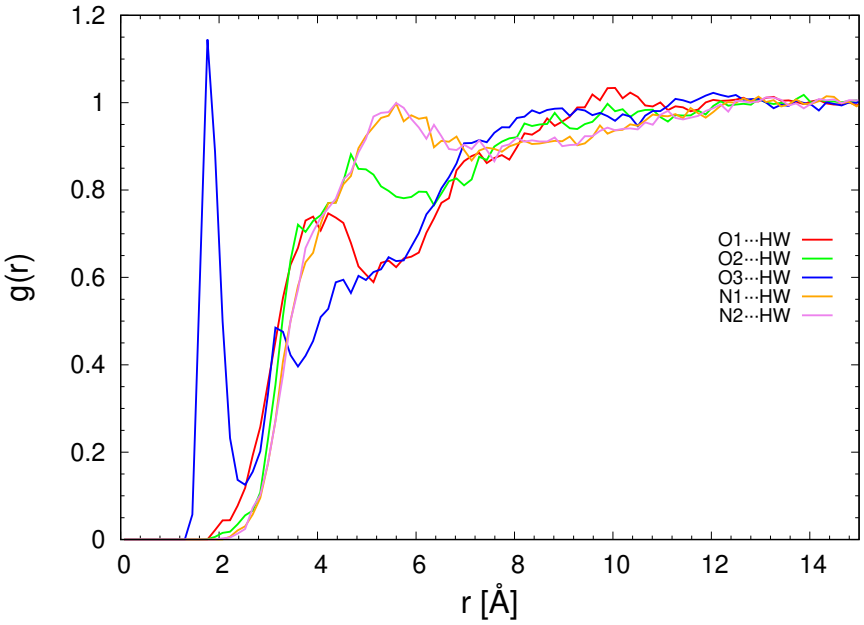


Figure 1.3: Radial distribution function $g(r)$ between selected sites of R6G and water molecules. The considered atomic sites are highlighted in Figure 1.2

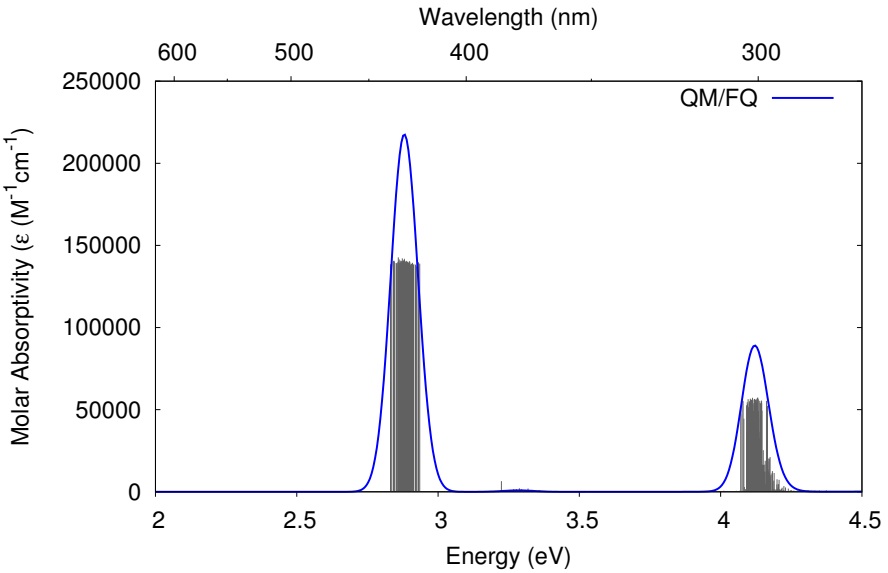


Figure 1.4: QM/FQ calculated OPA spectrum of R6G in aqueous solution reported as stick spectrum and convoluted with a Gaussian band shape (FWHM=0.1 eV)

approach is located at 4.1 eV (300 nm), and involves a combined transition between HOMO-1→LUMO+1 and HOMO-2→LUMO. The whole computed spectrum, involving also the higher transitions, is reported in the SI (Figure S3). [150]

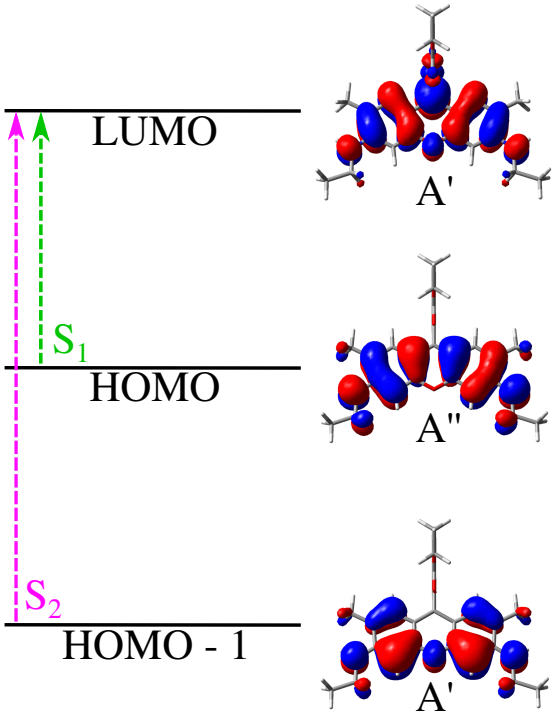


Figure 1.5: Graphical scheme of the first two transitions of R6G in aqueous solution. Molecular orbitals related to each transition are depicted.

We investigated the relevance of solvent effects by employing different computational approaches, ranging from continuum QM/PCM to nonpolarizable QM/TIP3P and polarizable QM/FQ descriptions. As a reference, the OPA spectrum of the solute *in vacuo* was also calculated. Vertical excitations and oscillator strengths as obtained by exploiting the different approaches summarized in Table 1.1 and graphically depicted in Figure 1.6. The experimental spectrum is also reported. [58]

Table 1.1: Vacuum and aqueous solution vertical excitation energies (eV) and oscillator strengths (a.u.) for the three lowest singlet excited states of R6G calculated at the CAM-B3LYP/6-31+G* level of theory. QM/PCM, QM/TIP3P and QM/FQ approaches were used to model environment effects in solution. The experimental vertical excitation energies are also reported. [a]Experimental values are taken from Ref. [58]

	Vacuum		QM/PCM		QM/TIP3P		QM/FQ		Exp. [a]
	E_{vert}	f	E_{vert}	f	E_{vert}	f	E_{vert}	f	E_{vert}
S1	3.07	0.89	2.98	1.05	2.82	0.84	2.81	0.90	2.27
S2	3.68	0.00	3.71	0.00	3.18	0.01	3.20	0.00	
S3	4.42	0.25	4.47	0.29	4.15	0.21	4.13	0.26	3.46

The experimental spectrum is dominated by an intense transition (S1) at about 2.3 eV (550 nm), characterized by an inhomogeneous band broadening, probably

due to a vibronic convolution. The second visible transition is instead at about 3.6 eV (350 nm) and can be associated to the S3 transition.

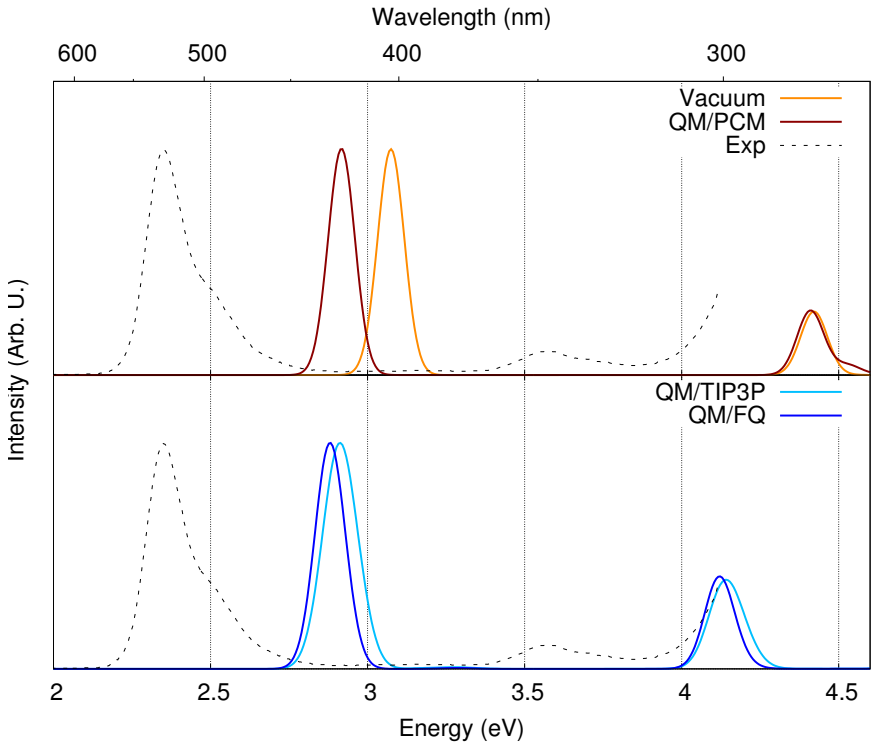


Figure 1.6: One-photon absorption spectra for the R6G chromophore at the CAM-B3LYP/6-31+G* level of theory. Computed vacuum, QM/PCM, QM/TIP3P and QM/FQ OPA spectra are reported. The experimental spectrum is reproduced from Ref. [58].

The first transition (S1) is predicted to be the most intense by all the different methods, with a redshift when solvent effects are taken into consideration. In particular, QM/PCM, QM/TIP3P and QM/FQ predict very similar vertical excitation energies for the S1 transition, with the largest redshift shown by QM/FQ. All the approaches considered correctly describe the S2 transition as being symmetry-forbidden, see Table 1.1. The models employed exhibit major differences for the third transition (S3), which is predicted at about 280 nm in vacuum and by QM/PCM, and at about 310 nm by both QM/TIP3P and QM/FQ approaches (see Fig. 1.6).

The atomistic description captures the explicit solvent-solute interactions that appear to be crucial in the modeling of this transition. Implicit solvation is unable to capture solvent effects, as it is shown by the negligible differences observed with respect to the vacuum results. Furthermore, the energy difference between S1 and S3 is correctly predicted by QM/MM approaches ($\sim$ 1.18 eV) as compared to the experimental value (1.19 eV), whereas it is overestimated by the QM/PCM approach (1.49 eV). Notice also that by considering energy differences, instead of absolute vertical excitation energies, systematic errors and biases due to the choice of QM method, and in particular of a specific DFT functional, should be avoided.

To further analyze the nature of the electronic transitions, their charge transfer (CT) nature was characterized by a simple index, denoted as D_{CT}. [151, 152] The

barycenters of the positive and negative density distributions are calculated as the difference of ground state (GS) and excited state (ES) densities. The CT length index (D_{CT}) is defined as the distance between the two barycenters. Calculated vacuum, QM/PCM, QM/TIP3P and QM/FQ D_{CT} values are reported in Table 1.2.

Table 1.2: D_{CT} indices (Å) for the first three excited states as obtained by exploiting the various models considered in this work.

	Vacuum	QM/PCM	QM/TIP3P	QM/FQ
S1	1.195	1.170	1.668	1.624
S2	1.318	1.336	1.587	1.574
S3	0.811	0.759	4.043	3.580

The first two transitions (S1 and S2) have little CT character. The largest difference between the considered approaches is shown by the S3 transition, of which the CT character is irrelevant *in vacuo* and at the QM/PCM level, whereas is huge at both QM/TIP3P and QM/FQ levels. This different behaviour was expected by considering the computed OPA spectra reported in Figure 1.6; in fact, the largest discrepancy between vacumm-QM/PCM and QM/MM approaches was indeed predicted for the S3 transition.

TPA Spectra of R6G in aqueous solution

Figure 1.7 reports computed QM/FQ TPA raw data and their Gaussian convolution. We adopted a Gaussian lineshape with FWHM of 0.1 eV in agreement with previous computational studies on TPA and suggested best practices, see Ref. [52].

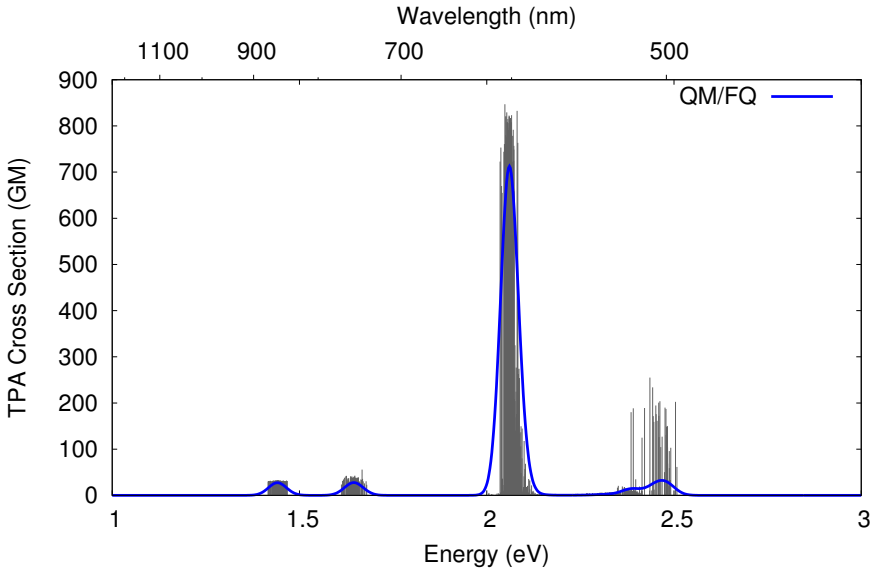


Figure 1.7: QM/FQ calculated TPA spectrum of R6G in aqueous solution reported as stick spectrum and convoluted with a Gaussian band shape (FWHM = 0.1 eV)

The first two visible transitions at about 1.4 and 1.6 eV have almost the same intensity and they are the S1 and S2 transitions, respectively. The S2 transition is not dark, due to the different symmetry selection rules in TPA compared to OPA

(see Figure 1.4). The computed QM/FQ TPA spectrum is dominated by an intense peak at about 2.06 eV (600 nm) associated to the S3 transition.

We compared different approaches to model solvent effects also for TPA spectra and considered QM/PCM, QM/TIP3P and QM/FQ models. In addition, TPA spectra of R6G *in vacuo* was computed as an additional reference point. Vacuum, QM/PCM, QM/TIP3P and QM/FQ TPA spectra are plotted in Figure 1.8 together with the experimental spectrum, reproduced from Ref. [115].

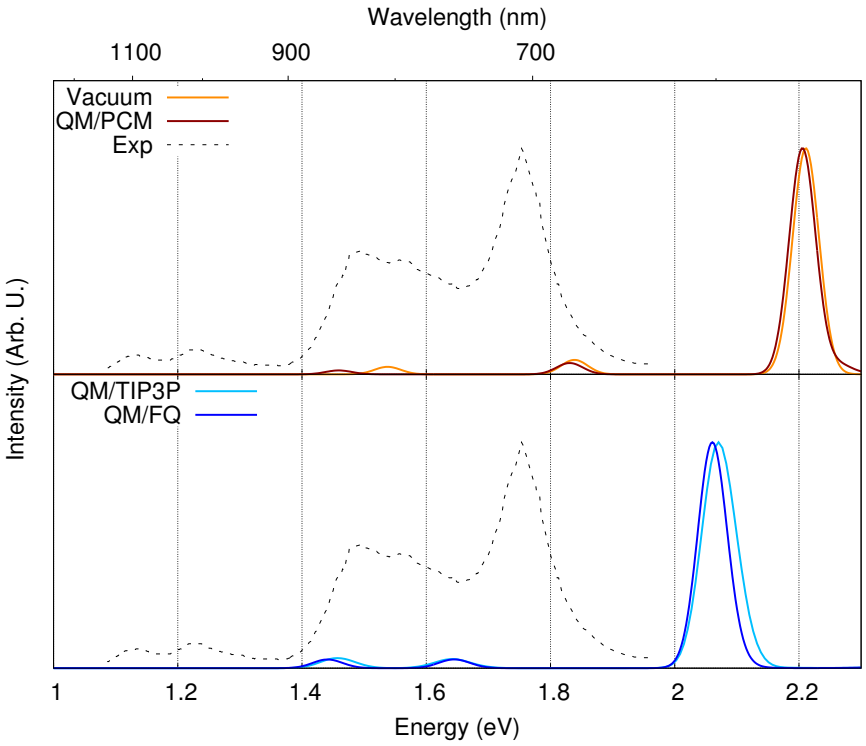


Figure 1.8: Two-photon absorption spectra for the R6G chromophore at the CAM-B3LYP/6-31+G* level of theory. Computed vacuum, QM/PCM, QM/TIP3P and QM/FQ OPA spectra are reported. The experimental spectrum is reproduced from Ref. [115].

The experimental TPA spectrum is characterized by three main peaks at 1.14, 1.52, and 1.73 eV. These correspond to the S1, S2 and S3 transitions, respectively. It is worth noting that the peak at about 1.25 eV was wrongly reported to be associated to the S2 transition in some experimental works due to the similar intensity with respect to the peak at 1.14 eV. [58] Milojevich, Silverstein, Jensen, and Camden recently showed that such a peak is instead a vibronic band due to Herzberg–Teller terms. [116] Notice in fact that such a vibronic peak exactly corresponds to the vibronic peak present also in the experimental OPA spectrum, see Figure 1.6. In this work, we are not considering any vibronic contributions because our main goal is to show the performance of QM/FQ approach in predicting solvent effects on multi-photon spectroscopies. As a consequence, the peak at 1.25 eV cannot be reproduced by the different approaches explored in this work.

The computed vacuum and aqueous solution transition energies and TPA cross sections (in GM units) of the first three transitions are reported in Table 1.3, together

with their experimental values. It is clear from Figure 1.8 and Table 1.3 that all the approaches considered in this work predict the S3 transition as the most intense in the TPA spectrum. The whole TPA spectra, including also higher energy transitions, are reported in Figure S4 of the SI. [150]

Table 1.3: Vacuum and aqueous solution TPA vertical excitation energies (eV) and cross sections (GM) for the six lowest singlet excited states of R6G calculated at the CAM-B3LYP/6-31+G* level of theory. QM/PCM, QM/TIP3P and QM/FQ approaches were used to model environment effects in aqueous solution. The experimental TPA vertical excitation energies and cross sections are also reported. [a]Experimental values from Ref. [115]

	Vacuum		QM/PCM		QM/TIP3P		QM/FQ		Exp. [a]	
	E_{TPA}	σ	E_{TPA}	σ	E_{TPA}	σ	E_{TPA}	σ	E_{TPA}	σ
S1	1.54	15.52	1.46	24.03	1.46	18.67	1.44	27.37	1.14	15
S2	1.84	29.83	1.83	69.55	1.64	16.92	1.64	27.59	1.52	65
S3	2.21	466.50	2.20	1390.33	2.07	417.39	2.06	712.69	1.73	150

Both QM/MM approaches outperform the vacuum and QM/PCM models in reproducing the experimental relative differences between the vertical transition energies. As already pointed out in case of OPA spectra (see Figure 1.6 and Table 1.1), this is particularly evident in the case of the S1-S3 difference. The relative intensity of the S1 and S3 transitions is correctly reproduced by all the different methods, with the best agreement with experiment shown by QM/TIP3P and QM/FQ.

The largest discrepancy is reported in case of the S2 transition for both relative energies and intensities. The data reported in Table 1.3 show that the S1-S2 energy difference is indeed wrongly reproduced even by QM/MM approaches ($\sim$ 0.2 eV vs. 0.38 eV in the experiment). Some differences between the considered methods are encountered in case of relative intensities between S1 and S2. In particular, QM/PCM is almost able to reproduce the experimental intensity ratio between the two peaks (2.9 vs. 4.3), whereas non-polarizable QM/TIP3P results are the worst (0.9 vs. 4.3). Such a huge discrepancies probably reflects the fact that for this transition, HB interactions play a minor role with respect to polarization effects.

Finally, the major effect of including polarization effects in the description of solvent effects is the increase of the S3 TPA cross section, which is reported by both QM/PCM and QM/FQ whereas QM/TIP3P is consistent with the vacuum counterpart. Remarkably, in passing from a continuum to a discrete approach (QM/FQ), the S3 TPA cross section decreases, thus moving closer to the experimental value.

To conclude our comparison between computed and experimental TPA spectra of R6G in water, it is worth pointing out that experimental TPA measurements are not a standard technique. [115] In fact, as reported in Figure S5 in the SI, [150] several experimental TPA spectra of R6G in aqueous solution have been reported previously in the literature, [115] showing large discrepancies even in the measured spectra. Therefore, a quantitative comparison is particularly challenging.

A polarizable embedding approach to second harmonic generation (SHG) of molecular systems in aqueous solutions

Non Linear Optics (NLO) [153–156] has been reported to impact many different research fields, such as material science, communications, medicine and the developoment of electronic devices. [157–159] The role of theoretical investigations have been relevant in this field, because such studies have demonstrated to provide valuable information guiding the design of new optical devices. [160] Through the variety of NLO properties, [161–163] the most basic one is the first hyperpolarizability, i.e. the quadratic response of a system to an external electric field. The frequency of the resulting exiting wave is the combination of incident frequencies of the laser beams: if the entering waves have the same frequency ω and the frequency of the resulting wave is 2ω, then the process takes the name of Second Harmonic Generation (SHG). SHG will be the topic of the present paper. Within the Born-Oppheneimer approximation, (hyper-)polarizabilities bear electronic and a vibrational terms. [76, 164–168] The first derives from the rearrangement of the electronic density in the presence of the external field, whereas the latter from the rearrangement of nuclear motions. In case of SHG, the electronic contribution is predominant over the vibrational one, which only accounts for the 5-20% of the total value. [169–172] Notice that this is not generally true for static properties, where vibrational contributions can be even larger than electronic terms. [173–175] Electronic properties can be calculated by resorting to standard Quantum-Mechanical (QM) methods. If compared to Hartree-Fock (HF), Density Functional Theory (DFT) generally increases the agreement between calculated and experimental values. [176–178] However, whenever electron correlation plays a crucial role, Coupled Cluster (CC) approaches are best suited, [179–184] although they cannot be applied to medium-large systems. In such cases, DFT represents a good compromise between accuracy, feasibility and computational cost. For these reasons, DFT is employed in this work. A proper choice of the QM level to treat a given system is not the only key for a good reproduction of experimental data. In fact, for condensed phase systems the inclusion of environmental effects is mandatory to achieve a good modeling of the system. [185–188] The standard approach to include such effects is to resort to continuum solvation approaches, such as the Polarizable Continuum Model (PCM). [2] This model has been successfully applied to the modeling of several properties and spectroscopies, and also to SHG of solvated systems. [176, 186, 189, 190] However, it may fail when specific solute-solvent interactions play a dominant role, and this is due to the lack of any atomistic description of solvent molecules. [25, 110, 111, 191] In all such cases, the best strategy has demonstrated to be to resort to the so called QM/Molecular Mechanics (MM) methods. [72, 74, 82, 97] Such approaches have already been employed for the description of SHG. [85, 192–195] In standard MM Force-Fields (FF) no mutual solute-solvent polarization effects are considered: this is due to the fact that the so-called electrostatic embedding approach is exploited, i.e. the charges placed on the MM atoms, which define the solute-solvent interaction, are fixed. The solute-solvent mutual polarization can be restored by employing polarizable force-fields, based on distributed multipoles, [16–19] induced dipoles, [21, 24, 25] Drude oscillators [26] or Fluctuating Charges (FQ). [27, 196]

This latter approach is exploited in this study, because it has already been successfully applied to the calculation of several molecular properties and spectroscopies for systems in aqueous solution. [29, 110, 111] In the present paper, the QM/FQ approach is applied for the first time to the calculation of SHG of molecular systems in aqueous solution. A purely electrostatic solute-solvent interaction will be considered, however quantum non-electrostatic effects, such as Pauli Repulsion and Dispersion, may play a role and preliminary models to account for them have been proposed in the recent literature. [11, 14, 197, 198] Notice also that polarizable embedding approaches based on induced dipoles have been applied to the modeling of SHG with promising results. [23, 183]

We present the extension to Second Harmonic Generation (SHG) of the atomistic fully polarizable QM/FQ method that treats the solvent atomistically and embeds each atom in the solvent with a fluctuating charge (FQ) which responds to the solute QM electrostatic potential in a self-consistent manner.The proposed approach is able to achieve an adequate modeling of solvent effects both in the quantum mechanical response equations and on the conformational properties of the system, which is sampled by resorting to MD simulations. The application of the model to selected organic acids in aqueous solution, for which the interaction with the surrounding environment is dominated by HB interactions, shows a good agreement in both the modeling of solvent effects and in the reproduction of experimental SHG data extracted from Hyper Raman Scattering experiments.

The QM/FQ model

In the FQ approach, polarization effects are considered by allowing point charges placed on the MM moiety to fluctuate according to differences in atomic electronegativities. The model is based on a set of parameters representing atomic hardnesses and electronegativities, whose physical origin can be rigorously defined within the so called "conceptual DFT". [123, 125] Through these parameters, atomic charges can be computed based on the difference of electronegativities between the atoms. [27, 28, 196] More in detail, the FQ FF describes polarization effects by endowing each MM atom with a fluctuating charge whose value depends on the electronegativity [27, 28, 196] according to the Electronegativity Equalization Principle (EEP) [123, 124] which states that, at equilibrium, the instantaneous electronegativity χ of each atom has the same value. [123, 124] The FQs ($\mathbf{q}$) can be defined as those minimizing the following functional: [126]

$$\begin{aligned} F(\mathbf{q},\boldsymbol{\lambda}) =& \sum_{\alpha,i} q_{\alpha i}\chi_{\alpha i} + \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j} q_{\alpha i}J_{\alpha i,\beta j}q_{\beta j} + \sum_{\alpha}\lambda_{\alpha}(\sum_{i} q_{\alpha i} - Q_{\alpha}) \\ =& \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} \end{aligned} \tag{1.23}$$

where $\mathbf{q}$ is a vector containing the FQs, the Greek indices α run over molecules and the Latin ones i over the atoms of each molecule. $\boldsymbol{\lambda}$ is a set of Lagrangian multipliers used to impose charge conservation constraints on each molecule. The charge interaction kernel $\mathbf{J}$ is, in our implementation, the Ohno kernel. [199] Atomic units are used throughout the article. The stationarity conditions of the functional in Equation1.23 are defined through the following equations: [126]

$$\begin{cases} \sum_{\beta,j} J_{\alpha i,\beta j} q_{\beta j} + \lambda_\alpha = -\chi_{\alpha i} \\ \\ \sum_i q_{\alpha_i} = Q_\alpha \end{cases} \tag{1.24}$$

The previous system of equations can be recast in a more compact formalism by introducing the extended **D** matrix:

$$\mathbf{D} = \begin{pmatrix} \mathbf{J} & \mathbf{1_\lambda} \\ \mathbf{1_\lambda^\dagger} & \mathbf{0} \end{pmatrix}$$

where $\mathbf{1}_\lambda$ is a rectangular matrix containing the Lagrangian multipliers. The linear system of equation then reads:

$$\mathbf{D q_\lambda} = -\mathbf{C_Q} \tag{1.25}$$

where $\mathbf{C_Q}$ collects atomic electronegativities and total charge constraints, whereas charges and Lagrange multipliers are collected in $\mathbf{q_\lambda}$.

The FQ FF can be effectively coupled to QM methods. The resulting QM/FQ approach [29, 126] has been shown to be especially suited for the modeling of response and spectral properties because, as it is shown below, its energy expression can be easily differentiated up to high orders. The QM/FQ method fully accounts for polarization effects: in fact the FQs placed in the MM moiety adjust to both the electrostatic potential generated by the QM portion and their electronegativities, while the QM core feels the presence of the FQs through specific additional terms in the QM Hamiltonian. In case of the calculation of response/spectroscopic properties, such terms propagate to the solute's response equations, so that polarization effects are fully considered also in the computed final spectral data. [107, 110, 111, 127, 200, 201]

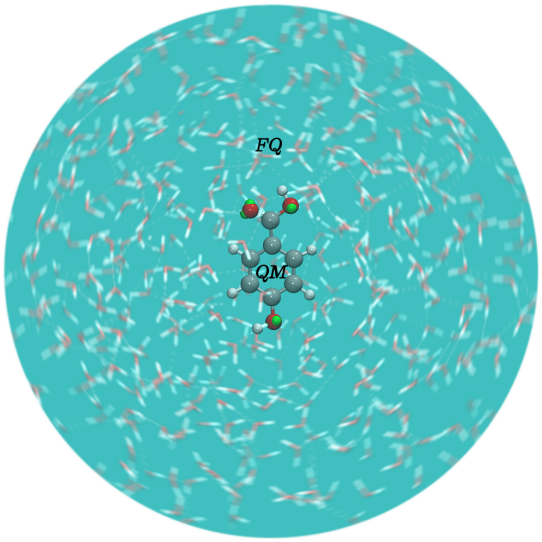


Figure 1.9: Schematic representation of the QM/FQ model.

We recall that the QM/FQ model system is constituted by a QM core region placed at the center of a spherical region defining the environment (see Figure 1.9), i.e. containing a number of solvent molecules, which are described classically and whose atoms carry fluctuating charges that can respond to the solute's electrostatic

potential. The size of this region is chosen so to yield converged final results, and the relative positions of the QM and MM atoms results from a sampling performed through a classical molecular dynamics (MD) simulation which explored the target-environment configuration space, as well as the target's different conformers.

The basic QM/FQ energy is defined as the classical electrostatic interaction between the FQs and the QM density: [200]

$$E_{\mathrm{QM/FQ}} = \sum_{i=1}^{N_q} V_{\mathrm{QM}}[\rho](\mathbf{r}_i) q_i \tag{1.26}$$

where $V_{\mathrm{QM}}[\rho](\mathbf{r}_i)$ is the electrostatic potential due to the QM density of charge at the i-th FQ q_i placed at $\mathbf{r}_i$. Notice that non-electrostatic interaction terms, which have been recently proposed by some of us, [11] will not be considered in this work. By exploiting a Self Consistent Field (SCF) description of the QM moiety, the global QM/MM energy functional reads: [200–202]

$$\mathcal{E}[\mathbf{P}, \mathbf{q}, \boldsymbol{\lambda}] = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{1}{2}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} + \mathbf{q}^{\dagger}\mathbf{V}(\mathbf{P}) \tag{1.27}$$

where $\mathbf{h}$ and $\mathbf{G}$ are the one and two electron contributions to the energy and Fock operator, and $\mathbf{P}$ is the density matrix. The FQs consistent with the QM density are obtained by solving the following equation:

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q - \mathbf{V}(\mathbf{P}) \tag{1.28}$$

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

First Hyperpolarizability in the QM/FQ approach

The theoretical framework sketched above can be further extended to electric response properties. The microscopic response of a molecular system to an external electric field $\mathbf{E}(t)$ can be represented by an induced dipole moment $\boldsymbol{\mu}(t)$:

$$\boldsymbol{\mu}(t) = \boldsymbol{\mu}^0 + \boldsymbol{\mu}^{\omega}\cos(\omega t) + \boldsymbol{\mu}^{2\omega}\cos(2\omega t) + ... \tag{1.29}$$

Each Fourier amplitude in Equation 1.29, can be rewritten as a Taylor expansion with respect to the external electric field. [155] In particular, SHG, i.e. the generation of a photon at 2ω as a result of the interaction with an incident ω photon reads:

$$\boldsymbol{\mu}^{2\omega} = \frac{1}{4}\boldsymbol{\beta}(-2\omega;\omega,\omega) : \mathbf{E}^{\omega}\mathbf{E}^{\omega} \tag{1.30}$$

The most important contribution to this quantity is due to the first term, that is namely the first hyperpolarizability $\boldsymbol{\beta}$. It is a third rank tensor that can be described by a 3 x 3 x 3 matrix, whose 27 components can be reduced to 10 assuming Kleinmann's symmetry, i.e. $\beta_{\alpha\beta\beta} = \beta_{\beta\alpha\beta} = \beta_{\beta\beta\alpha}$. [203]

By exploiting the response theory formalism, the first-order hyperpolarizability $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ can be calculated as:[204, 205]

$$\boldsymbol{\beta}(-2\omega;\omega,\omega) = 2\mathrm{Tr}\boldsymbol{\mu}\mathbf{P}^{(2)} \tag{1.31}$$

where $\boldsymbol{\mu}$ is the electric dipole moment integral matrix and $\mathbf{P}^{(2)}$ is the second-order density matrix. A generic second-order density matrix is obtained by solving perturbed equations up to the second order; however, when only one dynamic perturbation is involved, it is possible to avoid the solution of the second-order coupled perturbed equations by using an iterative procedure to reconstruct the density matrix. [204–206] Thus, in order to calculate the electrostatic QM/FQ contribution to $\boldsymbol{\beta}(-2\omega;\omega,\omega)$, the first-order perturbed density matrix $\mathbf{P}^{(1)}$ is constructed by resorting to linear response theory and by solving the first-order CPHF/CPKS equations. The right-hand side of the CPHF equations is real: hence, $\mathbf{Q}_X = \mathbf{Q}_Y$ and it is possible to reduce the response equations to a problem of half dimension solving for $\mathbf{X}+\mathbf{Y}$. By summing the CPHF equations, we obtain:

$$(\tilde{\mathbf{A}}+\tilde{\mathbf{B}})(\mathbf{X}+\mathbf{Y})+2\mathbf{Q}=0 \tag{1.32}$$

which can be used together with $(\tilde{\mathbf{A}}+\tilde{\mathbf{B}})(\mathbf{X}-\mathbf{Y})=0$. Notice that FQ contributions affect the orbital rotation Hessian $\tilde{\mathbf{A}}+\tilde{\mathbf{B}}$.

If one or more oscillating electric fields are applied, the response equations need to be generalized to the frequency dependent case. Frequency dependent CPHF equations (FD-CPHF) need to be solved and the proper transition densities are to be used to compute the desired properties:[200]

$$\begin{pmatrix} \tilde{\mathbf{Q}}_X \\ \tilde{\mathbf{Q}}_Y \end{pmatrix} + \begin{pmatrix} \tilde{\mathbf{A}}-\omega\mathbb{I} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}} & \tilde{\mathbf{A}}+\omega\mathbb{I} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = 0 \tag{1.33}$$

Computational Details

QM/FQ calculations of the SHG hyperpolarizabilities were performed by resorting to the following multi-step protocol:

1. *Definition of the system and calculation of atomic charges.* The six molecules depicted in Figure 1.10 were surrounded with a number of water molecules sufficient to represent all the relevant solute-solvent interactions. The atomic charges of the solutes were computed using the RESP approach.

2. *Classical MD equilibration, simulation and sampling.* Minimization runs were performed to yield the starting configurations used in the MD simulation production runs: minimization was accomplished by resorting to the NPT ensemble. MD production runs were carried out for each of the six molecules long enough for obtaining a sufficient sampling of a representative portion of the phase-space, so to correctly reproduce all possible system configurations and their relative energy. In order to refine the description of hydrogen bonding (HB) interactions, off-site charges (the so-called Virtual Sites (VS) or dummy atoms) with a fixed position with respect to the generating atom were added. In such a way, also the directionality of HBs was recovered and described. Two different classical MD simulation runs were performed for each molecule, i.e. with and without the inclusion of VS ($\mathrm{MD_{VS}}$ and $\mathrm{MD_{noVS}}$, respectively). From the MD runs, a set of snapshots was extracted to be used in the QM/FQ calculations.

3. *Definition of the different regions of the two-layer scheme and their boundaries.* Each snapshot extracted from the MD runs was cut into a sphere centered on the solute, retaining only the solvent molecules within the sphere. The radius of this region was chosen to include specific solute-solvent interactions.

4. *Running SHG QM/FQ calculations. Analysis of the results, comparison with experimental data.* SHG $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ calculations were performed on the set of structures obtained for the six molecules in the previous step of the protocol. The results obtained for each spherical snapshot were extracted and averaged to produce the final SHG value.

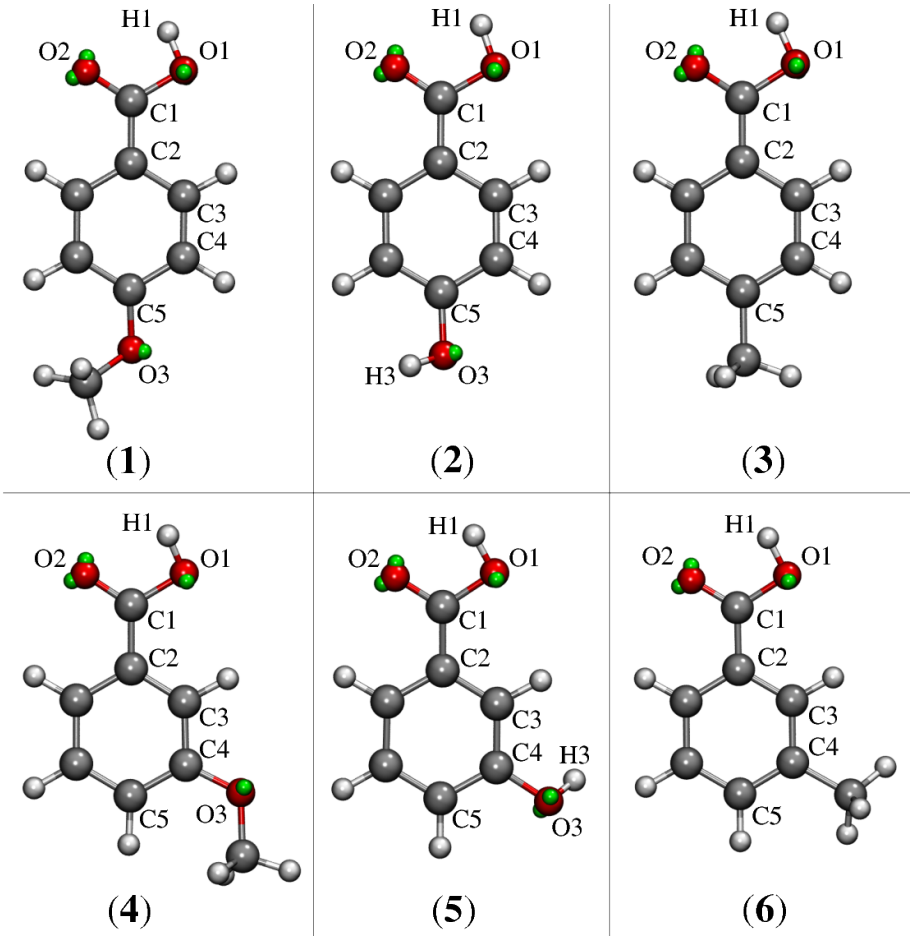


Figure 1.10: Structure of the molecules studied in this work. The green spheres depicted on each Oxygen atom represent the Virtual Sites (VS).

Molecular geometries of the molecules depicted in Figure 1.10 were optimized and the RESP atomic charges were calculated at the CAM-B3LYP/6-311++G** level of the theory. The Polarizable Continuum Model (PCM) was used to account for the aqueous environment in such optimizations. [2, 207] MD simulations were carried out by using GROMACS[141] with Amber 99ffSB-ILDN force field to describe intramolecular and intermolecular potentials. [208] RESP charges were used for electrostatic interactions. [209, 210] VS were possibly placed on the centroids of Boys orbitals. [211, 212] In particular a couple of VS was assigned to each carboxylic and hydroxylic Oxygen atom (see Figure 1.10). A single solute molecule was

solvated in a cubic box with a side length of approximately 4.7 nm containing a variable number of water molecules (3400-3900) modeled using the TIP3P parameter set, depending on the considered molecule **1**-**6**. [213] Electrostatic interactions were taken into account by means of the particle mesh Ewald method [145] using a cutoff radius of 1.2 nm in real space. A cutoff radius of 1.2 nm was also chosen for van der Waals interactions. Periodic boundary conditions were applied in all directions.

For each molecule, two single short (1.0 ns) NPT MD_{noVS} and MD_{VS} simulations were performed at 300 K for thermalization purposes. Consequently, two 10 ns NVT MD_{noVS} and MD_{VS} simulations were carried out for each molecule. The MD simulations were carried out using a constant temperature of 300 K and adopting the velocity-rescale method with a coupling constant of 0.1 ps and a time step of 0.5 fs. [214] The LINCS algorithm was used in order to constrain all bonds of the solute molecules. [144] The coordinates of each system were stored every 0.5 ps of simulation.

A total of 200 uncorrelated snapshots were extracted from the 10 ns MD (one snapshot every 50 ps). For each snapshot a 15 Å sphere centered at the solute's geometric center was cut. This radius assures the convergence of the computed data (see Figure S1 given as Supporting Informations (SI) [215]). All hyperpolarizabilities ($\boldsymbol{\beta}(-2\omega;\omega,\omega)$) were calculated within the QM/MM framework at the CAM-B3LYP/6-311++G** . For molecule **1** a comparison between CAM-B3LYP, B3LYP and ωB97X-D functionals and cc-pVDZ, aug-cc-pVDZ and 6-311++G** basis sets was also performed. The water molecules were modeled both with the SPC FQ parameters [27] and the parametrization proposed by some of the present authors. [128] The TIP3P [122] force-field was exploited in non-polarizable QM/MM calculations. The $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ convergence, as a function of the number of snapshots, was checked for each system. All QM/FQ calculation were performed by using a locally modified version of Gaussian 16. [146] Finally, the calculated values were compared with experimental Hyper Rayleigh Scattering (HRS) data taken from Ref. [216] Experimental HRS values were divided by the Lorentz local-field factor in order to be directly compared with our calculated data. [217–219]

In Ref. [216] a comparison between computed and experimental data was done by referring to the following quantity:

$$|\vec{\beta}| = \sqrt{\beta_x^2 + \beta_y^2 + \beta_z^2} \tag{1.34}$$

where:

$$\beta_i = \sum_{k=1}^{3} \beta_{ikk} + \beta_{kik} + \beta_{kki} \tag{1.35}$$

Therefore, our calculated data refer to Equation 1.34. We note, however, that alternative definitions for HRS values, giving computed results directly comparable with experimental data, can be found in the literature. [218, 220–222]

MD Analysis

The analysis of MD trajectories was performed by using the TRAVIS package. [149] Three different results are presented and discussed: the radial distribution function $g(r)$ (RDF), the spatial distribution function (SDF) and the dihedral distribu-

tion function (DDF). It is worth noticing that from a structural point of view, the molecules depicted in Figure 1.10 are strictly related. In fact, molecules **1**-**3** are characterized by the same electron-donor group as molecules **4**-**6**. The only difference between the two sets of molecules is that the first triplet is *para*-substituted, whereas the last is *meta*-substituted.

Hydration patterns

The hydration pattern was analyzed by means of the radial distribution function $g(r)$ computed taking as reference all Oxygen and hydroxil Hydrogen atoms of the solutes. **2** bears the largest number of potential HB sites: for this reason, we will focus on its RDF (Figure 1.11), whereas the data for the other molecules are given as SI (see Figure S2 [215]). The left and right panels of Figure 1.11 report the $g(r)$ obtained from the two MD runs performed with or without virtual sites (MD_{VS} and MD_{noVS}, respectively).

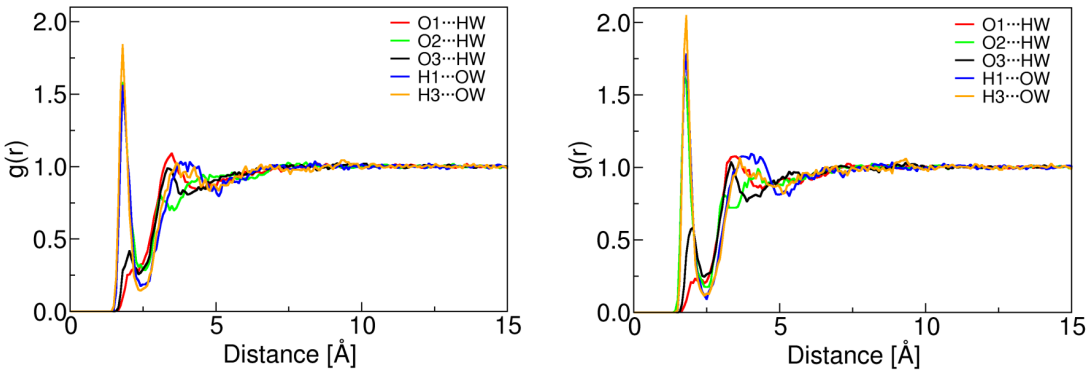


Figure 1.11: RDF of system **2**, obtained from the analysis of the MD run without (left panel) or with (right panel) virtual sites.

Focusing on both panels of Figure 1.11, it is worth pointing out that the most intense peaks of the $g(r)$ are presented by the carboxylic Oxygen (O2) and the water Hydrogen (HW) atoms, and by both the hydroxylic Hydrogen (H1 and H3) and water Oxygen (OW) atoms. Moreover, with the exception of O1, all $g(r)$ maxima increase if virtual sites are considered in the MD runs (right panel). Notice also that such an effect is accompanied by a thinning of the $g(r)$ width. The presence of VSs in the MD runs guarantees a better reproduction of the properties of HBs, in particular their directionality. This results in a smaller spatial spreading of the HBs and therefore a thinner maximum in the $g(r)$.

Table 1.4 reports the average number of HBs per site, as obtained from the analysis of both MD_{noVS} and MD_{VS}. Such numbers were calculated by integrating the first peaks of each $g(r)$, which refer to the first hydration sphere. The number of HBs exhibited by O1 and H1 is constant for most of the molecules, with the exception of O1 in **5** in the case of MD_{noVS}. The same behaviour is shown by H3, but not by O3, whose HB pattern varies a lot among the various molecules. Notice also that in case of MD_{VS} the number of HBs for O1 and O2 is generally lower than for MD_{noVS}. The opposite is noticed for O3, for which the presence of VSs in the MD

runs causes an increase in the number of HBs. Finally, it is worth pointing out that the smaller spatial spreading of HBs and the thinner maximum of the corresponding $g(r)$ already commented for MD_{VS} generally result in the decreasing of the average number of HBs (see column labelled HB_{mean} in Table 1.4).

		O1···HW	O2···HW	O3···HW	H1···OW	H3···OW	HB_{mean}
MD_{noVS}	**1**	0.64	2.21	0.76	1.03		1.16
	2	0.55	2.37	0.73	1.04	0.98	1.14
	3	0.55	2.19		1.04		1.26
	4	0.56	2.42	0.66	1.15		1.20
	5	0.33	2.13	1.21	1.01	0.99	1.14
	6	0.50	2.08		1.00		1.19
MD_{VS}	**1**	0.63	1.96	0.97	1.00		1.14
	2	0.51	2.02	1.06	1.00	1.01	1.12
	3	0.51	2.01		1.00		1.17
	4	0.49	1.91	1.00	1.01		1.10
	5	0.55	1.86	1.34	1.00	1.02	1.15
	6	0.57	1.90		1.01		1.16

Table 1.4: Number of Hydrogen Bonds for the different molecules depicted in Figure 1.10. HB_{mean} reports the average number of HBs reported in the other columns.

To refine the analysis on hydration patters, SDF were calculated from the MD_{VS}: the results are plotted in Figure 1.12 for all the molecules depicted in Figure 1.10. SDF calculated from MD_{noVS} are reported in Figure S3, as SI. [215] Calculated SDF isodensity values are equal to 70 and 100 nm^{-3} for water Hydrogen and Oxygen atoms, respectively. Figure 1.12 gives a pictorial view of HBs spatial distribution: red and white surfaces refer to water Oxygen and Hydrogen atoms, respectively. All investigated molecules present a common feature in their acceptor portion, with a strong O2-HW and H1-OW hydrogen bonding interaction. This is not surprising considering the results discussed before and reported in Table 1.4. **2** and **5** present an H3-OW HB which is symmetrically distributed. This suggests that the hydroxyl moiety (O3H3) can rotate during the MD simulation. These findings are confirmed by the dihedral distribution function depicted in Figure 1.13 for molecule **2**. In fact, the distribution of the C4-C5-O3-H3 dihedral angle shows two maxima at 0 and 180 degrees, thus confirming two most probable configurations. In Figure 1.13, the distribution related to the dihedral angle of the donor hydroxyl group (C2-C1-O1-H1) is plotted. Notice that in this case a single maximum occurs at 0 degrees, thus confirming the SDF sampling.

Convergence with respect to sampling extracted from the MD

Before analyzing the QM/FQ $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ results for the selected systems, we first analyze the dependence of the calculated data on the sampling extracted from the MD, i.e. on the number of selected uncorrelated snapshots. In Figure 1.14, the average values of $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ as a function of the number of snapshots for the six selected systems are plotted. Notice that panel (a) refers to the snapshots ex-

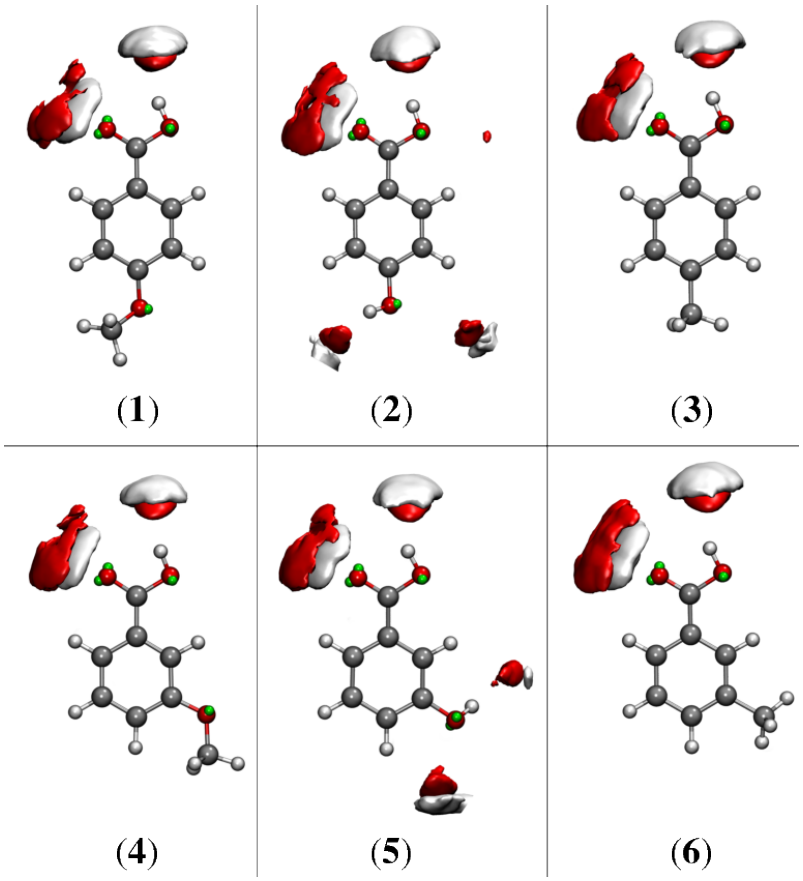


Figure 1.12: Spatial distribution functions extracted from MD_{VS}. Calculated SDF isodensity values are given in nm^{-3} and are equal to 70 and 100 nm^{-3} for water hydrogen and oxygen atoms, respectively.

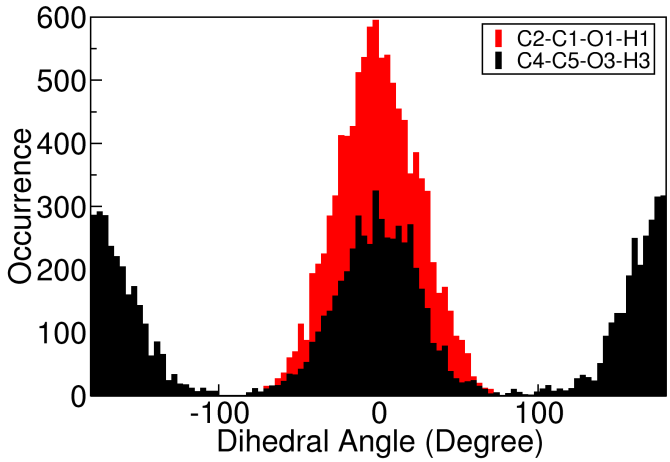


Figure 1.13: Dihedral distribution function of molecule **2** obtained with MD_{VS}.

tracted from the MD performed without the inclusion of VS, whereas (b) to the MD with the inclusion of VS. In both cases, the convergence in the average property is almost reached with 100 snapshots and completely guaranteed if the final values are obtained by averaging 200 snapshots. For this reason, all the data which will be reported in the next sections are obtained as a result of the averaging over 200 snapshots.

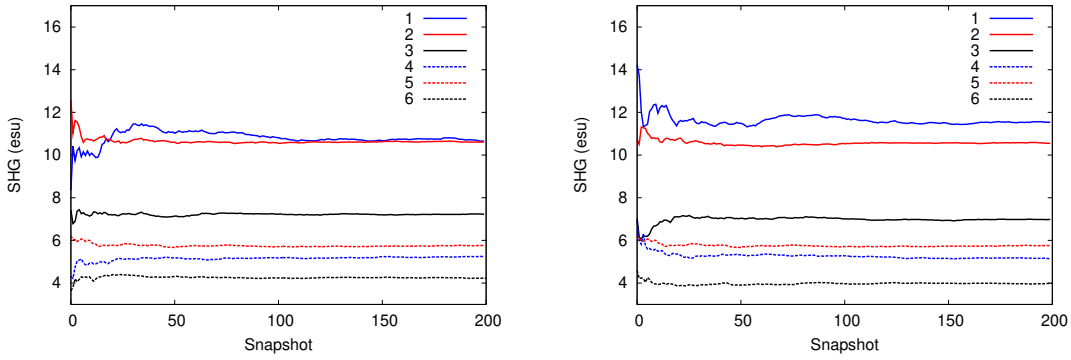


Figure 1.14: Average values of $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ as a function of the number of snapshots extracted from the MD for all the molecules depicted in Figure 1.10. (a) Snapshots extracted from MD_{noVS}. (b) Snapshots extracted from MD_{VS}.

Dependence on DFT Functional and Basis Set

The dependence of the $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ on the choice of the DFT functional and basis set is discussed for molecule **1**. The CAM-B3LYP, B3LYP and ωB97X-D DFT functionals were combined with the cc-pVDZ, aug-cc-pVDZ and 6-311++G** basis sets. The average $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ results are reported in Table 1.5.

	Functional	cc-pVDZ	aug-cc-pVDZ	6-311++G**
MD_{noVS}	B3LYP	8.65	13.31	13.59
	CAM-B3LYP	7.36	11.31	11.54
	ωB97x-D	7.04	10.84	11.08
MD_{VS}	B3LYP	8.63	12.03	12.39
	CAM-B3LYP	7.43	10.36	10.66
	ωB97x-D	7.09	9.91	10.22

Table 1.5: Calculated QM/FQ $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ of molecule **1** as varying the DFT functionals and basis set. Data refer to both the MD runs performed without or with the inclusion of VS. All data are reported in esu.

Table 1.5 shows some general trends. First, the calculated $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ increases going from the cc-pVDZ to the 6-311++G** basis sets for all functionals employed. By taking as reference the cc-pVDZ basis set, aug-cc-pVDZ shows an increase in the values of 39% (MD_{noVS}) and 54% (MD_{VS}), whereas the 6-311++G** set of 44% (MD_{noVS}) and 57% (MD_{VS}). Such a behaviour can be explained by considering that $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ values have already been reported to be extremely sensitive to

the inclusion of diffuse functions in the basis set. [223–225] The largest values are obtained with the 6-311++G** set, however these data are only slightly different from those calculated by using the aug-cc-pVDZ set. Let us pass to consider the dependence of $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ values on the choice of the DFT functional. For both MDs, the smallest value is given by ωB97x-D and the highest by B3LYP. With respect to ωB97x-D, CAM-B3LYP data are 4-5% greater, whereas B3LYP values are 22-23% greater. Such results are in agreement with previous studies, [186] which showed that B3LYP generally overestimates electric response properties.

On the basis of the data discussed so far, the further analyses are performed by exploiting the CAM-B3LYP/6-311++G** values.

QM/MM $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ results

In panels (a) and (b) of Figure 1.15 raw $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ QM/FQ values for molecule **1** are plotted. Again, the values obtained with or without the inclusion of VS are depicted. Mean values, obtained by averaging the property all over the snapshots, are also plotted as horizontal lines. The plots in Figure 1.15 clearly show the range of variability in time of the calculated property, i.e. the dependence of the calculated data on the spatial arrangement of the solvent molecules around the solute, and its conformation.

Panels (c) and (d) of Figure 1.15 report the same analysis, this time focused on solvent effects. The latter are calculated as the difference between QM/FQ and vacuum $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ at the same level of theory (CAM-B3LYP/6-311++G**). Both MD runs without and with VS are considered. The inspection of the plots shows positive, as well as negative solvent effects for the single snapshot. Therefore, although the average solvent effect is positive (see the horizontal lines), the contribution to the single snapshot can be either positive or negative. Our dynamical, atomistic approach to the solvation phenomenon is able to give insight into such a variability, whereas mean-field approaches (such as continuum solvation) would instead focus on the mean value. Similar findings can also be extracted from the analysis of molecules **2-6**, whose data are given as SI (see Figures S4-S8 [215]).

The calculated values in vacuo, the average QM/FQ and QM/TIP3P $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ values (together with their standard errors) are reported in Tables 1.6. Experimental data, taken from Ref. [216] are also reported. The calculated vacuum $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ data only qualitatively reproduce the experimental trend; in fact, absolute values are underestimated, with the average deviation being of the order of 50%.

Moving to QM/FQ data (Table 1.6), we note that both MD runs, with or without VS, give similar $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ results, being the average difference of the order of 3% . The major discrepancy between the two MD runs is given by molecule **1**: this is probably due to the larger number of water molecules bound to O3, via HBs. In fact, by chemical intuition, the more water molecules are bound to O3, the less π coordination occurs in the molecule, thus reducing the absolute value of $\boldsymbol{\beta}(-2\omega;\omega,\omega)$. Notice also that calculated $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ for molecule **5** are larger than the corresponding data for molecule **4**. Such a behaviour is not modelled by vacuum calculations. Therefore, it cannot be attribuited to the level of theory chosen, but instead to solvent effects described by coupling the MD to the atomistic description of the QM/FQ model.

Also, solvent effects (i.e.the difference between QM/FQ and vacuum values) reported as percentages in Table 1.6, follow indeed a specific trend. In fact, solvent

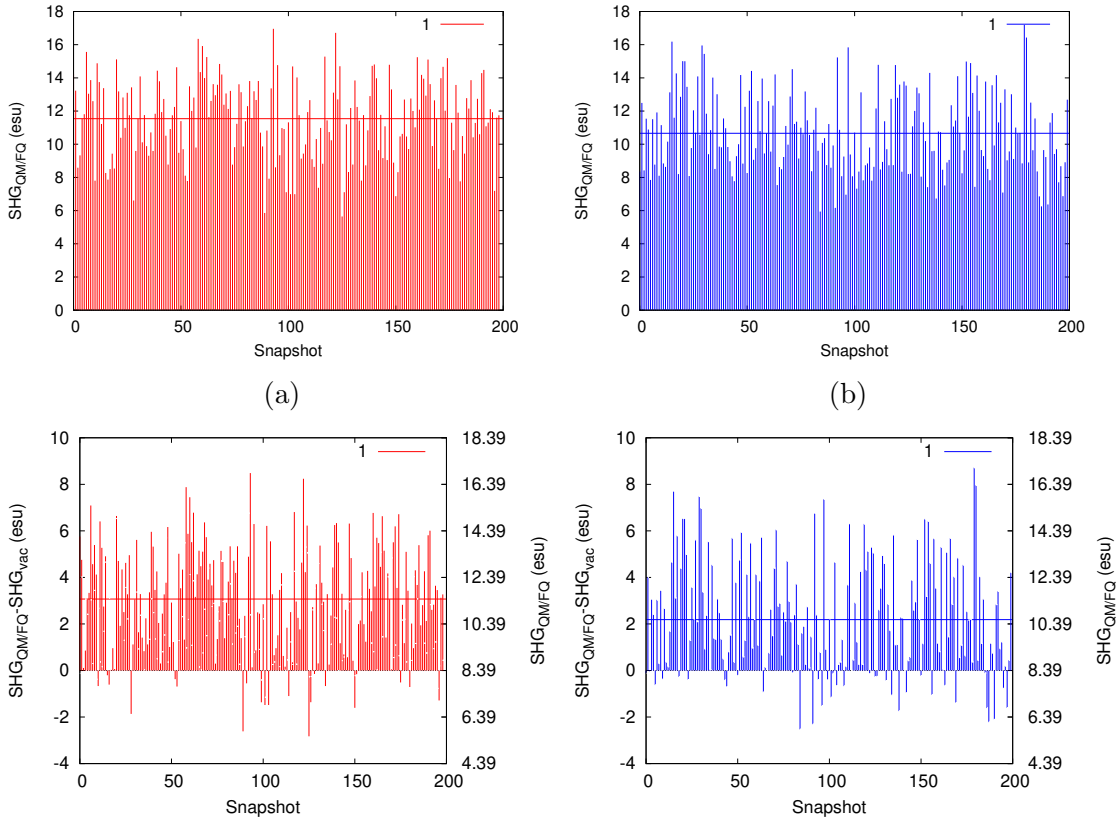


Figure 1.15: QM/FQ $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ values for molecule **1** calculated for the different snapshots extracted from the MD runs without ((a))-((c)) or with ((b))-((d)) VS. Panels (a) and (b) report $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ raw data, panels (c) and (d) report difference between QM/FQ and vacuum $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ data. All values are given in esu.

effects always increase moving from molecule **1** to **3** and from **4** to **6**. Thus, the observed smaller $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ value for **4** is numerically due to the similarity in the vacuum values for **4** and **5**. However, the final difference in the $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ values for **4** and **5** is very small. To end the discussion on solvent effects on $\boldsymbol{\beta}(-2\omega;\omega,\omega)$, we note that, on average, they increase vacuum values of about of 53 % and 60 % for $\mathrm{MD}_{\mathrm{noVS}}$ and $\mathrm{MD}_{\mathrm{VS}}$, respectively.

In order to evaluate the dependence of our findings upon the specific parametrization exploited for modeling the aqueous solution by means of the FQ approach, both the parametrization proposed by Rick et al. [27] and by Carnimeo et al. [128] are compared (see columns $\mathrm{QM/FQ}^a$ and $\mathrm{QM/FQ}^b$ in Table 1.6) . The most relevant discrepancy between the two parametrizations is the difference in electronegativities between water and oxygen and hydrogen atoms, which in case FQ^b is greater than case FQ^a. FQ charges are calculated by solving Equation 1.27, therefore a greater difference in electronegativities causes greater electrostatic interactions (i.e. larger absolute electric charges). For this reason, it is not surprising that $\mathrm{QM/FQ}^b$ results are larger than $\mathrm{QM/FQ}^a$, being the difference 5% in case of $\mathrm{MD}_{\mathrm{noVS}}$ and 3% in case of $\mathrm{MD}_{\mathrm{VS}}$.

Finally, calculated QM/FQ $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ are compared to experimental values (Table 1.6). Clearly, the inclusion of solvent effects, described by means of our

protocol, drastically reduces the error between calculations and experimental values. In fact, the error is 22% and 23% in case of MD_{noVS} and MD_{VS} respectively, with the minimal error (20%) being reported for MD_{noVS} and the FQ^b parametrization.

To end the discussion on calculated QM/MM $\boldsymbol{\beta}(-2\omega;\omega,\omega)$, it is worth noticing that the non polarizable QM/TIP3P approach underestimates experimental of about 37% on average. Thus, by taking as reference the calculations for the isolated molecules, indeed a non-polarizable atomistic description of the environment reduces the discrepancy with respect to experimental data. However, solvent effects are dramatically underestimated by the non-polarizable force field.

		Vacuum	QM/TIP3P	QM/FQ[a]	QM/FQ[b]
MD_{noVS}	**1**	8.47	9.63±0.15 (14%)	11.54±0.17 (36%)	12.11±0.18 (43%)
	2	7.10	8.61±0.07 (21%)	10.54±0.08 (48%)	10.96±0.09 (54%)
	3	4.15	6.00±0.08 (44%)	6.98 ±0.07 (68%)	7.48 ±0.09 (80%)
	4	3.71	4.44±0.06 (20%)	5.15 ±0.07 (39%)	5.45 ±0.07 (47%)
	5	3.63	4.83±0.05 (33%)	5.76 ±0.05 (58%)	5.98 ±0.06 (65%)
	6	2.35	3.43±0.05 (46%)	3.99 ±0.05 (70%)	4.28 ±0.06 (82%)
MD_{VS}	**1**		8.65±0.15 (2%)	10.66±0.17 (26%)	10.92±0.18 (29%)
	2		8.40±0.08 (18%)	10.61±0.09 (49%)	10.79±0.10 (52%)
	3		6.08±0.07 (46%)	7.23 ±0.07 (74%)	7.53 ±0.08 (81%)
	4		4.36±0.06 (17%)	5.24 ±0.06 (41%)	5.41 ±0.07 (46%)
	5		4.69±0.05 (29%)	5.73 ±0.05 (59%)	5.88 ±0.06 (65%)
	6		3.52±0.05 (50%)	4.23 ±0.04 (80%)	4.43 ±0.05 (88%)

Table 1.6: CAM-B3LYP/6-311++G** Vacuum, QM/TIP3P and QM/FQ $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ (± standard errors) for molecules **1-6**. Solvent effects, defined as the difference between the QM/FQ or QM/TIP3P and the vacuum data, are given as percentages in brackets. All values are reported in esu.
[a] FQ parametrization proposed by Rick et al. [27]
[b] FQ parametrization proposed by Carnimeo et al. [128]
Experimental values are reported from Ref. [216]:
1 13.06 **4** 9.91
2 10.93 **5** 6.78
3 10.28 **6** 6.57

Chapter 2

Towards non-aqueous environments

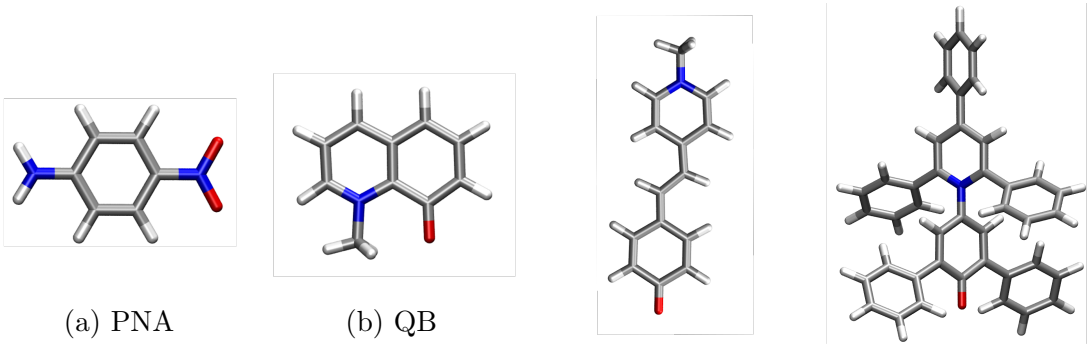
QM/Fluctuating Charge Protocol to Compute Solvatochromic Shifts

The study of electronic, optical properties of chromophores is of particular interest for many different applications, ranging from photochemistry to technology.[226] In most cases, such optical properties are tuned by dissolving the selected dye in different solvents, which can yield to substantial changes in the solute's properties.[227] When dealing with electronic excitations, solvent effects mainly manifest in a shift of the solute's absorption band,[228] which is usually referred to as solvatochromism, or solvatochromic shift. [227, 229–235] Depending on the nature of the transition, blue- or red-shifts are observed;[227] therefore, reliable computational approaches need to correctly reproduce both the 'sign" of the solvatochromic shift, and its magnitude as a function of the nature of the solvent.[234]

To this end, different methods have been proposed, generally focusing the attention to the solute, which is responsible for the spectral signal and is accurately described at the Quantum-Mechanical (QM) level. The solvent, which modifies but does not determine the spectral property, is instead treated at lower level of sophistication.[236, 237] The various approaches differ in the way they describe the solvent, which can be treated implicitly, as a continuum, or atomistically.[237, 238] In the latter case, QM/Molecular Mechanics (MM)[74, 97] or quantum embedding approaches [239, 240] may be used, where solvent molecules are described by a classical force field or a QM wavefunction, respectively. Another remarkable difference between the three aforementioned approaches is the way they model solute-solvent interactions. Continuum solvation is a "mean-field" approach, where specific, direction interaction (e.g. hydrogen bonding) are neglected.[236] The latter are instead considered in explicit, atomistic approaches, which are specified by the quality of the description of solvent molecules and their interaction with the QM solute.[74, 241–244] In most quantum embedding methods, electrostatic, polarization and Pauli repulsion solute-solvent interactions are accurately treated,[242, 245] whereas QM/MM approaches generally discard quantum repulsion,[246, 247] and sometimes also polarization effects.[248] However, being the computational cost of QM/MM methods much lower than most quantum embedding methods, the former are rapidly becoming the golden-standard for many applications,[241, 248, 249] especially those refined approaches which are able to correctly take into account solute-solvent mutual polarization effects (i.e. in the so-called polarizable QM/MM embedding methods).[23, 241, 249–252]

Polarizable QM/MM embedding approaches can be based on distributed multipoles, [253] induced dipoles, [248, 250–252] Drude oscillators, [26, 254] fluctuating charges (FQ) [29, 104] and possibily dipoles, [31, 255, 256] or AMOEBA.[19, 257, 258] In particular, the QM/FQ approach has been specifically developed to model spectral properties.[241] There, MM atoms are endowed with a charge that can vary as a function of differences in MM electronegativity or as a response to the electric potential generated by the QM density.[104]

Although the excellent performance of QM/FQ to describe aqueous solutions, mainly due to a reliable parametrization of the force field and the extension up to analytical energy third derivatives, [103, 110, 111, 215, 259–261] its application to non-aqueous solutions has been severely hampered by lack of reliable parametrizations. In fact, QM/FQ is a completely general approach, which can be applied to



(a) PNA (b) QB (c) MER (d) BET

Figure 2.1: Molecular structures of the studied dyes.

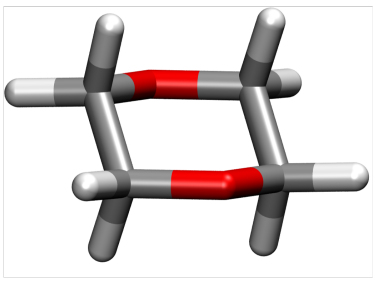
any embedded system, pending a reliable parametrization of the classical layer.[241] In this work, we extend for the first time the FQ force field to six solvents of different polarity and hydrogen-bonding capability, thus substantially increasing the applicability of QM/FQ beyond aqueous systems. The method is tested to reproduce solvatochromic shifts of four chromophores, which exhibit large solvatochromic shifts when dissolved in polar, protic and apolar, aprotic solvents, namely: para-nitroaniline (PNA), 1-methyl-8-oxyquinolinium betaine (QB), 1-methyl-4-[(oxocyclohexadienylidene)et 1,4-dihydropyridine (MER) and 2,6-Diphenyl-4-(2,4,6-triphenylpyridin-1-ium-1-yl) phenolate (BET). The molecular structures of the dyes are given in Fig. 2.1

The paper is organized as follows. The next section reports on the computational methods which have been used to parametrize the FQ force field for six selected solvents. The novel FQ parameters are then exploited to predict solvatochromic shifts of the four chromophores in the selected solvents. A summary and a discussion on the future perspectives of the method end the paper.

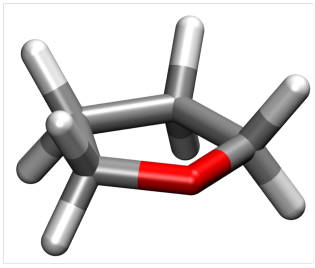
Parameterization Procedure

In QM/FQ, each solvent atom is endowed with a Fluctuating Charge (FQ), whose value can vary as a response to the QM electronic density. FQs are defined by solving a linear equation which is written in terms of atomic electronegativities (χ) and chemical hardnesses (η), which constitute the parameters of the FQ force field. In this paper we propose a novel parametrization for 1,4-dioxane (DIO), tetrahydrofuran (THF), acetonitrile (ACN), ethanol (ETH) and methanol (MET), and water (WTR), see Figure 2.2, where also polarity (ϕ) and dielectric constant (ϵ) of each solvent values are reported. Notice that ϕ is equivalent to the normalized E_N^T value usually exploited in Reichardt's scale.[234]

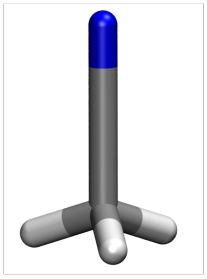
As already mentioned above, the application of QM/FQ to different environments needs appropriate specification of atomic electronegativities and chemical hardnesses.[27, 28, 196] In this work, we have decided to exploit atomic parameters independently of chemical makeups. For instance, in MET, the two Hydrogen atoms in C-H and O-H groups are characterized by the same atomic parameters. Notice however that our procedure is general and specific atom-types, defined similarly to other force fields, can also be considered. The parameterization workflow is sketched in Figure 2.3.



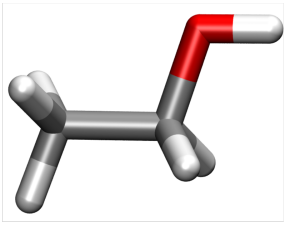
(a) 1,4-dioxane (DIO)
$\phi = 0.164$
$\epsilon = 2.2$



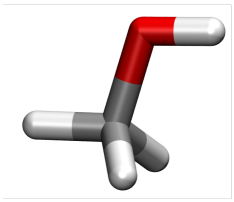
(b) tetrahydrofuran (THF)
$\phi = 0.207$
$\epsilon = 7.4$



(c) acetonitrile (ACN)
$\phi = 0.460$
$\epsilon = 35.7$



(d) ethanol (ETH)
$\phi = 0.654$
$\epsilon = 24.8$



(e) methanol (MET)
$\phi = 0.762$
$\epsilon = 32.6$

(f) water (WTR)
$\phi = 1.0$
$\epsilon = 78.3$

Figure 2.2: Geometries of the solvents studied in the present work ordered based on their polarity. Relative polarity (ϕ) as reported in Ref. [227] and dielectric constant (ϵ) used in QM/PCM calculations are also given.

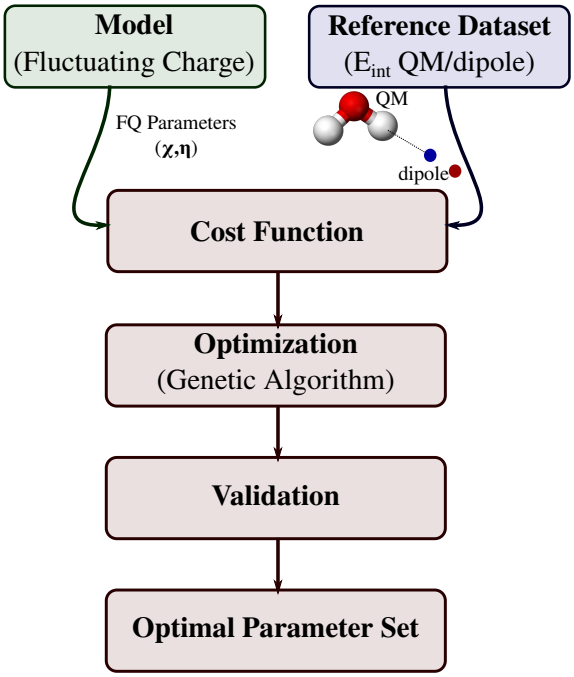


Figure 2.3: Graphical representation of the parameterization workflow.

Reference Dataset. The first step of the parametrization procedure involves the definition of a reliable reference data-set, that we built for the following four solvents

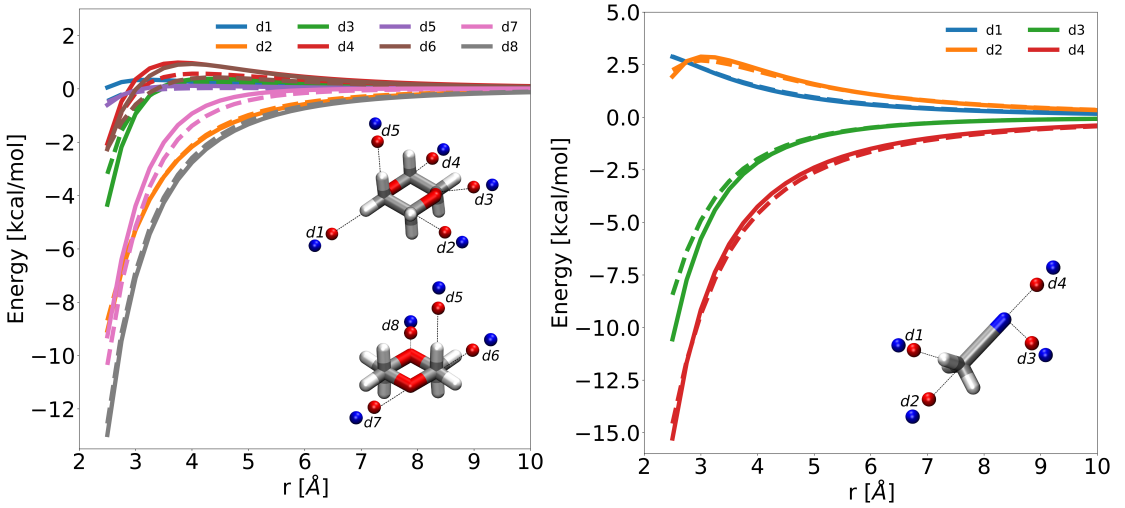
of increasing polarity: [227] DIO, ACN, MET and WTR. Notice that, although the parameterization procedure is solvent specific, the final parameters can be transferred to similar solvents, i.e. those that are constituted of similar geometrical structures or have similar physico-chemical properties. This is indeed the case of THF and ETH, for which the parameters were specified by transferring those obtained for DIO and MET, respectively (*vide infra*). The reference data-set was assembled by using solvent geometries optimized at the CCSD/aug-cc-pVTZ level of theory.
For all structures in the data-set, the interaction energy (E_{int}^{REF}) between each solvent molecule (treated at the QM level) and a dipolar probe was calculated. All QM/dipole reference calculations were performed at the Hartree-Fock (HF)/6-311++G** level, by using the electronic structure program $e^{\mathcal{T}}$. [262] In particular, we followed the strategy first proposed by Stern et al. [263] and recently exploited in Ref.[264]. In this approach, the dipolar probe is constituted by a pair of ± 1 a.u. point charges separated by 1 Å. The solvent-dipole interaction energy is then sampled along different symmetry axes and solvent-dipole distances. Depending on the solvent, a different number of points, ranging between 150 and 250, was exploited. The minimum/maximum distance between the probe and the solvent molecule is 2.0/10.0 Å, with a constant step of 0.25 Å. The results of the calculations for the reference data set obtained for the four solvents are plotted as solid line in Fig. 2.4.

Cost Function The loss function (ξ^2) used to train the model was defined as:

$$\xi^2 = \frac{1}{N}\sum_{i=1}^{N} w_i (E_{int,i}^{REF} - E_{int,i}^{FQ}(\boldsymbol{\chi}, \boldsymbol{\eta}))^2 \tag{2.1}$$

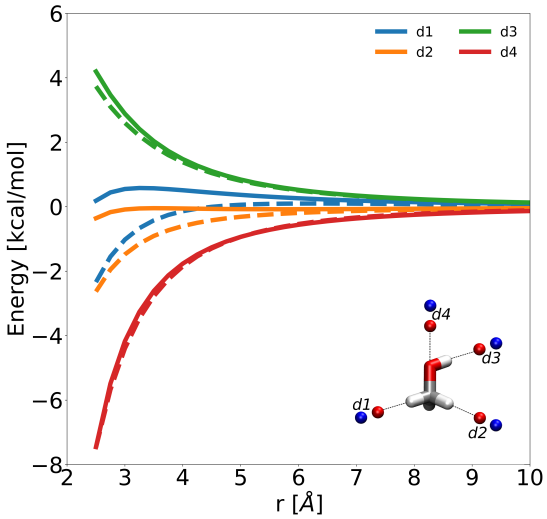
It is a Weighted Mean Squared Error (WMSE) loss function, where N represents the number of points used in the fitting, i.e. the solvent-dipole scan configurations obtained at the previous step. $E_{int,i}^{REF}$ is the QM/dipole reference interaction energy of the *i-th* geometry. On the same geometry, the FQ-dipole interaction energy ($E_{int,i}^{FQ}(\boldsymbol{\chi}, \boldsymbol{\eta})$) is computed by treating the solvent molecule at the FQ level. The FQ energy depends on $\boldsymbol{\chi}$ and $\boldsymbol{\eta}$, which are the adjustable parameters of the optimization. In Eq. 2.1, w_i is a weighting factor, which was set to the inverse of the solvent-dipole distance.

Optimization. FQ parametrization is a non-convex optimization problem, because the two parameter sets $\boldsymbol{\chi}$ and $\boldsymbol{\eta}$ are not linearly related to the interaction energy $E_{int,i}^{FQ}(\boldsymbol{\chi}, \boldsymbol{\eta})$. As a consequence, the function is characterized by several local minima, thus a gradient-based algorithm (e.g., steepest descent, conjugate gradient method), would not be effective. For this reason, we exploited a genetic algorithm (GA), which is an evolutionary algorithm that mimics the process of natural selection. [265, 266] In particular, we used the *Python* module *inspyred*, [267, 268] where the parameter space ($\{\boldsymbol{\chi}, \boldsymbol{\eta}\}$) was restrained to the [0,1] interval.
For each solvent molecule, a minimum of twenty parameter optimizations was performed, each constituted by a population of 100 individuals, corresponding to a randomly generated parameter set ($\boldsymbol{\chi}, \boldsymbol{\eta}$). Therefore, for each molecule 2000 different starting points in the parameter space were considered.

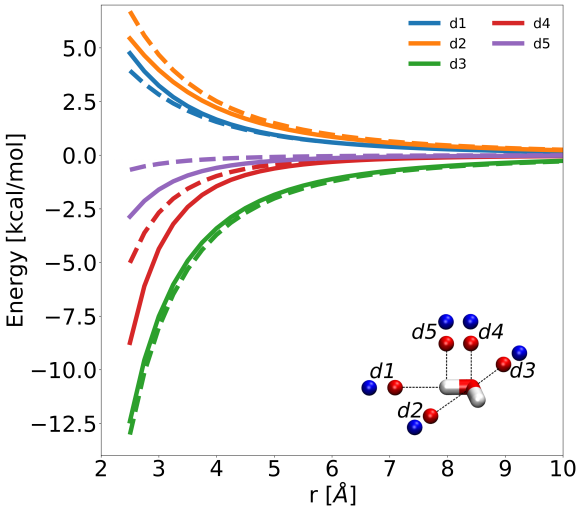


(a) Dioxane

(b) Acetonitrile



(c) Methanol



(d) Water

Figure 2.4: Reference QM/dipoles (solid lines) and FQ/dipoles (dashed lines) interaction energies (in kcal/mol) as a function of the solvent-dipole distance for all the parametrized solvents (DIO, ACN, MET, WTR).

Validation. GA is a stochastic optimization technique. At the end of the previous step, multiple sets of sub-optimal parameters, i.e. characterized by similar values of the loss function ξ^2, were obtained. In order to select the best parameter set, we performed an additional validation step, which is based on two physical observations. First, the parameter sets have to follow the electronegativity scale, i.e. atomic electronegativity need to increase as moving right along a row of the periodic table. Therefore, all sub-optimal parameter sets that do not follow this criterion were rejected.
Second, we imposed the parameter sets to correctly reproduce the molecular static isotropic polarizability α_{mol}, computed at the CCSD/aug-cc-pVTZ level, and the static isotropic bulk polarizability α_{bulk}, computed at the CAM-B3LYP/6-311++G**, of the selected solvent. Note in fact that in the FQ formalism, $\boldsymbol{\eta}$ and α are strictly related. [269]
Note that, although the aforementioned validation criteria could have been included in the loss function, we preferred to keep the loss function independent of model specific features. It is worth remarking that the procedure may provide different optimal parameter sets, also in case the constraints mentioned above are incorporated in the loss function.

Optimal parameter set. The optimal parameter set, i.e. that associated to the lowest value of the loss function and the lowest error for both α_{mol} and α_{bulk}, was selected. The values obtained for each solvent molecule are reported in Tab. S1 given in the Supporting Information (SI). In Figure 2.4, FQ solvent-dipole interaction energies as a function of the solvent-dipole distance are reported. Reference QM/dipole values are also plotted for the sake of comparison. The FQ optimal parameter sets correctly reproduce reference curves, although with some discrepancies, essentially due to the lack of out-of-plane polarization in the FQ model (see for instance, d4 and d5 in Fig. 2.4d). This limitation of FQ may be overcome by adding terms depending on atomic dipoles, as it has been recently shown by some of us. [255]. Finally notice that our procedure permits to select the best parameters to reproduce the most relevant interactions of a specific solvent (see for instance d3-d4 vs. d1-d2 in Fig. 2.4c).

Computational Protocol

In order to compute QM/FQ UV-Vis absorption spectra of the selected dyes in the aforementioned solvents, dynamical aspects of the solvation phenomenon were considered by resorting to classical MD simulations. PNA, QB and MER geometries were optimized at the CAM-B3LYP/aug-cc-pVDZ level, whereas we adopted the CAM-B3LYP/6-31+G(d,p) level for BET, according to Ref. [19]. Solvent effects on molecular geometries were modeled by using the PCM. [236] All MD runs were performed by using GROMACS. [270] The GAFF force field was exploited to describe both intramolecular and intermolecular interactions. [271, 272] Solute and solvent bonded and non-bonded parameters were generated by means of Antechamber package [273, 274] with the only exception of water for which the standard TIP3P force field was used. [275] Atomic charges of both solute and solvent molecules were calculated by using the RESP charge fitting method. [276] During each MD run solutes were constrained in their minimum energy structure.

The size of the simulation boxes ranged from 7 nm (PNA) to approximately 10 nm (BET). An integration time step of 1 fs was adopted for all MD runs. The temperature was kept constant to 300 K by adopting the velocity-rescaling method [214] with a coupling constant of 0.1 ps. Electrostatic interactions were taken into account by means of the particle mesh Ewald method [277] using a cut-off radius of 1.4 nm in real space. The same cut-off was also used for van der Waals interactions.

An NPT simulation of 1 ns (using the Berendsen barostat [278] and a coupling constant of 2.0 ps) was performed on each system for equilibration purposes. Then, a 2.5 ns NVT production run was performed in order to sample the system's configuration space. A total of 100 uncorrelated snapshots were extracted from the last 2 ns of the MD (one snapshot every 20 ps). For each snapshot, a sphere of variable radius (20-25 Å), depending on the solute size, centered at the solute center of mass, was cut and used in the following QM/FQ UV-Vis calculations.

QM/PCM and QM/FQ vertical excitation energies were calculated at the TD-DFT level by exploiting Linear Response (LR) and corrected Linear Response (cLR) regimes, which is a first-order state-specific approximation. [256, 260] For all the transitions, both LR and cLR shifts with respect to the frozen density approximation (ω_0) were summed, in agreement to Refs. [279] and [280]. Such an approach has been recently proposed in the literature and named cLR^2.[281] All QM/PCM and QM/FQ ω_0, LR, cLR and cLR^2 energies are reported in Tabs. S2-S3 in the SI. All TD-DFT calculations were performed by using the CAM-B3LYP functional in combination with the aug-cc-pVDZ as basis set (PNA, QB, MER) or 6-31+G(d,p) (BET), the latter according with Ref. [19]. All QM/FQ calculations were performed by using a locally modified version of the Gaussian 16 package. [146] Non-polarizable QM/Electrostatic Embedding (QM/EE) were also performed by exploiting TIP3P and GAFF charges, for water and the other solvents respectively (see Table S4 in the SI for QM/EE vertical excitation energies).

Numerical Results

In this section we apply the parameters obtained above to describe absorption spectra and solvatochromic shifts of the dyes in Fig.2.1. For each molecule, we analyze the transition involved in the spectral signal, and we compare QM/FQ values with electrostatic embedding (QM/EE) and continuum solvation approaches (QM/PCM), so to disentangle the role of explicit solute-solvent interactions (when relevant) and polarization effects. Finally, computed data are compared with experimental values (see Tables S5 in the SI) and general trends are commented.

Para-nitroaniline

Para-nitroaniline (PNA) belongs to the family of "push-pull" organic compounds, being characterized by an electron-donor amino group (NH_2) and a para electron acceptor nitro group (NO_2), which are connected by a π-conjugated phenyl ring (see Figure 2.1a). Different theoretical approaches (continuum and atomistic) have already been challenged to reproduce PNA solvatochromic shifts, by also exploiting correlated wavefunctions.[282–284] The absorption spectrum of PNA is characterized by a bright band, which is due to a Charge-Transfer (CT) transition from the donor to the acceptor moieties.[282–289] Similarly to most CT transitions, PNA absorption maximum exhibits a large red solvatochromic shift as the polarity of the solvent

increases. Therefore, PNA represents an ideal candidate to test the quality of the FQ parametrization discussed above. Here we discuss PNA absorption spectra in DIO, THF, ACN, MET and WTR. For all methods (QM/FQ, QM/EE and QM/PCM), the lowest bright excitation is predicted to be a HOMO-LUMO $\pi \rightarrow \pi^*$ transition, with clear CT character (see Fig. 2.5a).

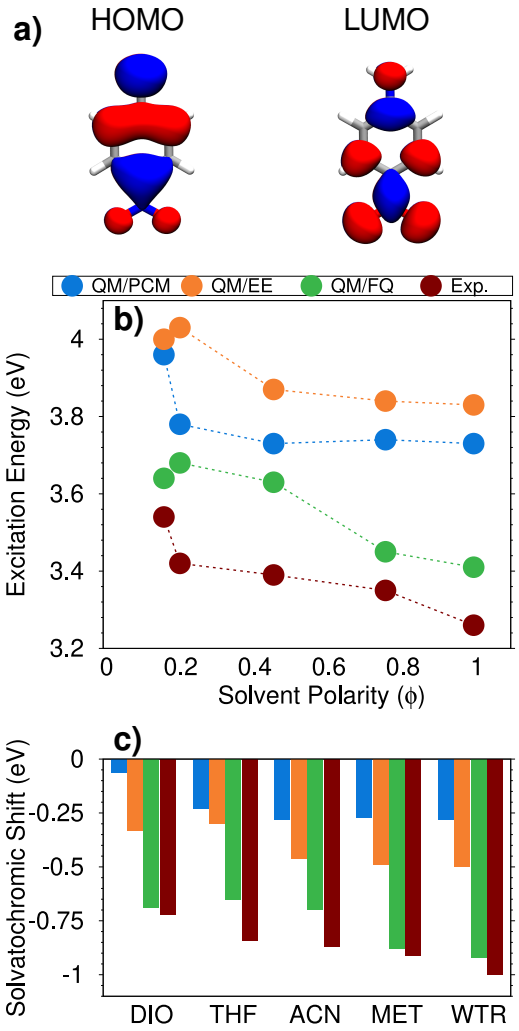


Figure 2.5: **a)** PNA HOMO and LUMO involved in the studied electronic transition. **b)** QM/PCM, QM/EE, QM/FQ and experimental PNA excitation energies as a function of the solvent polarity (ϕ). **c)** QM/PCM, QM/EE, QM/FQ and experimental PNA solvatochromic shifts computed with respect to gas-phase excitation energy.

In Fig. 2.5b, QM/FQ, QM/EE and QM/PCM excitation energies are compared as a function of the solvent polarity (ϕ). We first notice that, as expected, QM/PCM absorption energies remains almost constant when ϕ is larger than 0.4. This is not surprising, due to the well-known asymptotic behavior of PCM contributions for solvents' permittivity constants larger than 20.[236] Similar results are given by QM/EE, whereas QM/FQ excitation energies decrease when ϕ increases, thus matching experimental trends (see 2.5).[285, 288] For THF and DIO, the experimental trend is generally better reproduced by QM/PCM as compared to QM/FQ. This may be due to inaccurate calculation of the QM/FQ excitation energy of PNA dissolved in THF, because for THF the same FQ atomic parameters obtained for

DIO are used (see above).

In Fig. 2.5c QM/FQ, QM/EE and QM/PCM solvatochromic shifts (ΔE) computed as:

$$\Delta E = E^{vac} - E^{solv} \tag{2.2}$$

are compared with experimental values. Notice that Eq. 2.2 refers to vacuo-to-solvent solvatochromic shifts, defined according to Ref. [19]. In eq. 2.2, E^{vac} and E^{solv} are excitation energies in vacuo and in solution, respectively. Fig. 2.5 clearly shows that both QM/PCM and QM/EE strongly underestimate experimentally measured shifts, whereas QM/FQ gives very good values, in some cases in perfect agreement with experiments. More in details, QM/FQ gives errors for VAC $\rightarrow$ DIO, VAC $\rightarrow$ THF, VAC $\rightarrow$ ACN, VAC $\rightarrow$ MET and VAC $\rightarrow$ WTR shifts of 0.03, 0.19, 0.17, 0.03 and 0.08 eV, respectively, thus confirming the reliability of both the method and its parametrization for the different solvating environments. As expected, the largest error occurs for VAC $\rightarrow$ THF, because for THF the same parameters as DIO are exploited (see above).

It is worth noticing that our simulations disregard solute-solvent Pauli repulsion, dispersion effects, dynamical changes in solute configurations and vibronic effects, therefore a perfect agreement with experimental values is not expected.[290, 291] However, the exploited computational protocol permits to disentangle electrostatics/polarization effects on the solute response. The reduced error obtained by exploiting QM/EE and QM/FQ approaches as compared to the implicit (QM/PCM) method shows that an accurate, dynamic description of the solute-solvent interactions is needed, for both apolar (DIO) and polar solvents. Also, the polarizable QM/FQ outperforms the non-polarizable QM/EE when compared to the experimental data, thus showing that solute-solvent polarization plays a crucial role in determining the electronic properties of PNA in solution. Finally, the almost perfect agreement between QM/FQ and the experimental data clearly demonstrates the reliability of the novel FQ parametrization.

Quinolinium Betaine

Quinolinium betaine (QB, see Fig.2.1) has a zwitterionic character, therefore it is strongly hydrophilic. Its absorption spectrum is dominated by the transition from the dipolar ground state to an excited state of considerably reduced polarity.[292] The ground state is therefore stabilized in polar solvents, and this leads to an increase of the transition energy, i.e. to a positive solvatochromism. In this work, we have investigated the solvatochromic shift of QB dissolved in DIO, THF, ACN, MET and WTR.

For all solvents and embedding methods, the lowest bright excitation of QB corresponds to the HOMO-LUMO $\pi \rightarrow \pi^*$ transition, which has a partial CT character (see Fig. 2.6a). Calculated QM/PCM, QM/EE and QM/FQ excitation energies as a function of solvent polarity (ϕ) are reported in Figure 2.6b, together with experimental values taken from Ref. [292]. Similarly to PNA, QM/PCM is able to reproduce the experimental trend only for the less polar solvents (DIO, THF) while, as expected, the curve is substantially flat for the most polar solvents (ACN, MET, WTR). On the contrary, both QM/EE and QM/FQ well reproduce the experimental trend. Differently from PNA, in this case QM/EE values are closer to experiments

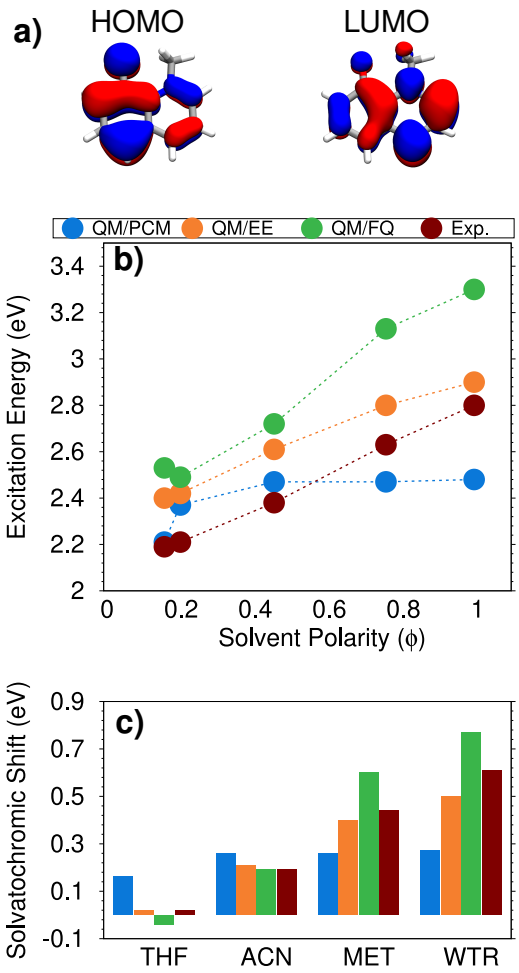


Figure 2.6: **a)** QB HOMO and LUMO involved in the studied electronic transition. **b)** QM/PCM, QM/EE, QM/FQ and experimental QB excitation energies as a function of the solvent polarity (ϕ). **c)** QM/PCM, QM/EE, QM/FQ and experimental QB solvatochromic shifts computed with respect to the excitation energy in dioxane.

than QM/FQ ones. Such a good agreement may be ascribed to systematic errors due to the selected level of theory;[19] therefore, in the following we will mainly focus on energy differences, i.e. solvatochromic shifts, which are less affected by systematic errors.

Finally, in fig. 2.6c computed and experimental solvatochromic shifts are reported. Note that they are calculated with respect to DIO, due to the lack of experimental spectra of QB in the gas phase reported in the literature. Fig. 2.6c clearly shows that QM/PCM cannot reproduce the experimental trend for the most polar solvents, whereas both QM/EE and QM/FQ computed solvatochromic shifts increase as the solvent polarity increases. In particular, we notice that for ACN the two atomistic approaches yield almost the same value (with QM/FQ being in almost perfect agreement with experiments). For MET and WTR, QM/FQ computed shifts are larger than their experimental counterparts, whereas the opposite occurs for QM/EE. This behavior can be due to the fact that Pauli repulsion effects, which might be large for MET and, especially, WTR,[290] are neglected in our calculations. The inclusion of such effects, which are always repulsive, would bring computed values towards the experimental findings because they act in opposite direction with respect to QM/FQ electrostatic and polarization contributions.

To conclude this section on QB, it is worth noticing that QM/FQ solvatochromic shift in THF is wrongly predicted in sign (negative instead of positive). However, the QM/FQ error is of about 0.06 eV ($\sim$ 1.3 kcal/mol), thus the absolute discrepancy can be considered satisfactory.

Brooker's Merocyanine

1-methyl-4-[(oxocyclohexadienylidene)ethylidene]-1,4-dihydropyridine also known as Brooker's merocyanine [293] (MER) has been amply studied both experimentally and theoretically due to the sensitivity of its absorption spectrum on the solvent polarity.[294–297] Similarly to PNA and QB, MER is characterized by two resonance structures, neutral (quinoid) and zwitterionic (benzenoid). The latter is stabilized by polar solvents, whereas the neutral form is predominant in apolar solvents. Also, the phenolate Oxygen atom can form hydrogen bonds with protic solvents, thus further increasing the weight of the zwitterionic form. [297–299] For the aforementioned reasons, MER is used as an indicator to measure solvents' polarity and hydrogen bond donor capability.

In this work, we focus on MER absorption spectra in THF, ACN, ETH, and WTR. MER lowest bright excitation corresponds to a HOMO-LUMO $\pi \rightarrow \pi^*$ transition, independently from the solvent and the solvation approach. Fig. 2.7 reports the pictures of MOs involved in the transition, which clearly has CT character, as the density moves from Nitrogen to Oxygen. In Figure 2.7b, experimental [295, 296] MER excitation energies as a function of the solvent polarity are compared with computed results obtained at the QM/PCM, QM/EE and QM/FQ levels. QM/PCM excitation energies are placed around 2.6 eV, in sharp contrast with experiments, which report an almost linear increase of transition energies by increasing the solvent polarity. Remarkably, this trend is reproduced by QM/FQ, whereas, similarly to QM/PCM, QM/EE excitation energies are almost constant in all solvents, with a slight increase when MER is dissolved in WTR.

Computed and experimental solvatochromic shifts with respect to DIO are reported in Fig. 2.7c. Although QM/FQ solvatochromic shifts are underestimated

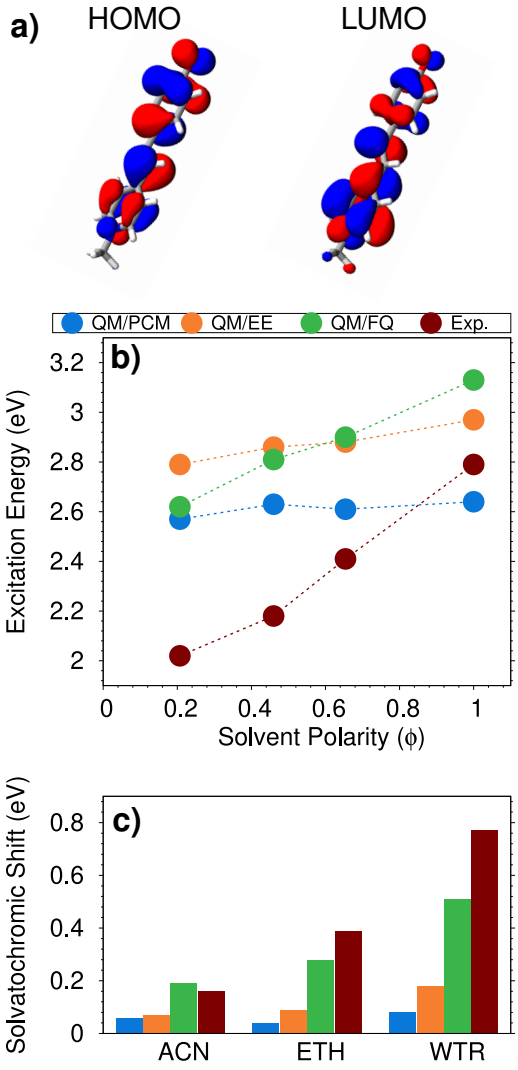


Figure 2.7: **a)** MER HOMO and LUMO involved in the studied electronic transition. **b)** QM/PCM, QM/EE, QM/FQ and experimental MER excitation energies as a function of the solvent polarity (ϕ). **c)** QM/PCM, QM/EE, QM/FQ and experimental MER solvatochromic shifts computed with respect to excitation energy in dioxane.

for ETH and WTR, QM/FQ results directly follow the experimental trend, whereas both QM/PCM and QM/EE solvatochromic shifts are almost unaffected by solvent polarity. Such a behavior demonstrates the important role of both specific and polarization solute-solvent effects. To conclude the discussion on MER, we also computed MER gas-phase vertical excitation energy, which is reported in Tab. S4 and S6, given as SI. QM/FQ is able to reproduce both the experimentally measured positive and negative solvatochromic shifts in highly polar (ETH and WTR) and medium-to-low polarity solvents, respectively;[298] remarkably, such a behavior cannot be described by QM/PCM.

Reichardt's Betaine

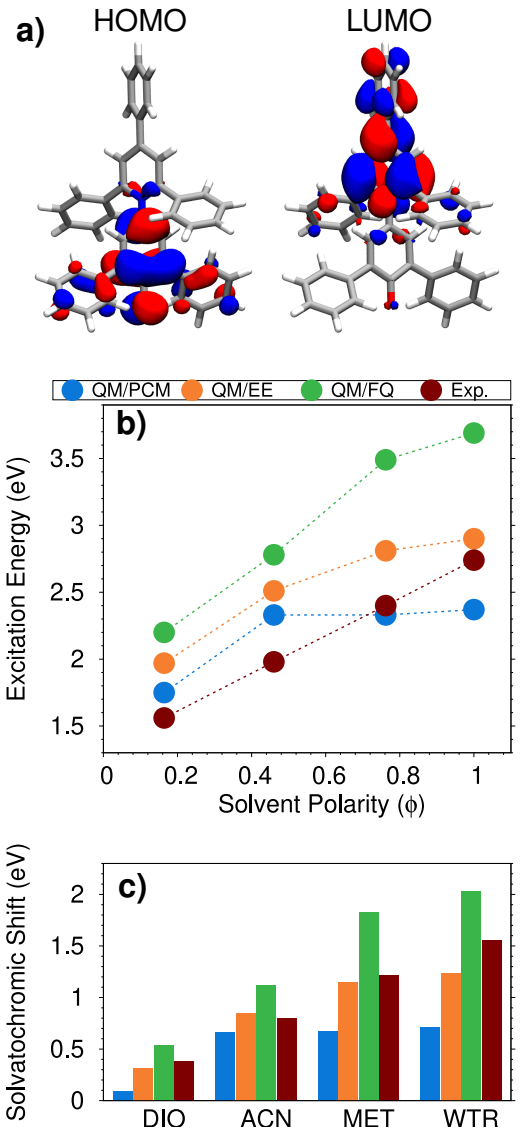


Figure 2.8: **a)** BET HOMO and LUMO involved in the studied electronic transition. **b)** QM/PCM, QM/EE, QM/FQ and experimental BET excitation energies as a function of the solvent polarity (ϕ). **c)** QM/PCM, QM/EE, QM/FQ and experimental BET solvatochromic shifts computed with respect to gas-phase excitation energy.

Solvent effects induce dramatic shifts on BET absorption spectrum, and for this reason it has been used to develop the Reichardt's polarity index ET(30), which is probably the most popular solvent polarity scale [227, 234, 300, 301].

The absorption spectrum of BET as dissolved in DIO, ACN, MET and WTR has been computed by modelling the solvent effects by means of the QM/PCM, QM/EE or QM/FQ approaches. BET lowest excitation, of clear CT character (see 2.8a), corresponds to the HOMO-LUMO $\pi \rightarrow \pi^*$ transition for all considered solvents and solvation approaches. Experimental [295, 296] excitation energies as a function of the solvent polarity are given in Fig. 2.8b, together with computed values. Clearly, experimental data linearly depend on the solvent polarity, whereas both QM/EE and QM/PCM values flatten out by increasing the solvent polarity. Remarkably, among the tested approaches, QM/FQ is the only one that is able to model the experimentally observed trend.

In Fig. 2.8c we finally report computed and experimental solvatochromic shifts. Reference excitation energies in vacuo are 1.66 (experimental) and 1.18 (calculated) eV. BET values confirm what has already been discussed for the previous molecules. QM/PCM cannot reproduce experimental solvatochromic shifts for the most polar and protic solvents, whereas QM/EE results are in good agreement with experiments, especially for DIO, ACN and MET. QM/FQ always overestimates experimental values, similarly to the case of QB (see above). However, also for BET the good performance of QM/EE can be ascribed to a fortunate error cancellation, due to the fact that both Pauli repulsion and solute-solvent polarization effects are neglected (the latter are instead considered by QM/FQ, and are large).

Discussion

In this section, we further study the quality of the computational approach, by focusing on general trends for all four molecules. First, we investigate band broadening by moving from apolar to polar protic solvents. To this end stick spectra, displayed as histograms, of the four chromophores in the different solvents are reported in Fig. 2.9. Clearly, by moving from DIO to WTR the absorption band broadens out. By fitting the different absorption spectra with gaussian functions, we see that their Full Width Half Maximum (FWHM) values, which are a direct measure of band broadening, are substantially influenced by the solvent. The increase of band broadening is particularly evident for BET, for which FWHM in WTR is 65% larger than in DIO (0.71 vs 0.43 eV). Because BET is kept frozen during MD runs, broadening arises from fluctuations of solvent molecules around the solute. From the physico-chemical point of view, this is not unexpected. In fact, for strongly interacting solvents, such as MET and WTR which can interact with BET via HB interactions, such fluctuations yield to a larger scattering of absorption energies as compared to apolar non-protic solvents, such as DIO.

The reliability of the method can also be investigated by studying the dependence of ground state (GS) and excited state (ES) dipole moments (in Debye) as a function of the solvent polarity ϕ (see Fig. 2.10). We first notice that PNA differs from other solutes, because its GS dipole moment is lower than the ES one, independently of the solvent polarity. This is clearly reflected by the negative solvatochromic shifts reported for PNA only. However, for all studied systems, an almost linear dependence of both GS and ES dipole moments as a function of the solvent polarity

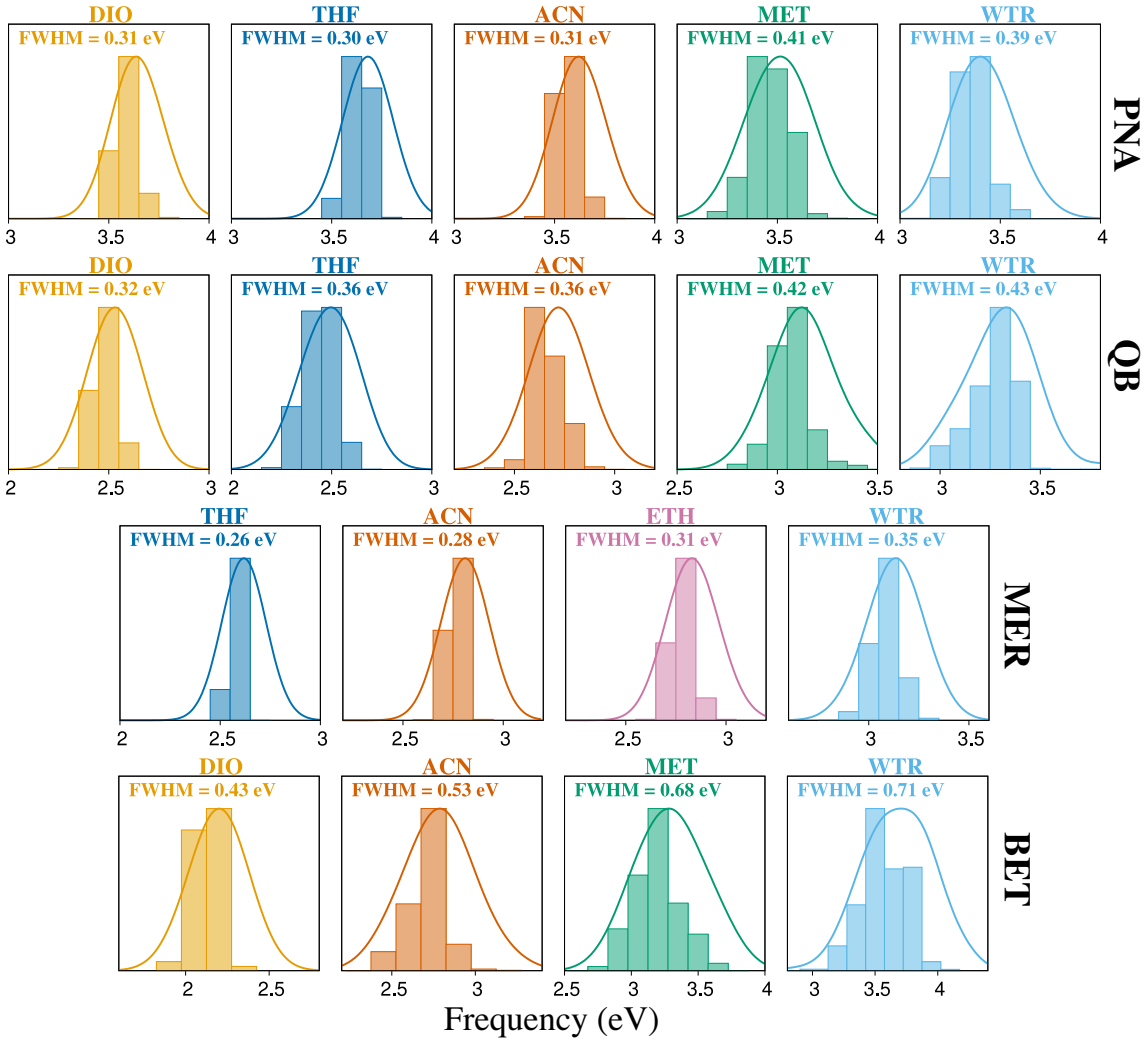


Figure 2.9: QM/FQ excitation energy distributions of PNA, QB, MER and BET dissolved in different solvents of increasing polarity. The FWHM of each band is also reported (in eV).

is predicted. Remarkably, ETH, MET and WTR values (see for instance MER, QB, and BET) deviate from linearity, being such a behaviour possibly explained by the fact that protic solvents may not be barely described by the solvent polarity index, which does not take into account strong, directional solute-solvent interactions.

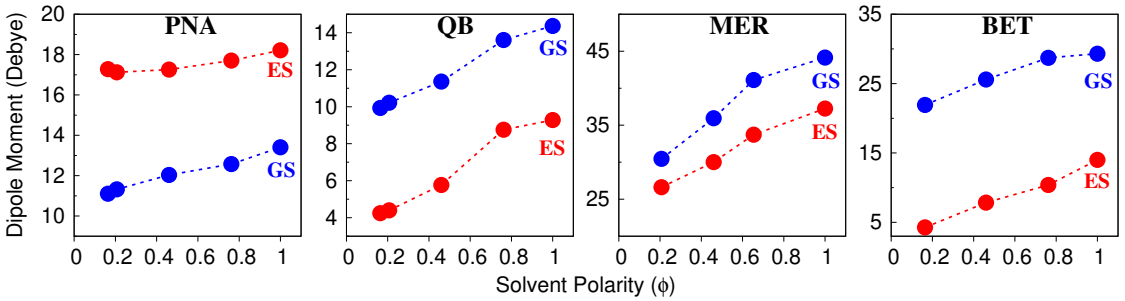


Figure 2.10: QM/FQ PNA, QB, MER and BET ground (GS, blue) and excited (ES, red) dipole moments (in Debye) as a function of the solvent polarity index, (ϕ).

Finally, in Fig. 2.11, correlation maps between experimental (ΔE_{exp}) solvatochromic shifts and calculated QM/PCM (left panel, $\Delta E_{QM/PCM}$), QM/EE (middle panel, $\Delta E_{QM/EE}$) and QM/FQ (right panel, $\Delta E_{QM/FQ}$) values are reported.

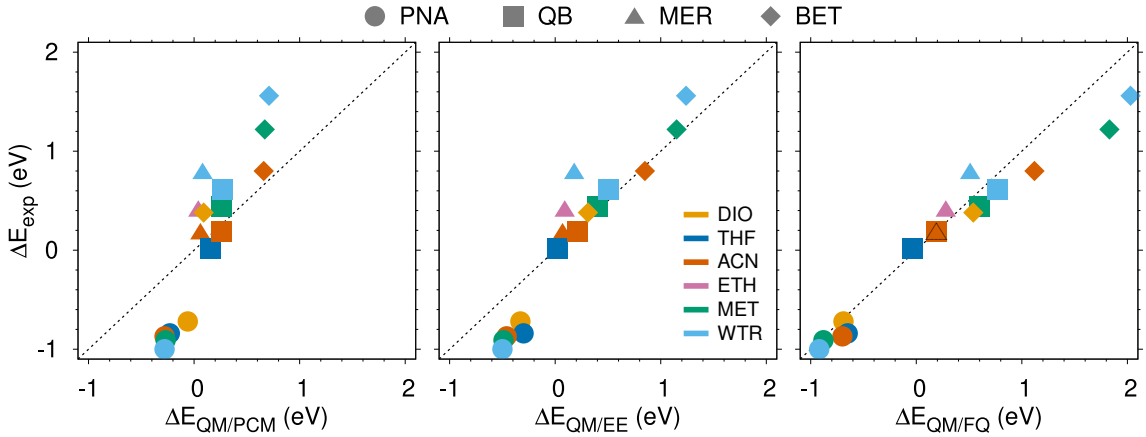


Figure 2.11: PNA, QB, MER and BET QM/PCM (left), QM/EE (middle) and QM/FQ (right) solvatochromic shifts (ΔE) with respect to experimental values (ΔE_{exp}).

For PNA (circles), QM/FQ deviations with respect to experimental values oscillate from a minimum value of 0.03 eV for DIO and MET to a maximum value of 0.19 eV for THF. A similar behaviour is observed for QB, where the largest discrepancy is observed for WTR (0.17 eV), whereas for MER, the largest error increases up to 0.26 eV (WTR). As already commented above, the largest deviations are observed for BET. However, we remark that such discrepancies can attributed to the solute geometry being kept frozen during MD runs, and also to the fact that solute-solvent interactions are limited to electrostatic (and polarization) contributions, and other quantum forces may be in place for some of the studied molecules. This is for instance the case of BET in WTR; in fact, if Pauli repulsion is included by resorting to the approach that we have presented in Ref. [246] and extended to TD-DFT in Ref. [290], for aqueous BET the error moves from 0.46 eV to 0.1 eV. We finally remark

that vibronic effects may affect computed accurate solvatochromic shifts[291]. Gas-phase vibronic UV-Vis spectra of the studied systems within the vertical gradient approximation[302] are reported in Tab. S6 in the SI, showing that such terms can be as large as 0.1 eV.

Finally, let's compare QM/FQ to QM/EE and QM/PCM. QM/FQ outperforms the other two approaches, resulting in mean error of about 0.18 eV with respect to 0.25 and 0.43 eV, as it is obtained by using QM/EE and QM/PCM, respectively. Not surprisingly, QM/PCM gives the worst values, thus showing the limitations of continuum solvation and the necessity of explicit treatment of protic, polar solvents. On the other hand, the reduced errors of QM/FQ compared to QM/EE demonstrate that polarization effects are relevant to the description of solvatochromic shifts.

Chapter 3

Beyond the Fluctuating Charge model: FQFμ

Polarizable QM/MM Approach with Fluctuating Charges and Fluctuating Dipoles: The QM/FQFμ Model

The problem of describing the interaction between a molecule and its embedding environment is a challenge in Quantum Chemistry. The interplay between the molecule and the environment can in fact dramatically alter both the structure and the electronic response to external electromagnetic fields. The most successful answer to this problem has been found within the realm of multiscale approaches: [2, 3, 72, 73, 86, 87] there, the focus is always the molecule and the key is to accurately capture the molecule/environment interactions and their effects on the molecular structure and properties, while neglecting to simulate the intrinsic properties of the environment. Such an approach is based on the assumption that molecular energetic and response properties are local properties of the molecule, which are modified but not determined by the presence of the environment.

In the last years, much effort has been devoted to develop multiscale QM/MM approaches, which keep an atomistic description of all the system under study and are therefore able to model specific molecule-environment interactions, such as hydrogen bonding (HB). [74, 97] Most QM/MM approaches developed so far focus on describing the electrostatic interactions between the QM and MM portions. The most physically consistent of such methods are those in which the mutual polarization between the QM and MM portions of the system is recovered. This has led to the development of the so-called polarizable QM/MM approaches, which can be based on distributed multipoles, [16–20] induced dipoles, [21, 24, 25] Drude oscillators [26] or Fluctuating Charges (FQ). [27–29] In the latter approach, the electrostatic interaction is described by endowing each MM atom with a charge that can vary as a response to both the differences in electronegativity between MM atoms and in electric potential generated by the QM density.

The classical FQ force field is described only in terms of charges. This poses some conceptual issues because only monopoles, i.e. zeroth order of the electrostatic Taylor expansion, are taken into consideration. As a consequence, the intrinsic anisotropy of some specific molecule-environment interactions, such as HB, is not explicitly taken into account. To overcome this problem, the electrostatic description of the FQ force field can be refined by including an additional source of polarization. This can be done by adding induced point dipoles, [263] Drude Oscillators (Polarizable Charge Equilibration PQEq), [264, 303] or Gaussian-like induced atomic dipoles (Q+P iso $[R, \alpha_{iso}]$ model, [304] Discrete Interaction Model DIM, [305] Capacitance Polarization Model CMM [306]). Differently from the basic formulation of the FQ force field, in the last two approaches gaussian distributions representing the charges, the Drude Oscillators or the induced atomic dipoles are considered, so that the Coulomb law divergence at zero distance, i.e. the so-called "polarization catastrophe", is avoided. [304, 305]

In this work, we present a novel polarizable force field, which we will call Fluctuating Charge Fluctuating Dipoles (FQFμ), in which both monopoles (charges) and dipoles can vary as a response to the external Maxwell sources, i.e. electric potential/field. The proposed model founds its fundamental basis on Ref. [304] and overcomes the limitations of FQ at describing anisotripic electrostatic terms. FQFμ

is then coupled to a QM description, following the general structure of QM/MM approaches, yielding the novel QM/FQFμ method. Therefore, QM/FQFμ can be seen as a refinement of our previously developed QM/FQ method. [29, 103, 104, 112, 113, 127]

An important difference between QM/FQFμ (and QM/FQ) and other polarizable QM/MM approaches, is that the latter only adjust the first order of the electrostatic Taylor expansion (i.e. dipole terms) to the QM density, but they keep the monopole (and higher orders) terms fixed. However, it has been proven that charges indeed give the main contribution to the electrostatic interaction energy. [307, 308]

As stated before, similar polarizable QM/MM approaches, in which both charges and dipoles are polarizable, have been proposed (see Refs. [305, 306]). However, our approach presents several differences with respect to them. The most relevant is that the widths of the gaussian charge/dipoles distributions are defined in terms of atomic chemical hardnesses and polarizabilities, which are the quantities entering the definition of FQFμ. Also, to the best of our knowledge, we report on the first application of a QM/fluctuating charges+fluctuating dipoles approach to molecular systems in a molecular environment (not on surfaces/nanoparticles, as for instance in Ref. [304–306]). Another relevant novelty of the present work stands in the strategy which is exploited for model parametrization and testing. In fact, parametrization is tuned to get an accurate reproduction of electrostatic interaction energies (vide infra), whereas other approaches rely on atomic parameters defining the specific MM polarizable force field which is actually used.

The manuscript is organized as follows. In the next section, the FQFμ force field is proposed and then coupled to a QM SCF description (QM/FQFμ). A parametrization for aqueous solutions is proposed and applied to the the calculation of electrostatic and total interaction energies of a water dimer as a function of the intermolecular distance. Then, QM/FQFμ is tested against solute-solvent electrostatic interactions of four selected systems in aqueous solution. Some conclusions and future perspectives end the manuscript.

FQFμ force field

In the FQFμ force field each MM atom is endowed with both a charge q and an atomic dipole $\boldsymbol{\mu}$, that can vary according to the external electric potential and electric field. Both charges and dipoles are described as s-type gaussian distribution functions:

$$\begin{aligned}\rho_{q_i}(\mathbf{r}) &= \frac{q_i}{\pi^{\frac{3}{2}} R_{q_i}^3} \exp\left(-\frac{|\mathbf{r}-\mathbf{r}_i|^2}{R_q^2}\right) \\ \rho_{\boldsymbol{\mu}_i}(\mathbf{r}) &= \frac{|\boldsymbol{\mu}_i|}{\pi^{\frac{3}{2}} R_{\boldsymbol{\mu}_i}^3} \hat{\mathbf{n}}_i \cdot \boldsymbol{\nabla}\left[\exp\left(-\frac{|\mathbf{r}-\mathbf{r}_i|^2}{R_{\boldsymbol{\mu}_i}^2}\right)\right]\end{aligned} \tag{3.1}$$

where R_{q_i} and $R_{\boldsymbol{\mu}_i}$ are the width of the Gaussian distributions ρ_{q_i} and $\rho_{\boldsymbol{\mu}_i}$, respectively. $\hat{\mathbf{n}}_i$ is a unit vector pointing to the dipole direction $\boldsymbol{\mu}_i$.

The total energy E associated with a distribution of charges and dipoles is equal to: [304]

$$E(\mathbf{q},\boldsymbol{\mu}) = \sum_i q_i\chi_i + \frac{1}{2}\sum_i q_i\eta_i q_i + \frac{1}{2}\sum_i\sum_{j\neq i} q_i \mathrm{T}_{ij}^{qq} q_j + \sum_i\sum_{j\neq i} q_i \mathbf{T}_{ij}^{q\mu}\boldsymbol{\mu}_j + \\ - \frac{1}{2}\sum_i\sum_{j\neq i}\boldsymbol{\mu}_i^{\dagger}\mathbf{T}_{ij}^{\mu\mu}\boldsymbol{\mu}_j + \frac{1}{2}\sum_i \boldsymbol{\mu}_i^{\dagger}\alpha_i^{-1}\boldsymbol{\mu}_i \tag{3.2}$$

where χ is the atomic electronegativity, η the chemical hardness and α the atomic polarizability. T_{ij}^{qq}, $\mathbf{T}_{ij}^{q\mu}$ and $\mathbf{T}_{ij}^{\mu\mu}$ are the charge-charge, charge-dipole and dipole-dipole interaction kernels, respectively. If the gaussian distributions in Equation 3.1 are adopted, the functional form of the interaction kernels provided by Mayer [304] can be exploited. T_{ij}^{qq} term reads:

$$\mathrm{T}_{ij}^{qq} = \frac{1}{|\mathbf{r}_{ij}|}\mathrm{erf}\left(\frac{|\mathbf{r}_{ij}|}{R_{q_i-q_j}}\right) \tag{3.3}$$

where $R_{q_i-q_j}$ is equal to $\sqrt{R_{q_i}^2 + R_{q_j}^2}$. When $\mathbf{r}_i$ tends to $\mathbf{r}_j$, the use of gaussian distributions avoids any issues which are related to the typical divergence of Coulomb kernels (i.e. the so-called "polarization catastrophe"): [304, 305]

$$\lim_{\mathbf{r}_{ij}\to 0}\mathrm{T}_{ij}^{qq} = \mathrm{T}_{ii}^{qq} = \frac{2}{\sqrt{\pi}}\frac{1}{R_{q_i-q_i}} \tag{3.4}$$

In order to collect all the quadratic terms in the charges, the diagonal elements of $\mathbf{T}^{qq}$ can be imposed to be equal to the atomic chemical hardnesses η, so that the width of the charge distribution R_q is defined without the need of any parametrization:

$$\mathrm{T}_{ii}^{qq} = \eta_i \Rightarrow R_{q_i} = \sqrt{\frac{2}{\pi}}\frac{1}{\eta_i} \tag{3.5}$$

where it is assumed $R_{q_i-q_i} = \sqrt{2}R_{q_i}$.

The charge-dipole and dipole-dipole interaction kernels are obtained as first and second derivatives of the charge-charge interaction kernel in Equation 3.3 [304]:

$$\mathbf{T}_{ij}^{q\mu} = -\boldsymbol{\nabla}_{r_i}\mathrm{T}_{ij}^{qq} = -\frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^3}\left[\mathrm{erf}\left(\frac{|\mathbf{r}_{ij}|}{R_{q_i-\mu_j}}\right) - \frac{2|\mathbf{r}_{ij}|}{\sqrt{\pi}R_{q_i-\mu_j}}\exp\left(\frac{-|\mathbf{r}_{ij}|}{R_{q_i-\mu_j}}\right)^2\right] \tag{3.6}$$

$$\mathbf{T}_{ij}^{\mu\mu} = \boldsymbol{\nabla}_{r_j}\mathbf{T}_{ij}^{q\mu} = \frac{3\mathbf{r}_{i,j}\otimes\mathbf{r}_{i,j} - |\mathbf{r}_{i,j}|^2\mathbf{I}}{|\mathbf{r}_{i,j}|^5}\left[\mathrm{erf}\left(\frac{|\mathbf{r}_{ij}|}{R_{\mu_i-\mu_j}}\right) - \frac{2}{\sqrt{\pi}}\frac{|\mathbf{r}_{ij}|}{R_{\mu_i-\mu_j}}\exp\left(-\frac{\mathbf{r}_{ij}}{R_{\mu_i-\mu_j}}\right)^2\right] + \\ - \frac{4}{\sqrt{\pi}R_{\mu_i-\mu_j}^3}\frac{\mathbf{r}_{ij}\otimes\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^2}\exp\left(-\frac{|\mathbf{r}_{ij}|}{R_{\mu_i-\mu_j}}\right)^2 \tag{3.7}$$

where $R_{x_i-x_j} = \sqrt{R_{x_i}^2 + R_{x_j}^2}(x = q,\mu)$ and $\mathbf{I}$ is the identity matrix. Similarly to what was done before for T_{ij}^{qq}, the limits for $\mathbf{r}_{ij}\to 0$ in the case of $\mathbf{T}_{ij}^{q\mu}$ and $\mathbf{T}_{ij}^{\mu\mu}$ are:

$$\lim_{\mathbf{r}_{ij}\to 0} \mathbf{T}^{q\mu}_{ij} = \mathbf{T}^{q\mu}_{ii} = 0 \tag{3.8}$$

$$\lim_{\mathbf{r}_{ij}\to 0} \mathbf{T}^{\mu\mu}_{ij} = \mathbf{T}^{\mu\mu}_{ii} = -\sqrt{\frac{2}{\pi}}\frac{\mathbf{I}}{3R^3_{\mu_i}} \tag{3.9}$$

From Equation 3.9, R_{μ_i} can be defined in terms of the atomic polarizability α_i:

$$\alpha_i^{-1} = \sqrt{\frac{2}{\pi}}\frac{1}{3}\frac{1}{R^3_{\mu_i}} \Rightarrow R_{\mu_i} = \left(\sqrt{\frac{2}{\pi}}\frac{1}{3}\alpha_i\right)^{\frac{1}{3}} \tag{3.10}$$

The definition of the gaussian width R_{q_i} and R_{μ_i} in terms of η_i and α_i limits the number of parameters which enter the definition of FQFμ to electronegativity, chemical hardness and polarizability for each atom type. Therefore, Equation 3.2 can be formally rewritten as:

$$\begin{aligned} E(\mathbf{q},\boldsymbol{\mu}) &= \frac{1}{2}\sum_i\sum_j q_i \mathrm{T}^{qq}_{ij} q_j + \frac{1}{2}\sum_i\sum_j \boldsymbol{\mu}_i^\dagger \mathbf{T}^{\mu\mu}_{ij}\boldsymbol{\mu}_j + \sum_i\sum_j q_i \mathbf{T}^{q\mu}_{ij}\boldsymbol{\mu}_j^\dagger + \sum_i q_i\chi_i = \\ &= \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} \end{aligned} \tag{3.11}$$

where a matrix notation has been adopted. Notice that FQFμ can be further expanded to consider also polarizable quadrupolar terms, by only defining the appropriate interaction kernels.

In Equation 3.11, the sum of charge values is not forced by any external constrain. However, the equilibrium condition is reached when the Electronegativity Equalization Principle (EEP) is satisfied. Such a principle states that at equilibrium each atom has the same electronegativity. Thus, an energy functional to be minimized can be written for instance by adopting Lagrangian multipliers. Notice that we can in principle assume:

- The entire system is constrained to have charge $\mathrm{Q_{tot}}$, and no constraint is imposed on single molecules. This permits inter-molecular Charge Transfer (CT) and makes, at the equilibrium, the electronegativity of each atom to be the same.

- Each molecule is constrained to assume a fixed, total charge Q_α, which sums to $\mathrm{Q_{tot}}$). Therefore, the electronegativity of each atom in the same molecule is the same but generally has different values among different molecules.

We report here the equations obtained by adopting the first assumption. Consistently with what has been done for FQ by some of the present authors, [104, 126] similar equations can be derived under the second assumption : they are given in Section S1 of the Supporting Information (SI). [255] Notice that our implementation is general and can treat both cases. The energy functional F can be written by exploiting the Lagrangian multiplier (λ):

$$
\begin{aligned}
F(\mathbf{q},\boldsymbol{\mu},\lambda) &= E(\mathbf{r},\mathbf{q},\boldsymbol{\mu}) + \lambda\left[\sum_i (q_i) - Q_{\text{tot}}\right] = \\
&= \frac{1}{2}\sum_i\sum_j q_i \mathrm{T}^{qq}_{ij} q_j + \frac{1}{2}\sum_i\sum_j \boldsymbol{\mu}_i^\dagger \mathbf{T}^{\mu\mu}_{ij}\boldsymbol{\mu}_j + \sum_i\sum_j q_i \mathbf{T}^{q\mu}_{ij}\boldsymbol{\mu}_j^\dagger + \sum_i q_i\chi_i \\
&+ \lambda\left[\sum_i (q_i) - Q_{\text{tot}}\right] = \\
&= \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \lambda\mathbf{q}
\end{aligned}
\tag{3.12}
$$

where λ is meant to preserve the total charge Q_{tot} of the MM portion. Therefore, the conditions for the constrained minimum are found by imposing the derivatives of F with respect to all the variables to be zero, resulting in the following linear problem:

$$
\begin{cases}
\sum_j \mathbf{T}^{qq}_{i,j} q_j + \lambda + \sum_j \mathbf{T}^{q\mu}_{i,j}\boldsymbol{\mu}_j = -\chi_i \\
\sum_j \mathbf{T}^{\mu\mu}_{i,j}\boldsymbol{\mu}_j + \sum_j \mathbf{T}^{q\mu}_{i,j} q_j = 0 \\
\sum_i q_i = Q_{\text{tot}}
\end{cases}
\tag{3.13}
$$

The whole system can be recast in a more compact form as: [263, 306]

$$
\begin{pmatrix}
\mathbf{T}^{qq} & \mathbf{1}_\lambda & \mathbf{T}^{q\mu} \\
\mathbf{1}_\lambda^\dagger & 0 & \mathbf{0} \\
-\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu}
\end{pmatrix}
\begin{pmatrix}
\mathbf{q} \\ \lambda \\ \boldsymbol{\mu}
\end{pmatrix}
=
\begin{pmatrix}
-\boldsymbol{\chi} \\ Q_{\text{tot}} \\ \mathbf{0}
\end{pmatrix}
\qquad \Rightarrow \qquad
\mathbf{D}\mathbf{Q}_\lambda = -\mathbf{C}_Q
\tag{3.14}
$$

where $\mathbf{1}_\lambda$ is a vector which accounts for the Lagrangian. $\mathbf{C}_Q$ is a vector containing atomic electronegativities and total charge constraint, whereas $\mathbf{Q}_\lambda$ is a vector containing charges, dipoles and the Lagrange multiplier.

The QM/FQFμ model

In order to couple FQFμ to a QM wavefunction in a QM/MM framework, the first step is to define an extended energy functional, which is composed of three terms:

$$
\mathcal{E} = E_{QM} + E_{MM} + E_{QM/MM}
\tag{3.15}
$$

where E_{MM} is defined in Equation 3.12. If the QM term is a variational functional itself, the resulting, coupled equations are derived following the same procedure as for the uncoupled case. The QM density interacts as a classical density of charge with both charges and dipoles:

$$
E_{QM/MM} = \sum_i V[\rho_{QM}](\mathbf{r}_i) q_i - \boldsymbol{\mu}_i^\dagger \mathbf{E}[\rho_{QM}](\mathbf{r}_i)
\tag{3.16}
$$

where $V[\rho_{QM}](\mathbf{r}_i)$ and $\mathbf{E}[\rho_{QM}](\mathbf{r}_i)$ are the electric potential and electric field, respectively, calculated at the i-th charge and i-th dipoles placed at $\mathbf{r}_i$. The QM potential and the electric field are composed by an electronic ($\mathbf{V}^e$, $\mathbf{E}^e$) and a nuclear ($\mathbf{V}^N$, $\mathbf{E}^N$) contribution:

$$V[\rho_{QM}](\mathbf{r}_i) = V_i[\mathbf{P}] = V_i^N(\mathbf{P}) + V_i^e(\mathbf{P}) = \sum_\zeta \frac{Z_\zeta}{|\mathbf{r}_i - \mathbf{R}_\zeta|} + \int_{\mathbb{R}^3} \frac{\rho_{el}(\mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|}\,\mathrm{d}\mathbf{r} \tag{3.17}$$

$$\mathbf{E}[\rho_{QM}](\mathbf{r}_i) = \mathbf{E}_i(\mathbf{P}) = \mathbf{E}_i^N(\mathbf{P}) + \mathbf{E}_i^e(\mathbf{P}) = \sum_\zeta \frac{Z_\zeta(\mathbf{R}_\zeta - \mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{R}_\zeta|^3} - \int_{\mathbb{R}^3} \frac{\rho_{el}(\mathbf{r})(\mathbf{r}_i - \mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|^3}\,\mathrm{d}\mathbf{r} \tag{3.18}$$

where ρ_{el} is the electron density. ζ index runs over the QM nuclei, whose charges are named Z_ζ and whose positions are $\mathbf{R_1}$.

If the electronic density $\rho_{el}(\mathbf{r})$ is expanded in an atomic basis set $\{\chi_\mu\}$, the second terms in Equations 3.17 and 3.18 become:

$$V_i^e(\mathbf{P}) = -\sum_{\nu\mu} P_{\mu\nu} \int_{\mathbb{R}^3} \mathrm{d}\mathbf{r} \frac{\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|} = \sum_{\mu\nu} P_{\mu\nu} V_{\mu\nu,i} \tag{3.19}$$

$$\mathbf{E}_i^e(\mathbf{P}) = -\sum_{\mu\nu} P_{\mu\nu} \int_{\mathbb{R}^3} \frac{\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})(\mathbf{r}_i - \mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|^3} = \sum_{\mu\nu} P_{\mu\nu} E_{\mu\nu,i} \tag{3.20}$$

where we have introduced the "uncontracted" potential $V_{\mu\nu}$ and the "uncontracted" field $E_{\mu\nu}$. $P_{\mu\nu}$ are elements of the QM density matrix. Finally, the global QM/MM energy functional for a SCF-like description of the QM portion is:

$$F(\mathbf{P}, \mathbf{q}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{1}{2}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q} + \\ + \mathbf{q}^\dagger\mathbf{V}(\mathbf{P}) - \boldsymbol{\mu}^\dagger\mathbf{E}(\mathbf{P}) \tag{3.21}$$

where

$$h_{\mu\nu} = \langle\psi_\mu|-\frac{\nabla^2}{2} - \sum_\zeta \frac{Z_\zeta}{|\mathbf{r} - \mathbf{R}_\zeta|}|\psi_\nu\rangle$$

$$G_{\mu\nu} = \sum_{\sigma\tau} P_{\sigma\tau}\left(\langle\mu\sigma|\nu\tau\rangle - c_x\langle\mu\sigma|\tau\nu\rangle\right) + c_l\langle\psi_\mu|v^{xc}|\psi_\nu\rangle$$

are the usual one- and two-electron matrices. The coefficients c_x and c_l define whether Hartree–Fock ($c_x = 1, c_l = 0$), pure DFT ($c_x = 0, c_l = 1$), or hybrid DFT are exploited. For the sake of brevity, we will refer to both the HF and KS matrices as Fock matrix. The effective Fock matrix is defined as the derivative of the energy with respect to the density matrix:

$$\tilde{F}_{\mu\nu} = \frac{\partial\mathcal{E}}{\partial P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{V}_{\mu\nu}^\dagger\mathbf{q} - \mathbf{E}_{\mu\nu}^\dagger\boldsymbol{\mu} \tag{3.22}$$

where the interaction of the electron density with both charges and dipoles are included through the coupling electrostatic terms. Charges and dipoles are obtained by imposing the global functional to be stationary with respect to charges, dipoles and Lagrangian multiplier.

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_\lambda & \mathbf{T}^{q\mu} \\ \mathbf{1}_\lambda^\dagger & 0 & \mathbf{0} \\ -\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \lambda \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ Q_{\mathrm{tot}} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} -\mathbf{V}(\mathbf{P}) \\ 0 \\ \mathbf{E}(\mathbf{P}) \end{pmatrix} \quad \Rightarrow \quad \mathbf{D}\mathbf{Q}_\lambda = -\mathbf{C}_Q - \mathbf{S}(\mathbf{P}) \tag{3.23}$$

Notice that, with respect to Equation 3.14, a new source term $\mathbf{S}(\mathbf{P})$ arises. Such a term, which represents the coupling of both charges and dipoles with the SCF density, permits to determine them for a given density matrix.

QM/FQFμ introduces two polarization sources: fluctuating charges and fluctuating dipoles. From Equation 3.23 both QM/FQ and QM/Induced Dipoles can be recovered by considering only charge-charge or dipole-dipole blocks in the linear system. QM/FQFμ response matrix is four times bigger than the QM/FQ one (T^{qq} block). As a consequence, QM/FQ can treat four times bigger systems than QM/FQFμ at the same computational cost. Similarly to QM/Induced Dipoles, QM/FQFμ introduces two contributions in Fock matrix (see Equation 3.22). However, in QM/FQFμ both the zeroth order monopoles and the first order dipoles are indeed dependent on the QM density. This only causes a small increase in the computational cost with respect to QM/Induced Dipoles, because the response matrix need to be enlarged so to include the T^{qq} block (which is squared the number of MM atoms).

As pointed out in the Introduction, QM/FQFμ finds its fundamentals in Ref. [304], similarly to QM/DIM and QM/CMM. [306, 309] However, the definition of the gaussian widths, which in both QM/DIM and QM/CMM are external parameters, is automatically obtained in QM/FQFμ from chemical hardnesses and polarizabilities (see Equations 3.5 and 3.10). As a consequence, QM/FQFμ is defined only in terms of three parameters for each atom type: electronegativity, chemical hardness and polarizability. A second relevant difference stands in the formulation of the interaction between QM and MM portions. In QM/CMM, the electrostatic interactions is expressed in terms of a Coulomb integral of the gaussian distributions of both charges and dipoles with the QM density. [306] In this way, Coulomb repulsion is also taken into account. In QM/FQFμ, instead, MM charges and dipoles are seen as point charges and point dipoles by the QM density, as it is generally assumed in most polarizable QM/MM approaches. However, QM/FQFμ can be reformulated in a similar way to QM/DIM and QM/CMM, for instance by following what has been proposed in other contexts. [88] In addition, differently from DIM and CMM, FQFμ has been formulated in terms of a variational functional (see Equations 3.12 and 3.21), which guarantees its rigorous further extension to molecular properties. [89] It is worth remarking that the application of both QM/DIM and QM/CMM has been so far limited to the study of molecular properties of systems adsorbed on a metal surface, in which the metal surface is described by DIM or CMM force fields. [310–313] The heterogeneous environment is instead modeled in terms of non-polarizable force fields. [310] QM/FQFμ is not limited to any specific kind of environment (pending an accurate parametrization is obtained), thus being of broader applicability with respect to other approaches.

Fluctuating Dipoles vs Drude Oscillators

FQFμ describes the first order of electrostatic Taylor expansion in terms of fluctuating dipoles. As an alternative, Drude oscillators can be employed, as it has been recently proposed in the PQEq force field. [264] PQEq combines the Charge Equilibration model (QEq) [314] with the Drude Oscillator approach. [315] Each MM atom is seen as composed of a core and a shell, on which gaussian charge distributions are placed. In particular, both a fluctuating charge (q) and a fixed charge

$(+Z)$ are placed on the core. The fixed charge is connected trough an isotropic harmonic spring to the shell fixed but mobile charge $(-Z)$, thus allowing variable charge displacements. PQEq can be coupled to a QM description by following the same strategy adopted above for QM/FQFμ, yielding the QM/PQEq model. In this approach fluctuating charges result from the solution of a modified FQ system, whereas the positions of the shell mobile charges are obtained by imposing the total electric force acting on them to be zero (see Section S2, given as SI [255]). QM/PQEq Fock operator reads:

$$\tilde{F}_{\mu\nu} = \frac{\partial \mathcal{E}}{\partial P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{V}^{\dagger}_{\mu\nu,c}\mathbf{q} + \mathbf{V}^{\dagger}_{\mu\nu,c}\mathbf{Z} - \mathbf{V}^{\dagger}_{\mu\nu,s}\mathbf{Z} \tag{3.24}$$

where, $\mathbf{q}$ and $\mathbf{Z}$ are the vectors containing fluctuating and fixed charges, respectively, wheareas c and s subscripts indicate core and shell positions, where the QM potential $\mathbf{V}$ is calculated. Thus, differently from QM/FQFμ, QM/PQEq is defined only in terms of the QM electric potential. The equation which defines the equilibrium positions of shell mobile charges (see Equation S11 given as SI [255]) introduces a non-linearity in the problem, which can be solved only by exploiting iterative techniques. [26, 254, 316, 317] Also, due to its non-linearity, QM/PQEq could present some issues in the definition of the response property of the QM portion. On the contrary, this does not apply to QM/FQFμ. Due to its linearity and variational nature, QM/FQFμ can be extended to the calculation of molecular properties [19, 103, 104, 110, 111, 128, 318, 319] by using the standard techniques of quantum chemistry. [129] PQEq can indeed be mapped into the FQFμ approach, similarly to what has been done in the case of basic Drude Oscillator and Induced Dipole force fields. [320]

Computational Details

QM/FQFμ was implemented in a locally modified version of Gaussian16. [146] All QM/FQ and QM/FQFμ calculations were performed by treating the QM portion at the HF or DFT levels of theory, combined with selected basis sets. Three different parametrizations to treat the FQ electrostatic component in QM/FQ calculations were exploited, taken from Ref. [27], Ref. [128] and Ref. [321]. Non-electrostatic contributions, i.e. repulsion and dispersion, were modeled as reported in Ref. [11] All the classical Molecular Dynamics (MD) simulations were performed with the Gromacs package, [322–325] by keeping the same settings as previously reported by some of the present authors. [105, 106, 113] Details on MD simulations are given as SI. [255] The Kitaura-Morokuma Energy Decomposition Analysis (KM-EDA) [326, 327] was performed by using the GAMESS package. [328, 329] Symmetry Adapted Perturbation Theory (SAPT) [330, 331] calculations were performed by using Psi4 1.1. [332]

Numerical Results

In this section, the parametrization of the QM/FQFμ approach to treat aqueous solutions is presented and discussed. Then, the resulting parameters are tested to reproduce electrostatic energies of a water dimer as a function of the oxygen-oxygen distance as computed at the KM-EDA/6-31+G* level. Then, the total

interaction energy E_{int}, i.e. the sum of electrostatic (E_{ele}), repulsion (E_{rep}) and dispersion (E_{dis}) contributions, is compared to SAPT2+3(CCD)/aug-cc-pVTZ and CCSD(T)/aug-cc-pVTZ with Counter-Poise corrections values for the same dimer. The dependence of the QM/FQFμ electrostatic energy on the level of theory, i.e. the combination of HF/DFT with several basis sets is also discussed. Finally, in order to test the transferability of our parameters to other systems, four molecules (Methyloxirane, Acrolein, N-Methyl Acetamide and Methanol) in aqueous solution are studied. In such cases, QM/FQFμ and QM/FQ electrostatic energies are compared to SAPT0/6-311++G** values.

Model Parametrization

QM/FQFμ is general enough to model any kind of external environment, pending an appropriate parametrization of the quantities entering Equations 3.21 and 3.22. Such a parametrization is a crucial step towards the routinely application of the method to real cases. In this section we will focus on aqueous solutions, which will also allow for a quantitative comparison with QM/FQ, thus highlighting the effect of including atomic fluctuating dipoles in QM/MM electrostatic energies.

In order to set the parameters entering Equation 3.21, selected water clusters taken from Kratz et al. [333] (see Figure 3.1), were studied. Reference full QM electrostatic energy values of such clusters were calculated by performing a KM-EDA [326, 327] calculation on each structure in Figure 3.1 at the HF/6-31+G* level, according to what has already been proposed in the literature. [8, 9, 11]

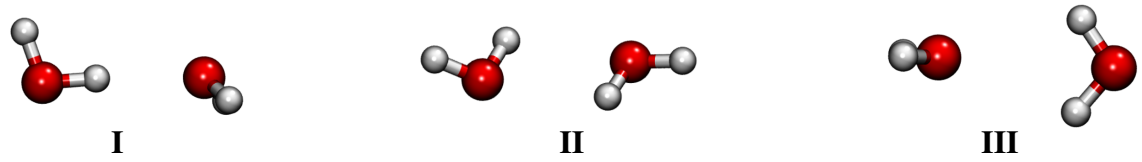


Figure 3.1: Structures of water dimers exploited in the parametrization of QM/FQFμ.

KM-EDA values were compared to electrostatic energies obtained with the QM/FQFμ model. In the latter, one water molecule was treated at the QM level (HF/6-31+G*), whereas the second molecules was described by means of the FQFμ force field. For each dimer structure two calculations were performed, by exchanging the QM and FQFμ water molecules. Remarkably, electrostatic interaction is by definition symmetric if the two water molecules are interchanged. Thus, differently to what some of the present authors reported for the parametrization of non-electrostatic contributions, [11] we imposed the two calculations (i.e. the MM water molecule acts as HB donor or acceptor) to give the same results. In such a way the transferability of the final parameters should be guaranteed. Notice that in the QM/FQFμ approach, electrostatic and polarization terms cannot be separated, because the electrostatic charge contribution is partially due to QM polarization (see Equation 3.23). Thus, QM/FQFμ electrostatic energies are compared with the sum of electrostatic and polarization KM-EDA energy contributions.

For the studied dimers, Equation 3.21 depends on six parameters (electronegativities, chemical hardnesses and polarizabilities of hydrogen and oxygen atoms,

respectively): their best values were defined by performing a least square roots fitting on full KM-EDA data, by allowing the parameters to vary freely. Their best fitted values are reported in Table S1 in the SI. [255] Electrostatic energies of the selected water dimers are reported in Table S2 in SI. [255] The reliability of the parametrization protocol is confirmed by the agreement between QM/FQFμ and KM-EDA data. The effects due to the introduction of atomic dipoles can also be quantified. The zeroth order monopoles, i.e. fluctuating charges, account for almost 70∼72 % of the total electrostatic energy, whereas the first order dipoles for 28∼30%, i.e. they give a minor, but not negligible contribution.

Interaction energy of a water dimer as a function of O-O distance

In this section, the dependence of E_{ele} and E_{int} on the water-water intermolecular distance is investigated. To this end, the water dimer depicted in Figure 3.2 (optimized at the MP2/aug-cc-pVQZ level) was exploited, and the distance d between the oxygen atoms was taken as reference.

Figure 3.2: Structure of the water dimer used to study the dependence of electrostatic and interaction energies as a function of O-O intermolecular distance.

In Figure 3.3, QM/FQFμ E_{ele} is reported as a function of d. The plot was constructed by increasing the O-O distance from 2.54 Å to 6.49 Å by a step of 0.05 Å (80 points). E_{ele} was calculated by treating the QM moiety at the HF/6-31+G* level. Again, QM and MM moieties were interchanged, and the average values were taken. Charge and dipole electrostatic contributions QM/FQF$\mu|_q$ and QM/FQF$\mu|_\mu$, are also depicted, showing that also in this case the charge contribution is dominating at all distances (70 $\sim$ 72 %).

In Figure 3.4, computed QM/FQFμ electrostatic energies are compared with KM-EDA full-QM reference electrostatic (summed with polarization) energies. An almost perfect superposition is observed, the average computed error being of about 7% and the computed Root Mean Squared Deviation (RMSD) being only 0.29 kcal/mol ($\sim$ 0.47 mH). The excellent reproduction of E_{ele} is not unexpected, because the dimer structure under study is very similar to structure **I** exploited in the parametrization step (see Figure 3.1).

The inset in Figure 3.4 shows the difference between calculated QM/FQFμ E_{ele} values obtained by assuming the QM water molecule to act as H-bond donor or acceptor.The two curves are almost superimposed, as expected by considering the parametrization protocol that we have followed (see previous section).

To end this discussion, the total B3LYP/aug-cc-pVTZ QM/FQFμ interaction energy as a function of d is plotted in Figure 3.5 and compared with SAPT2+3(CCD)/aug-cc-pVTZ or CCSD(T)/aug-cc-pVTZ data (counterpoise corrections are included). To this end, QM/FQFμ is coupled to the approach proposed by some of the present authors to model non-electrostatic repulsion/dispersion contributions, [11] which formulates repulsion in terms of an auxiliary density on the MM portion, whereas

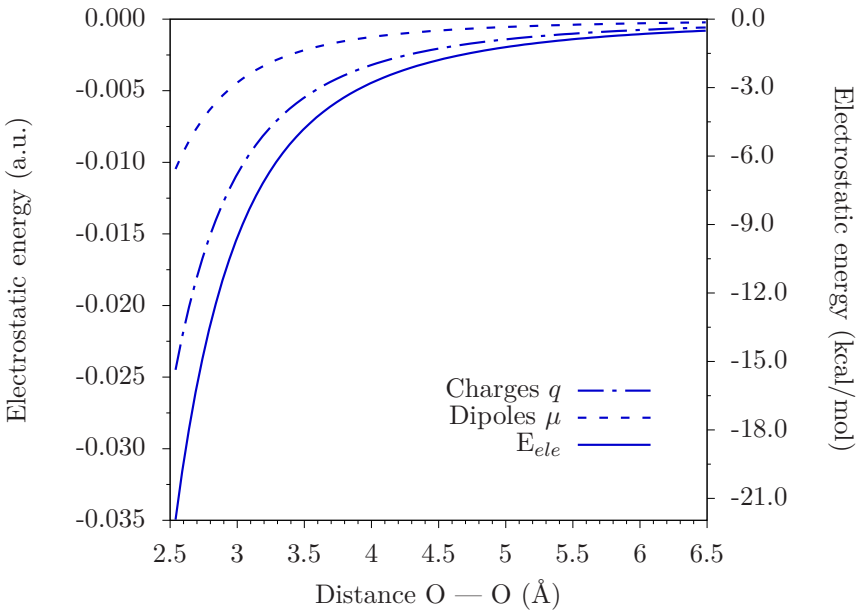


Figure 3.3: Plot of the QM/FQFμ electrostatic energy as a function of the O-O intermolecular distance for the water dimer depicted in Figure 3.2. Charge and dipole contributions to E$_e$*le* are also plotted.

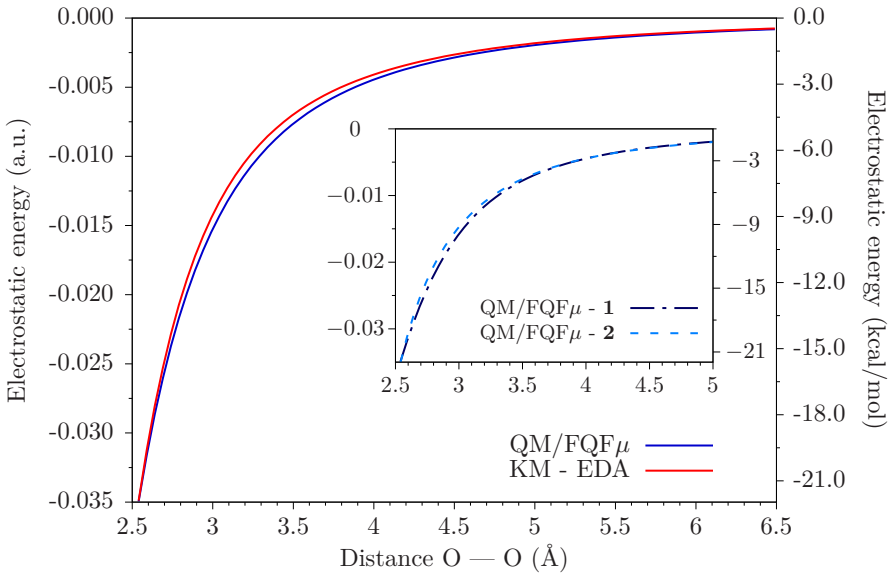


Figure 3.4: Plot of the electrostatic energy as a function of the O-O intermolecular distance for the water dimer depicted in Figure 3.2. QM/FQFμ values (HF/6-31+G* for the QM moiety) are compared to KM-EDA (HF/6-31+G*) calculations. In KM-EDA calculations, electrostatic and polarization contributions are summed up. In the inset QM/FQFμ E$_{ele}$ as a function of the O-O distance is depicted for the two structures (**1** and **2**) in Figure 3.2.

QM/MM dispersion is obtained by extending the Tkatchenko-Scheffler approach to DFT. [12, 334–337]

Clearly, QM/FQFμ+dis/rep is able to correctly reproduce both CCSD(T) equilibrium distance (2.99 Å vs. 2.99 Å) and CCSD(T) interaction energy at the equilibrium distance (-4.56 vs. -4.65 kcal/mol). The RMSD calculated over all 80 structures is 0.34 kcal/mol.

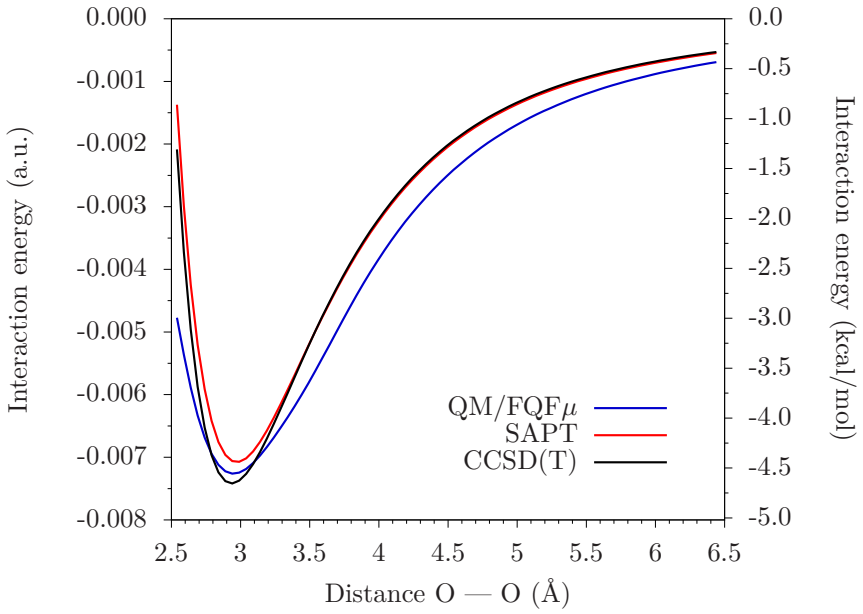


Figure 3.5: QM/FQFμ (B3LYP/aug-cc-pVTZ for the QM moiety), SAPT2+3(CCD) and CCSD(T)/aug-cc-pVTZ total interaction energies for the water dimer in Figure 3.2 as a function of the O-O distance.

Dependence on the QM level of theory

In this section, the dependence of calculated QM/FQFμ E_{ele} values on the level used to model the QM moiety is studied. To this end, the water dimer depicted in Figure 3.2 with $d = 2.94$ Å is exploited. Thirteen different methods were used by following the recent literature, [107, 338] ranging from HF to pure DFT functionals (LDA, PBE, [339] B97D, [340, 341] R-TPSS [342]), to different classes of hybrid functionals (BLYP, [343] M06, [344] PBE0, [345] B3LYP, [346] M062X, [344] SOGGA11-X, [347] mPW1PW91 [348]), also including long-range (CAM-B3LYP [349]). Each functional was coupled to several Pople-type basis sets (see Figure 3.6), in order to separate the contributions arising from polarization and diffuse functions. In addition, correlation-consistent and augmented correlation-consistent basis sets were employed, up to aug-cc-pVQZ. [350]

Figure 3.6 schematically reports the observed trends. Numerical values are given in Table S3 in the SI. [255]

All employed QM methods predict very similar E_{ele} values as varying the basis set, with HF always showing the highest absolute values for a given basis set (on average, HF values are about 5% higher than the absolute average value of the other methods). The lowest absolute values are instead shown by PBE and BLYP functionals. However, the difference between HF and PBE/BLYP functionals (i.e.

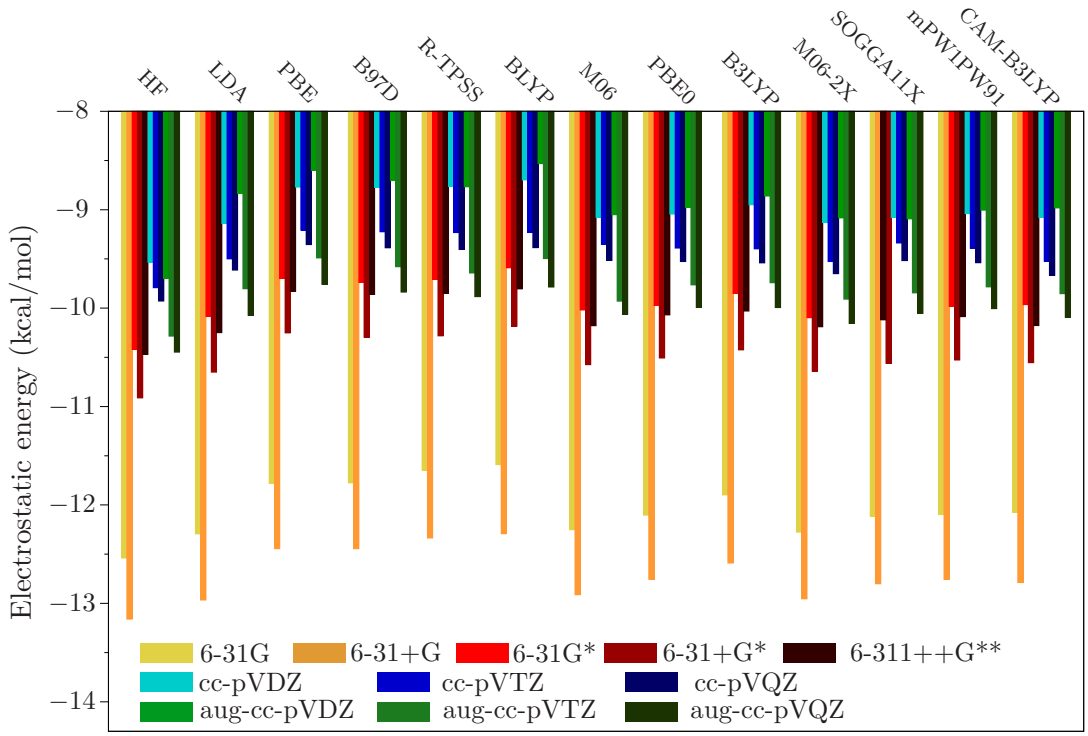


Figure 3.6: Dependence of E_{ele} on the choice of basis set and QM method for the water dimer depicted in Figure 3.2 with $d = 2.94$ Å.

the limit values of the computed E_{ele} for a given basis set) is 0.75 kcal/mol on average, being the maximum value 1.15 kcal/mol for aug-cc-pVDZ. This clearly shows that the computed QM/FQFμ energy values are almost unaffected by the choice of the QM description.

Let us focus on the dependence of E_{ele} on the choice of the basis set. First, we notice that electrostatic energy absolute values increase (of about 5% on average) with adding diffuse functions, that probably due to the spreading of the QM density. The addition of polarization functions has instead an opposite effect, in fact absolute values decrease of about 17% on average. Such trends are almost constant for all QM descriptions. Calculated E_{ele} obtained by exploiting correlation consistent basis sets are always smaller than Pople-calculated values. Moving from cc-pVDZ to cc-pVQZ, the QM/FQFμ electrostatic energy increases in absolute value, and the same trend is reported if augmented basis sets are considered. aug-cc-PVQZ gives very similar results with respect to 6-311++G**, being the average difference of about 0.5%.

In conclusion, stable values of E_{ele} are obtained by adding both diffuse and polarization functions, so that their inclusion appears mandatory. For this reason, in the following section the 6-31+G* basis set is exploited, being a good compromise between accuracy and computational cost.

Molecules in aqueous solution

In order to show the applicability of QM/FQFμ to the study of molecular systems, and to investigate on the reliability of its parametrization, in this section the method is applied to four selected molecules in aqueous solution: (R)-Methyloxirane

(MOXY), acrolein (ACRO), N-methyl acetamide (NMA) and methanol (MeOH). In the first three molecules solute-solvent Hydrogen Bonding (HB) can occur, however the surrounding water molecules can only act as HB donor. For aqueous MeOH, water molecules can instead act as both H-donor and H-acceptor, due to the presence of the O-H group in solute structure. Therefore, the chosen set of systems can appropriately represent the main solute-solvent interactions which are in place in aqueous solutions.

For each of the selected molecules, we ran classical MD simulations (see Section S4.1 for further details) to sample the phase space. From each MD run, we extracted 10 representative structures. Spheres of 5 Å centered in the oxygen atom in case of MOXY, ACRO, and MeOH, and in the nitrogen atom in case of NMA were cut. Sample final structures are reported in Figure 3.7, where solute-solvent HBs are sketched. All extracted structures are reported in Figures S1-S4 in SI. [255] The choice of the spheres' radius is justified by the analysis of the Radial Distribution Functions $g(r)$ (see Section S4.1 in SI [255]), which show that a cutting radius of 5.0 Å guarantees that all water molecules in the first two solvation shells are included. The actual number of water molecules in each of the ten considered snapshots for each system is reported in Table 3.1.

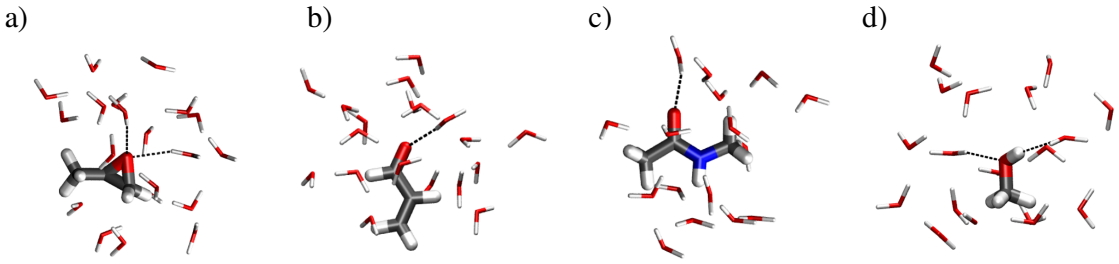


Figure 3.7: Sample structures obtained by cutting a sphere of 5.0 Å around **a)** (R)-methyloxirane; **b)** acrolein; **c)** N-methyl acetamide; **d)** methanol.

Structure	MOXY	ACRO	NMA	MeOH
1	22	19	21	26
2	18	23	17	25
3	19	19	19	23
4	19	18	18	19
5	20	19	16	21
6	19	20	17	23
7	20	14	16	18
8	13	19	15	24
9	19	20	16	17
10	20	21	20	25

Table 3.1: Number of water molecules included in each of the ten considered snapshots for each studied molecule in aqueous solution, obtained by using a cutting radius of 5 Å

For each of the extracted structures, solute-solvent E_{ele} was calculated by exploiting both QM/FQ and QM/FQFμ. In case of QM/FQ calculations, three different

parametrizations, namely QM/FQ^a, [27] QM/FQ^b [128] and QM/FQ^c [321] were considered. QM/FQ and QM/FQFμ were compared with full-QM electrostatic energies calculated by exploiting SAPT0/6-311++G** (see Figure 3.8, the corresponding raw data are given in Tables S4-S5 in SI [255]). In both QM/FQ and QM/FQFμ calculations, the QM portion was described at the HF/6-311++G** level, and the charge constraint in Equation 3.21 is imposed so to fix the total charge of the solvent molecules to zero. This implies that Charge Transfer (CT) between different water molecules is allowed. Such a choice is justified by the fact that reference full-QM data implicitly take into account CT between solvent molecules. Additional calculations on the same structures were performed by fixing the total charge of the single MM water to zero; the corresponding results are given in Section S5.4 in SI. [255]

The comparison between polarizable QM/MM and SAPT0 E_{ele} are graphically depicted in Figure 3.8. RMSD, Maximum Absolute Error (MAE) and Relative Error (RE) on the ten selected structures are reported in Table 3.2. Let us focus on the results obtained for MOXY in aqueous solution. SAPT0 values range from -17 to -30 kcal/mol, thus showing large electrostatic interactions due to HBs, which are reported for all the ten selected structures (see Figure S1 in SI [255]). QM/FQ^b values are always larger than QM/FQ^a: this is related to the difference between atomic electronegativities of the two parametrizations. Such a difference is larger in FQ^b. On the other hand, QM/FQ^c predicts the greatest absolute E_{ele} values, because polarization is promoted by smaller values of chemical hardnesses. The largest discrepancy between QM/FQ and SAPT0 is observed for the pristine FQ parametrization by Rick et al., [27] i.e. QM/FQ^a, whereas the best agreement is given by our recent parametrization, [321] i.e. QM/FQ^c (see also Table 3.2). This is not surprising, because FQ^c was tuned to reproduce the total interaction energy calculated at the CCSD(T) level, whereas FQ^a and FQ^b were set to reproduce bulk water properties (FQ^a, Ref. [27]) or QM atomic charges (FQ^b, Ref. [128]). As depicted in Figure 3.8, QM/FQFμ over-performs QM/FQ. This is also confirmed by the data reported in Table 3.2, where a RMSD of only 1.62 kcal/mol and an RE of 6.47 % are reported. QM/FQFμ E_{ele} are dominated by charge contributions (on average 75%), that in agreement with what has shown above for the water dimer. Raw data of charge/dipoles contributions to QM/FQFμ E_{ele} are reported in Table S4 in SI. [255] To further confirm the quality of QM/FQFμ, the same analysis was applied to a snapshots of MOXY in aqueous solution constructed by using a cutting radius of 7 Å. Such an analysis is discussed in Section S5.7 in SI. [255]

The same behavior highlighted for MOXY also applies to the other selected molecules (ACRO, MeOH, NMA). In fact, QM/FQFμ always overperforms QM/FQ. This is particularly evident in case of MeOH, where SAPT0 values range from -40 to -15 kcal/mol, thus moving from weak solute-solvent interactions to strong HBs. This is due to the fact the MeOH is the only chosen molecule in which solvent water molecules can act both as H-donor and H-acceptor. Figure 3.8 clearly shows that at small E_{ele} values all four approaches predict similar energy values, whereas as energy increases, the differences between the methods increases. On the other hand, QM/FQFμ correctly reproduces SAPT0 values in the whole range of energies (i.e. for both weak and strong HBs configurations), as can be seen both from Table 3.2 and Figure 3.8, where QM/FQFμ values lie almost perfectly on the diagonal. This can be particularly appreciated from the data shown in the last column of Table 3.2, which reports a statistical analysis over the whole set of 40 structures. It is also

remarkable that QM/FQFμ, as well as all the three QM/FQ parametrizations, give errors with respect to full QM calculation by far lower than what has been recently reported for QM/AMOEBA calculations on different aqueous systems. [351]

To end the discussion, QM/FQFμ charge and dipole contributions for two representative structures of MOXY and MeOH in aqueous solution (structures $MOXY_1$ and MeOH-2 in Figures S1 and S2, given as SI [255]) are analyzed. In Figures 3.9 and 3.10, each water molecule is colored as a function of the contribution to E_{ele}. Such an analysis is done according to what has been recently proposed for Functional group-SAPT (FSAPT). [352, 353] $MOXY_1$ is characterized by one HB, where a single water molecule acts as H-donor, whereas MeOH-2 is involved in two HBs, in which one water molecules acts as H-donor and a second one as H-acceptor.

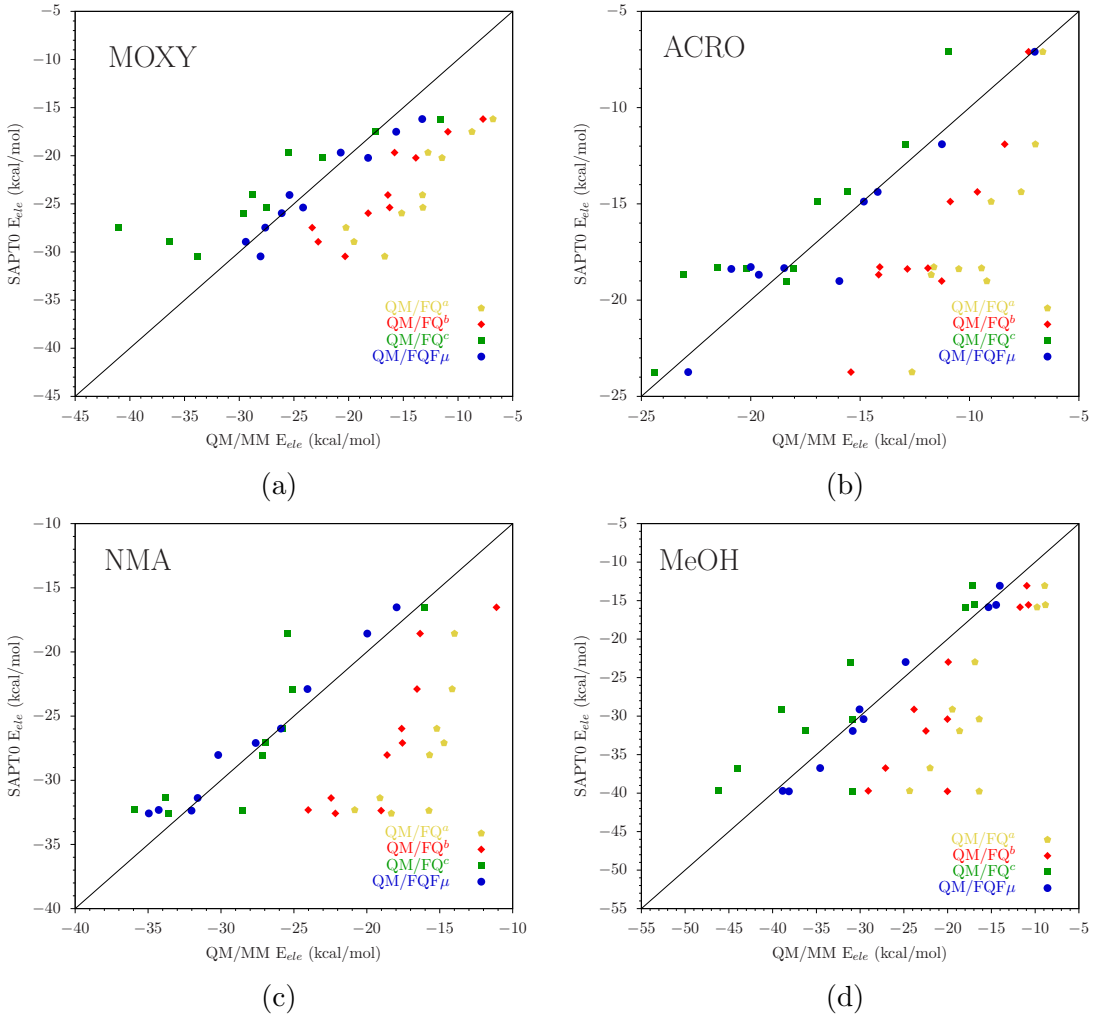


Figure 3.8: Comparison between calculated QM/FQ^a, QM/FQ^b, QM/FQ^c, QM/FQFμ E_{ele} (HF/6-311++G** level for the QM portion) and SAPT0/6-311++G** data. In case of SAPT0 calculations electrostatic and induction energy contributions are summed up. Raw data are given in Table S5 in SI. [255] All data are reported in kcal/mol.
[a] FQ parametrization taken from Ref. [27]
[b] FQ parametrization taken from Ref. [128]
[c] FQ parametrization taken from Ref. [321]

		MOXY	ACRO	NMA	MeOH	TOT
QM/FQ[a]	RMSD	10.02	7.46	11.53	12.66	10.60
	MAE	13.76	11.11	16.62	23.38	23.38
	RE	42.54%	39.72%	40.56%	39.85%	40.67%
QM/FQ[b]	RMSD	7.29	5.38	8.71	9.35	7.83
	MAE	10.14	8.32	13.34	19.75	19.75
	RE	30.85%	28.02%	30.07%	27.09%	29.01%
QM/FQ[c]	RMSD	5.92	2.36	2.96	6.16	4.67
	MAE	13.59	4.41	6.88	9.90	13.59
	RE	19.67%	14.39%	8.79%	19.54%	15.60%
QM/FQFμ	RMSD	1.62	1.44	1.41	1.29	1.45
	MAE	2.92	3.06	2.37	2.18	3.06
	RE	6.47%	5.68%	4.65%	4.72%	5.38%

Table 3.2: Root Mean Squared Deviation (RMSD), Maximum Absolut Error (MAE) and Relative Error (RE) of ten selected structures of MOXY, ACRO, MeOH and NMA in aqueous solution extracted from aqueous solution. SAPT0/6-311++G** E_{ele} values are taken as reference. TOT indicates statistical parameters calculated on all 40 structures extracted from MD runs. RMSD and MAE are given in kcal/mol.
[a] FQ parametrization taken from Ref. [27]
[b] FQ parametrization taken from Ref. [128]
[c] FQ parametrization taken from Ref. [321]

Figures 3.9 and 3.10 clearly show that in both cases HB water molecules give the largest contributions to E_{ele}. However, other water molecules, which are not directly involved in HB with the QM portion, give non-negligible contributions to the total electrostatic energy. This has a practical consequence: in fact, cluster approaches, in which only few, geometrically close, water molecules are included in the QM portion, can inappropriately model solvent effects, because such relevant contributions will be most probably neglected.

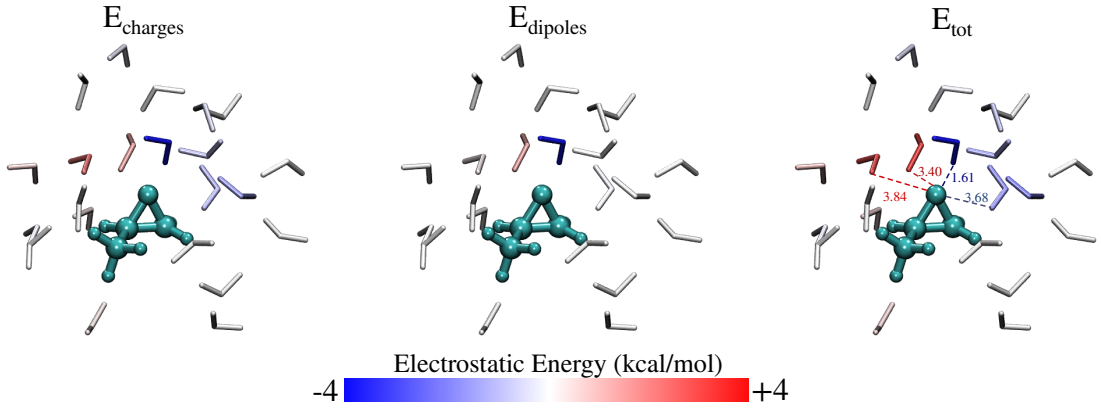


Figure 3.9: QM/FQFμ electrostatic energy contributions (kcal/mol) for $MOXY_1$. $E_{charges}$ and $E_{dipoles}$ indicate charge and dipole contributions to the total E_{ele}. All atoms in each water molecule are colored according to their contribution. The color maps saturate at ± 4 kcal/mol.

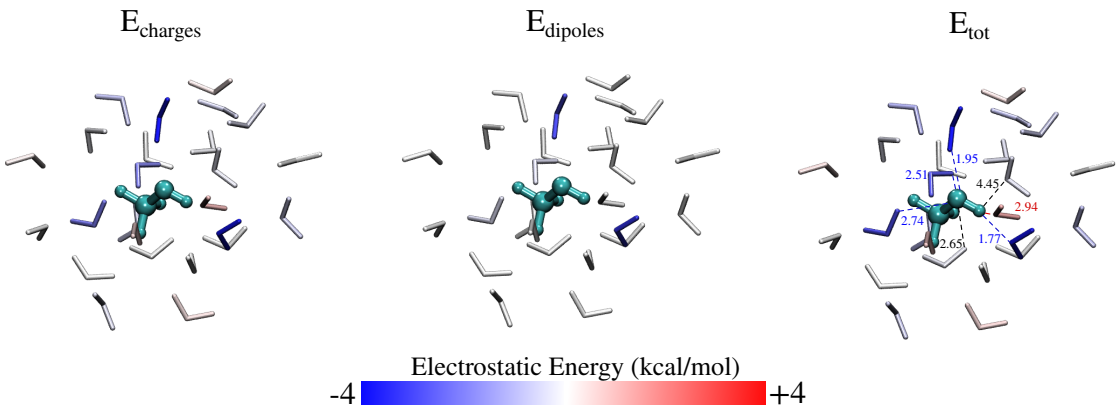


Figure 3.10: QM/FQFμ electrostatic energy contributions (kcal/mol) for MeOH-2. $E_{charges}$ and $E_{dipoles}$ indicate charge and dipole contributions to the total E_{ele}. All atoms in each water molecule are colored according to their contribution. The color maps saturate at ± 4 kcal/mol.

Electronic transitions for a fully polarizable QM/MM approach based on fluctuating charges and fluctuating dipoles: Linear and corrected linear response regimes

Excited-state phenomena play a crucial role in many application fields, as for instance photocatalysis, optical information storage and solar cells. In the past decades, theoretical modeling of excited-state properties of molecules in the gas-phase has become a widespread strategy of investigation, [129] giving precious information on, for instance, the nature of the electronic excitation, [151, 354, 355] nuclei-electron coupling effects [356–358] and excited state electron dynamics. [359–361]

However, for electronic phenomena taking place in the condensed phase, [234, 238, 279, 362–367] the interplay between the molecule and its environment can substantially alter the electronic response to external electromagnetic fields. Therefore, any accurate modeling of excited states of solvated systems asks for reliable theoretical approaches to include the effects of the environment at all levels of the excitation phenomenon.

Most of the currently available approaches to describe the effects of the external environment on molecular properties belong to the class of the so-called focused models; [2, 3, 72, 73, 86, 87] the attention is focused on the molecule and the environment is treated a lower level of sophistication as it modifies, but not determines, the molecular response to the external radiation. In order to keep the atomistic description of the environment, thus substantially overcoming well-known and amply used continuum solvent descriptions, [2–7] multiscale QM/Molecular Mechanics (MM) approaches have been developed. [74, 97] In such models, the molecule (solute) is treated at the QM level, whereas the environment (solvent) is modelled by means of classical MM force fields (FF). The interaction between the QM and MM portions is usually described in terms of electrostatic forces, although approaches to include non-electrostatic QM/MM interactions have been proposed. [8–15] In order to fully capure the physics of solute-solvent interactions, mutual polarization effects between the QM and MM moieties need to be accounted for. Therefore, several polarizable QM/MM approaches have been proposed, based on distributed multipoles, [16–20] induced dipoles, [21–25] Drude oscillators, [26] Fluctuating Charges (FQ) [27–29] and the recently developed approach based on both Fluctuating Charges and Fluctuating Dipoles (FQFμ). [30, 31]

The latter can be seen as a refinement of QM/FQ, previously developed by some of us, [29, 103, 104, 108, 110–113, 127, 368, 369] where the MM portion is described by means of electric charges, which vary as a function of differences in MM atomic electronegativities and as a response to the electric potential generated by the QM density. Therefore, in QM/FQ only monopoles, i.e. zeroth order of the electrostatic Taylor expansion, are taken into consideration. This means that the intrinsic anisotropy and out-of-plane contributions of molecule-environment interactions are not explicitly taken into account. To overcome this problem, the electrostatic description given by FQ has been refined by including fluctuating dipoles (Fμ), thus resulting in the QM/FQFμ approach. Notice that similar approaches have been developed in other contexts, [263, 264, 303–306] however they are not specifically

intended to model molecular systems/properties in solution.

The potentialities of QM/FQFμ at predicting electrostatic interaction energies of organic molecules in solution have been previously highlighted. [30] In this work, this model is further extended to the calculation of vertical excitation energies, by resorting to linear response (LR) theory for Self Consistent Field (SCF) methods. Excitation energies are determined by solving the so-called Casida equations, [370] for the QM portion, appropriately modified so to include extra terms due to QM/MM mutual polarization effects. [19, 22–24, 104, 250, 306, 371, 372] The LR approach may reliably describe electronic transitions involving excited states associated with large transition dipole moments. However, it lacks any accounting of the relaxation of the classical portion as a response to charge equilibration following the excitation of the QM density. For this reason, several state specific (SS) approaches have been proposed in the literature both in case of continuum models [279, 280, 373–383] and polarizable QM/MM approaches. [19, 384]

In this work, the perturbative SS corrected linear response (cLR) scheme [373] originally developed for the Polarizable Continuum Model (PCM) is exploited and extended to both QM/FQ and QM/FQFμ transition energies for the first time. In this way, the classical MM portion is adjusted to the relaxed density matrix of a given QM excited state. Notice that a similar approach has been applied to polarizable QM/AMOEBA calculations. [19]

To show both the potentialities of QM/FQFμ at describing excitation energies and the differences arising from resorting to LR and cLR regimes, the method is applied to the calculation of vacuo-to-water solvatochromic shifts of para-nitroaniline (pNA), pyridine and pyrimidine. From such an analysis, it will be possible to (i)highlight differences in the description of the excitation phenomenon moving from continuum PCM to QM/FQ and QM/FQFμ fully atomistic approaches; (ii) directly quantify the separate role fluctuating dipoles and fluctuating charges, and their coupling; (iii) evaluate effects arising from the specification of LR or cLR regimes.

The manuscript is organized as follows. In the next section, the FQFμ force field is presented and its coupling with a QM description at the SCF level (QM/FQFμ) is detailed. LR and cLR working equations are then presented and discussed. After a brief section discussing on the computational protocol which is adopted, numerical results are presented. Some drawn conclusions and a discussion on the future perspectives of the approach end the manuscript.

QM/FQFμ Approach

In the FQFμ force field each MM atom is endowed with both a charge (q) and an atomic dipole ($\boldsymbol{\mu}$), that are not fixed but can vary according to the external electric potential/ electric field.

The total energy $\mathcal{E}$ associated with a gaussian distribution of charges and dipoles reads: [304]

$$\begin{aligned}\mathcal{E}(\mathbf{q},\boldsymbol{\mu}) = &\sum_i q_i\chi_i + \frac{1}{2}\sum_i q_i\eta_i q_i + \frac{1}{2}\sum_i\sum_{j\neq i} q_i \mathrm{T}^{qq}_{ij} q_j + \\ &+ \sum_i\sum_{j\neq i} q_i \mathbf{T}^{q\mu}_{ij}\boldsymbol{\mu}_j + \frac{1}{2}\sum_i\sum_{j\neq i}\boldsymbol{\mu}_i^\dagger \mathbf{T}^{\mu\mu}_{ij}\boldsymbol{\mu}_j - \frac{1}{2}\sum_i \boldsymbol{\mu}_i^\dagger \alpha_i^{-1}\boldsymbol{\mu}_i \end{aligned} \tag{3.25}$$

where χ is the atomic electronegativity, η the chemical hardness and α the atomic polarizability. T_{ij}^{qq}, $\mathbf{T}_{ij}^{q\mu}$ and $\mathbf{T}_{ij}^{\mu\mu}$ are the charge-charge, charge-dipole and dipole-dipole interaction kernels, respectively. Their functional form can be found in Refs. [30, 304]

In order to collect all the charge quadratic terms, the diagonal elements of $\mathbf{T}^{qq}$ and $\mathbf{T}^{\mu\mu}$ can be imposed to be related to atomic chemical hardness η and atomic polarizability α, respectively. [30] Therefore, the parameters entering the definition of FQFμ are limited to electronegativity, chemical hardness and polarizability of each atom type. As a result, Equation 3.25 can be formally rewritten as:

$$\begin{aligned}\mathcal{E}(\mathbf{q},\boldsymbol{\mu}) &= \frac{1}{2}\sum_i\sum_j q_i \mathrm{T}_{ij}^{qq} q_j + \frac{1}{2}\sum_i\sum_j \boldsymbol{\mu}_i^\dagger \mathbf{T}_{ij}^{\mu\mu}\boldsymbol{\mu}_j + \\ &+ \sum_i\sum_j q_i \mathbf{T}_{ij}^{q\mu}\boldsymbol{\mu}_j^\dagger + \sum_i q_i\chi_i = \\ &= \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q}\end{aligned} \tag{3.26}$$

where a vector notation has been adopted. Similarly to QM/FQ, the equilibrium condition of Equation 3.26 is reached when the Electronegativity Equalization Principle (EEP) is satisfied, i.e. when each atom has the same electronegativity. [124] If each MM moiety is constrained to assume a fixed, total charge Q_α, Equation 3.26 can be written by exploiting a set of Lagrangian multipliers (λ_α), whose number is equal to the total number of moieties in the MM portion:

$$\begin{aligned}\mathcal{E}\left(\mathbf{q},\boldsymbol{\mu},\boldsymbol{\lambda}\right) &= E\left(\mathbf{q},\boldsymbol{\mu}\right) + \sum_\alpha\left[\lambda_\alpha\sum_i\left(q_{\alpha i}\right) - Q_\alpha\right] = \\ &= \frac{1}{2}\sum_{i\alpha}\sum_{j\beta} q_{i\alpha}\mathrm{T}_{i\alpha,j\beta}^{qq}q_{j\beta} + \frac{1}{2}\sum_i\sum_j \boldsymbol{\mu}_{i\alpha}^\dagger\mathbf{T}_{i\alpha,j\beta}^{\mu\mu}\boldsymbol{\mu}_{j\beta} + \\ &+ \sum_i\sum_j q_{i\alpha}\mathbf{T}_{i\alpha,j\beta}^{q\mu}\boldsymbol{\mu}_{j\beta}^\dagger + \sum_{i\alpha} q_{i\alpha}\chi_{i\alpha} + \\ &+ \sum_\alpha\lambda_\alpha\left[\sum_i q_{\alpha i} - Q_\alpha\right] = \\ &= \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q}\end{aligned} \tag{3.27}$$

where α and β run over MM moieties. By minimizing $\mathcal{E}$ with respect all variables, the following linear system is defined: [30]

$$\begin{pmatrix}\mathbf{T}^{qq} & \mathbf{1}_\lambda & \mathbf{T}^{q\mu} \\ \mathbf{1}_\lambda^\dagger & 0 & \mathbf{0} \\ -\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu}\end{pmatrix}\begin{pmatrix}\mathbf{q} \\ \lambda \\ \boldsymbol{\mu}\end{pmatrix} = \begin{pmatrix}-\boldsymbol{\chi} \\ Q_{\mathrm{tot}} \\ \mathbf{0}\end{pmatrix} \qquad \Rightarrow \qquad \mathbf{D}\mathbf{Q}_\lambda = -\mathbf{C}_Q \tag{3.28}$$

where $\mathbf{1}_{\boldsymbol{\lambda}}$ is a rectangular matrix, which accounts for the Lagrangians. $\mathbf{C}_Q$ is a vector containing atomic electronegativities and total charge constraints, whereas $\mathbf{L}_\lambda$ is a vector containing charges, dipoles and Lagrange multipliers.

FQFμ FF can be effectively coupled to a QM SCF description, within a QM/MM framework:

$$\mathcal{E}(\mathbf{P},\mathbf{q},\boldsymbol{\mu},\boldsymbol{\lambda}) = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{1}{2}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q} + \\ + \mathbf{q}^\dagger\mathbf{V}(\mathbf{P}) - \boldsymbol{\mu}^\dagger\mathbf{E}(\mathbf{P}) \tag{3.29}$$

where $\mathbf{h}$ and $\mathbf{G}$ are the usual one- and two-electron matrices, and $\mathbf{P}$ is the density matrix. $\mathbf{q}^\dagger\mathbf{V}(\mathbf{P})$ and $\boldsymbol{\mu}^\dagger\mathbf{E}(\mathbf{P})$ are the electrostatic couplings between the QM density and the FQs and Fμs, respectively. The effective Fock matrix, i.e. the derivative of the energy with respect to the density matrix, reads:

$$\tilde{F}_{\mu\nu} = \frac{\partial\mathcal{E}}{\partial P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{V}^\dagger_{\mu\nu}\mathbf{q} - \mathbf{E}^\dagger_{\mu\nu}\boldsymbol{\mu} \tag{3.30}$$

Charges and dipoles are obtained by imposing the global functional to be stationary with respect to charges, dipoles and Lagrangian multipliers:

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_{\boldsymbol{\lambda}} & \mathbf{T}^{q\mu} \\ \mathbf{1}^\dagger_{\boldsymbol{\lambda}} & \mathbf{0} & \mathbf{0} \\ -\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ \mathbf{Q}_{\mathrm{tot}} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} -\mathbf{V}(\mathbf{P}) \\ \mathbf{0} \\ \mathbf{E}(\mathbf{P}) \end{pmatrix}$$

$$\mathbf{D}\mathbf{L}_\lambda = -\mathbf{C}_Q - \mathbf{R}(\mathbf{P}) \tag{3.31}$$

Notice that, with respect to Equation 3.28, a new term, $\mathbf{R}(\mathbf{P})$, appears, which collects QM polarization sources. $\mathbf{L}_\lambda$ is the vector containing charges, dipoles and Lagrangian multipliers.

QM/FQFμ Electronic Transition Energies

Linear Response Regime

LR for polarizable QM/MM approaches has been already discussed in the previous literature. [24, 104, 306] Electronic excitation energies ω for SCF Hamiltonians can be obtained by solving Casida's equations:[370]

$$\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^* & \tilde{\mathbf{A}}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{3.32}$$

$\tilde{A}$ and $\tilde{B}$ matrices are defined as:

$$\tilde{A}_{ai,bj} = (\epsilon_a - \epsilon_i)\delta_{ab}\delta_{ij} + (ai|bj) - c_x(ab|ij) + c_l f^{xc}_{ai,bj} + C^{pol}_{ai,bj} \tag{3.33}$$

$$\tilde{B}_{ai,bj} = (ai|bj) - c_x(aj|ib) + C^{pol}_{ai,bj} \tag{3.34}$$

where $(pq|rs)$ are two electron integrals, and ε are molecular orbital (MO) energies. c_x and c_l are coefficients of which the definition depends on the SCF level adopted (c_x =1, c_l = 0 for HF wavefunctions, c_x =0, c_l = 1 for DFT).

In case of polarizable QM/MM approaches, $\tilde{A}$ and $\tilde{B}$ in Equations 3.33 and 3.34 account for an extra term, C^{pol}, which is specified according to a particular method. In particular, for QM/FQ it reads:

$$C^{FQ}_{ai,bj} = \sum_{P}^{N_q} \left(\int_{\mathbb{R}^3} \phi_a(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \phi_i(\mathbf{r}) \, \mathrm{d}\mathbf{r} \right) \cdot q_P^T(\phi_b, \phi_i) \tag{3.35}$$

where, q^T are the perturbed fluctuating charges adjusted to the transition density $\mathbf{P}_K^T = \mathbf{X}_K + \mathbf{Y}_K$. [104] For QM/FQF$\mu$, an extra term appears n the definition of C^{pol}, accounting for the presence of fluctuating dipoles:

$$\begin{aligned} C^{FQF\mu}_{ai,bj} = & \sum_{P}^{N_q} \left(\int_{\mathbb{R}^3} \phi_a(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \phi_i(\mathbf{r}) \, \mathrm{d}\mathbf{r} \right) \cdot q_P^T(\phi_b, \phi_i) + \\ & - \sum_{P}^{N_\mu} \left(\int_{\mathbb{R}^3} \phi_a(\mathbf{r}) \frac{(\mathbf{r} - \mathbf{r}_p)}{|\mathbf{r} - \mathbf{r}_p|^3} \phi_i(\mathbf{r}) \, \mathrm{d}\mathbf{r} \right) \cdot \boldsymbol{\mu}_P^T(\phi_b, \phi_i) \end{aligned} \tag{3.36}$$

Perturbed charges (q^T) and perturbed dipoles ($\boldsymbol{\mu}^T$) are calculated by solving the following system of equations:

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_{\boldsymbol{\lambda}} & \mathbf{T}^{q\mu} \\ \mathbf{1}_{\boldsymbol{\lambda}}^{\dagger} & \mathbf{0} & \mathbf{0} \\ -\mathbf{T}^{q\mu^{\dagger}} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\mathbf{V}(\mathbf{P}_K^T) \\ \mathbf{0} \\ \mathbf{E}(\mathbf{P}_K^T) \end{pmatrix}$$

$$\mathbf{D}\mathbf{L}_{\lambda}^T = -\mathbf{R}(\mathbf{P}_K^T) \tag{3.37}$$

where:

$$V(\mathbf{P}_K^T) = -\sum_{ai} \mathbf{P}_{K,ai}^T \int_{R^3} \phi_a(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \phi_i(\mathbf{r}) \, \mathrm{d}\mathbf{r} \tag{3.38}$$

$$\mathbf{E}(\mathbf{P}_K^T) = \sum_{ai} \mathbf{P}_{K,ai}^T \int_{R^3} \phi_a(\mathbf{r}) \frac{(\mathbf{r} - \mathbf{r}_p)}{|\mathbf{r} - \mathbf{r}_p|^3} \phi_i(\mathbf{r}) \, \mathrm{d}\mathbf{r} \tag{3.39}$$

The right hand side of Equation 3.37 contains both the electric potential and field due to the perturbed density matrix $\mathbf{P}_K^T$. Notably, the constant term in Equation 3.31 due to atomic electronegativities vanishes. Therefore, polarization arises from the QM density, and not from differences in electronegativity, that similarly to the QM/FQ approach. [104, 260]

Corrected Linear Response Regime

LR and SS approaches provide the same result in case of gas-phase systems. [370] However, when polarizable QM/classical methods are exploited, the two approaches differ. [279] This is essentially due to the non-linear term introduced in the Hamiltonian (see e.g. Equation 3.29). Different approximations to solve this problem have been proposed. [373, 374, 377] In this work, the SS cLR approach, firstly proposed in for QM/PCM Hamiltoniands and successfully extended to the polarizable QM/AMOEBA method, [19] is exploited. To the best of our knowledge, this the

first time that the extension of cLR to QM/FQ and QM/FQFμ-like methods is proposed.

The common way to discuss the physical differences between LR and cLR regimes in condensed phase is to resort to a two-step protocol, in which the first step is the same in both approaches and consists in the electronic excitation of the solute to the K-th excited state by keeping frozen the MM response to the solute ground state. From a computational point of view, such a picture can be achieved by imposing the extra C^{pol} term in Equations 3.33 and 3.34 to be equal to zero. This means that MM contributions are only included in the specification of MOs through Equation 3.30. The excitation energy obtained in this approximation is refereed to as ω_0^K. In the second step of the process, MM electronic degrees of freedom adjust to the K-th excited state density. Due to the fact that in the LR approach the MM variables are determined as a response of the whole transition densities, LR is unable to catch the energy differences associated to the relaxation of the QM electron density, whereas it only accounts for a correction due to dynamic solute-environment interactions. QM/FQ and QM/FQFμ cLR excitation energies from ground to K-th excited state can be written as:

$$\omega_{K,FQ}^{cLR} = \omega_K^0 + \frac{1}{2}\sum_P^{Nq} q_p(\mathbf{r}_p; \mathbf{P}_K^{\Delta}) V(\mathbf{r}_p, \mathbf{P}_K^{\Delta}) \tag{3.40}$$

$$\begin{aligned}\omega_{K,FQF\mu}^{cLR} = \omega_K^0 &+ \frac{1}{2}\sum_P^{Nq} q_p(\mathbf{r}_p; \mathbf{P}_K^{\Delta}) V(\mathbf{r}_p, \mathbf{P}_K^{\Delta}) + \\ &- \frac{1}{2}\sum_P^{N_\mu} \boldsymbol{\mu}_p(\mathbf{r}_p; \mathbf{P}_K^{\Delta}) \mathbf{E}(\mathbf{r}_p, \mathbf{P}_K^{\Delta})\end{aligned} \tag{3.41}$$

$\mathbf{P}_K^{\Delta}$ is the so-called relaxed-density matrix, computed through the so-called Z-vector approach, as:

$$\mathbf{P}_K^{\Delta} = \mathbf{P}_K^T + \mathbf{Z}_K \tag{3.42}$$

where $\mathbf{P}_K^T$ is the unrelaxed density matrix adopted in the LR approach, whereas $\mathbf{Z}_K$ is the Z-vector contribution which accounts for orbital relaxations. Fluctuating Charges and Fluctuating Dipole in Equation 3.41 are calculated as in Equation 3.37, by substituting $\mathbf{P}_K^T \rightarrow \mathbf{P}_K^{\Delta}$.

Computational Details

The following computational protocol was exploited: [13, 112, 113, 260]

1. *Definition of the system:* The solute (pNA, pyridine and pyrimidine) was put at the center of a cubic box containing a number of water molecules large enough to account for solute-solvent interactions (vide infra). pNA, pyridine and pyrimidine geometries were optimized at the B3LYP/aug-cc-pVDZ level of theory, and RESP charges were calculated by including aqueous solvent effects by means of the PCM approach. [2].

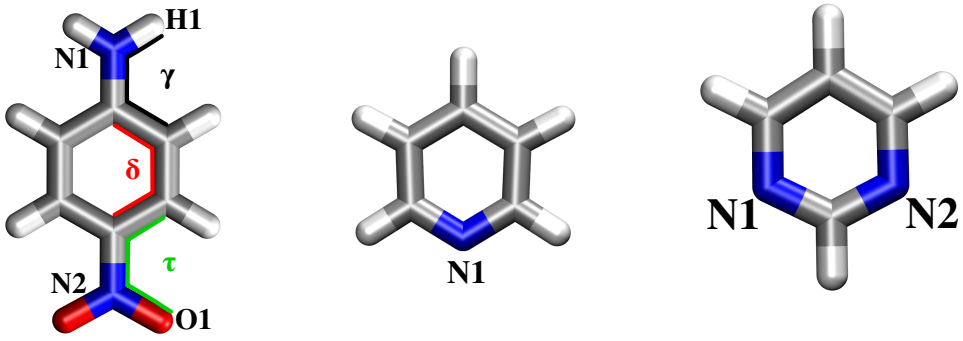
2. *Classical Molecular Dynamics (MD) runs:* For each system, MD simulations were performed by imposing periodic boundary conditions (PBC) on the cubic box. Each MD run was preceded by an equilibration step. From each MD run, a set of 100 uncorrelated snapshots was extracted.

3. *Definition of the the two-layer scheme:* For each snapshot, a solute-centered sphere was cut. The radius of the sphere was chosen so to describe all specific water-solute interactions (vide infra).

4. *QM/MM calculations:* QM/FQ and QM/FQFμ vertical excitation energies were calculated on the spherical frames obtained at the previous step. The results obtained for each frame were convoluted with a gaussian band-shape and averaged to produce the final spectrum:

$$\varepsilon_i(\tilde{\nu}) = \frac{\sqrt{\pi}e^2 N_A}{1000\ln(10)c^2 m_e}\frac{f_i}{\sigma}\exp\left[-\left(\frac{\tilde{\nu}-\tilde{\nu}_i}{\sigma}\right)^2\right] \tag{3.43}$$

where, ε is the molar absorptivity in L mol^{-1}cm^{-1}, $\tilde{\nu}_i$ is the excitation energy of the $i-th$ state, f_i is its oscillator strength, whereas σ is the standard deviation of the gaussian convolution. e, N_A, c and m_e are the electron charge, the Avogadro number, the speed of light, and the electron mass, respectively.



(a) (b) (c)

Figure 3.11: pNA (**a**), pyridine (**b**) and pyrimidine (**c**) molecular structures and atom labelling. For pNA, the atoms involved in the definition of γ, δ and τ dihedral angles are highlighted.

MD simulations were performed using GROMACS, [141] by exploiting GAFF [142] to describe both intra- and inter-molecular interactions. RESP charges were used to account for electrostatic interactions, and TIP3P FF was used to describe water molecules. [122] A single molecule of different solutes was dissolved in a cubic box containing 5781, 5342 and 5324 water molecules in case of pNA, pyridine and pyrimidine, respectively. pyridine and pyrimidine structures were kept fixed during all the steps of the MD run because they exhibit a single conformer in aqueous solution. PNA geometry was kept fixed during the equilibration step only, whereas it was free to move in the production run. The different solutes were brought to 0 K

with the steepest descent minimization procedure and then heated to 298.15 K in an NVT ensemble using the velocity-rescaling [143] method with an integration time step of 1.0 fs and a coupling constant of 0.1 ps for 100 ps. NPT simulations (using the Parrinello-Rahman barostat and a coupling constant of 1.0 ps) for 1 ns were performed to obtain a uniform distribution of molecules in the box. 5 ns production runs in the NVT ensemble were finally carried out, fixing the fastest internal degrees of freedom, i.e. hydrogen atoms, by means of the LINCS algorithm (δt=2.0 fs). [144] Electrostatic interactions were treated by using particle-mesh Ewald (PME) [145] method with a grid spacing of 1.2 Å and a spline interpolation of order 4. Intramolecular interactions between atom pairs separated up to three bonds were excluded. A snapshot every 50 ps was extracted in order to obtain a total of 100 uncorrelated snapshots for each system.

For each snapshot a solute-centered sphere with a radius of 15 Å was cut (containing approximately 400 water molecules). On each droplet QM/FQ and QM/FQFμ vertical excitation energies were calculated by exploiting both LR and cLR regimes. In case of QM/FQ, water molecules were modeled by exploiting polarizable FQ SPC parametrization proposed by Carnimeo *et al.* [128], whereas in case of QM/FQFμ calculations the parametrization proposed in Ref. [30] was adopted.

For each snapshot, additional reference calculations, labelled as QM/QM$_{\mathrm{w}}$/FQFμ, were also performed by including in the QM portion those water molecules placed at a distance lower than 3.5 Åwith respect to any QM atom. This resulted in clusters with 17 QM water molecules in case of pyridine and pyrimidine, and 23 for pNA. All remaining solvent molecules were treated with FQFμ. For the sake of comparison, additional QM calculations on the chromophores treated at the PCM level were performed. PNA was described at the CAM-B3LYP/aug-cc-pVDZ level of theory in agreement with previous studies. [76] pyridine and pyrimidine were instead treated at the M06/6-311+G(2df,2p) level, according to Ref. [375]. All computed absorption spectra were convoluted with a gaussian band-shape with a Full Width at Half Maximum (FWHM) of 0.3 eV.

All QM/FQ and QM/FQFμ LR and cLR calculations were performed by using a locally modified version of the Gaussian 16 package. [146]

Gas-Phase Vertical Excitation Energies

In this section, computed UV-Vis spectra for aqueous solutions of pNA, pyridine and pyrimidine [250, 288, 289, 374, 385–402] (see Figure 3.11 for their structures and atom labelling) are discussed. First, excitation energies of the three isolated molecules are discussed and compared to experimental values. Then, QM/FQFμ, QM/FQ, QM/QM$_{\mathrm{w}}$/FQFμ and QM/PCM LR and cLR solvatochromic shifts are presented and compared to experimental findings.

Computed gas phase vertical excitation energies for the three systems are reported in Table 3.3, together with experimental values, taken from the literature. Calculated ground and excited dipole moments are also reported.

The molecular orbitals (MOs) involved in the transitions are graphically depicted in Figure 3.12. The investigated excitation is a pure $\pi \rightarrow \pi^*$ transition in case of pNA, whereas in case of both pyridine and pyrimidine we focus our attention to the dark $n \rightarrow \pi^*$ transition. pNA $\pi \rightarrow \pi^*$ transition is often considered to have a Charge Transfer (CT) character, due to the fact that it involves a huge variation of

	Excitation	μ_g	μ_{ex}	E_{vert}	Exp.
pNA	$\pi \rightarrow \pi^*$	7.4	12.9	4.33	4.25 [a]
pyridine	$n \rightarrow \pi^*$	2.3	0.4	4.67	4.59,[b] 4.63,[c] 4.74[d]
pyrimidine	$n \rightarrow \pi^*$	2.4	0.6	4.17	4.18,[e] 4.19[f]

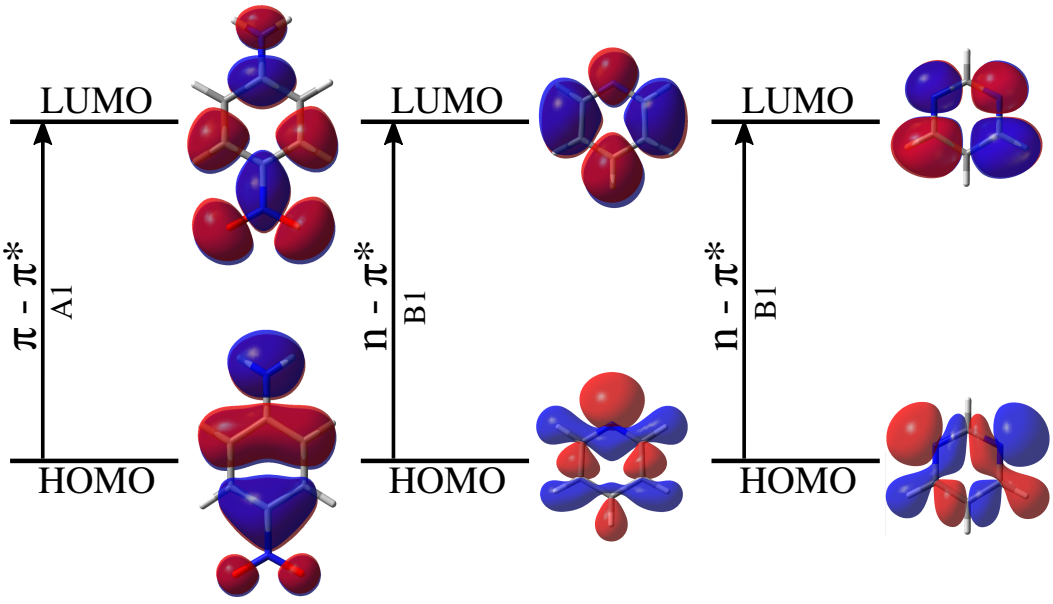
Table 3.3: pNA (CAM-B3LYP/aug-cc-pVDZ), pyridine and pyrimidine (M06/6-311+G(2df,2p) computed vertical excitation energies in gas phase (E_{vert}) given in eV. The character of the transition is reported, together with ground (μ_g) and excited (μ_{ex}) molecular dipole moments (Debye). Experimental excitation energies (eV) are also shown. [a]Ref. [396] [b]Ref. [398] [c]Ref. [374] [d]Ref. [397] [e]Ref. [399] [f]Ref. [395]

the dipole moment between the ground and excited state; this is confirmed by the computed values, see Table 3.3. The same also occurs of pyridine and pyrimidine, which differently from pNA show a decrease of the electric dipole moment in their excited states with respect to their ground state.

The comparison between computed and experimental vertical excitation energies in Table 3.3, confirms that our choice of DFT functionals/basis sets give computed data that are nicely in agreement with experiments for all systems.

MD Analysis

In this section, MD runs for pNA, pyridine and pyrimidine in aqueous solution are analyzed by using the TRAVIS package, [149] so to extract hydration patterns and analyze the role of hydrogen bonding (HB) interactions, which are crucial to rationalize the trends of vertical excitation energies (vide infra). In particular, the radial distribution function $g(r)$, together with the running coordination number (RCN) between the solute atoms highlighted in Figure 3.11 and solvent atoms (HW



(a) (b) (c)

Figure 3.12: pNA (**a**), pyridine (**b**) and pyrimidine (**c**) Molecular Orbitals (MOs) involved in the investigated electronic transitions.

and OW) are studied. the RCN for a specific site is defined as the integral of the $g(r)$ as a function of the distance. The coordination number of a specific hydration shell is the RCN at the corresponding $g(r)$ minima. In addition, in case of pNA, the dihedral distribution function (ddf) is also discussed.

$g(r)$ and RCN for the three studied molecules are reported in Figs. 3.13 and 3.14. In case of both pyridine and pyrimidine, only the nitrogen atom(s) can be involved in HB interactions with water molecules. For this reason, in Figure 3.13, panels b and c, the $g(r)$ for N1$\cdots$OW (and N2$\cdots$OW in case of pyrimidine) are reported. As it can be evinced, the $g(r)$ of both pyridine and pyrimidine present a sharp peak related to the first solvation shell at about 2 Å. The coordination numbers associated with the first solvation shell are approximately 2 and 1, for pyridine and pyrimidine, respectively. This indicates that two water molecules are involved in HB interactions with the nitrogen atom(s) of this two solutes.

In case of pNA (see Figure 3.13, panel a), H1$\cdots$OW and O1$\cdots$HW $g(r)$ are both characterized by a sharp peak at about 2 Å, whereas N1 and N2 are not involved in HB interactions with the solvent. The coordination numbers of the first shells are around 1, showing that both the push and pull groups are involved in HB interactions with one water molecule, on average. Notice that pNA is symmetric, thus two water molecules undergo HB interactions with both push and pull groups of the solute.

To end the discussion, the ddf of γ, δ and τ pNA dihedral angles are reported

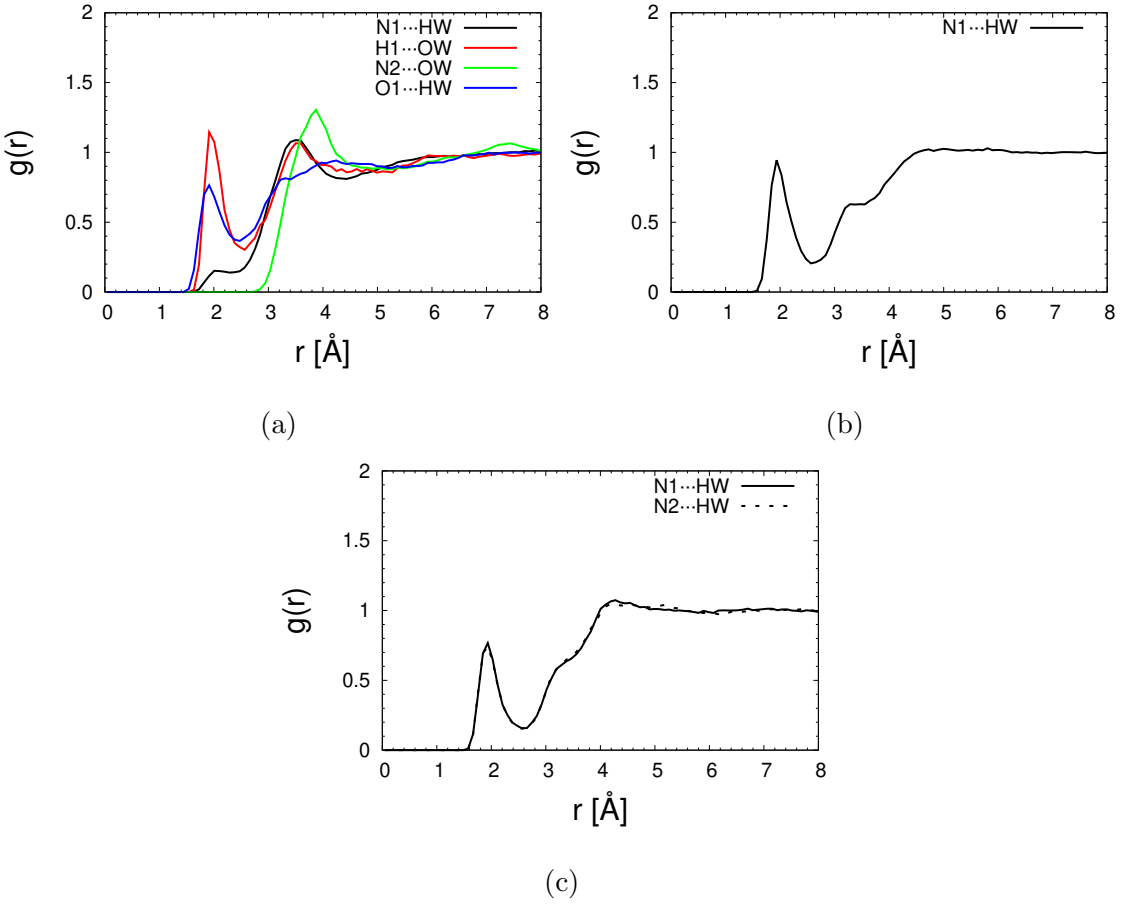


Figure 3.13: pNA (**a**), pyridine (**b**) and pyrimidine (**c**) radial distribution functions for selected atomic sites. See Figure 3.11 for labeling.

in Figure 3.15 (see Figure 3.11 for atoms involved in the definition of the dihedral angles). As it can be seen, all the considered dihedral angles fluctuate around 0 degrees (or 180 degrees in case of τ), thus indicating that pNA oscillates around its planar configuration during the MD simulation. Specifically, out-of-plane motions of the benzene ring were monitored by analyzing δ dihedral angle, which only marginally varies during all MD simulations, thus indicating an almost rigid structure of this ring. In addition, the analysis of γ and θ dihedral angles shows that

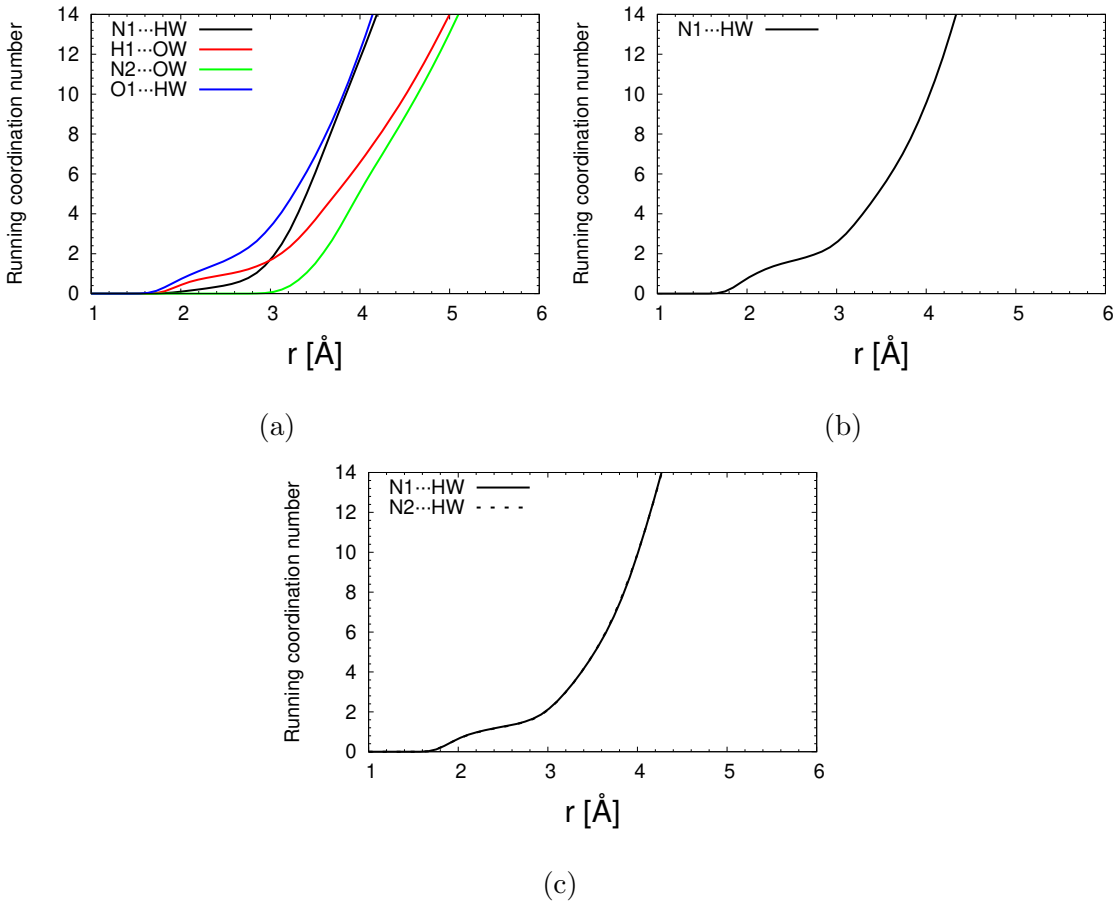


Figure 3.14: pNA (**a**), pyridine (**b**) and pyrimidine (**c**) running coordination number (RCN). See Figure 3.11 for atom labeling..

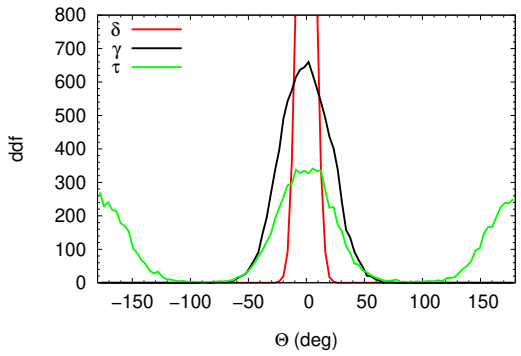


Figure 3.15: pNA dihedral distribution function (ddf). See Figure 3.11 for atom labelling.

the oxygen atoms of the nitro group can pass the rotational barrier (e.g. they can exchange their positions), whereas the same does not apply to the hydrogen atoms of the amine group. Indeed, they can oscillate in the molecule's plane, but they do not exchange their positions during the MD runs.

Vacuo-to-Water Solvatochromism

QM/FQ, QM/FQFμ and QM/QM$_w$/FQFμ vertical excitation energies of pNA (CAM-B3LYP/aug-cc-pVDZ), pyridine and pyrimidine (M06/6-311+G(2df,2p) in aqueous solution were computed on 100 uncorrelated snapshots extracted from MD runs.

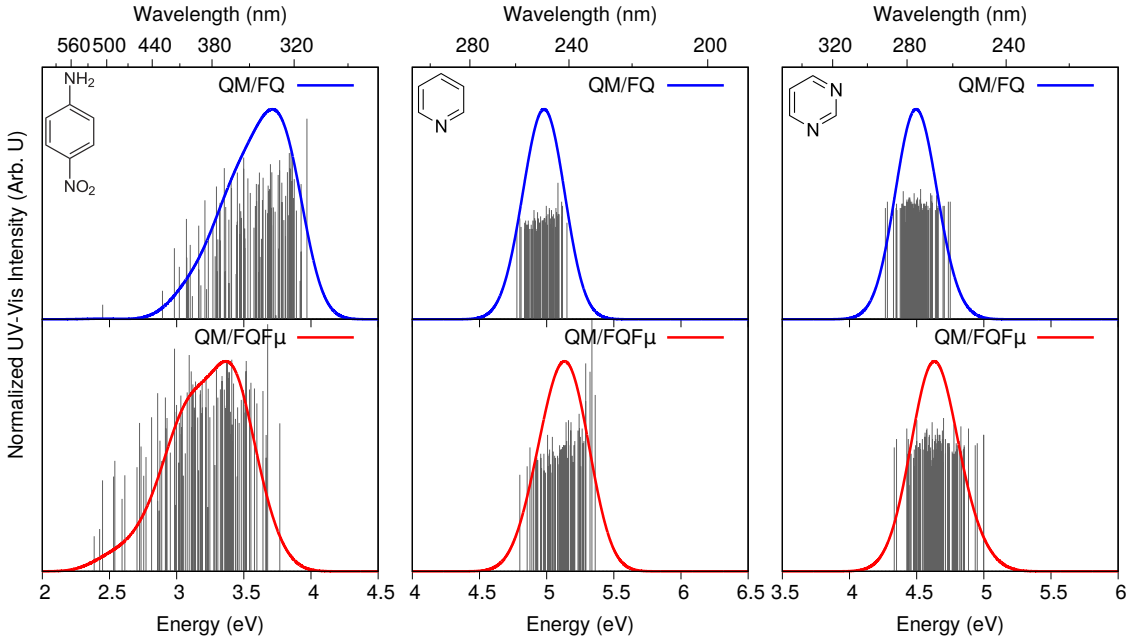


Figure 3.16: pNA (left), pyridine (middle) and pyrimidine (right) QM/FQ and QM/FQFμ UV-Vis cLR raw data (sticks) together with their gaussian convolution (FWHM=0.3 eV)

In Figure 3.16, pNA, pyridine and pyrimidine cLR absorption intensities as a function of the excitation energy, obtained by exploiting QM/FQ (top) and QM/FQFμ (bottom), for all 100 snapshots extracted from the MD simulations are plotted. Notice that stick spectra are obtained by plotting the raw data extracted from QM/MM calculations performed on each snapshot. ω_0 and LR stick spectra are reported in Figs. S1 and S2 in the Supplementary Informations (SI). [256]

In all cases, the inspection of Figure 3.16 shows that QM/FQFμ data (sticks) present a larger variability with respect to QM/FQ, which results in a greater broadening of the transition band, and in a different description of vertical excitation energies. To quantify such a variability, the standard deviation of both energies and intensities was computed (see Table S1 given as SI [256]). As expected, QM/FQFμ energy standard deviations are always larger than QM/FQ ones. Such findings are due to the inclusion of fluctuating dipoles in the MM portion, which increase the molecular dipole moment of MM water molecules. As a consequence, QM/FQFμ convoluted spectra in Figure 3.16 show a larger inhomogeneous broadening with respect to QM/FQ. To further investigate the origin of QM/FQFμ larger inhomogeneous broadening, in Figure 3.17 pNA QM/FQ (top) and QM/FQFμ (bottom)

excitation energies as a function of γ and τ dihedral angles (see Figure 3.11) are reported. Such a plot permits to understand whether the observed trend in broadening is related to the solute's geometry fluctuations or to solvent fluctuations, i.e. to the different description of the electrostatic interactions given by QM/FQ and QM/FQFμ. As it is clearly shown by Figure 3.17, the two approaches give almost identical excitation energies as a function of the dihedral angles. Therefore, it is confirmed that the different broadening is only due to a different description of the QM/MM electrostatic coupling.

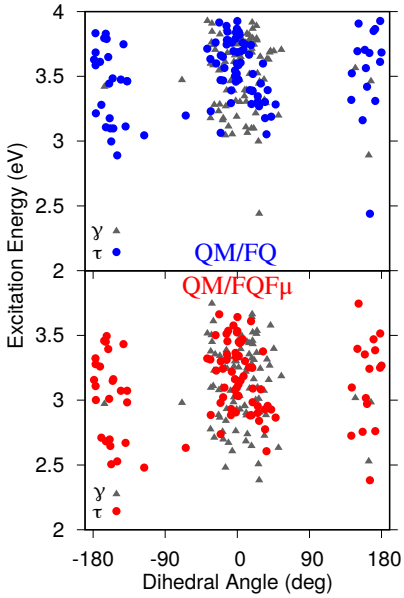


Figure 3.17: QM/FQ (top) and QM/FQFμ (bottom) cLR excitation energies for pNA as a function of the dihedral angles γ and τ (see Figure 3.11, panel a for their definition).

PNA, pyridine and pyrimidine ω_0, LR and cLR vertical excitation energies in aqueous solution, as calculated by exploiting QM/PCM, QM/FQ and QM/FQFμ, and the corresponding vacuo-to-water solvatochromic shifts are reported in Table 3.4. The associated spectra are graphically depicted in Figure 3.18, where gas-phase excitation energies are also given as stick bars. By moving from a continuum description (QM/PCM) to an atomistic approach (QM/FQ and QM/FQFμ) a gradual increase of the solvatochromic shift is reported, independently from which vertical excitation energy (ω_0, LR or cLR) is taken in consideration (see Table 3.4 and Figure 3.18). In particular, in case of pNA, QM/FQ redshifts with respect to QM/PCM of 0.11 eV, whereas QM/FQFμ of almost 0.50 eV, thus doubling the computed solvatochromism with respect to continuum QM/PCM values. In case of pyridine and pyrimidine, a shift in the opposite direction with respect to pNA is reported by all the solvation models. In particular, QM/FQ and QM/FQFμ predict an increase of vacuo-to-water solvatochromism of 0.18/0.21 and 0.33/0.35 eV with respect to QM/PCM, respectively. Notice that such an increase corresponds to a relative difference with respect to QM/PCM of almost 300 and 500 %, in case of QM/FQ and QM/FQFμ, respectively. This huge blueshift can be related to the dynamical atomistic description of the solvent water molecules given by QM/MM approaches coupled to MD simulations. The difference between QM/FQ

		E_{vert} (eV)			
	Method	QM/PCM	QM/FQ	QM/FQFμ	QM/QM$_w$/FQFμ
pNA	ω_0	3.92	3.74±0.03	3.47±0.03	3.66±0.03
	LR	**3.81**	**3.70±0.03**	**3.31±0.03**	
	cLR	3.84	3.71±0.03	3.37±0.03	
pyridine	ω_0	4.84	5.00±0.01	5.21±0.01	5.02±0.01[a]
	LR	4.84	5.00±0.01	5.20±0.01	
	cLR	**4.79**	**4.97±0.01**	**5.12±0.01**	
pyrimidine	ω_0	4.31	4.52±0.01	4.67±0.01	4.49±0.01[a]
	LR	4.30	4.51±0.01	4.66±0.01	
	cLR	**4.28**	**4.49±0.01**	**4.63±0.01**	

E_{vert} Exp. (eV): pNA 3.25[b], pyridine 4.94[c], pyrimidine 4.57[d]

		ΔE (eV)			
	Method	QM/PCM	QM/FQ	QM/FQFμ	QM/QM$_w$/FQFμ
pNA	ω_0	0.41	0.59±0.03	0.86±0.03	0.67±0.03
	LR	**0.52**	**0.63±0.03**	**1.02±0.03**	
	cLR	0.49	0.62±0.03	0.96±0.03	
pyridine	ω_0	-0.17	-0.33±0.01	-0.54±0.01	-0.35±0.01
	LR	-0.17	-0.33±0.01	-0.53±0.01	
	cLR	**-0.12**	**-0.30±0.01**	**-0.45±0.01**	
pyrimidine	ω_0	-0.14	-0.34±0.01	-0.49±0.01	-0.31±0.01
	LR	-0.13	-0.34±0.01	-0.48±0.01	
	cLR	**-0.11**	**-0.31±0.01**	**-0.45±0.01**	

ΔE Exp. (eV): pNA 0.99, pyridine -0.31, pyrimidine -0.35±0.06[e]

Table 3.4: pNA, pyridine and pyrimidine QM/PCM, QM/FQ and QM/FQFμ ω_0, LR and cLR vertical excitation energies (E_{vert}) and vacuo-to-water solvatochromic shifts (ΔE). "Best results" (see text) are in bold. Calculated QM/FQ and QM/FQFμ standard errors are also reported. They are calculated as $\sigma/\sqrt{N}$, where σ is the standard deviation and N is the number of the snapshots used to obtain the average property. Experimental data taken from the literature are reported for comparison's sake. [a]Ref. [403], [b]Ref. [288], [c]Ref. [404], [d]Ref. [395], [e]Ref. [375].

and QM/FQFμ results reported by all the selected molecular systems, shows instead the relevance of the inclusion of fluctuating dipoles in addition to fluctuating charges.

By deepening the analysis of Table 3.4, the different pictures given by LR and cLR approaches emerge. First, let's focus on pNA ω_0, LR and cLR results. The continuum QM/PCM and atomistic QM/FQ and QM/FQFμ LR and cLR excitation energies are very similar, and shifted with respect to the frozen density approximation (ω_0) of 0.11, 0.04 and 0.16 eV, respectively. The similarity between LR and cLR energies indicates that in this particular case, the relaxation of the solvent (Which is considered in the cLR approach), is as relevant as the dynamical solute-solvent interactions taken into account by the LR regime. This is not surprising, by considering that the studied excitation is a pure $\pi \rightarrow \pi^*$ transition, for which the LR regime is generally successfull. [383] Furthermore, the major differences with respect to ω_0 are reported for LR energies, which are therefore selected as "best" results in

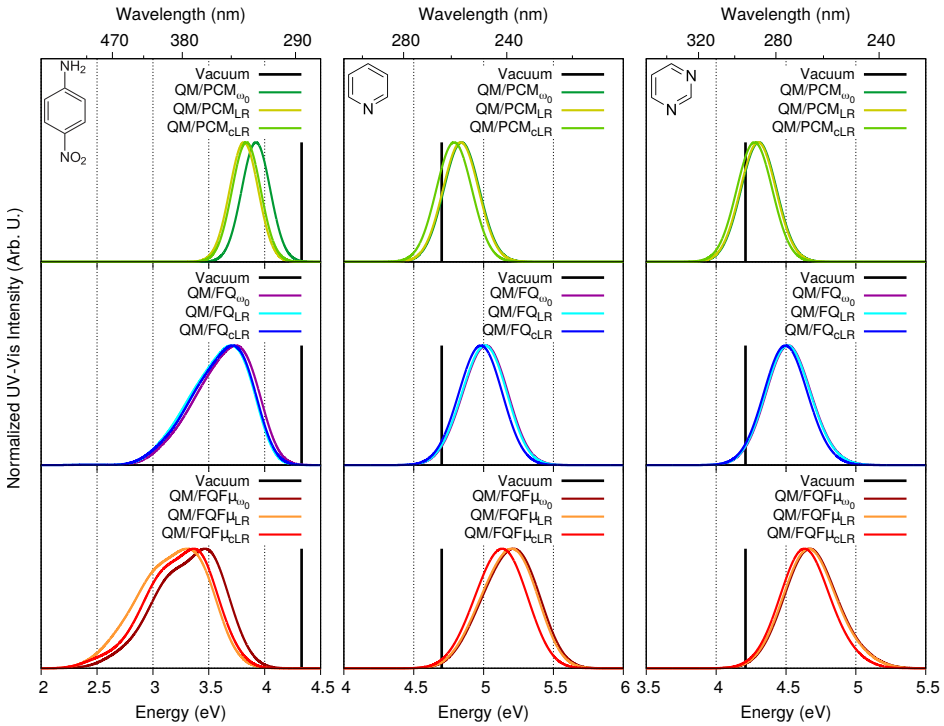


Figure 3.18: pNA (left), pyridine (middle) and pyrimidine (right) QM/PCM (top) QM/FQ (middle) and QM/FQFμ (bottom) ω_0, LR and cLR spectra. (FWHM=0.3 eV). The vertical excitation energy in gas-phase is also reported.

case of pNA (and reported in bold in Table 3.4).

Differently from pNA, the $n \rightarrow \pi^*$ transition is studied in case of pyridine and pyrimidine. For such a transition, which is associated to a large variation of the dipole moment together with a transition dipole moment close to zero (i.e. the transition is dark), cLR regime has been previously reported to be essential in order to correctly account for the relaxation of the solvent both in QM/PCM [373] and polarizable QM/MM calculations. [19] This is confirmed by the data reported in Table 3.4, where ω_0 and LR excitation energies are predicted to be almost the same by all investigated solvation approaches, whereas a shift in case of cLR is reported. In particular, QM/PCM cLR corrections shift ω_0 (and LR) energies of about 0.05 and 0.03 eV, for pyridine and pyrimidine, respectively; this corresponds to 33% deviation in both cases. The magnitude of the QM/FQ shift reported for both molecules, and of QM/FQFμ shift for pyrimidine are very similar to QM/PCM values. This is not surprising, because it has been recently reported to also apply to QM/AMOEBA cLR absorption energies. [19] QM/FQFμ cLR correction is particularly relevant for pyridine in aqueous solution, for which a shift of 0.09 (20%) with respect to the frozen density approximation ω_0 (and LR) is obtained.

Best estimates of QM/PCM, QM/FQ and QM/FQFμ vacuo-to-water solvatochromic shifts for the three systems are highlighted in Table 3.4. They are obtained by exploiting the LR regime for pNA and cLR for both pyridine and pyrimidine. In case of QM/FQ and QM/FQFμ, standard errors are also given, which are calculated as $\sigma/\sqrt{N}$, where σ is the standard deviation and N is the number of the snapshots used to obtain the average values of the shifts. We note that the continuum QM/PCM is generally unable to recover the experimental vacuo-to-water solva-

tochromism, that reasonably due to the improper account of specific solute-solvent interactions and to the "static" description of the solvation phenomenon. [19, 111, 257] As a consequence, the error on pNA solvatochromism is of 50%, whereas in case of pyridine and pyrimidine QM/PCM only recovers 35% of the experimental shift. A much better agreement with the experimental values is achieved by exploiting the polarizable QM/FQ coupled to MD simulations; in fact, this approach is able to recover the experimental vacuo-to-water solvatochromism in case of pyridine and pyrimidine, whereas it fails at reproducing the value for pNA. pNA is correctly described by QM/FQFμ, which, however, overestimates of about 0.10 eV the values for pyridine and pyrimidine. Remarkably, in case of QM/FQFμ, cLR shifts ω_0 and LR vertical excitation energies towards their experimental counterparts (see Table 3.4).

In Table 3.4, we also report computed QM/QM$_{\rm w}$/FQFμ vertical excitation energies and vacuo-to-water solvatochromic shifts performed on the 100 snapshots extracted from the MD runs. At a first glance, QM/FQ results are in very good agreement with values computed on clusters, whereas larger discrepancies are reported for QM/FQFμ . However, it is worth noticing that cluster QM/QM$_{\rm w}$/FQFμ results include not only solute-solvent mutual polarization effects, but also non-electrostatic interactions (Pauli repulsion). As some of the present authors have recently shown, [403] solute-solvent Pauli repulsion cannot be neglected when comparing QM/MM values with supermolecular (cluster) approaches. The inclusion of Pauli repulsion in QM/FQ or QM/FQFμ would change the agreement with QM/QM$_{\rm w}$/FQFμ, up to the point of increasing the agreement for QM/FQFμ and reducing it for QM/FQ. Remarkably, however, in any case cLR shifts ω_0 and LR absorption energies towards QM/QM$_{\rm w}$/FQFμ values.

	Density	Transition	Vacuo	QM/PCM	QM/FQ	QM/FQFμ
pNA		$\pi \rightarrow \pi^*$	3.14	3.11	2.60	3.35
pyridine	$\mathbf{P}_K^T$	$n \rightarrow \pi^*$	1.60	1.51	1.48	1.40
pyrimidine		$n \rightarrow \pi^*$	0.99	0.99	1.00	1.00
pNA		$\pi \rightarrow \pi^*$	2.05	2.26	1.66	2.22
pyridine	$\mathbf{P}_K^\Delta$	$n \rightarrow \pi^*$	0.80	1.00	0.82	0.85
pyrimidine		$n \rightarrow \pi^*$	0.54	0.73	0.58	0.64

Table 3.5: pNA, pyridine, pyrimidine vacuo, QM/PCM, QM/FQ and QM/FQFμ D$_{CT}$ values (Å) for the studied transitions. The values are calculated by using both the unrelaxed density $\mathbf{P}_K^T$ and relaxed $\mathbf{P}_K^\Delta$ densities .

To investigate the reasons of the overestimation of vacuo-to-water solvatochromic shifts reported for QM/FQFμ (which, as stated above, is particularly evident for pyridine and pyrimidine) the charge transfer (CT) nature of the considered transitions was characterized by a quantitative index, denoted as $D_{\rm CT}$. [151, 152] The barycenters of the positive and negative density distributions are calculated as the difference of the ground state (GS) and excited state (ES) densities. The CT length index ($D_{\rm CT}$) is defined as the distance between the two charge barycenters. $D_{\rm CT}$ indexes, as calculated by considering both the unrelaxed $\mathbf{P}_K^T$ and relaxed densities $\mathbf{P}_K^\Delta$, are reported in Table 3.5.

The data in Table 3.5, confirm that all the considered transitions are associated with an intramolecular CT. In particular, the largest $D_{\rm CT}$ values are reported in case of pNA (both in gas phase and in solution), independently from the specific

solvation model. Notice that the CT nature of each transition is different (lower) if the calculation is performed by exploiting the relaxed density ($\mathbf{P}_K^{\Delta}$) instead of the unrelaxed one ($\mathbf{P}_K^{T}$), in line with what reported in previous studies. [405] Such an observation supports the necessity of taking into account the relaxation of the solvent polarization, i.e. to adopt a cLR regime rather than the common LR approach for the solvated systems. However, in order to correctly decide the best regime for our systems, molecular dipole moments need to be analyzed. In particular, differences in cLR and LR vertical excitation energies can be related to a different description of the change in the dipole moment moving from GS to ES ($\Delta\mu = |\mu_{ex} - \mu_g|$) and the transition dipole moment ($\mu_{g,ex}$). In fact, it has been shown that if $\Delta\mu^2$ is smaller than $2\mu_{g,ex}^2$, the LR excitation energy is smaller than cLR one, and vice versa. [373] Therefore, the relevant values are reported in Table 3.6 for the three investigated solvation approaches.

QM/PCM	μ_g	μ_{ex}	$\Delta\mu^2$	$2\mu_{g,ex}^2$
pNA	10.5	17.6	50.4	72.5
pyridine	3.1	0.2	8.4	0.5
pyrimidine	3.3	0.8	6.2	0.7
QM/FQ				
pNA	10.3	15.0	22.1	38.1
pyridine	3.6	0.9	7.3	0.4
pyrimidine	3.6	1.9	2.9	0.6
QM/FQFμ				
pNA	12.9	19.0	37.2	68.0
pyridine	4.4	1.6	7.8	0.5
pyrimidine	4.4	2.7	2.9	0.7

Table 3.6: pNA, pyridine and pyrimidine QM/PCM, QM/FQ and QM/FQFμ GS(μ_g) and ES (μ_{ex}) electric dipole moments and their squared difference ($\Delta\mu^2 = |\mu_{ex} - \mu_g|^2$). ($2\mu_{g,ex}^2$), with $\mu_{g,ex}$ being the transition dipole moment, is also reported. All dipole data are given in Debye.

We notice that moving from GS to ES the dipole moment decreases for both pyridine and pyrimidine, whereas it increases for pNA. Remarkably, this is related to the negative solvatochromism reported in case of pyridine and pyrimidine, and positive in case of pNA. [375] Also, the results obtained in Table 3.4 are coherent with the fact that for pNA $2\mu_{g,ex}^2$ is larger than $\Delta\mu^2$, whereas the opposite applies to both pyridine and pyrimidine. As a consequence, LR excitation energies are smaller than cLR for pNA, whereas cLR ones are smaller than LR for the other two systems (see Table 3.4). Furthermore, the analysis based on dipole moments (Table 3.6) suggest that in case of pNA, QM/FQFμ outperforms QM/FQ because it predicts a larger difference (of almost 25%) between ES and GS dipole moments, thus resulting in a larger vacuo-to-solvent solvatochromism. The same does not apply to pyridine and pyrimidine, because QM/FQ and QM/FQFμ predict almost the same difference in dipole moments. The computed larger solvatochromism for QM/FQFμ is thus probably due to the fact that both GS ans ES dipole moments are increased of almost 50% with respect to QM/FQ. Howevef, it is worth noticing that molecular dipole moments are highly affected by non-electrostatic interactions, in particular repulsion

energy terms, as recently reported by some of the present authors. [11] The inclusion of such contributions can in principle reduce the molecular dipole moments of both GS and ES. As a consequence, non-electrostatic interactions should reasonably shift QM/FQFμ values towards the experimental data. On the other hand, in light of such considerations, the good performance of QM/FQ for pyridine and pyrimidine might be ascribed to error cancellations between the description of electrostatic effects and the lack of explicit consideration of exchange-repulsion effects. [11–13]

Calculation of IR Spectra with a Fully Polarizable QM/MM Approach Based on Fluctuating Charges and Fluctuating Dipoles

Vibrational spectroscopy, in particular infrared spectroscopy, is one of the most common techniques to study structural and dynamical features of molecular systems. Experimental spectra can be affected by a combination of effects, ranging from anharmonicity to solvent effects, the latter playing a relevant role because most experiments are conducted in the condensed phase. [75, 406–410]

In this work, we especially focus on the development of a method to account for the mutual interaction between a molecular system and its environment and its effect on the prediction of IR spectra. In fact, the presence of the environment can alter the electronic response of the target molecule to the external electric field and the vibrational frequecies associated with the normal modes. Therefore, approaches able to accurately describe environmental effects are required to obtain computed spectra directly comparable with experiments.

In the computational practice, the effects of the environment on a given spectral property are usually included by resorting to focused models, [2, 3, 72, 73, 86, 87] which are based on the assumption that the spectral signal is essentially due to the target molecule (e.g. the solute in case of solutions) and the environment (e.g. the solvent) only modifies but not determines it.

Besides the commonly used continuum solvation approaches, [6, 7] the family of QM/MM methods, [72–74, 97] may be exploited.

Their quality is connected to the specific force field which is exploited to treat the MM portion, and on the approach which is used to define the QM/MM interaction. The latter can be modeled by means of the basic mechanical embedding approach or by resorting to the so-called electrostatic embedding, where the coupling between QM and MM portions is described in terms of the Coulomb law, i.e. the electrostatic interaction between the potential of the QM density and the fixed charges which are placed at MM atoms. This picture is refined in the so-called polarizable QM/MM approaches, in which the mutual QM/MM polarization is taken into account; it can be modeled in different ways, e.g. by resorting to distributed multipoles, [16–20] induced dipoles, [21, 24, 25, 79] Drude oscillators, [26] Fluctuating Charges (FQ) [27–29] and the recently developed approach based on both Fluctuating Charges and Fluctuating Dipoles (FQFμ). [30]

In QM/FQFμ the MM portion is described in terms of both fluctuating charges and fluctuating dipoles, which are placed at MM atoms positions and can vary as a response to the QM electric potential and MM atomic electronegativities (the FQs) and QM electric field (the Fμs), respectively. Such an approach is a pragmatical extension of the QM/FQ approach, previously developed by some of the present authors, [29, 104, 110–113, 127] in which the MM portion is described by means of electric charges which can be polarized by the QM density and viceversa. As a consequence, QM/FQFμ takes into account both the out-of-plane and anisotropy contributions to polarization thanks to the inclusion of the electric dipoles in the MM portion. [411] It is worth remarking that similar approaches have been proposed in other contexts, [263, 264, 303–306, 309] however they are not based on a variational formalism and therefore are not specifically intended to model molecu-

lar properties/spectra. Also, to the best of our knowledge, they have never been extended to the calculation of molecular properties/spectra determined by the nuclear response to external fields. An additional comparison between QM/FQFμ and similar approaches can be found in Ref. [30]

The quality of QM/FQFμ at predicting electrostatic interaction energies has been recently discussed [30] and some of the present authors have recently extended this approach to the calculation of electronic vertical transition energies of organic molecules in solution at the TD-DFT level. [412] In this work, QM/FQFμ is further extended to the calculation of IR spectra, through its extension to energy nuclear derivatives. Remarkably, other QM/MM approaches have been extended to the calculation of energy gradients, [94, 318, 333, 413–417] however the only previous polarizable QM/MM approach extended to vibrational spectroscopy is the QM/FQ method developed by some of us. [103, 110–112]

The manuscript is organized as follows. In the next section the FQFμ force field is briefly presented and its coupling with a QM wavefunction is specified at the SCF level (QM/FQFμ) (see Ref. [30] for more details). Equations for analytical first and second energy derivatives are then presented and discussed. After a brief section discussing on the computational protocol which is adopted, numerical results are presented. In particular, QM/FQFμ is challenged against the description of IR spectra of three organic molecules in aqueous solution, namely methyloxirane, glycidol and gallic acid. IR spectra are computed by exploiting a hierarchy of polarizable embedding approaches, namely QM/PCM, QM/FQ and QM/FQFμ. Computed spectra are compared to their experimental counterparts, which are taken from the literature. [187, 418, 419]

Also, we will show a comparison between different polarizable approaches (PCM to QM/FQ and QM/FQFμ fully atomistic approaches), and we will underline the effects of adding fluctuating dipoles to fluctuating charges in the MM force filed. Some drawn conclusions and the discussion on future perspectives of this approach end the manuscript.

QM/FQFμ Approach

In the FQFμ force field each MM atom is endowed with both a charge q and an atomic dipole $\boldsymbol{\mu}$, that can vary according to the external electric potential and electric field.

The total energy E associated with a distribution of charges and dipoles can be written as: [30, 304, 306]

$$\begin{aligned}\mathcal{E}(\mathbf{q},\boldsymbol{\mu}) = &\sum_i q_i\chi_i + \frac{1}{2}\sum_i q_i\eta_i q_i + \frac{1}{2}\sum_i\sum_{j\neq i} q_i \mathrm{T}_{ij}^{qq} q_j + \sum_i\sum_{j\neq i} q_i \mathbf{T}_{ij}^{q\mu}\boldsymbol{\mu}_j + \\ &+ \frac{1}{2}\sum_i\sum_{j\neq i}\boldsymbol{\mu}_i^\dagger \mathbf{T}_{ij}^{\mu\mu}\boldsymbol{\mu}_j - \frac{1}{2}\sum_i \boldsymbol{\mu}_i^\dagger \alpha_i^{-1}\boldsymbol{\mu}_i \end{aligned} \tag{3.44}$$

where χ is the atomic electronegativity, η the chemical hardness and α the atomic polarizability. T_{ij}^{qq}, $\mathbf{T}_{ij}^{q\mu}$ and $\mathbf{T}_{ij}^{\mu\mu}$ are the charge-charge, charge-dipole and dipole-dipole interaction kernels, respectively. If charges and dipoles are represented as s-type gaussian distributions, the functional form of the interaction kernels provided by Mayer [304] can be exploited (see also Ref. [30]).

In order to collect all quadratic terms in the charges, the diagonal elements of $\mathbf{T}^{qq}$ can be imposed to be equal to the atomic chemical hardness η, so that the width of the charge gaussian distribution (R_q) is defined without the need of any parametrization. [30] The same holds for the diagonal elements of $\mathbf{T}^{\mu\mu}$ and the dipole gaussian distribution (R_μ), which can be defined in terms of the atomic polarizabilities (α). [30] The definition of the gaussian width R_{q_i} and R_{μ_i} in terms of η_i and α_i limits the number of parameters which enter the definition of FQFμ to electronegativity, chemical hardness and polarizability for each atom type.

In case of a molecular system, Equation 3.44 reads: [30]

$$\begin{aligned}\mathcal{E}\left(\mathbf{q},\boldsymbol{\mu},\boldsymbol{\lambda}\right) &= \mathcal{E}\left(\mathbf{q},\boldsymbol{\mu}\right) + \sum_{\alpha}\left[\lambda_\alpha\left(\sum_i\left(q_{\alpha i}\right) - Q_\alpha\right)\right] = \\ &= \frac{1}{2}\sum_{i\alpha}\sum_{j\beta} q_{i\alpha}\mathrm{T}^{qq}_{i\alpha,j\beta}q_{j\beta} + \frac{1}{2}\sum_i\sum_j \boldsymbol{\mu}^\dagger_{i\alpha}\mathbf{T}^{\mu\mu}_{i\alpha,j\beta}\boldsymbol{\mu}_{j\beta} + \sum_i\sum_j q_{i\alpha}\mathbf{T}^{q\mu}_{i\alpha,j\beta}\boldsymbol{\mu}^\dagger_{j\beta} + \\ &+ \sum_{i\alpha} q_{i\alpha}\chi_{i\alpha} + \sum_\alpha \lambda_\alpha\left[\sum_i q_{\alpha i} - Q_\alpha\right] = \\ &= \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q} \end{aligned} \tag{3.45}$$

where α and β run over molecules, whereas i and j run over the atoms of each molecule. The lagrangian multipliers λ_α are meant to preserve the total charge Q_α of each molecule. Therefore, the constrained minimum is found by imposing all derivatives of $\mathcal{E}$ with respect all variables to be equal to zero, thus resulting in the following linear system: [30, 306, 309]

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_{\boldsymbol{\lambda}} & \mathbf{T}^{q\mu} \\ \mathbf{1}^\dagger_{\boldsymbol{\lambda}} & \mathbf{0} & \mathbf{0} \\ -\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix}\begin{pmatrix} \mathbf{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ \mathbf{Q} \\ \mathbf{0} \end{pmatrix} \quad \Rightarrow \quad \mathbf{D}\mathbf{L}_\lambda = -\mathbf{C}_Q \tag{3.46}$$

where $\mathbf{1}_{\boldsymbol{\lambda}}$ is a rectangular matrix containing Lagrangian multipliers. $\mathbf{C}_Q$ is a vector containing atomic electronegativities and total charge constraints, whereas $\mathbf{L}_\lambda$ is a vector containing charges, dipoles and Lagrangian multipliers.

FQFμ can be effectively coupled to a QM SCF wavefunction in a QM/MM framework.

The global QM/MM energy functional for a SCF-like description of the QM portion reads: [30]

$$\begin{aligned}\mathcal{E}(\mathbf{P},\mathbf{q},\boldsymbol{\mu},\boldsymbol{\lambda}) = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{1}{2}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q} + \\ + \mathbf{q}^\dagger\mathbf{V}(\mathbf{P}) - \boldsymbol{\mu}^\dagger\mathbf{E}(\mathbf{P})\end{aligned} \tag{3.47}$$

where $\mathbf{h}$ and $\mathbf{G}$ are the usual one- and two-electron matrices, and $\mathbf{P}$ is the density matrix. $\mathbf{q}^\dagger\mathbf{V}(\mathbf{P})$ and $\boldsymbol{\mu}^\dagger\mathbf{E}(\mathbf{P})$ represent the coupling between charges-QM electric potential and dipoles-QM electric field, respectively. The effective Fock matrix is defined as the derivative of the energy with respect to the density matrix:

$$\tilde{F}_{\mu\nu} = \frac{\partial \mathcal{E}}{\partial P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{V}^{\dagger}_{\mu\nu}\mathbf{q} - \mathbf{E}^{\dagger}_{\mu\nu}\boldsymbol{\mu} \tag{3.48}$$

where the interaction of the electron density with both charges and dipoles is included through the electrostatic coupling terms.

Charges and dipoles are obtained by imposing the global functional to be stationary with respect to charges, dipoles and Lagrangian multipliers.

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_{\boldsymbol{\lambda}} & \mathbf{T}^{q\mu} \\ \mathbf{1}^{\dagger}_{\boldsymbol{\lambda}} & \mathbf{0} & \mathbf{0} \\ -\mathbf{T}^{q\mu^{\dagger}} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ \mathbf{Q}_{\text{tot}} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} -\mathbf{V}(\mathbf{P}) \\ \mathbf{0} \\ \mathbf{E}(\mathbf{P}) \end{pmatrix} \quad \Rightarrow \quad \mathbf{D}\mathbf{L}_{\lambda} = -\mathbf{C}_{Q} - \mathbf{R}(\mathbf{P}) \tag{3.49}$$

Notice that, with respect to Equation 3.46, a new source term, $\mathbf{R}(\mathbf{P})$, due to the coupling of both charges and dipoles with the SCF density, arises. Again, $\mathbf{L}_{\lambda}$ is a vector containing charges, dipoles and Lagrangian multipliers. A similar equation, but involving only a non-QM source of the external electric field, has already been proposed in Ref. [305]

Analytical Energy Derivatives

In this section QM/FQFμ analytical first and second energy derivatives with respect to nuclear coordinates are presented and discussed. The following equations are defined in the so-called Partial Hessian Vibrational Approach (PHVA), [420–422] which has been amply exploited to treat vibrational phenomena of complex systems. [103, 110–112] Within such a framework, it is assumed that the geometrical perturbation only acts on the QM portion of the system, whereas MM atoms are unaffected. Remarkably, the PHVA is fully consistent with a focused approach. For the sake of completeness, however, equations for first and second derivatives of FQFμ MM atoms are given in the Appendix section. The following derivation directly follows what already reported by some of the present authors in case of QM/FQ. [103] This allows to directly identify the additional terms which depend on the presence of fluctuating dipoles in the MM portion.

Energy first derivatives

The energy first derivative of Equation 3.47 with respect to the x nuclear displacement can be expressed by means of the chain rule: [103, 423]

$$\mathcal{E}^{x}(\mathbf{P}, \mathbf{q}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \frac{\partial \mathcal{E}}{\partial x} + \frac{\partial \mathcal{E}}{\partial \mathbf{P}}\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathcal{E}}{\partial \mathbf{q}}\frac{\partial \mathbf{q}}{\partial x} + \frac{\partial \mathcal{E}}{\partial \boldsymbol{\mu}}\frac{\partial \boldsymbol{\mu}}{\partial x} + \frac{\partial \mathcal{E}}{\partial \boldsymbol{\lambda}}\frac{\partial \boldsymbol{\lambda}}{\partial x}$$

The last three terms vanish because of the stationarity conditions. The first term, which is the partial derivative of the energy with respect to the position of a QM nucleus, reads:

$$\frac{\partial \mathcal{E}}{\partial x} = \text{tr } \mathbf{h}^{x}\mathbf{P} + \frac{1}{2}\text{tr } \mathbf{G}^{(x)}(\mathbf{P})\mathbf{P} + \mathbf{q}^{\dagger}\mathbf{V}^{(x)}(\mathbf{P}) - \boldsymbol{\mu}^{\dagger}\mathbf{E}^{(x)}(\mathbf{P}) \tag{3.50}$$

where: [103]

$$
\begin{aligned}
V_i^{(x)}(\mathbf{P}) &= \sum_{\mu\nu} P_{\mu\nu} V^x_{\mu\nu,i} + \text{nuclear contribution} \\
&= \frac{Z_\zeta}{|\mathbf{R}_\zeta - \mathbf{r}_i|^2} - \sum_{\mu\nu} \left\langle \frac{\partial(\chi_\mu \chi_\nu)}{\partial \mathbf{R}_\zeta} \middle| \frac{1}{\mathbf{r} - \mathbf{r}'} \middle| \delta(\mathbf{r}' - \mathbf{r}_i) \right\rangle P_{\mu\nu} \qquad (3.51) \\
E_i^{(x)}(\mathbf{P}) &= \nabla_{r_i} V_i^{(x)}(\mathbf{P}) \qquad (3.52)
\end{aligned}
$$

The term involving the derivatives of the density matrix can be computed starting from the idempotency condition: [202]

$$
-\mathbf{P}\tilde{\mathbf{F}}\mathbf{P}\mathbf{S}^x_{oo} = -\tilde{\mathbf{W}}\mathbf{S}^x_{oo}
$$

where the subscript *oo* denotes the occupied–occupied block of the matrix in the MO basis, and $\mathbf{W}$ is the energy-weighted density matrix contribution. By collecting all the terms:

$$
\mathcal{E}^x(\mathbf{P}, \mathbf{q}, \boldsymbol{\lambda}) = \text{tr } \mathbf{h}^x\mathbf{P} + \frac{1}{2}\text{tr } \mathbf{G}^{(x)}(\mathbf{P})\mathbf{P} + \mathbf{q}^\dagger \mathbf{V}^{(x)}(\mathbf{P}) - \boldsymbol{\mu}^\dagger \mathbf{E}^{(x)}(\mathbf{P}) - \text{tr}\tilde{\mathbf{W}}\mathbf{S}^x_{oo} \quad (3.53)
$$

Notice that the term $(\mathbf{q}^\dagger \mathbf{V}^{(x)}(\mathbf{P}))$ is the same as computed in the QM/FQ approach. [103] Therefore, the inclusion of fluctuating dipoles gives rise to the additional term $\boldsymbol{\mu}^\dagger \mathbf{E}^{(x)}(\mathbf{P})$.

Energy second derivatives

The energy second derivative with respect to nuclear displacements x and y is obtained by differentiating Equation 3.53 and by exploiting once again the chain rule:

$$
\begin{aligned}
\mathcal{E}^{xy} &= \sum_{\mu\nu} \left[h^{xy}_{\mu\nu} + \frac{1}{2} G^{(xy)}_{\mu\nu}(\mathbf{P}) + \mathbf{q}^\dagger \mathbf{V}^{xy}_{\mu\nu} - \boldsymbol{\mu}^\dagger \mathbf{E}^{xy}_{\mu\nu} \right] P_{\mu\nu} - \text{tr } \mathbf{W}\mathbf{S}^{xy} - \text{tr } \mathbf{W}^y \mathbf{S}^x \\
&+ \sum_{\mu\nu} \left[h^x_{\mu\nu} + G^{(x)}_{\mu\nu}(\mathbf{P}) + \mathbf{q}^\dagger \mathbf{V}^x_{\mu\nu} - \boldsymbol{\mu}^\dagger \mathbf{E}^x_{\mu\nu} \right] P^y_{\mu\nu} + \sum_{\mu\nu} \mathbf{L}^{y\dagger} \mathbf{R}^x_{\mu\nu} P_{\mu\nu} \qquad (3.54)
\end{aligned}
$$

Thus, the derivatives of the off-diagonal blocks of the density matrix and charges/dipoles need to be calculated. Density matrix derivatives can be obtained through a Coupled Perturbed Hartree–Fock (CPHF) or Kohn-Sham (CPKS) procedure. [202]

FQFμ charge and dipole derivatives can be calculated by differentiating Equation 3.49:

$$
\mathbf{D}\mathbf{L}^x = -\mathbf{R}^{(x)}(\mathbf{P}) - \mathbf{R}(\mathbf{P}^x) \qquad (3.55)
$$

The Fock matrix derivative is defined as:

$$
\widetilde{F}^x_{\mu\nu} = \widetilde{F}^{(x)}_{\mu\nu} + G_{\mu\nu}(\mathbf{P}^x) - \mathbf{R}^\dagger_{\mu\nu} \mathbf{D}^{-1} \mathbf{R}(\mathbf{P}^x) \qquad (3.56)
$$

where $\widetilde{F}^{(x)}_{\mu\nu}$, which collects all explicit derivatives of the Fock matrix, reads:

$$
\widetilde{F}^{(x)}_{\mu\nu} = h^x_{\mu\nu} + G^{(x)}_{\mu\nu}(\mathbf{P}) + \mathbf{L}^\dagger \mathbf{R}^x_{\mu\nu} + \mathbf{R}^\dagger_{\mu\nu} \mathbf{L}^{(x)}
$$

By rearranging the terms, the CPHF/CPKS equations are obtained (MO basis):

$$\epsilon_i P^x_{ia} - \epsilon_a P^x_{ia} = -\tilde{Q}_{ia} \quad + \quad \sum_{jb} \left[\langle aj||ib\rangle - \mathbf{R}^\dagger_{ia}\mathbf{D}^{-1}\mathbf{R}_{jb} \right] P^x_{jb} + \quad \sum_{jb} \left[\langle ab||ij\rangle - \mathbf{R}^\dagger_{ia}\mathbf{D}^{-1}\mathbf{R}_{bj} \right] P^x_{bj} \tag{3.57}$$

By taking the adjunct equation and introducing the following matrices (we assume orbitals to be real):

$$\tilde{A}_{ia,jb} = (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab} + \langle aj||ib\rangle - \mathbf{R}^\dagger_{ia}\mathbf{D}^{-1}\mathbf{R}_{jb} \tag{3.58}$$

$$\tilde{B}_{ia,jb} = \langle ab||ij\rangle - \mathbf{R}^\dagger_{ia}\mathbf{D}^{-1}\mathbf{R}_{bj} \tag{3.59}$$

the following equation is obtained:

$$\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^* & \tilde{\mathbf{A}}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{Q} \\ \mathbf{Q}^* \end{pmatrix} \tag{3.60}$$

where:

$$\tilde{A}_{ia,jb} = (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab} + \langle aj||ib\rangle - \mathbf{R}^\dagger_{ia}\mathbf{D}^{-1}\mathbf{R}_{jb} \tag{3.61}$$

$$\tilde{B}_{ia,jb} = \langle ab||ij\rangle - \mathbf{R}^\dagger_{ia}\mathbf{D}^{-1}\mathbf{R}_{bj} \tag{3.62}$$

$$\tilde{Q}_{ia} = \tilde{F}^{(x)}_{ia} - G_{ia}(S^x_{oo}) - \tilde{\mathbf{F}}S^x_{ia} + \mathbf{R}^\dagger_{ia}\mathbf{D}^{-1}\mathbf{R}(\mathbf{S}^x_{oo}) \tag{3.63}$$

Therefore, the derivatives of the density matrix and FQFμ charge/dipole derivatives with respect to QM region nuclear positions are obtained by solving Equations 3.60 and 3.55, respectively. Notice that such a derivation is coherent with what has been reported for linear response in the zero-frequency limit. [412]

To summarize, FQFμ contributions to analytical second derivatives can be grouped into three categories:

1. explicit terms:
$$\mathbf{q}^\dagger\mathbf{V}^{(xy)} - \boldsymbol{\mu}^\dagger\mathbf{E}^{(xy)} + \mathbf{L}^{(x)\dagger}\mathbf{R}^{(y)}$$

2. contributions to Fock matrix derivatives:
$$\mathbf{L}^\dagger\mathbf{R}^x_{\mu\nu} + \mathbf{L}^{(x)\dagger}\mathbf{R}_{\mu\nu}$$

3. additional terms to the CPHF/CPKS matrix:
$$-\mathbf{R}^\dagger_{ia}\mathbf{D}^{-1}\mathbf{R}_{jb}$$

Notice that only the last term is required in case of electric perturbations.Also, similarly to what has already been discussed for energy first derivatives, QM/FQFμ second derivatives differ from QM/FQ ones because additional terms depending on fluctuating dipoles need to be included. [103]

Computational Details

The computational protocol exploited for calculating QM/MM IR spectra involves the following steps: [112]

1. *Definition of the system:* the model system is composed by the target molecule (solute) surrounded by a sufficiently large number of solvent molecules, chosen so that both the dynamics and the subsequent QM/FQ calculations can capture all the relevant solute-solvent interactions.

2. *Classical MD simulations and sampling:* this step is required to sample the phase space of the system. Simulations are run long enough to sample a sufficiently large portion of the phase-space and such that the simulation parameters correctly reproduce all possible system configurations and their relative energy (and thus population). From the MD simulations a number of uncorrelated snapshots are extracted to be used later in the QM/FQ calculations.

3. *Definition of the different regions of the two-layer scheme and their boundaries:* for each snapshot extracted from the dynamics, a sphere centered on the solute is cut, retaining all solvent molecules within the sphere.

4. *Running the QM/FQ calculations on the snapshots:* for each of the spherical snapshots (droplets), IR spectra are calculated, after the geometry of the QM solute is optimized in each snapshot, by keeping fixed the positions/geometries of all the solvent molecules.

5. *Extraction of the average spectra and analysis of the results:* the spectra obtained for each snapshot are extracted and the final IR spectra for the system are obtained as the arithmetic mean of the spectra for all the snapshots.

R-methyloxirane (MOXY), (S)-Glycidol (GLY) and Gallic Acid (GA) geometries were optimized at the B3LYP/ aug-cc-pVDZ level. Solvent effects on solutes' geometries were included through the Polarizable Continuum Model (PCM). MOXY and GLY Molecular Dynamics (MD) simulations were carried out as detailed in previous works by some of the present authors [106, 112] by using GROMACS. [141]

A 25 ns MD simulation of GA in aqueous solution was performed by using a similar computational protocol. A GA molecule was placed at the center of a cubic box and solvated with 6025 TIP3P water molecules. [122] The parameters used to describe the GA inter-/intra-molecular interactions were taken from the GAFF force field [424] by using the ANTECHAMBER package. [425] GA molecule was kept fixed during all the steps of the simulation. To model intermolecular solute-solvent interactions, GA partial charges were computed by relying on the Hirshfeld population analysis, [426] as previously done by some of the present authors. [113, 260] Partial charges were computed at the B3LYP/6-311++G** level of theory by including solvent effects by means of PCM. Two subsequent 100 ps runs were performed for equilibration purposes by using NVT and NPT ensembles, respectively. A 25 ns NVT MD simulation was then performed to sample the configurational space at time steps of 1 fs and by saving coordinates every 10 ps. The system was simulated by using three-dimensional periodic boundary conditions; non-bonded interactions cutoff was set to 10 Å. A particle mesh Ewald (PME) correction for the long-range

electrostatics was applied and the temperature was maintained at 300 K by using the velocity rescale algorithm.

A total of 200 uncorrelated snapshots were extracted from the last 20 ns, 50 ns and 25 ns of MD simulations in case of MOXY, GLY and GA, respectively. For each snapshot, a sphere centered at the solute's geometric center was cut.A cutting radius of 12 Å was used for MOXY and GLY, whereas a cutting radius of 15 Å was used for GA.

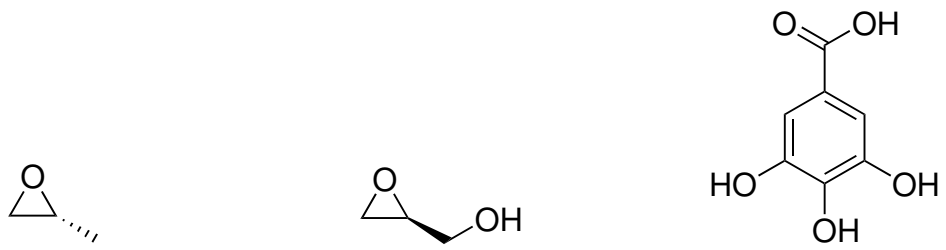
QM/FQ and QM/FQFμ partial geometry optimization of the solute moiety on each snapshot was performed according to the default settings of Gaussian16, [146] by keeping all water molecules fixed. Finally, infrared (IR) spectra were calculated on each partially optimized snapshot with the QM/FQ and QM/FQFμ models at the B3LYP/aug-cc-pVDZ level (MOXY and GLY) and B3LYP/6-311++G** level (GA), the latter in agreement with Ref. [187] SPC parameters for water were used for FQ calculations. [27] The set of parameters recently developed by some of us for QM/FQFμ calculations in aqueous solutions were exploited. [30]. IR data were averaged to obtain final spectra for the three solutes; peaks were convoluted with a Lorentzian lineshape, with Full Width at Half Maximum (FWHM) of 4 cm^{-1}. For the sake of comparison, QM/PCM IR spectra were also computed.

All QM/FQ and QM/FQFμ calculations were performed by using a locally modified version of the Gaussian 16 package, [146] where QM/FQFμ analytical energy first derivatives were implemented.

Numerical Results

In this section, the results obtained by applying QM/FQFμ to the calculation of IR spectra of MOXY, GLY and GA in aqueous solutions are reported (see Figure 3.19, panels a-c for molecular structures). MOXY is a widely exploited test system for computational models. [427–433] GLY, which bears an additional hydroxil group, has been previously studied both theoretically and experimentally. [112, 418] In particular, it has been shown that eight different GLY conformers exist in aqueous solution, thus its theoretical modeling is challenging. [112, 418, 434] GA is an organic acid characterized by the presence of three hydroxil groups linked to the aromatic ring. The modeling of IR spectra of GA in aqueous solution is also challenging, because it has previously been reported by one of the present authors that the implicit PCM fails at reproducing the experimental IR spectrum, and that experimental spectral features can be recovered by adopting a supramolecular approach which includes eight explicit water molecules in the QM portion. [187] Therefore, it appears to be an ideal test-bed for QM/FQFμ.

QM/FQFμ spectra will be compared to QM/FQ, which only considers fluctuating charges on MM atoms. The reason for such a comparison is twofold: (i) QM/FQFμ is formally an extension of QM/FQ, and (ii) QM/FQ has been successfully applied to vibrational spectra of molecular systems in aqueous solution. [110–112] Thus, the comparison between the two approaches can directly quantify the relevance of fluctuating dipoles in the description of vibrational spectra and report on the performance of the novel QM/FQFμ method.



(a) MOXY (b) GLY (c) GA

Figure 3.19: MOXY (**a**), GLY (**b**) and GA (**c**) molecular structures.

Methyloxirane in Aqueous Solution

Figure 3.20 compares QM/FQ (top) and QM/FQFμ (bottom) stick and convoluted spectra for 200 snapshots extracted from the MD simulation. Such a number of snapshots is enough to reach convergence. [111, 112] Stick spectra are obtained by plotting raw data extracted from each frequency calculation. Figure 3.20 clearly shows that both QM/FQ and QM/FQFμ IR wavenumbers and dipole strengths depend on the snapshot, i.e. on the arrangement of water molecules around the solute. As compared to QM/FQ, QM/FQFμ exhibits a larger spread, in particular in the vibrational wavenumbers. The largest variability of QM/FQFμ sticks occurs in the region between 1450 and 1500 cm^{-1}, which is associated to methyl and CH bending modes (see Figures S1 in the Supporting Information (SI) for a pictorial view of the normal modes for a randomly chosen snapshot [435]).

Clearly, band broadening is automatically considered in both QM/FQ and QM/FQFμ approaches coupled with the dynamical description given by the MD simulation, which samples over the solute-solvent phase-space. Therefore, solvent inhomogeneous broadening (due to the fluctuations of the solvent molecules) does not need to be artificially considered by imposing a pre-defined (and arbitrary) band-width, which is instead necessary when other static approaches, such as PCM, are used. The lorentzian convolution obtained by using a FWHM of 4 cm^{-1} is plotted for both QM/FQ and QM/FQFμ spectra. Notice that QM/FQ IR spectrum is identical to what we reported in case of the three layer QM/FQ/PCM approach, where the PCM is used as third layer and coupled to both QM and FQ portions. Such a similarity means that water molecules explicitly included in snapshots are able to account also for water bulk effects.

We now move to the comparison between computed and experimental spectra (see Figure 3.21). Notice that the experimental spectrum was measured for neat liquid MOXY rather than aqueous solution; [419] therefore, some discrepancies with our computed results need to be expected.

The computed and experimental IR spectra are dominated by an intense band at about 850 cm^{-1}. This signal is given by the symmetric stretching of the C-O bond of the epoxyl group. QM/FQFμ band is blueshifted with respect to QM/FQ, thus the inclusion of fluctuating dipoles increases solute-solvent interactions. In fact, an increase in the solute-solvent interactions is reflected in a decrease in the solute intramolecular bond strengths, therefore, resulting in a blueshift. Similar considerations apply also to the regions between 900-1100 cm^{-1}, where the normal modes involve vibrations of the MOXY oxygen atom.

In the other regions of the spectra, such a blueshift is not recorded. This can be

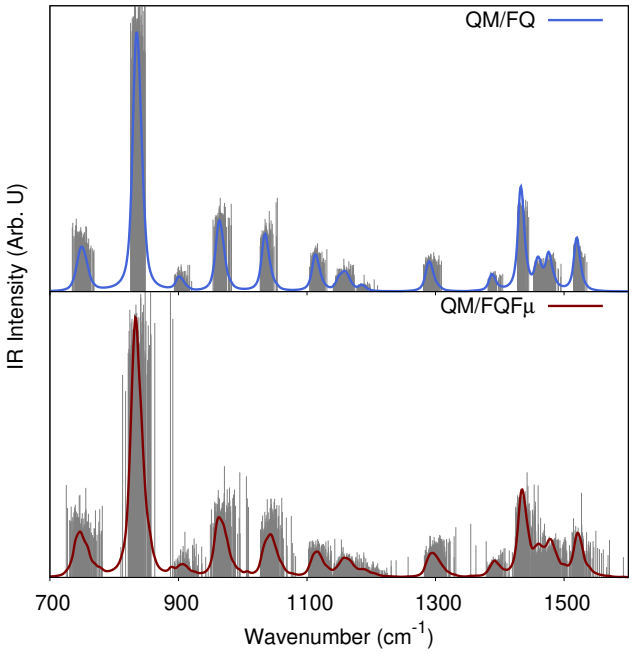


Figure 3.20: Calculated QM/FQ (top) and QM/FQFμ (bottom) stick and convoluted IR spectra of MOXY in aqueous solution. FWHM = 4 cm^{-1}.

explained by the fact that the normal modes do not involve vibrations of the oxygen atom, which is the only MOXY atom potentially exhibiting an Hydrogen Bond with the solvent molecules.

Further differences between QM/FQ and QM/FQFμ are computed for the inhomogeneous band broadening, which is almost absent in case of QM/FQ, whereas it affects almost all QM/FQFμ bands. This is not unexpected if the raw data depicted in Figure 3.20 are considered. In fact, as already pointed out, QM/FQFμ generally spreads a larger energy region for each vibrational normal mode.

The major discrepancies between QM/FQ and FQ/FQFmu and the experiment are reported in case of the inhomogenous band broadening, which is almost absent in the experiment. Thus, the main feature added by aqueous solution seems to be a larger broadening of vibrational bands.

Glycidol in Aqueous Solution

Similarly to MOXY, QM/FQ and QM/FQFμ IR spectra of GLY in aqueous solution were calculated by averaging over 200 snapshots extracted from the MD simulation. [112] QM/FQ and QM/FQFμ raw data are graphically plotted in Figure 3.22, together with their lorentzian convolution. As stated before, the case of GLY in aqueous solution is far more complicated than MOXY, because GLY is a flexible molecule, i.e. it exists in different conformations. This is reflected in the stick spectra depicted in Figure 3.22, which show a larger variability both in intensities and wavenumbers as compared to MOXY (see Figure 3.20); this applies to both QM/FQ and QM/FQFμ calculations with the exception of the region around 1110 cm^{-1}. There, QM/FQ shows a substantial variation in intensity, whereas QM/FQFμ spreads a larger wavenumber range. The spreading of the intensities is instead larger for QM/FQFμ in the region between 1400-1600 cm^{-1}. Despite such discrepancies

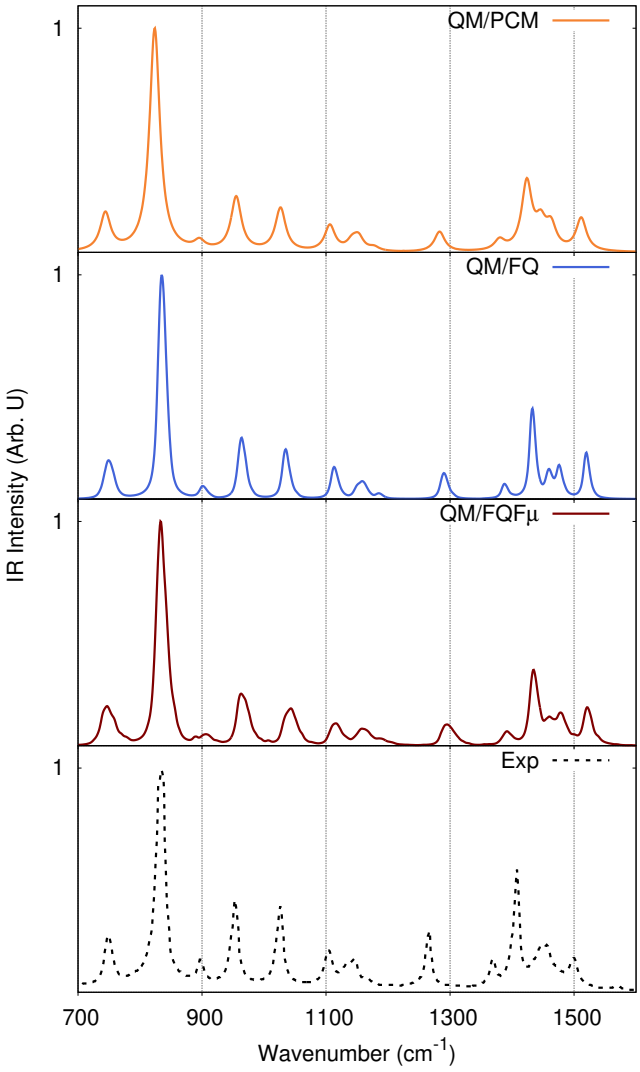


Figure 3.21: Calculated QM/PCM (orange), QM/FQ (blue) and QM/FQFμ (maroon) IR spectra of MOXY in aqueous solution (QM/PCM FWHM = 10 cm^{-1}; QM/MM FWHM = 4cm^{-1}). The experimental spectrum (dashed black) is reproduced from Ref. [419]

between QM/FQ and QM/FQFμ, inhomogeneous broadening is described by both approaches.

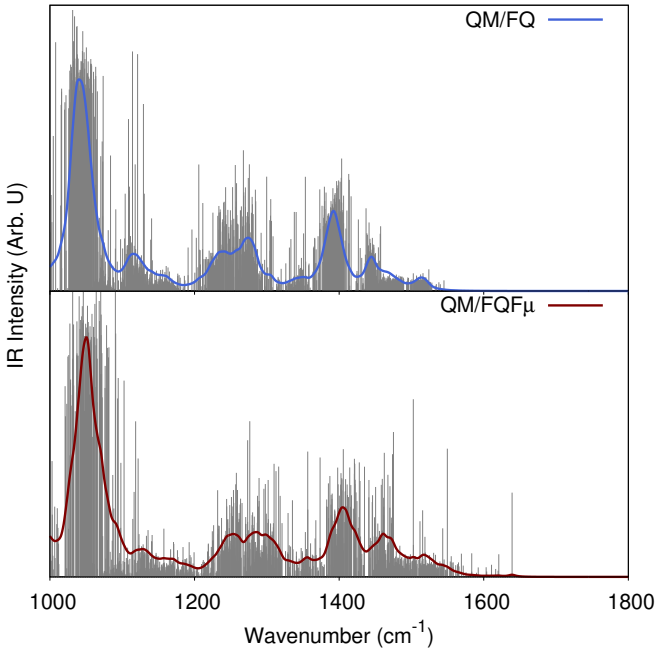


Figure 3.22: Calculated QM/FQ (top) and QM/FQFμ (bottom) stick and convoluted IR spectra of GLY in aqueous solution (FWHM = 4 cm^{-1}).

QM/PCM, QM/FQ and QM/FQFμ convoluted IR spectra are compared to experiments in Figure 3.23. The experimental IR spectrum is reproduced from Ref. [418] Normal modes for a randomly chosen snapshot extracted from the MD are depicted in Figure S2 in the SI for the region 700-1800 cm^{-1}. [435]

Computed and experimental IR spectra are dominated by an intense band at about 1050 cm^{-1}, assigned to a diffuse stretching/bending normal mode, involving the hydroxyl group.

QM/FQFμ spectrum is generally blueshifted with respect to both QM/FQ and QM/PCM, probably due to the fact it predicts larger solute-solvent interactions.

Moving to the comparison with experimental data, both the experimental and computed spectra are characterized by a two-peak-shaped band between 1200 and 1300 cm^{-1}, which is assigned to the C-OH and C-CH bending modes (at about 1230 and 1270 cm^{-1}, respectively). Furthermore, above 1400 cm^{-1} a three-peak-shaped band can be identified, due to the C-OH bending (1395 cm^{-1}), a diffuse C-CH bending (1440 cm^{-1}) and a CH_2 bending (1465 cm^{-1}).

QM/FQ, QM/FQFμ and the experimental IR spectra are nicely in agreement. In fact, most of relative intensities and the band broadening are correctly reproduced. In particular, QM/FQ accurately predicts the two-peak band between 1200 and 1300 cm^{-1}, whereas QM/FQFμ is able to catch the inhomogeneity of the three-peak-shaped band between 1400 and 1500 cm^{-1}. Some discrepancies are reported in case of the peak at about 1100 cm^{-1} (due to the CH scissoring), which is predicted to have a very low intensity by both atomistic QM/MM approaches, whereas it is the second most intense peak in the experimental spectrum. Notice that it has been reported that the relative intensity of this peak can be correctly reproduced

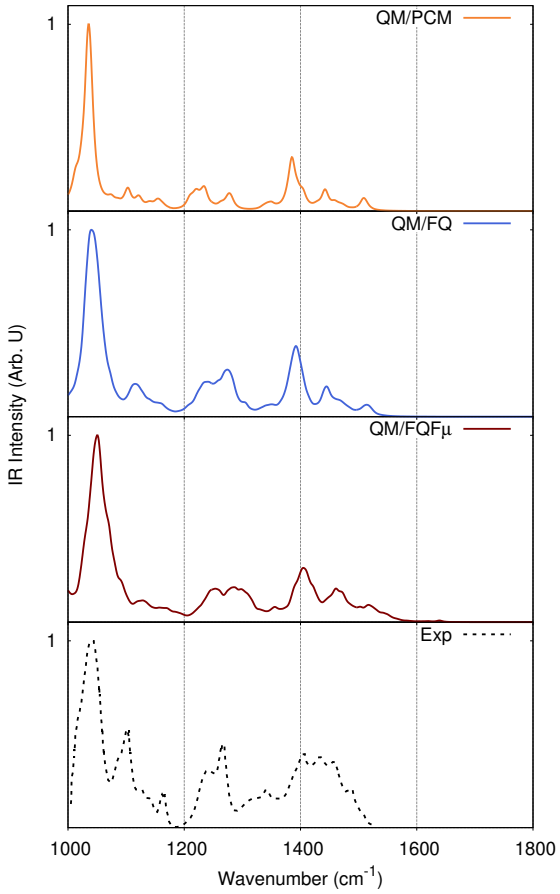


Figure 3.23: Calculated QM/PCM (orange), QM/FQ (blue) and QM/FQFμ (maroon) IR spectra of GLY in aqueous solution (QM/PCM FWHM = 10 cm^{-1}; QM/MM FWHM = 4cm^{-1}). The experimental spectrum (dashed black line) is reproduced from Ref. [418]

if a supermolecule approach is adopted, i.e. if water molecules are included in the QM region. [418] These findings, together with the results obtained by adopting our QM/classical modeling, suggest that the inclusion of non-electrostatic interactions, which are considered in the full QM supermolecule approach, can play a relevant role to improve the quality of the computed spectrum in this region.

Overall, it is worth noticing that the continuum QM/PCM approach cannot reproduce the experimental spectrum (see top of Figure 3.23), thus remarking once again the huge potentialities of our atomistic QM/FQ and QM/FQFμ approaches to model vibrational spectra of solutes strongly interacting with the aqueous environment.

This is particularly evident by analyzing integrated intensities, which have been extracted from the computed and experimental spectra by performing a spectral deconvolution in terms of lorentzian functions (see Figure S3 and Tabs. S2-S3 in the SI [435]). In particular, both QM/FQ and QM/FQFμ overperform QM/PCM, thus confirming the qualitative findings based on relative intensities.

To end the discussion on the IR spectrum of GLY in aqueous solution, we point out that the broad band measured between 1600-1700 cm^{-1} in the experiment is not reproduced by any of the selected QM/classical approaches. As already reported by some of the present authors [110, 112] and in Refs. [418, 436, 437], such band is due to the OH bending mode of water molecules linked to GLY; therefore, it cannot be modeled by our approaches, in which the normal modes/frequencies of the environment are not computed.

Gallic Acid in Aqueous Solution

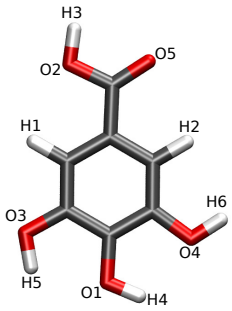


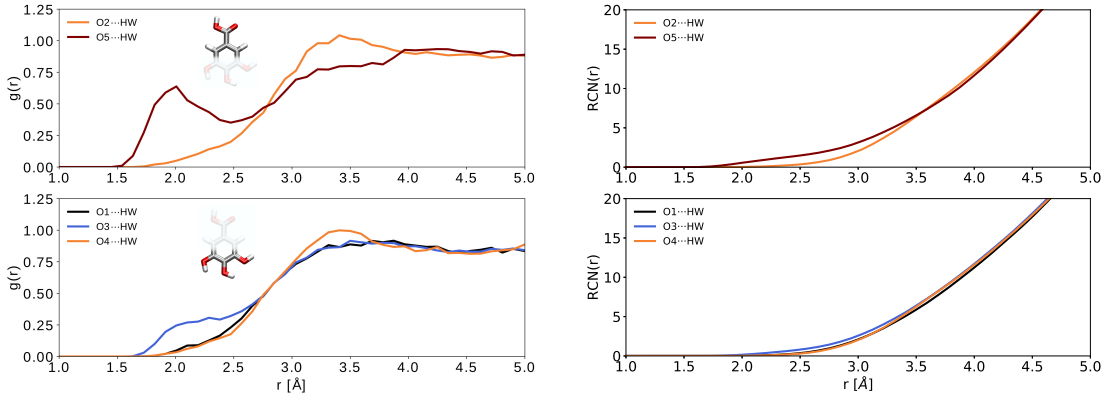
Figure 3.24: Gallic Acid (GA) molecular structure with atom labelling.

MD Analysis Before discussing GA IR spectra (see Figure 3.24 for atom labelling), in this section the MD trajectory is examined in terms of both radial distribution functions (rdfs) and running coordination numbers (RCNs). In particular, in order to obtain a description of the solvent local structure and to analyze hydrogen bonding patterns between GA and water molecules, intermolecular H(GA)$\cdots$OW and O(GA)$\cdots$HW rdfs and the corresponding running coordination numbers were calculated and are reported in Figures 3.26 and 3.25 respectively. The coordination number of a specific site is defined by combining the distance at which the first/second minimum of rdf occurs with the corresponding distance in the running coordination number.

The oxygen atoms of the three hydroxyl groups, i.e. O1, O3 and O4 (see Figure 3.24), present a radial distribution function with a similar shape, consisting

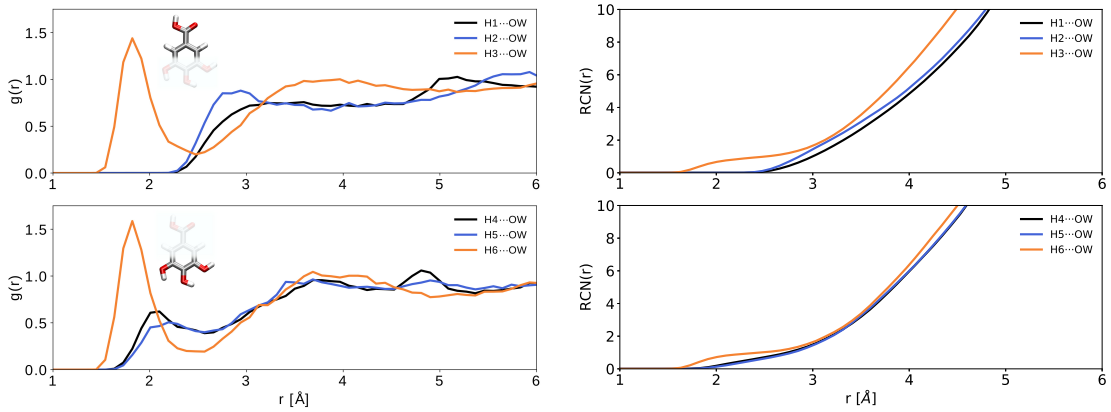
of a broad peak at about 3.4 Å, which is associated to the second solvation shell. The coordination number of these sites is equal to 16.7, 12.8 and 12.2, respectively. Remarkably, O3 rdf shows a peak at about 2.0 Å, due to the fact that this is the only one among the hydroxyl groups that can form intermolecular HB with water molecules. O1 and O4 are involved in an intramolecular HB. Moving to the acid group, O5 rdf presents a sharp peak related to the first solvation shell, with a maximum at 2.0 Å and a coordination number corresponding to 1.6; the plot for O2 is very similar to what has already been discussed in case of O1 and O4.

The specific arrangement of the hydroxyl groups in the meta and para positions (with respect to the acid) of GA induces a weak but detectable anisotropy on the rdfs of both hydrogen and oxygen atoms. In fact, both H4 and H5 rdfs present a similar shape, with a first peak at about 2.1 and 2.2 Å, respectively. They correspond to coordination numbers of 0.8 and 0.7, respectively. A second peak related to the second solvation shell is placed at 3.68 Å , which corresponds to coordination



(a) (b)

Figure 3.25: GA radial distribution functions between GA oxygen atoms and water hydrogen atoms (HW). See Figure 3.24 for atom labeling.



(a) (b)

Figure 3.26: GA radial distribution functions between GA hydrogen atoms and water oxygen atoms (OW). See Figure 3.24 for atom labeling.

numbers of 7.5 and 6.3, respectively. The H3 and H6, instead, show a pronounced first narrow peak at about the same distance (1.73 Å), with a coordination number of 1.0 and a second peak around 3.97 Å with coordination numbers of 15.2 and 13.8, respectively. Again, the differences between H3 and H4/H5 are due to the fact that H4 and H5 are involved in the intramolecular HB, whereas H6 is free to form intermolecular HB with water molecules. Remarkably, the results here discussed are similar to the findings previously reported by one of the present authors. [187]

IR spectrum of GA in Aqueous Solution QM/FQ and QM/FQFμ IR spectra of GA were obtained by averaging over 200 snapshots extracted from the MD run. Similarly to MOXY and GLY, we checked that such a number of snapshots produce converged spectra. The raw data extracted from the calculations, i.e. stick spectra are reported in Figure 3.27 for the region 1000-1800 cm^{-1} (i.e. the region of interest for the experimental investigation, vide infra); lorentzian convolution (FWHM = 4 cm^{-1}) is also depicted. Both QM/FQ and QM/FQFμ stick spectra show a large spreading in intensities and frequencies. Such a feature is reported for most computed bands, in particular in the region between 1100-1400 cm^{-1}, in which single bands are not easily detectable (see for comparison MOXY and GLY raw spectra in Figures 3.20 and 3.22).

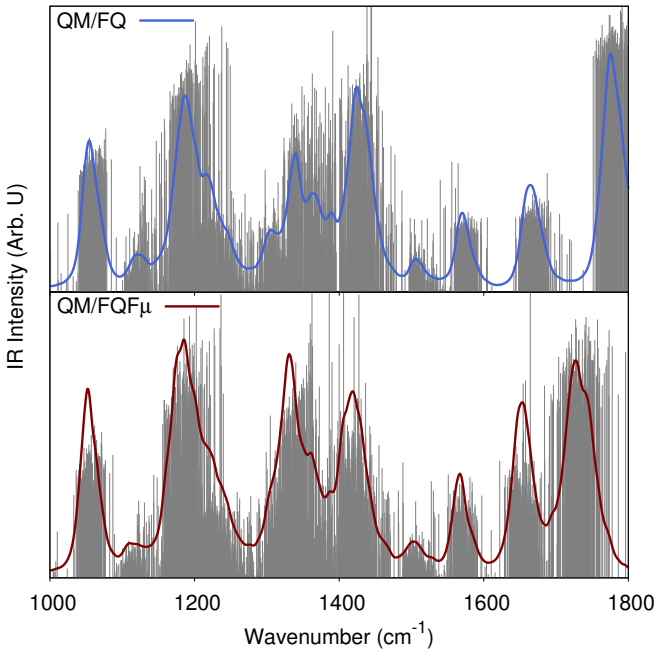


Figure 3.27: Calculated QM/FQ (top) and QM/FQFμ (bottom) stick and convoluted IR spectra of GA in aqueous solution (FWHM = 4cm^{-1}).

The comparison between QM/PCM, QM/FQ, QM/FQFμ and the experimental IR spectrum [438] shows that the latter is dominated by a three-band broad structure between 1200 and 1500 cm^{-1}, which probably involves more than one normal mode. A well-separated peak is present at 1000 cm^{-1} and it is associated to the C-OH bending (see Figure S3 in the SI for a graphical depiction of the normal modes [435]). Moreover, three small bands of the same intensity are reported in the region 1500-1800 cm^{-1}, which are mainly due to composite C-OH bending modes of the hydroxyl

groups and the acidic C=O stretching.

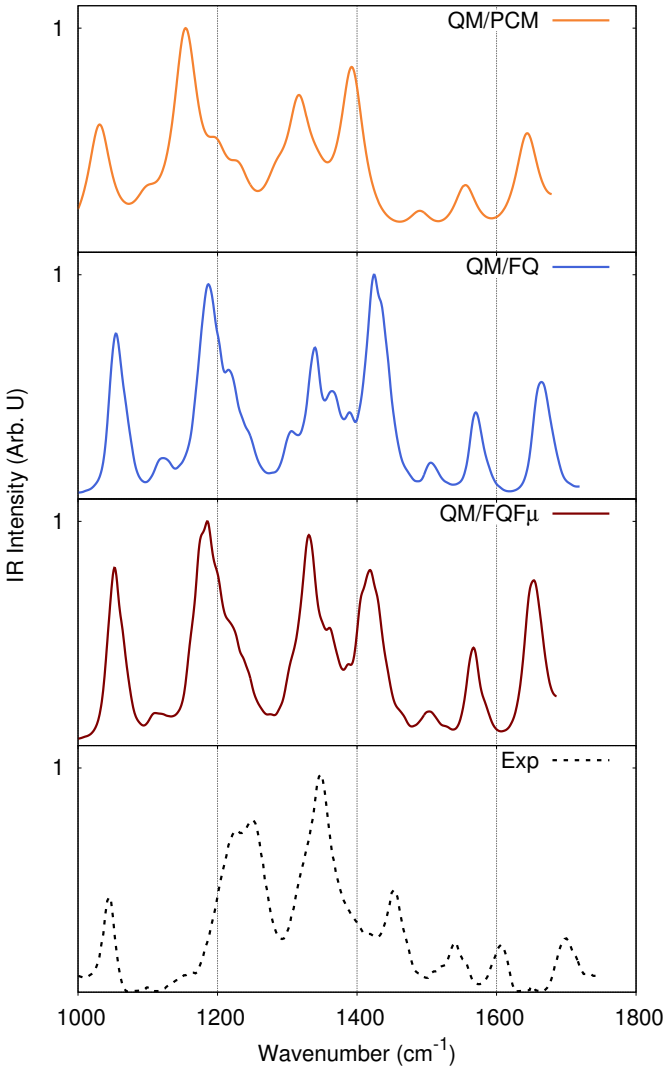


Figure 3.28: Calculated QM/PCM (orange), QM/FQ (blue) and QM/FQFμ (maroon) IR spectra of GA in aqueous solution (QM/PCM FWHM = 10 cm^{-1}; QM/MM FWHM = 4cm^{-1}). The experimental spectrum (dashed black line) is reproduced from Ref. [187]

The QM/PCM spectrum is dominated by two peaks, placed at about 1170 cm^{-1} and 1750 cm^{-1}, respectively. Such peaks are related to a composite C-OH bending and to the C=O stretching. A similar spectrum is predicted by adopting the atomistic QM/FQ approach, in which the most intense peak is predicted in the case of the C=O stretching at about 1750 cm^{-1}, whereas the peaks at about 1200 and 1400 cm^{-1} have almost the same intensity. It is worth noticing that in QM/FQ spectra most bands present an inhomogenous broadening that is related once again to the dynamical picture given by the sampling of the phase-space through MD. In addition, similarly to MOXY and GLY, most of the computed QM/FQ bands are blueshifted with respect to their QM/PCM counterparts, thus reflecting the stronger solute-solvent interaction.

Most QM/FQFμ bands are blueshifted with respect to QM/PCM, whereas they

are redshifted with respect to QM/FQ values, thus highlighting the different electrostatic description given by the two explicit approaches. Remarkably, similarly to GLY in aqueous solution, vibrational frequencies are not completely in agreement with the experimental ones, that probably due to the lack of anharmonicity effects and the use of DFT. Moreover, inhomogenous band broadening is correclty repoduced by both atomistic approaches, thus resulting in a very good agreement with the experiments, in particular for the experimentally most intense band at about 1350 cm^{-1}.

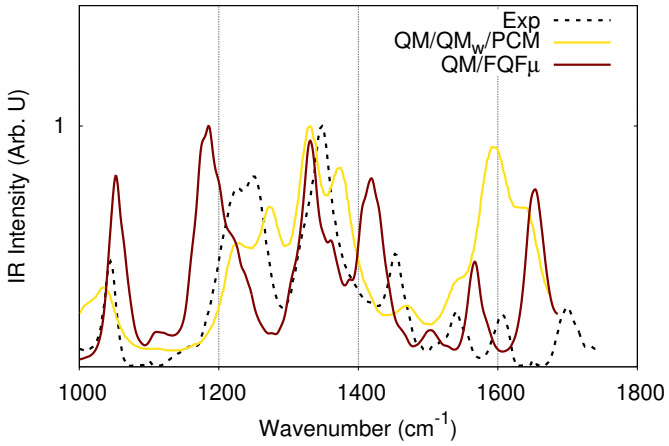


Figure 3.29: Calculated QM/FQFμ (maroon) and QM/QM$_w$/PCM (yellow) reproduced form Ref. [187] The experimental spectrum (dashed black line) is reproduced from Ref. [187]

By further deepening the analysis of computed spectra, we see that QM/FQFμ IR spectrum is dominated by three bands at about 1200, 1350 and 1700 cm^{-1}, which are predicted almost with the same intensity. This is a big improvement with respect to both QM/PCM and QM/FQ approaches, because the most intense band in the experimental spectrum is correctly predicted only by QM/FQFμ. It is worth noticing that a correct reproduction of the intensity of this peak was achieved by some of the present authors by resorting to a supermolecule approach (here called QM/QM$_w$/PCM), i.e. by including 8 QM water molecules in the definition of the QM solute in QM/PCM calculations (see Figure 3.29). Such an observation indeed indicates that an explicit solvation approach is needed to recover the experimental features in this region. In addition, due to the fact that QM/FQFμ appropriately reproduces the most intense band of the experimental spectrum, in a similar way as the supermolecule approach, we can conclude that the electrostatic description of the HB interaction is the most relevant solute-solvent contribution, and that the electrostatic description given by QM/FQFμ in this case overcomes that modeled by the QM/FQ.

The good reproduction of the experimental spectrum obtained by both QM/QM$_w$/PCM and QM/FQFμ is also confirmed by the analysis of integrated intensities, which are obtained by performing a spectral deconvolution in terms of lorentzian functions (see Figure S5 and Tabs. S3-S4 in the SI [435]). The largest discrepancies between QM/FQFμ and QM/QM$_w$/PCM are reported for normal mode frequencies, which are better reproduced by the supramolecule approach, in particular in the region between 1200 and 1400 cm^{-1}. Such an improvement can be related to the fact that

the supermolecule accounts for non-electrostatic interactions (in particular Pauli repulsion) which can therefore play a relevant role in the determination of vibrational frequencies.

To conclude the discussion on GA, it is worth noticing that all the considered computational approaches predict very large intensities for the three modes in the 1500-1800 cm^{-1} region, even the supramolecule QM/QM_w/PCM (see also integrated intensities in Tab. S4 in the SI [435]). This is probably due to both the lack of vibrational anharmonicity, which has been reported to affect not only frequencies but also intensities. [75, 406, 407]

Chapter 4

Non-Electrostatic Effects

Quantum Confinement Effects on Solvatochromic Shifts of Molecular Solutes

Solvatochromism, i.e. the shift in absorption energies which is evidenced by measuring absorption spectra of a given system as vapor phase or dissolved in a polar solvent, is an amply studied phenomenon. [234, 439] Its rationalization in terms of molecular structural patterns, electronic structure and solute-solvent interactions has benefited much in recent years by the interplay between experiments and computational approaches. [19, 374, 375, 440–442] Due to the nature of solvatochromism, reliable computational methods need to accurately describe the solute's geometry, but especially its electronic structure both in the ground and excited state(s), which generally means that electron correlation needs to be accounted for. However, the interactions between the solute and the surrounding solvent need to be modeled at the same level of accuracy, so to reproduce all interaction forces between the two players in a physically consistent way.

An iconic case is the vacuo-to-water solvatochromism exhibited by the dark $n \rightarrow \pi^*$ and bright $\pi \rightarrow \pi^*$ transitions of acrolein, which are shifted to opposite directions due to the presence of the surrounding water molecules. [280, 374, 443–447] Different theoretical solvation approaches, either based on purely continuum descriptions [2] or atomistic Quantum Mechanics/Molecular Mechanics (QM/MM) [72, 74] coupled to Molecular Dynamics (MD) simulations, have been challenged to reproduce the experimental findings. [280, 374, 444–453] While continuum approaches fail, explicit QM/MM methods based either on electrostatic or polarizable schemes, yield computed results in good agreement with respect to experimental values only when explicit water molecules (either 12 or 25) [444, 447] are included in the QM moiety. These findings demonstrate that a proper account of electrostatic interactions is insufficient to reproduce the experimental behavior. In addition, due to the ability of atomistic QM/MM approaches to account for specific (hydrogen bonding) solvent effects, the need to enlarge the QM portion so to include several water molecules clearly shows that solute-solvent non-electrostatic interactions, which are neglected in standard continuum or QM/MM approaches play a crucial role on this property.

In this letter, we demonstrate the pivotal role of QM density confinement effects on solvatochromic shifts. In particular, by resorting to a QM/MM approach able to account for confinement effects we successfully reproduce vacuo-to-water solvatochromic shifts for $n \rightarrow \pi^*$ and bright $\pi \rightarrow \pi^*$ transitions of acrolein and $n \rightarrow \pi^*$ transitions of pyridine and pyrimidine without the need of including explicit water molecules in the QM portion. It is worth remarking that the selected systems present $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ electronic transitions, which are of high chemical interest because among the most common in a plethora of organic molecules. Remarkably, our approach reduces the computational effort required to obtain reliable results and is also capable of dissecting the effects of the single forces acting on the solute-solvent couple, so to allow for a rationalization of the experimental findings in terms of physico-chemical quantities.

The interaction energy $E^{\text{int}}_{\text{QM/MM}}$ between the QM and MM layers within a QM/MM framework can be written as:

$$E^{\text{int}}_{\text{QM/MM}} = E^{\text{ele}}_{\text{QM/MM}} + E^{\text{pol}}_{\text{QM/MM}} + E^{\text{rep}}_{\text{QM/MM}} \tag{4.1}$$

where $E^{\text{ele}}_{\text{QM/MM}}$ and $E^{\text{pol}}_{\text{QM/MM}}$ are the electrostatic and polarization contributions, respectively.

$E^{\text{rep}}_{\text{QM/MM}}$ is Pauli repulsion energy term, which accounts for the QM density confinement. It will be modelled by resorting to the approach which has recently been proposed by some of us, [11] and that is able to reproduce intermolecular interaction energies and hyperfine coupling constants of aqueous solutes, [13] in almost perfect agreement with full-QM approaches and experiments. [11, 13, 30] Our approach formulates $E^{\text{rep}}_{\text{QM/MM}}$ as the opposite of the exchange integral between QM and MM densities: [11, 202, 454]

$$E^{\text{rep}}_{\text{QM/MM}} = \frac{1}{2} \int \frac{\mathrm{d}\mathbf{r_1}\, \mathrm{d}\mathbf{r_2}}{r_{12}} \rho_{\text{QM}}(\mathbf{r_1}, \mathbf{r_2}) \rho_{\text{MM}}(\mathbf{r_2}, \mathbf{r_1}) \tag{4.2}$$

where (ρ_{QM}) is the QM density, and the MM density (ρ_{MM}) is defined in terms of s-type gaussian functions placed at selected points in the MM portion. [11] Coefficients and exponents of such functions have been determined in a previous work of some of us, [11] and successfully coupled to different description of QM/MM coupling. [11, 13, 30] In our approach repulsion affects the molecular density and molecular orbitals, as well as electronic transition energies. Notice that, differently to approximated expressions, [8, 9, 14, 15, 197, 247, 390, 455, 456] the repulsion energy definition in Equation 4.2 is formally exact, [202] although approximations are introduced in the definition of ρ_{MM}. [11] In addition, as it is clear from Equation 4.2, its definition is totally independent of the choice of a specific QM/MM approach, and can thus be coupled to any QM/MM description of electrostatic interactions.

Our modeling of repulsion interaction is coupled in this letter with a hierarchy of QM/MM approaches (see Scheme 4.1) so to finally show that density confinement needs to be included to get a physically consistent picture of solvatochromism. The most basic approach we exploit is the electrostatic QM/MM embedding model (non-polarizable QM/TIP3P [122]), in which each MM solvent atom is endowed with a fixed charge q. To refine this basic picture, we adopt a polarizable QM/MM based on Fluctuating Charges (FQ), [27] in which the charges are not fixed, but can vary as a response to the QM electric potential (as a function of electronegativity, χ, and chemical hardness, η). [29, 104, 113, 260] In addition to FQs, in QM/Fluctuating Charges and Fluctuating Dipoles (FQFμ), each MM atom is also endowed with electric dipoles and both charges/dipoles vary as a response to the QM electric potential and field, respectively (as a function of χ, η and polarizability, α). [30, 412, 457] Both QM/FQ and QM/FQFμ are based on charge equilibration, thus Charge-Transfer between MM solvent molecules can be allowed (QM/FQ$_{\text{CT}}$ and QM/FQFμ_{CT}).

In particular, in this work, we allow intermolecular MM CT between water molecules that are placed at a distance lower than 5 Å to any QM atom. The dependence of acrolein excitation energies on the size of the CT region (from 3.5 Å to 8 Å) is reported in Table S1 in the Supporting Information (SI). [290] Notice that charge-equilibration based approaches, such as FQ and FQFμ, suffer from the so-called "CT catastrophe". [458] For this reason, CT is allowed across a relatively small region ($\leq$ 8 Å). In all calculations, recently reported FQ and FQFμ parametrizations are exploited. [13, 30]

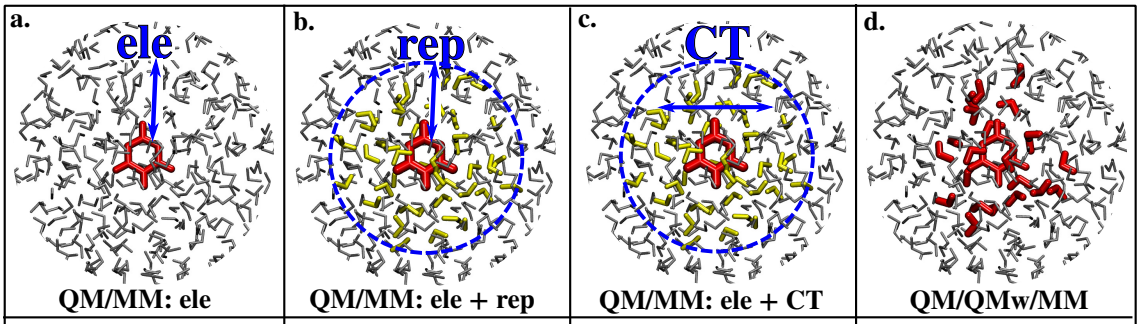


Figure 4.1: Hierarchy of QM/MM approaches adopted in this work. **a.** electrostatic/polarizable embedding: QM/MM interaction described by means of electrostatic-only terms. **b.** In addition to the electrostatic coupling, Pauli repulsion contribution is considered between the QM solute and the water molecules which are placed at distance lower than 5.0 Å. **c.** Intermolecular CT between solvent water molecules (which are placed at a distance lower than 5.0 Å from the QM) is allowed. **d.** Super-molecule QM/QM_w/MM approach: the first solvation shell (distance lower than 3.5 Å) is included in the QM portion, whereas the remaining solvent molecules are described by means of FQ and FQFμ force fields.

The methodological approach sketched above is applied to QM/MM vertical excitation energies of acrolein, pyridine and pyrimidine in aqueous solution. To this end, in order to account for solvent fluctuactions and to sample the phase-space, 100 uncorrelated snapshots were extracted from MD simulations (see Sections S1 and S2.1 given as SI for more details [290]). Vertical excitation energies were also compared with the results obtained by including in the QM portion the water molecules in the first solvation shell (17, 18 and 18 on average for acrolein, pyridine and pyrimidine, respectively), and by treating the remaining solvent molecules with FQ or FQFμ force field (QM/QM_w/FQ and QM/QM_w/FQFμ). Absorption energies were calculated on the same 100 snapshots extracted from MD. The results obtained by exploiting the above super-molecule approaches are taken as a reference (in addition to experiments) to further demonstrate the reliability and robustness of our approach. Notice that the results obtained in this way automatically include Pauli repulsion and intermolecular CT among QM water molecules. Additional calculations at the QM/PCM level, and on full QM clusters (QM-QM_w) made of pyridine+1w (water molecule), acrolein+2w and pyrimidine+2w were also performed.

Let us first focus on the vacuo-to-water solvatochromism exhibited by acrolein. Figure 4.2 shows computed vacuo-to-water solvatochromic shifts for the $n \rightarrow \pi^*$ (bottom) and $\pi \rightarrow \pi^*$ (top) transitions of acrolein (the corresponding values are reported in Table S2 in SI [290]).

The bottom part of Figure 4.2, reporting the results for $n \rightarrow \pi^*$ transition, shows that the continuum QM/PCM cannot reproduce the vacuo-to-water solvatochromism, whereas the cluster-like QM-QMw slightly overestimates it. By moving to QM/MM approaches, the non-polarizable QM/TIP3P well reproduces the experimental shift, but the inclusion of Pauli repulsion shifts the computed value to the wrong direction, so that the experimental value is underestimated of 20% (absolute

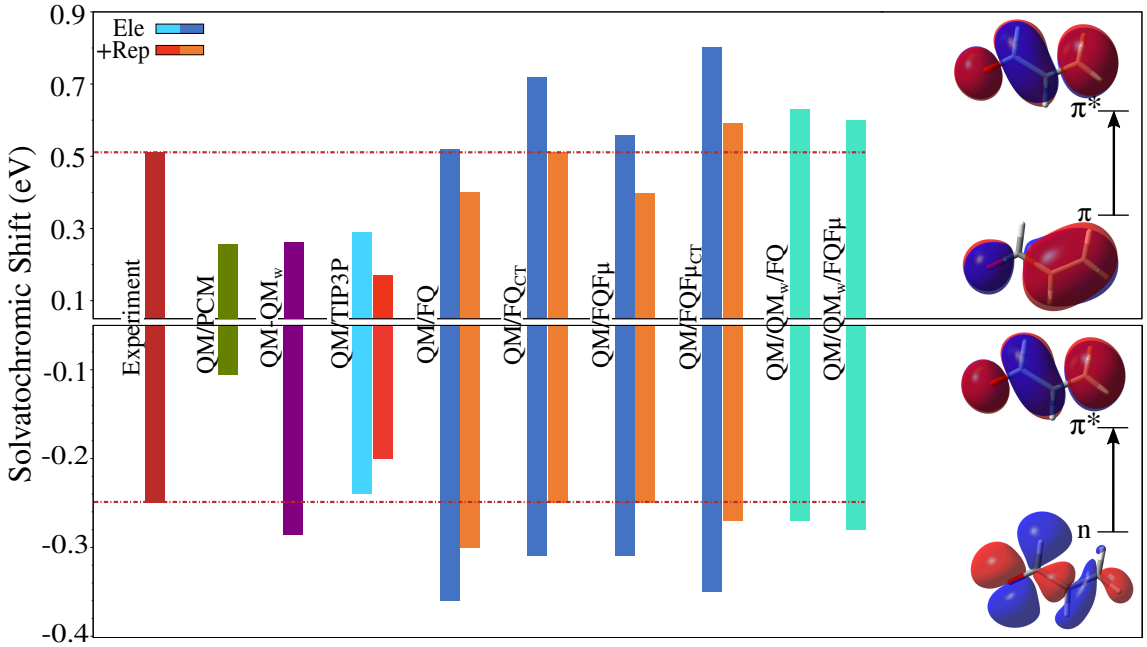


Figure 4.2: Computed and experimental $n \rightarrow \pi^*$ (bottom) and $\pi \rightarrow \pi^*$ (top) vacuo-to-water solvatochromic shifts of acrolein. Reference computed (CAM-B3LYP/aug-cc-pVDZ) vacuo vertical excitation energies: $n \rightarrow \pi^*$ 3.78 eV (Exp. [444] 3.69 eV); $\pi \rightarrow \pi^*$ 6.40 eV (Exp. [444] 6.42 eV). Horizontal dashed lines following the experimental data are drawn to guide the eye. Values are given in Table S2 in the SI [290]. All data are reported in eV.

value). A totally different picture is reported for polarizable QM/MM approaches. In fact, both QM/FQ and QM/FQFμ (by either including or neglecting CT effects) overestimate the experimental shift (33% on average). However, similarly to QM/TIP3P, computed vertical transition energies decrease when confinement (repulsion) effects are considered. Remarkably, the effect of quantum repulsion is substantial, decreasing the solvatochromic shift of 18% on average.

In order to get more insight into the role of repulsion, in Figure 4.3 ground and excited state molecular dipole moments are reported (the corresponding data are given in Table S8 in the SI [290]). The reference value for the isolated acrolein at the same level of calculation are 3.5 and 1.0 Debye for ground and $n \rightarrow \pi^*$ excited state, respectively. We see that both the ground and excited states are stabilized by the solvent; however, such a stabilization is larger in case the overall attractive electrostatic solute-solvent interactions are considered. On the other hand, repulsion acts as a confinement of the electron density and therefore the computed dipole moments decrease for all methods. These findings are reflected by computed transition energies (Figure 4.2). In fact, the stabilization of the ground state resulting from attractive electrostatic forces, yields a higher transition energy, which is instead reduced when the repulsive term is included. The differences reported by the various polarizable QM/MM approaches are due to a delicate balance between stabilization/destabilization effects of the two (ground/excited) states.

The inclusion of CT effects between MM water molecules has an opposite effect on values computed with QM/FQ or QM/FQFμ. In fact, solvatochromic shift decreases and shifts towards the experiment of almost 15% for QM/FQ$_{CT}$, whereas it increases (and shifts towards the supermolecule QM/QM$_w$/MM value) for

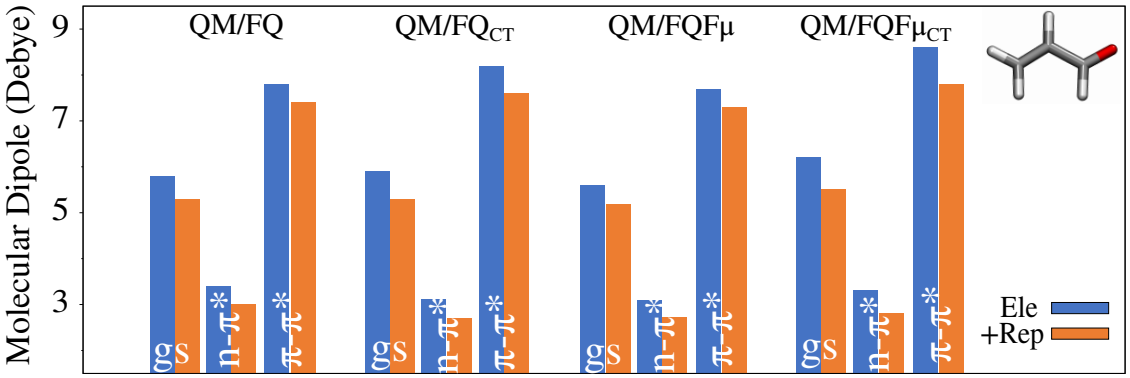


Figure 4.3: Computed QM/FQ, QM/FQ$_{CT}$, QM/FQFμ and QM/FQFμ_{CT} acrolein ground and excited state molecular dipole moments. Values are reported in Table S8 in the [290]. All data are given in Debye.

QM/FQFμ_{CT}. Such an opposite behavior can be rationalized once again by looking at excited state dipole moments (Figure 4.3), which decrease moving from QM/FQ to QM/FQ$_{CT}$, whereas the opposite trend is reported moving from QM/FQFμ to QM/FQFμ_{CT}. This means that in the first case (QM/FQ) the intermolecular CT acts as repulsive force, whereas in QM/FQFμ its effects is to stabilize the excited state. Thus, the larger vacuo-to-water solvatochromism shown by QM/FQFμ is primarly due to a greater stabilization of the ground state with respect to the $n \rightarrow \pi^*$ excited state.

Experimental values are exactly reproduced by QM/FQ$_{CT}$+rep and QM/FQFμ+rep, whereas they are overestimated (in absolute value) by QM/FQ+rep and QM/FQFμ_{CT}+rep.

We now move to the bright $\pi \rightarrow \pi^*$ transition of acrolein. The results depicted in Figure 4.2 (top) show that both QM/PCM and cluster-like QM-QMw approach cannot properly model this transition. The same applies to QM/TIP3P, which totally fails to reproduce the experiment, giving an error of 44% and exceedingly 67% if repulsion effects are considered.

The results obtained in case of polarizable QM/MM approaches retrace those already discussed for the dark transition, with QM/FQ$_{CT}$ giving a perfect agreement with the experiment when confinement effects are included; again, QM/FQFμ_{CT} +rep shows the best agreement with QM/QM$_w$/MM solvatochromic shifts. The effect of Pauli repulsion is in this case larger than for $n \rightarrow \pi^*$ transition, with a contribution that reaches 41% in case of QM/TIP3P. It is worth noticing that in this case confinement effects act in opposite direction with respect to $n \rightarrow \pi^*$ excitation. These findings may be explained by the fact that in case of $\pi \rightarrow \pi^*$ transition, the destabilization of the excited state due to the density confinement is predominant with respect to that of the ground state (see also dipole moment values reported in Table S8 in the SI, which confirm this trend [290]); therefore, excitation energies shift to higher values.

Remarkably, the role of MM intermolecular CT effects is opposite with respect to the dark $n \rightarrow \pi^*$ transition; also, such effects seem to be crucial, shifting the values of 35%, on average. In fact, the overall effect of CT is to stabilize the excited state for QM/FQ and QM/FQFμ calculations, thus resulting in larger vacuo-to-water solvatochromism. Therefore, Pauli repulsion and intermolecular CT act for this transition in opposite directions, so that their effects tend to cancel out (see

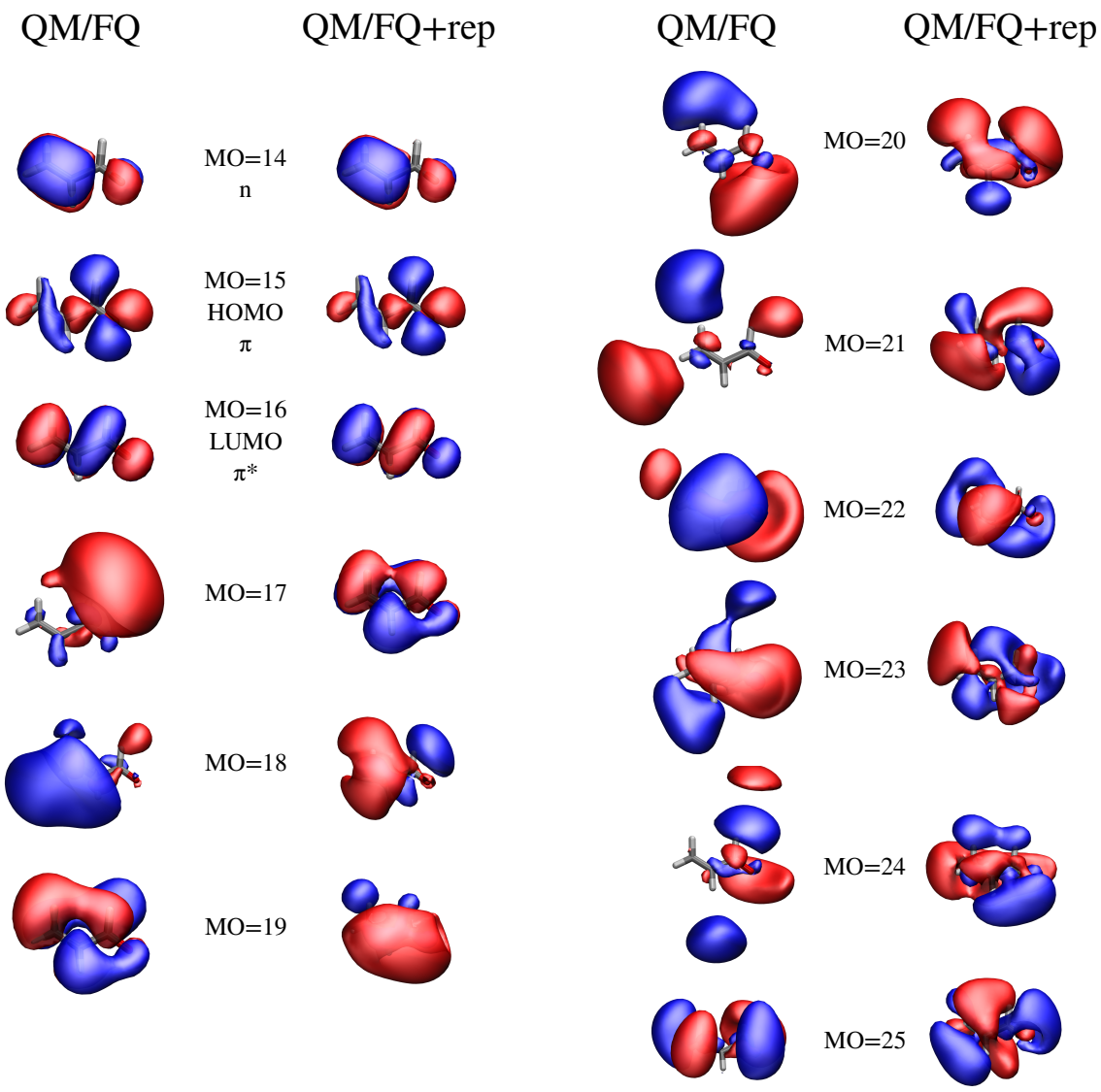


Figure 4.4: Selected acrolein molecular orbitals for a randomly chosen snapshot extracted from the MD simulation. QM/FQ and QM/FQ+rep values are depicted. Isovalue = 0.02.

Figure 4.2 (top) and Figure 4.3). We also note that CT effects are formulated in FQ and FQFμ in a similar way. Numerical differences are hidden in the actual numerical values of the elements of the linear equations that need to be solved (see Refs. [30] and [29]). For this reason, the numerical entity of CT and electrostatic interactions is difficult to quantify.

To end the discussion on acrolein, in Figure 4.4, 12 selected molecular orbitals (both occupied and virtual) of a randomly selected snapshot extracted from the MD are depicted. Both molecular orbitals obtained by either including or discarding Pauli repulsion are considered. As it is evident, the impact of quantum confinement on molecular orbitals is impressive. Also, virtual orbitals (from 17 on) are shrunk when Pauli repulsion is added; therefore, the aqueous solution has, as expected, a confining effect on the QM density. In light of such findings, remarkably our approach has the potentiality to solve some of the common issues related to the study of Rydberg states.

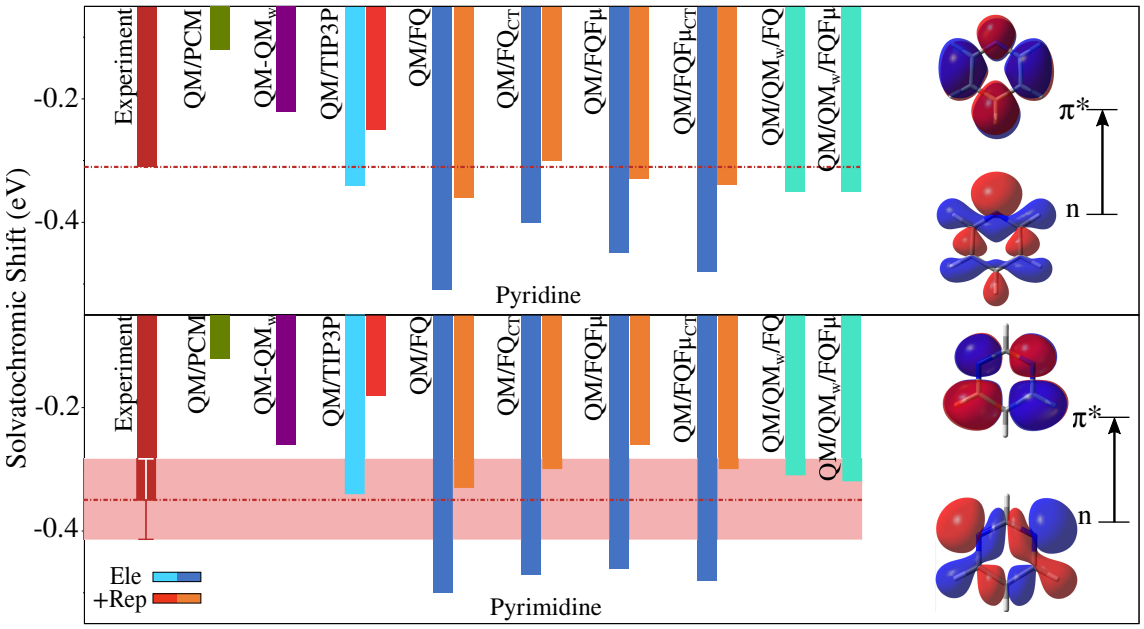


Figure 4.5: Computed (M06/6-311+G(2df,2p)) and experimental vacuo-to-water solvatochromic shifts for pyridine (top) and pyrimidine (bottom). Reference computed (M06/6-311+G(2df,2p)) vacuo vertical excitation energies: for pyridine $n \rightarrow \pi^*$ 4.67 eV (Exp. [375] 4.63 eV); for pyrimidine $n \rightarrow \pi^*$ 4.17 eV (Exp. [375] 4.18±0.01 eV). Horizontal dashed lines following experimental values are drawn to guide the eye. Values are given in Tables S5 and S7 in the SI [290]. All data are reported in eV.

To further demonstrate the reliability and robustness of our approach, it is also applied to pyridine and pyrimidine, of which vacuo-to-water solvatochromism has been investigated both by exploiting QM/continuum and atomistic QM/MM approaches. [375, 392, 393, 397, 401, 412]

Solvatochromic shifts as obtained by exploiting different computational approaches are shown in Figures 4.5 (values are given in Tables S5 and S7 in the SI [290]). Computed data show a similar behaviour as for acrolein. For both molecules, the continuum QM/PCM and the cluster-like approach (QM-QM_w) cannot reproduce experimental data. Therefore, it comes out that the inclusion of a minimal number of water molecules in the QM moiety is not a suitable approach to solvation, because dynamical and bulk effects are inappropriately modeled.

Moving to QM/MM approaches, non-polarizable QM/TIP3P apparently yields solvatochromic shifts in agreement with experiments and QM/QM_w/MM data, however the agreement decreases when confinement effects are considered. Both polarizable QM/FQ and QM/FQFμ approaches overestimate solvatochromic shifts, however the inclusion of Pauli repulsion shifts the results towards the experimental values. For all considered approaches, quantum confinement has a disruptive impact on the computed solvatochromisms, shifting the values of almost 30/40% on average. The same trend is also shown by computed dipole moments (see Figure S4 and Table S8 in the SI [290]). In particular, excited state dipole moments substantially decrease (30% on average) due to Pauli repulsion effects, thus also explaining the reported trends in Figure 4.5. Similarly to acrolein $n \rightarrow \pi^*$ transition, also in this case intermolecular CT plays an opposite role in QM/FQ and QM/FQFμ. Again, such

findings can be explained by the destabilization that occurs moving from QM/FQ to QM/FQ$_{\mathrm{CT}}$, in which intermolecular CT has overall a repulsive effect (see Figure S4 in SI [290]).

To summarize, for pyridine QM/FQ$_{\mathrm{CT}}$ is in best agreement with experiments (deviation of 3%), whereas QM/FQFμ (regardless the inclusion of CT effects) shows an almost perfect reproduction of reference calculations. For pyrimidine, the best agreement with the experimental value is shown by QM/FQ (6%), however, due to the error bar associated to experimental data, all polarizable QM/MM approaches (with the only exception of QM/FQFμ) give solvatochromic shifts lying within the experimental error.

Conclusions

In this thesis work, a theoretical framework and the related computational tools to describe the energy and response properties of complex molecular systems has been developed, implemented and tested.

The thesis has been divided into three main topics. First, the QM/Fluctuating Charge (FQ) model has been applied and extended to the calculation of different spectroscopies of molecular systems in aqueous solution. Second, the QM/FQ model has been parameterized for non aqueous systems, thus allowing the study of molecular properties/spectroscopies in different solvents. Third, the QM/FQ model has been extended so that an additional source of polarization, described in terms of fluctuating dipoles, is considered.

In Chapter 1, the extension of the QM/FQ approach to linear and quadratic response in the quasienergy formalism is presented and applied to the calculation of the Two-Photon Absorption (TPA). A computational protocol is developed and applied to TPA spectra of Rhodamine 6G (R6G) in aqueous solution. Such a molecule is characterized by a transition (S2) that is dark in the One-Photon Absorption (OPA) spectra due to symmetry selection rules, however, it is visible in TPA spectra. One of the main findings is that the explicit inclusion of water solvent molecules is essential to increase the agreement between theory and experiment. However, some discrepancies still remain and are principally related to the OPA-dark S2 transition for which polarization effects included by QM/FQ seem to be crucial.
In the same chapter, the extension of QM/FQ to Second Harmonic Generation (SHG) is presented. The modeling which is proposed focuses in achieving an appropriate representation of solvent effects both in the quantum mechanical response equations and in the conformational properties of the system, which is sampled by resorting to MD simulations. In the selected systems, the interaction with the surrounding environment is dominated by Hydrogen Bond (HB) interactions. QM/FQ shows a good performance at reproduction experimental SHG data extracted from Hyper Rayleigh Scattering (HRS) experiments. In particular, also in the case of SHG, polarization effects have a key role in getting a good agreement with experimental data.

In chapter 2, the extension of QM/FQ to different solvents of various polarity and hydrogen-bonding capability is presented. In particular, the model is challenged to reproduce solvatochromic shifts of four dyes dissolved in different solvents. The results are compared to experimental shift and other computational approaches. In most cases QM/FQ results are much better in agreement with experimental values than other approaches, which lack either the description of polarization effects (QM/EE) or any atomistic description of the solvent molecules (QM/PCM). Remarkably, non-electrostatic interactions play a crucial role to correctly reproduce experiments.

In chapter 3, a new polarizable force field, FQFμ, is proposed and coupled with a QM Hamiltonian. FQFμ models polarization in terms of both charges and dipoles that can vary as a response to the external electric potential/field. From the theoretical point of view, QM/FQFμ extends the QM/FQ model, where only fluctuating charges are used to describe the polarization of the environment. QM/FQFμ is parametrized so to reproduce electrostatic energies of aqueous solutions, and then tested to compute electrostatic energy of a water dimer as a function of the O-O distance, as well as the total interaction energy. QM/FQFμ is also coupled with a model to account for non-electrostatic energy terms, showing to be able to appropri-

ately reproduce CCSD(T) equilibrium geometry and the corresponding interaction energy for the same water dimer. Furthermore, QM/FQFμ is applied to the calculation of electrostatic energies of four molecules in aqueous solution, which can interact with water through hydrogen bonding interactions. The results show that charge polarization is indeed crucial to lower the errors with respect to full QM values, whereas the inclusion of dipole refines the numerical quality of the results. Thanks to its variational formulation, QM/FQFμ is also extended to molecular properties and spectroscopies, by following the same strategy which has been proposed for QM/FQ. Two different developments are reported in this thesis, namely the extension to vertical excitation energies and to IR spectra.

QM/FQFμ is extended to vertical excitation energies first by resorting to linear response (LR) theory: then, the corrected Linear Response (cLR) regime is considered are reported for the first time for both QM/FQ and QM/FQFμ. In fact, LR adequately describes environmental effects in case of transitions involving bright states associated with large transition dipole moments, but it is not able to catch the relaxation of the environment as a response to charge equilibration of the QM density on the specific excited state. cLR can instead describe such a relaxation, which is particularly relevant in case of transitions with low/null transition dipole moments, and with a large difference between the ground and excited dipole moments. Three systems, namely para-nitroaniline, pyridine and pyrimidine, in aqueous solution, are tested to proof the capabilities of the model. For all the studied systems, both QM/FQ and QM/FQFμ outperform QM/PCM, due to the fact it cannot properly describe specific solute-solvent interactions. Furthermore, QM/FQFμ correctly reproduces the experimental solvachromic shift in case of pNA, whereas it overestimates the experimental value for pyridine and pyrimidine. The analysis of ground and excited state molecular dipole moments shows that such a discrepancy is related to the overestimation of molecular dipole moments, probably related to the electrostatic-only description of the solvation.

QM/FQFμ is also extended to energy analytical first and second derivatives, thus extending the model to IR spectra. Three systems in aqueous solution are considered, namely methyloxirane, glycidol and gallic acid, which interact with water by means of strong specific interactions. As a result, computed IR spectra obtained using the atomistic QM/FQ and QM/FQFμ methods outperform QM/PCM. In particular, the bands which are mostly affected by the atomistic description of the environment, are those involving the polar moieties . In case of both methyloxirane and glycidol in aqueous solution, QM/FQ and QM/FQFμ predict similar spectra, whereas for gallic acid the inclusion of anisotropic terms in the MM modeling, i.e. the inclusion of fluctuating dipoles, yields better agreement with experimental data. It is worth pointing out that the development and implementation of analytical first energy derivatives, i.e. energy gradients, is not only the basic ingredient for computing vibrational spectra, but it also allows the further extension of the model to QM/MM MD.

In the last chapter the role of QM density confinement in the description of solvatochromic shifts is demonstrated. In particular, the results show that excellent agreement with experimental data, not biased by error cancellation, can be obtained by resorting to QM/MM approaches capable of accounting for quantum confinement effects. QM/MM calculations are performed on structures extracted from MD simulations. Therefore, the quality of the results also depends on the reliability of

the force field, and on the quality of the MD trajectory sampling. The proposed approach differs from alternative formulations of non-electrostatic terms and, remarkably, is so general that it can be coupled to any kind of QM/MM approach, either based on electrostatic- or polarizable-embedding.

Bibliography

[1] M. Karplus, M. Levitt, and A. Warshel, "The nobel prize in chemistry 2013," *Nobel Media AB 2014*, 2013.

[2] J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," *Chem. Rev.*, vol. 105, no. 8, pp. 2999–3094, 2005.

[3] J. Tomasi and M. Persico, "Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent," *Chemical Reviews*, vol. 94, no. 7, pp. 2027–2094, 1994.

[4] B. Mennucci and R. Cammi, Eds. Wiley, New York, 2007.

[5] B. Mennucci, "Continuum solvation models: What else can we learn from them?" *The Journal of Physical Chemistry Letters*, vol. 1, no. 10, pp. 1666–1674, 2010.

[6] B. Mennucci, "Modeling environment effects on spectroscopies through qm/classical models," *Phys. Chem. Chem. Phys.*, vol. 15, no. 18, pp. 6583–6594, 2013.

[7] F. Lipparini and B. Mennucci, "Perspective: Polarizable continuum models for quantum-mechanical descriptions," *J. Chem. Phys.*, vol. 144, no. 16, p. 160 901, 2016.

[8] J. H. Jensen and M. S. Gordon, "An approximate formula for the intermolecular pauli repulsion between closed shell molecules," *Mol. Phys.*, vol. 89, no. 5, pp. 1313–1325, 1996.

[9] J. H. J. and M. S. Gordon, "An approximate formula for the intermolecular pauli repulsion between closed shell molecules. ii. application to the effective fragment potential method," *J. Chem. Phys.*, vol. 108, no. 12, pp. 4772–4782, 1998.

[10] M. Ben-Nun and T. J. Martínez, "Direct evaluation of the pauli repulsion energy usingclassical'wavefunctions in hybrid quantum/classical potential energy surfaces," *Chem. Phys. Lett.*, vol. 290, no. 1-3, pp. 289–295, 1998.

[11] T. Giovannini, P. Lafiosca, and C. Cappelli, "A general route to include pauli repulsion and quantum dispersion effects in qm/mm approaches," *J. Chem. Theory Comput.*, vol. 13, no. 10, pp. 4854–4870, 2017.

[12] C. Curutchet, L. Cupellini, J. Kongsted, S. Corni, L. Frediani, A. H. Steindal, C. A. Guido, G. Scalmani, and B. Mennucci, "Density-dependent formulation of dispersion–repulsion interactions in hybrid multiscale quantum/molecular mechanics (qm/mm) models," *J. Chem. Theory Comput.*, vol. 14, no. 3, pp. 1671–1681, 2018.

[13] T. Giovannini, P. Lafiosca, B. Chandramouli, V. Barone, and C. Cappelli, "Effective yet reliable computation of hyperfine coupling constants in solution by a qm/mm approach: Interplay between electrostatics and non-electrostatic effects," *J. Chem. Phys.*, vol. 150, p. 124 102, 2019.

[14] L. J. Nåbo, J. M. H. Olsen, N. Holmgaard List, L. M. Solanko, D. Wüstner, and J. Kongsted, "Embedding beyond electrostatics–the role of wave function confinement," *J. Chem. Phys*, vol. 145, no. 10, p. 104 102, 2016.

[15] P. Reinholdt, J. Kongsted, and J. M. H. Olsen, "Polarizable density embedding: A solution to the electron spill-out problem in multiscale modeling," *J. Phys. Chem. Lett.*, vol. 8, no. 23, pp. 5949–5958, 2017.

[16] P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, H. Basch, and D. Cohen, "An effective fragment method for modeling solvent effects in quantum mechanical calculations," *J. Chem. Phys.*, vol. 105, no. 5, pp. 1968–1986, 1996.

[17] V. Kairys and J. H. Jensen, "Qm/mm boundaries across covalent bonds: A frozen localized molecular orbital-based approach for the effective fragment potential method," *J. Phys. Chem. A*, vol. 104, no. 28, pp. 6656–6665, 2000.

[18] Y. Mao, O. Demerdash, M. Head-Gordon, and T. Head-Gordon, "Assessing ion–water interactions in the amoeba force field using energy decomposition analysis of electronic structure calculations," *J. Chem. Theory Comput.*, vol. 12, no. 11, pp. 5422–5437, 2016.

[19] D. Loco, É. Polack, S. Caprasecca, L. Lagardere, F. Lipparini, J.-P. Piquemal, and B. Mennucci, "A qm/mm approach using the amoeba polarizable embedding: From ground state energies to electronic excitations," *J. Chem. Theory Comput.*, vol. 12, no. 8, pp. 3654–3661, 2016.

[20] D. Loco and L. Cupellini, "Modeling the absorption lineshape of embedded systems from molecular dynamics: A tutorial review," *Int. J. Quantum Chem.*, DOI: 10.1002/qua.25726, 2018.

[21] B. T. Thole, "Molecular polarizabilities calculated with a modified dipole interaction," *Chemical Physics*, vol. 59, no. 3, pp. 341–350, 1981.

[22] C. Curutchet, A. Muñoz-Losa, S. Monti, J. Kongsted, G. D. Scholes, and B. Mennucci, "Electronic energy transfer in condensed phase studied by a polarizable qm/mm model," *J. Chem. Theory Comput.*, vol. 5, no. 7, pp. 1838–1848, 2009.

[23] J. M. H. Olsen and J. Kongsted, "Molecular properties through polarizable embedding," *Adv. Quantum Chem.*, vol. 61, pp. 107–143, 2011.

[24] A. H. Steindal, K. Ruud, L. Frediani, K. Aidas, and J. Kongsted, "Excitation energies in solution: The fully polarizable qm/mm/pcm method," *J. Phys. Chem. B*, vol. 115, no. 12, pp. 3027–3037, 2011.

[25] S. Jurinovich, C. Curutchet, and B. Mennucci, "The fenna–matthews–olson protein revisited: A fully polarizable (td) dft/mm description," *ChemPhysChem*, vol. 15, no. 15, pp. 3194–3204, 2014.

[26] E. Boulanger and W. Thiel, "Solvent boundary potentials for hybrid qm/mm computations using classical drude oscillators: A fully polarizable model," *J. Chem. Theory Comput.*, vol. 8, no. 11, pp. 4527–4538, 2012.

[27] S. W. Rick, S. J. Stuart, and B. J. Berne, "Dynamical fluctuating charge force fields: Application to liquid water," *The Journal of chemical physics*, vol. 101, no. 7, pp. 6141–6156, 1994.

[28] S. W. Rick and B. J. Berne, "Dynamical Fluctuating Charge Force Fields: The Aqueous Solvation of Amides," *J. Am. Chem. Soc.*, vol. 118, no. 3, pp. 672–679, 1996. DOI: 10.1021/ja952535b. eprint: http://pubs.acs.org/doi/pdf/10.1021/ja952535b. [Online]. Available: http://pubs.acs.org/doi/abs/10.1021/ja952535b.

[29] C. Cappelli, "Integrated qm/polarizable mm/continuum approaches to model chiroptical properties of strongly interacting solute-solvent systems," *Int. J. Quantum Chem.*, vol. 116, no. 21, pp. 1532–1542, 2016.

[30] T. Giovannini, A. Puglisi, M. Ambrosetti, and C. Cappelli, "Polarizable qm/mm approach with fluctuating charges and fluctuating dipoles: The qm/fqfμ model," *J. Chem. Theory Comput.*, vol. 15, no. 4, pp. 2233–2245, 2019. DOI: 10.1021/acs.jctc.8b01149.

[31] T. Giovannini, L. Grazioli, M. Ambrosetti, and C. Cappelli, "Calculation of ir spectra with a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles," *J. Chem. Theory Comput.*, DOI: 10.1021/acs.jctc.9b00574, 2019, PMID: 31436976. DOI: 10.1021/acs.jctc.9b00574. eprint: https://doi.org/10.1021/acs.jctc.9b00574. [Online]. Available: https://doi.org/10.1021/acs.jctc.9b00574.

[32] F. Maseras and K. Morokuma, "Imomm: A new integrated ab initio+ molecular mechanics geometry optimization scheme of equilibrium structures and transition states," *Journal of Computational Chemistry*, vol. 16, no. 9, pp. 1170–1179, 1995.

[33] M. Svensson, S. Humbel, R. D. Froese, T. Matsubara, S. Sieber, and K. Morokuma, "Oniom: A multilayered integrated mo+ mm method for geometry optimizations and single point energy predictions. a test for diels- alder reactions and pt (p (t-bu) 3) 2+ h2 oxidative addition," *The Journal of Physical Chemistry*, vol. 100, no. 50, pp. 19 357–19 363, 1996.

[34] U. C. Singh and P. A. Kollman, "A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the ch3cl+ cl- exchange reaction and gas phase protonation of polyethers," *Journal of Computational Chemistry*, vol. 7, no. 6, pp. 718–730, 1986.

[35] S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma, and M. J. Frisch, "A new oniom implementation in gaussian98. part i. the calculation of energies, gradients, vibrational frequencies and electric field derivatives," *Journal of Molecular Structure: THEOCHEM*, vol. 461, pp. 1–21, 1999.

[36] W. J. Mortier, K. Van Genechten, and J. Gasteiger, "Electronegativity equalization: Application and parametrization," *Journal of the American Chemical Society*, vol. 107, no. 4, pp. 829–835, 1985.

[37] W. J. Mortier, S. K. Ghosh, and S. Shankar, "Electronegativity-equalization method for the calculation of atomic charges in molecules," *Journal of the American Chemical Society*, vol. 108, no. 15, pp. 4315–4320, 1986.

[38] R. S. Mulliken, "A new electroaffinity scale; together with data on valence states and on valence ionization potentials and electron affinities," *The Journal of Chemical Physics*, vol. 2, no. 11, pp. 782–793, 1934.

[39] R. G. Parr and R. G. Pearson, "Absolute hardness: Companion parameter to absolute electronegativity," *Journal of the American chemical society*, vol. 105, no. 26, pp. 7512–7516, 1983.

[40] K. Nishimoto and N. Mataga, "Electronic structure and spectra of some nitrogen heterocycles," *Zeitschrift für Physikalische Chemie*, vol. 12, no. 5_6, pp. 335–338, 1957.

[41] K. Ohno, "Some remarks on the pariser-parr-pople method," *Theoretica chimica acta*, vol. 2, no. 3, pp. 219–227, 1964.

[42] J. N. Louwen and E. T. Vogt, "Semi-empirical atomic charges for use in computational chemistry of molecular sieves," *Journal of Molecular Catalysis A: Chemical*, vol. 134, no. 1-3, pp. 63–77, 1998.

[43] G. S. He, L.-S. Tan, Q. Zheng, and P. N. Prasad, "Multiphoton absorbing materials: Molecular designs, characterizations, and applications," *Chem. Rev.*, vol. 108, no. 4, pp. 1245–1330, 2008.

[44] M. Göppert-Mayer, "Über elementarakte mit zwei quantensprüngen," *Ann. Phys-Berlin*, vol. 401, no. 3, pp. 273–294, 1931.

[45] W. Kaiser and C. Garrett, "Two-photon excitation in ca f 2: Eu 2+," *Phys. Rev. Lett.*, vol. 7, no. 6, p. 229, 1961.

[46] F. Terenziani, C. Katan, E. Badaeva, S. Tretiak, and M. Blanchard-Desce, "Enhanced two-photon absorption of organic chromophores: Theoretical and experimental assessments," *Adv. Mater.*, vol. 20, no. 24, pp. 4641–4678, 2008.

[47] M. Pawlicki, H. A. Collins, R. G. Denning, and H. L. Anderson, "Two-photon absorption and the design of two-photon dyes," *Angew. Chem. Int. Ed.*, vol. 48, no. 18, pp. 3244–3266, 2009.

[48] D. A. Parthenopoulos and P. M. Rentzepis, "Three-dimensional optical storage memory," *Science*, vol. 245, no. 4920, pp. 843–845, 1989.

[49] A. Gorman, J. Killoran, C. O'Shea, T. Kenna, W. M. Gallagher, and D. F. O'Shea, "In vitro demonstration of the heavy-atom effect for photodynamic therapy," *J. Am. Chem. Soc.*, vol. 126, no. 34, pp. 10 619–10 631, 2004.

[50] R. H. Köhler, J. Cao, W. R. Zipfel, W. W. Webb, and M. R. Hanson, "Exchange of protein molecules through connections between higher plant plastids," *Science*, vol. 276, no. 5321, pp. 2039–2042, 1997.

[51] K. D. Belfield, X. Ren, E. W. Van Stryland, D. J. Hagan, V. Dubikovsky, and E. J. Miesak, "Near-ir two-photon photoinitiated polymerization using a fluorone/amine initiating system," *J. Am. Chem. Soc.*, vol. 122, no. 6, pp. 1217–1218, 2000.

[52] M. T. Beerepoot, D. H. Friese, N. H. List, J. Kongsted, and K. Ruud, "Benchmarking two-photon absorption cross sections: Performance of cc2 and camb3lyp," *Phys. Chem. Chem. Phys.*, vol. 17, no. 29, pp. 19 306–19 314, 2015.

[53] D. H. Friese, C. Hättig, and K. Ruud, "Calculation of two-photon absorption strengths with the approximate coupled cluster singles and doubles model cc2 using the resolution-of-identity approximation," *Phys. Chem. Chem. Phys.*, vol. 14, no. 3, pp. 1175–1184, 2012.

[54] M. Guillaume, K. Ruud, A. Rizzo, S. Monti, Z. Lin, and X. Xu, "Computational study of the one-and two-photon absorption and circular dichroism of (l)-tryptophan," *J. Phys. Chem. B*, vol. 114, no. 19, pp. 6500–6512, 2010.

[55] C. Toro, L. De Boni, N. Lin, F. Santoro, A. Rizzo, and F. E. Hernandez, "Two-photon absorption circular dichroism: A new twist in nonlinear spectroscopy," *Chem–Eur. J.*, vol. 16, no. 11, pp. 3504–3509, 2010.

[56] L. Ferrighi, L. Frediani, E. Fossgaard, and K. Ruud, "Two-photon absorption of [2.2] paracyclophane derivatives in solution: A theoretical investigation," *J. Chem. Phys.*, vol. 127, no. 24, p. 244 103, 2007.

[57] A. H. Steindal, J. M. H. Olsen, K. Ruud, L. Frediani, and J. Kongsted, "A combined quantum mechanics/molecular mechanics study of the one-and two-photon absorption in the green fluorescent protein," *Phys. Chem. Chem. Phys.*, vol. 14, no. 16, pp. 5440–5451, 2012.

[58] A. Nag and D. Goswami, "Solvent effect on two-photon absorption and fluorescence of rhodamine dyes," *J. Photoch. Photobio. A*, vol. 206, no. 2-3, pp. 188–197, 2009.

[59] M. J. Paterson, O. Christiansen, F. Pawłowski, P. Jørgensen, C. Hättig, T. Helgaker, and P. Sałek, "Benchmarking two-photon absorption with cc3 quadratic response theory, and comparison with density-functional response theory," *J. Chem. Phys*, no. 5, p. 054 322, 2006.

[60] C. Hattig, O. Christiansen, and P. Jorgensen, "Multiphoton transition moments and absorption cross sections in coupled cluster response theory employing variational transition moment functionals," *J. Chem. Phys.*, vol. 108, no. 20, pp. 8331–8354, 1998.

[61] M. Albota, D. Beljonne, J.-L. Brédas, J. E. Ehrlich, J.-Y. Fu, A. A. Heikal, S. E. Hess, T. Kogej, M. D. Levin, S. R. Marder, D. McCord-Maughon, J. W. Perry, H. Röckel, M. Rumi, G. Subramaniam, W. W. Webb, X.-L. Wu, and C. Xu, "Design of organic molecules with large two-photon absorption cross sections," *Science*, vol. 281, no. 5383, pp. 1653–1656, 1998.

[62] K. D. Nanda and A. I. Krylov, "Two-photon absorption cross sections within equation-of-motion coupled-cluster formalism using resolution-of-the-identity and cholesky decomposition representations: Theory, implementation, and benchmarks," *J. Chem. Phys.*, vol. 142, no. 6, p. 064 118, 2015.

[63] H. Y. Woo, D. Korystov, A. Mikhailovsky, T.-Q. Nguyen, and G. C. Bazan, "Two-photon absorption in aqueous micellar solutions," *J. Am. Chem. Soc.*, vol. 127, no. 40, pp. 13 794–13 795, 2005.

[64] H. Y. Woo, J. W. Hong, B. Liu, A. Mikhailovsky, D. Korystov, and G. C. Bazan, "Water-soluble [2.2] paracyclophane chromophores with large two-photon action cross sections," *J. Am. Chem. Soc.*, vol. 127, no. 3, pp. 820–821, 2005.

[65] H. Y. Woo, B. Liu, B. Kohler, D. Korystov, A. Mikhailovsky, and G. C. Bazan, "Solvent effects on the two-photon absorption of distyrylbenzene chromophores," *J. Am. Chem. Soc.*, vol. 127, no. 42, pp. 14 721–14 729, 2005.

[66] A. Abbotto, L. Beverina, R. Bozio, A. Facchetti, C. Ferrante, G. A. Pagani, D. Pedron, and R. Signorini, "Novel heterocycle-based two-photon absorbing dyes," *Org. Lett.*, vol. 4, no. 9, pp. 1495–1498, 2002.

[67] S. Chakrabarti and K. Ruud, "Large two-photon absorption cross section: Molecular tweezer as a new promising class of compounds for nonlinear optics," *Phys. Chem. Chem. Phys.*, vol. 11, no. 15, pp. 2592–2596, 2009.

[68] M. T. Beerepoot, M. M. Alam, J. Bednarska, W. Bartkowiak, K. Ruud, and R. Zalesny, "Benchmarking the performance of exchange-correlation functionals for predicting two-photon absorption strengths," *J. Chem. Theory Comput.*, vol. 14, pp. 3677–3685, 2018.

[69] P. Sałek, O. Vahtras, J. Guo, Y. Luo, T. Helgaker, and H. Ågren, "Calculations of two-photon absorption cross sections by means of density-functional theory," *Chem. Phys. Lett.*, vol. 374, no. 5-6, pp. 446–452, 2003.

[70] S.-i. Kato, T. Matsumoto, M. Shigeiwa, H. Gorohmaru, S. Maeda, T. Ishi-i, and S. Mataka, "Novel 2, 1, 3-benzothiadiazole-based red-fluorescent dyes with enhanced two-photon absorption cross-sections," *Chem–Eur. J.*, vol. 12, no. 8, pp. 2303–2317, 2006.

[71] F. Terenziani, M. Morone, S. Gmouh, and M. Blanchard-Desce, "Linear and two-photon absorption properties of interacting polar chromophores: Standard and unconventional effects," *ChemPhysChem*, vol. 7, no. 3, pp. 685–696, 2006.

[72] A. Warshel and M. Levitt, "Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme," *Journal of molecular biology*, vol. 103, no. 2, pp. 227–249, 1976.

[73] A. Warshel and M. Karplus, "Calculation of ground and excited state potential surfaces of conjugated molecules. i. formulation and parametrization," *J. Am. Chem. Soc.*, vol. 94, no. 16, pp. 5612–5625, 1972.

[74] H. M. Senn and W. Thiel, "Qm/mm methods for biomolecular systems," *Angewandte Chemie International Edition*, vol. 48, no. 7, pp. 1198–1229, 2009.

[75] C. Cappelli, F. Lipparini, J. Bloino, and V. Barone, "Towards an accurate description of anharmonic infrared spectra in solution within the polarizable continuum model: Reaction field, cavity field and nonequilibrium effects," *J. Chem. Phys*, vol. 135, no. 10, p. 104 505, 2011.

[76] F. Egidi, T. Giovannini, M. Piccardo, J. Bloino, C. Cappelli, and V. Barone, "Stereoelectronic, vibrational, and environmental contributions to polarizabilities of large molecular systems: A feasible anharmonic protocol," *J. Chem. Theory Comput.*, vol. 10, no. 6, pp. 2456–2464, 2014.

[77] E. Boulanger and J. N. Harvey, "Qm/mm methods for free energies and photochemistry.," *Curr. Opin. Struc. Biol.*, vol. 49, pp. 72–76, 2018.

[78] P. Lahiri, K. B. Wiberg, P. H. Vaccaro, M. Caricato, and T. D. Crawford, "Large solvation effect in the optical rotatory dispersion of norbornenone," *Angew. Chem. Int. Ed.*, vol. 53, no. 5, pp. 1386–1389, 2014.

[79] D. Loco, S. Jurinovich, L. Cupellini, M. F. Menger, and B. Mennucci, "The modeling of the absorption lineshape for embedded molecules through a polarizable qm/mm approach," *Photochem. Photobiol. Sci.*, vol. 17, no. 5, pp. 552–560, 2018.

[80] R. D. Remigio, M. Repisky, S. Komorovsky, P. Hrobarik, L. Frediani, and K. Ruud, "Four-component relativistic density functional theory with the polarisable continuum model: Application to epr parameters and paramagnetic nmr shifts," *Mol. Phys.*, vol. 115, no. 1-2, pp. 214–227, 2017.

[81] A. Warshel, "Computer simulations of enzyme catalysis: Methods, progress, and insights," *Ann. Rev. Bioph. Biom.*, vol. 32, no. 1, pp. 425–443, 2003.

[82] J. Gao and X. Xia, "A priori evaluation of aqueous polarization effects through monte carlo qm-mm simulations," *Science*, vol. 258, no. 5082, pp. 631–635, 1992. DOI: 10.1126/science.1411573. eprint: http://www.sciencemag.org/content/258/5082/631.full.pdf. [Online]. Available: http://www.sciencemag.org/content/258/5082/631.abstract.

[83] P. Arora, L. V. Slipchenko, S. P. Webb, A. DeFusco, and M. S. Gordon, "Solvent-induced frequency shifts: Configuration interaction singles combined with the effective fragment potential method," *J. Phys. Chem. A*, vol. 114, no. 25, pp. 6742–6750, 2010. DOI: 10.1021/jp101780r. eprint: http://pubs.acs.org/doi/pdf/10.1021/jp101780r.

[84] V. Barone, M. Biczysko, and G. Brancato, "Extending the Range of Computational Spectroscopy by QM/MM Approaches: Time-Dependent and Time-Independent Routes," in *Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods*, ser. Advances in Quantum Chemistry, J. R. Sabin and S. Canuto, Eds., vol. 59, Academic Press, 2010, pp. 17–57. DOI: 10.1016/S0065-3276(10)59002-6. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0065327610590026.

[85] C. B. Nielsen, O. Christiansen, K. V. Mikkelsen, and J. Kongsted, "Density functional self-consistent quantum mechanics/molecular mechanics theory for linear and nonlinear molecular properties: Applications to solvated water and formaldehyde," *J. Chem. Phys*, vol. 126, no. 15, p. 154 112, 2007.

[86] S. Miertuš, E. Scrocco, and J. Tomasi, "Electrostatic interaction of a solute with a continuum. a direct utilizaion of ab initio molecular potentials for the prevision of solvent effects," *Chemical Physics*, vol. 55, no. 1, pp. 117–129, 1981.

[87] M. Orozco and F. J. Luque, "Theoretical methods for the description of the solvent effect in biomolecular systems," *Chem. Rev.*, vol. 100, no. 11, pp. 4187–4226, 2000.

[88] G. Scalmani and M. J. Frisch, "Continuous surface charge polarizable continuum models of solvation. I. general formalism," *J. Chem. Phys.*, vol. 132, no. 11, p. 114 110, 2010.

[89] F. Lipparini, G. Scalmani, B. Mennucci, E. Cancès, M. Caricato, and M. J. Frisch, "A variational formulation of the polarizable continuum model," *J. Chem. Phys.*, vol. 133, no. 1, p. 014 106, 2010.

[90] L. Frediani, Z. Rinkevicius, and H. Ågren, "Two-photon absorption in solution by means of time-dependent density-functional theory and the polarizable continuum model," *J. Chem. Phys.*, vol. 122, no. 24, p. 244 104, 2005.

[91] K. Zhao, L. Ferrighi, L. Frediani, C.-K. Wang, and Y. Luo, "Solvent effects on two-photon absorption of dialkylamino substituted distyrylbenzene chromophore," *J. Chem. Phys.*, vol. 126, no. 20, p. 204 509, 2007.

[92] L. Frediani, H. Ågren, L. Ferrighi, and K. Ruud, "Second-harmonic generation of solvated molecules using multiconfigurational self-consistent-field quadratic response theory and the polarizable continuum model," *J. Chem. Phys.*, vol. 123, no. 14, p. 144 117, 2005.

[93] R. Di Remigio, M. T. Beerepoot, Y. Cornaton, M. Ringholm, A. H. Steindal, K. Ruud, and L. Frediani, "Open-ended formulation of self-consistent field response theory with the polarizable continuum model for solvation," *Phys. Chem. Chem. Phys.*, vol. 19, no. 1, pp. 366–379, 2017.

[94] M. J. Field, P. A. Bash, and M. Karplus, "A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations," *Journal of Computational Chemistry*, vol. 11, no. 6, pp. 700–733, 1990.

[95] J. Gao, "Hybrid quantum and molecular mechanical simulations: An alternative avenue to solvent effects in organic chemistry," *Acc. Chem. Res.*, vol. 29, no. 6, pp. 298–305, 1996.

[96] R. A. Friesner and V. Guallar, "Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (qm/mm) methods for studying enzymatic catalysis," *Annu. Rev. Phys. Chem.*, vol. 56, pp. 389–427, 2005.

[97] H. Lin and D. G. Truhlar, "QM/MM: What have we learned, where are we, and where do we go from here?" *Theor. Chem. Acc.*, vol. 117, no. 2, pp. 185–199, 2007.

[98] A. Monari, J.-L. Rivail, and X. Assfeld, "Theoretical modeling of large molecular systems. advances in the local self consistent field method for mixed quantum mechanics/molecular mechanics calculations," *Acc. Chem. Res.*, vol. 46, no. 2, pp. 596–603, 2012.

[99] A. Monari, T. Very, J.-L. Rivail, and X. Assfeld, "A qm/mm study on the spinach plastocyanin: Redox properties and absorption spectra," *Comput. Theor. Chem.*, vol. 990, pp. 119–125, 2012.

[100] Z. Rinkevicius, X. Li, J. A. R. Sandberg, and H. Ågren, "Non-linear optical properties of molecules in heterogeneous environments: a quadratic density functional/molecular mechanics response theory," en, *Phys. Chem. Chem. Phys.*, vol. 16, no. 19, pp. 8981–8989, May 2014, ISSN: 1463-9076, 1463-9084. DOI: 10.1039/c4cp00992d. [Online]. Available: http://dx.doi.org/10.1039/c4cp00992d.

[101] Z. Rinkevicius, X. Li, J. A. R. Sandberg, K. V. Mikkelsen, and H. Ågren, "A Hybrid Density Functional Theory/Molecular Mechanics Approach for Linear Response Properties in Heterogeneous Environments," en, *J. Chem. Theory Comput.*, vol. 10, no. 3, pp. 989–1003, Mar. 2014, ISSN: 1549-9618, 1549-9626. DOI: 10.1021/ct400897s. [Online]. Available: http://dx.doi.org/10.1021/ct400897s.

[102] X. Li, Z. Rinkevicius, and H. Ågren, "Two-Photon Absorption of Metal-Assisted Chromophores," en, *J. Chem. Theory Comput.*, vol. 10, no. 12, pp. 5630–5639, Dec. 2014, ISSN: 1549-9618, 1549-9626. DOI: 10.1021/ct500579n. [Online]. Available: http://dx.doi.org/10.1021/ct500579n.

[103] F. Lipparini, C. Cappelli, G. Scalmani, N. De Mitri, and V. Barone, "Analytical first and second derivatives for a fully polarizable QM/classical hamiltonian," *J. Chem. Theory Comput.*, vol. 8, no. 11, pp. 4270–4278, 2012.

[104] F. Lipparini, C. Cappelli, and V. Barone, "Linear response theory and electronic transition energies for a fully polarizable QM/classical hamiltonian," *J. Chem. Theory Comput.*, vol. 8, no. 11, pp. 4153–4165, 2012.

[105] T. Giovannini, M. Macchiagodena, M. Ambrosetti, A. Puglisi, P. Lafiosca, G. Lo Gerfo, F. Egidi, and C. Cappelli, "Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling," *Int. J. Quantum Chem.*, vol. 119, e25684, 2019.

[106] F. Lipparini, F. Egidi, C. Cappelli, and V. Barone, "The optical rotation of methyloxirane in aqueous solution: A never ending story?" *J. Chem. Theory Comput.*, vol. 9, no. 4, pp. 1880–1884, 2013.

[107] F. Egidi, I. Carnimeo, and C. Cappelli, "Optical rotatory dispersion of methyloxirane in aqueous solution: Assessing the performance of density functional theory in combination with a fully polarizable QM/MM/PCMapproach," *Opt. Mater. Express*, vol. 5, no. 1, pp. 196–209, 2015.

[108] F. Egidi, T. Giovannini, G. Del Frate, P. M. Lemler, P. H. Vaccaro, and C. Cappelli, "A combined experimental and theoretical study of optical rotatory dispersion for (r)-glycidyl methyl ether in aqueous solution," *Phys. Chem. Chem. Phys.*, DOI: 10.1039/C8CP04445G, 2019.

[109] F. Egidi, R. Russo, I. Carnimeo, A. D'Urso, G. Mancini, and C. Cappelli, "The electronic circular dichroism of nicotine in aqueous solution: A test case for continuum and mixed explicit-continuum solvation approaches," *The Journal of Physical Chemistry A*, vol. 119, no. 21, pp. 5396–5404, 2015.

[110] T. Giovannini, M. Olszowka, and C. Cappelli, "Effective fully polarizable qm/mm approach to model vibrational circular dichroism spectra of systems in aqueous solution," *J. Chem. Theory Comput.*, vol. 12, no. 11, pp. 5483–5492, 2016.

[111] T. Giovannini, M. Olszòwka, F. Egidi, J. R. Cheeseman, G. Scalmani, and C. Cappelli, "Polarizable embedding approach for the analytical calculation of raman and raman optical activity spectra of solvated systems," *J. Chem. Theory Comput.*, vol. 13, no. 9, pp. 4421–4435, 2017.

[112] T. Giovannini, G. Del Frate, P. Lafiosca, and C. Cappelli, "Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution," *Phys. Chem. Chem. Phys.*, vol. 20, pp. 9181–9197, 2018.

[113] T. Giovannini, M. Ambrosetti, and C. Cappelli, "A polarizable embedding approach to second harmonic generation (shg) of molecular systems in aqueous solutions," *Theor. Chem. Acc.*, vol. 137, no. 6, p. 74, May 2018.

[114] A. H. Steindal, M. T. Beerepoot, M. Ringholm, N. H. List, K. Ruud, J. Kongsted, and J. M. H. Olsen, "Open-ended response theory with polarizable embedding: Multiphoton absorption in biomolecular systems," *Physical Chemistry Chemical Physics*, vol. 18, no. 40, pp. 28 339–28 352, 2016.

[115] N. S. Makarov, M. Drobizhev, and A. Rebane, "Two-photon absorption standards in the 550–1600 nm excitation wavelength range," *Opt. Express*, vol. 16, no. 6, pp. 4029–4047, 2008.

[116] C. B. Milojevich, D. W. Silverstein, L. Jensen, and J. P. Camden, "Surface-enhanced hyper-raman scattering elucidates the two-photon absorption spectrum of rhodamine 6g," *J. Phys. Chem. C*, vol. 117, no. 6, pp. 3046–3054, 2013.

[117] C. B. Milojevich, D. W. Silverstein, L. Jensen, and J. P. Camden, "Probing two-photon properties of molecules: Large non-condon effects dominate the resonance hyper-raman scattering of rhodamine 6g," *J. Am. Chem. Soc.*, vol. 133, no. 37, pp. 14 590–14 592, 2011.

[118] J. Guthmuller and B. Champagne, "Resonance raman scattering of rhodamine 6g as calculated by time-dependent density functional theory: Vibronic and solvent effects," *J. Phys. Chem. A*, vol. 112, no. 14, pp. 3215–3223, 2008.

[119] J. Guthmuller and B. Champagne, "Resonance raman spectra and raman excitation profiles of rhodamine 6g from time-dependent density functional theory," *ChemPhysChem*, vol. 9, no. 12, pp. 1667–1669, 2008.

[120] J. A. Dieringer, K. L. Wustholz, D. J. Masiello, J. P. Camden, S. L. Kleinman, G. C. Schatz, and R. P. Van Duyne, "Surface-enhanced raman excitation spectroscopy of a single rhodamine 6g molecule," *J. Am. Chem. Soc*, vol. 131, no. 2, pp. 849–854, 2008.

[121] P. A. Weiss, D. W. Silverstein, and L. Jensen, "Non-condon effects on the doubly resonant sum frequency generation of rhodamine 6g," *J. Phys. Chem. Lett.*, vol. 5, no. 2, pp. 329–335, 2014.

[122] P. Mark and L. Nilsson, "Structure and dynamics of the tip3p, spc, and spc/e water models at 298 k," *J. Phys. Chem. A*, vol. 105, no. 43, pp. 9954–9960, 2001.

[123] W. J. Mortier, K. Van Genechten, and J. Gasteiger, "Electronegativity equalization: application and parametrization," *J. Am. Chem. Soc.*, vol. 107, no. 4, pp. 829–835, 1985. DOI: 10.1021/ja00290a017. eprint: http://pubs.acs.org/doi/pdf/10.1021/ja00290a017. [Online]. Available: http://pubs.acs.org/doi/abs/10.1021/ja00290a017.

[124] R. Sanderson, "An interpretation of bond lengths and a classification of bonds," *Science*, vol. 114, no. 2973, pp. 670–672, 1951.

[125] R. Chelli and P. Procacci, "A transferable polarizable electrostatic force field for molecular mechanics based on the chemical potential equalization principle," *J. Chem. Phys.*, vol. 117, no. 20, pp. 9175–9189, 2002.

[126] F. Lipparini and V. Barone, "Polarizable force fields and polarizable continuum model: A fluctuating charges/pcm approach. 1. theory and implementation," *J. Chem. Theory Comput.*, vol. 7, no. 11, pp. 3711–3724, 2011.

[127] F. Lipparini, C. Cappelli, and V. Barone, "A gauge invariant multiscale approach to magnetic spectroscopies in condensed phase: General three-layer model, computational implementation and pilot applications," *J. Chem. Phys.*, vol. 138, no. 23, p. 234 108, 2013.

[128] I. Carnimeo, C. Cappelli, and V. Barone, "Analytical gradients for mp2, double hybrid functionals, and td-dft with polarizable embedding described by fluctuating charges," *J. Comput. Chem.*, vol. 36, no. 31, pp. 2271–2290, 2015.

[129] T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, and K. Ruud, "Recent advances in wave function-based methods of molecular-property calculations," *Chem. Rev.*, vol. 112, no. 1, pp. 543–631, 2012.

[130] A. J. Thorvaldsen, K. Ruud, K. Kristensen, P. Jørgensen, and S. Coriani, "A density matrix-based quasienergy formulation of the Kohn-Sham density functional response theory using perturbation- and time-dependent basis sets," *J. Chem. Phys.*, vol. 129, no. 21, p. 214 108, Dec. 2008, ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.2996351. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/19063545.

[131] M. Ringholm, D. Jonsson, and K. Ruud, "A general, recursive, and open-ended response code," *J. Comput. Chem.*, vol. 35, no. 8, pp. 622–633, 2014, ISSN: 0192-8651, 1096-987X. DOI: 10.1002/jcc.23533. [Online]. Available: http://dx.doi.org/10.1002/jcc.23533.

[132] H. Larsen, P. Joørgensen, J. Olsen, and T. Helgaker, "Hartree–Fock and Kohn–Sham atomic-orbital based time-dependent response theory," *J. Chem. Phys.*, vol. 113, no. 20, p. 8908, Nov. 2000, ISSN: 0021-9606. DOI: 10.1063/1.1318745. [Online]. Available: http://scitation.aip.org/content/aip/journal/jcp/113/20/10.1063/1.1318745.

[133] T. Kjaergaard, P. Jørgensen, J. Olsen, S. Coriani, and T. Helgaker, "Hartree-Fock and Kohn-Sham time-dependent response theory in a second-quantization atomic-orbital formalism suitable for linear scaling," *J. Chem. Phys.*, vol. 129, no. 5, p. 054 106, Aug. 2008, ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.2961039. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/18698887.

[134] R. E. Stratmann, G. E. Scuseria, and M. J. Frisch, "An efficient implementation of time-dependent density-functional theory for the calculation of excitation energies of large molecules," *J. Chem. Phys.*, vol. 109, no. 19, pp. 8218–8224, Nov. 1998, ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.477483. [Online]. Available: https://doi.org/10.1063/1.477483.

[135] D. H. Friese, M. T. P. Beerepoot, M. Ringholm, and K. Ruud, "Open-Ended Recursive Approach for the Calculation of Multiphoton Absorption Matrix Elements," en, *J. Chem. Theory Comput.*, vol. 11, no. 3, pp. 1129–1144, Mar. 2015, ISSN: 1549-9618, 1549-9626. DOI: 10.1021/ct501113y. [Online]. Available: http://dx.doi.org/10.1021/ct501113y.

[136] P. Cronstrand, Y. Luo, and H. Ågren, "Multi-photon absorption of molecules," *Adv. Quantum Chem.*, vol. 50, pp. 1–21, 2005.

[137] R. Di Remigio, A. H. Steindal, K. Mozgawa, V. Weijo, H. Cao, and L. Frediani, "Pcmsolver: An open-source library for solvation modeling," *Int. J. Quantum Chem.*, vol. 119, no. 1, e25685, 2019.

[138] K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, U. Ekström, T. Enevoldsen, J. J. Eriksen, P. Ettenhuber, B. Fernández, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hättig, H. Heiberg, T. Helgaker, A. C. Hennum, H. Hettema, E. Hjertenæs, S. Høst, I.-M. Høyvik, M. F. Iozzi, B. Jansıék, H. J. A. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjærgaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. B. Lutnæs, J. I. Melo, K. V. Mikkelsen, R. H. Myhre, C. Neiss, C. B. Nielsen, P. Norman, J. Olsen, J. M. H. Olsen, A. Osted, M. J. Packer, F. Pawlowski, T. B. Pedersen, P. F. Provasi, S. Reine, Z. Rinkevicius, T. A. Ruden, K. Ruud, V. V. Rybkin, P. Sałek, C. C. M. Samson, A. S. de Merás, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. Sneskov, A. H. Steindal, K. O. Sylvester-Hvid, P. R. Taylor, A. M. Teale, E. I. Tellgren, D. P. Tew, A. J. Thorvaldsen, L. Thøgersen, O. Vahtras, M. A. Watson, D. J. D. Wilson, M. Ziolkowski, and H. Agren, "The Dalton quantum chemistry program system," en, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, vol. 4, no. 3, pp. 269–284, May 2014, ISSN: 1759-0876, 1759-0884. DOI: 10.1002/wcms.1172. [Online]. Available: http://dx.doi.org/10.1002/wcms.1172.

[139] M. Bugeanu, R. Di Remigio, K. Mozgawa, S. S. Reine, H. Harbrecht, and L. Frediani, "Wavelet formulation of the polarizable continuum model. ii. use of piecewise bilinear boundary elements," *Phys. Chem. Chem. Phys.*, vol. 17, no. 47, pp. 31 566–31 581, 2015.

[140] L. S. Dodda, J. Z. Vilseck, K. J. Cutrona, and W. L. Jorgensen, "Evaluation of cm5 charges for nonaqueous condensed-phase modeling," *J. Chem. Theory Comput.*, vol. 11, no. 9, pp. 4273–4282, 2015.

[141] M. J. Abrahama, T. Murtola, R. Schulz, S. Pálla, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1-2, pp. 19–25, 2015.

[142] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren, "A biomolecular force field based on the free enthalpy of hydration and solvation: The gromos force-field parameter sets 53a5 and 53a6," *J. Comput. Chem.*, vol. 25, no. 13, pp. 1656–1676, 2004, ISSN: 1096-987X. DOI: 10.1002/jcc.20090. [Online]. Available: http://dx.doi.org/10.1002/jcc.20090.

[143] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *J. Chem. Phys.*, vol. 126, no. 1, 014101, 2007. DOI: http://dx.doi.org/10.1063/1.2408420. [Online]. Available: http://scitation.aip.org/content/aip/journal/jcp/126/1/10.1063/1.2408420.

[144] B. Hess, H. Bekker, H. J. Berendsen, and J. G. Fraaije, "Lincs: A linear constraint solver for molecular simulations," *Journal of computational chemistry*, vol. 18, no. 12, pp. 1463–1472, 1997.

[145] T. Darden, D. York, and L. Pedersen, "Particle mesh ewald: An nlog(n) method for ewald sums in large systems," *J. Chem. Phys.*, vol. 98, no. 12, pp. 10 089–10 092, 1993.

[146] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, *Gaussian 16 Revision A.03*, Gaussian Inc. Wallingford CT, 2016.

[147] A. Bondi, "van der Waals Volumes and Radii," *J. Phys. Chem.*, vol. 68, no. 3, pp. 441–451, Mar. 1964, ISSN: 0022-3654. DOI: 10.1021/j100785a001. [Online]. Available: http://pubs.acs.org/cgi-bin/doilookup/?10.1021/j100785a001.

[148] M. Mantina, A. C. Chamberlin, R. Valero, C. J. Cramer, and D. G. Truhlar, "Consistent van der Waals radii for the whole main group," en, *J. Phys. Chem. A*, vol. 113, no. 19, pp. 5806–5812, May 2009, ISSN: 1089-5639, 1520-5215. DOI: 10.1021/jp8111556. [Online]. Available: http://dx.doi.org/10.1021/jp8111556.

[149] M. Brehm and B. Kirchner, "Travis - a free analyzer and visualizer for monte carlo and molecular dynamics trajectories," *J. Chem. Inf. Model.*, vol. 51, no. 8, pp. 2007–2023, 2011, PMID: 21761915. DOI: 10.1021/ci200217w. eprint: http://dx.doi.org/10.1021/ci200217w. [Online]. Available: http://dx.doi.org/10.1021/ci200217w.

[150] R. Di Remigio, T. Giovannini, M. Ambrosetti, C. Cappelli, and L. Frediani, "Fully polarizable qm/fluctuating charge approach to two-photon absorption of aqueous solutions," *Journal of chemical theory and computation*, vol. 15, no. 7, pp. 4056–4068, 2019.

[151] T. Le Bahers, C. Adamo, and I. Ciofini, "A qualitative index of spatial extent in charge-transfer excitations," *J. Chem. Theory Comput.*, vol. 7, no. 8, pp. 2498–2506, 2011.

[152] F. Egidi, G. Lo Gerfo, M. Macchiagodena, and C. Cappelli, "On the nature of charge-transfer excitations for molecules in aqueous solution: A polarizable qm/mm study," *Theor. Chem. Acc.*, vol. 137, no. 6, p. 82, 2018.

[153] J. J. Wolff and R. Wortmann, "Organic materials for second-order non-linear optics," *Adv. Phys. Org. Chem.*, vol. 32, pp. 121–217, 1999.

[154] P. Innocenzi and B. Lebeau, "Organic–inorganic hybrid materials for non-linear optics," *J. Mater. Chem.*, vol. 15, no. 35-36, pp. 3821–3831, 2005.

[155] R. W. Boyd, *Nonlinear optics*. Academic press, 2003.

[156] Y.-R. Shen, "The principles of nonlinear optics," *New York, Wiley-Interscience, 1984, 575 p.*, 1984.

[157] S. R. Marder, J. E. Sohn, and G. D. Stucky, "Materials for nonlinear optics chemical perspectives," American Chemical Society Washington DC, Tech. Rep., 1991.

[158] H. S. Nalwa, "Organometallic materials for nonlinear optics," *Appl. Organomet. Chem.*, vol. 5, no. 5, pp. 349–377, 1991.

[159] D. S. Chemla, *Nonlinear optical properties of organic molecules and crystals*. Elsevier, 2012, vol. 1.

[160] J. Zyss, *Molecular nonlinear optics: materials, physics, and devices*. Academic press, 2013.

[161] C. Cappelli, B. Mennucci, R. Cammi, and A. Rizzo, "Quantum mechanical polarizable continuum model approach to the kerr effect of pure liquids," *J. Phys. Chem. B*, vol. 109, no. 39, pp. 18 706–18 714, 2005.

[162] A. Rizzo, C. Cappelli, B. Jansıék, D. Jonsson, P. Sałek, S. Coriani, D. J. Wilson, T. Helgaker, and H. Ågren, "Density-functional theory study of electric and magnetic properties of hexafluorobenzene in the vapor phase," *J. Chem. Phys*, vol. 122, no. 23, p. 234 314, 2005.

[163] C. Cappelli, A. Rizzo, B. Mennucci, J. Tomasi, R. Cammi, G. L. Rikken, R. Mathevet, and C. Rizzo, "The cotton–mouton effect of furan and its homologues in the gas phase, for the pure liquids and in solution," *J. Chem. Phys*, vol. 118, no. 23, pp. 10 712–10 724, 2003.

[164] D. M. Bishop and B. Kirtman, "A perturbation method for calculating vibrational dynamic dipole polarizabilities and hyperpolarizabilities," *J. Chem. Phys*, vol. 95, no. 4, pp. 2646–2658, 1991.

[165] B. Champagne, "Vibrational polarizability and hyperpolarizability of p-nitroaniline," *Chem. Phys. Lett.*, vol. 261, no. 1-2, pp. 57–65, 1996.

[166] D. M. Bishop, J. M. Luis, and B. Kirtman, "Additional compact formulas for vibrational dynamic dipole polarizabilities and hyperpolarizabilities," *J. Chem. Phys*, vol. 108, no. 24, pp. 10 013–10 017, 1998.

[167] B. Kirtman and D. M. Bishop, "Evaluation of vibrational hyperpolarizabilities," *Chem. Phys. Lett.*, vol. 175, no. 6, pp. 601–607, 1990.

[168] D. M. Bishop and P. Norman, "Effects of vibration on the polarizability and the first and second hyperpolarizabilities of hf, hcl, and hbr," *J. Chem. Phys.*, vol. 111, no. 7, pp. 3042–3050, 1999.

[169] D. M. Bishop, B. Kirtman, H. A. Kurtz, and J. E. Rice, "Calculation of vibrational dynamic hyperpolarizabilities for h2o, co2, and nh3," *J. Chem. Phys*, vol. 98, no. 10, pp. 8024–8030, 1993.

[170] O. Quinet, B. Kirtman, and B. Champagne, "Analytical time-dependent hartree-fock evaluation of the dynamic zero-point vibrationally averaged (zpva) first hyperpolarizability," *J. Chem. Phys*, vol. 118, no. 2, pp. 505–513, 2003.

[171] B. Champagne and B. Kirtman, "Vibrational versus electronic first hyperpolarizabilities of π-conjugated organic molecules: An ab initio hartree–fock investigation upon the effects of the nature of the linker," *Chem. Phys.*, vol. 245, no. 1, pp. 213–226, 1999.

[172] J. Kongsted and O. Christiansen, "Automatic generation of force fields and property surfaces for use in variational vibrational calculations of anharmonic vibrational energies and zero-point vibrational averaged properties," *J. Chem. Phys*, vol. 125, no. 12, p. 124 108, 2006.

[173] D. M. Bishop, B. Champagne, and B. Kirtman, "Relationship between static vibrational and electronic hyperpolarizabilities of π-conjugated push–pull molecules within the two-state valence-bond charge-transfer model," *J. Chem. Phys*, vol. 109, no. 22, pp. 9987–9994, 1998.

[174] C. Castiglioni, M. Del Zoppo, and G. Zerbi, "Molecular first hyperpolarizability of push-pull polyenes: Relationship between electronic and vibrational contribution by a two-state model," *Phys. Rev. B*, vol. 53, no. 20, p. 13 319, 1996.

[175] J. M. Luis, B. Champagne, and B. Kirtman, "Calculation of static zero-point vibrational averaging corrections and other vibrational curvature contributions to polarizabilities and hyperpolarizabilities using field-induced coordinates," *Int. J. Quantum Chem.*, vol. 80, no. 3, pp. 471–479, 2000.

[176] E. Benassi, F. Egidi, and V. Barone, "General strategy for computing nonlinear optical properties of large neutral and cationic organic chromophores in solution," *J. Phys. Chem. B*, vol. 119, no. 7, pp. 3155–3173, 2015.

[177] D. L. Silva, R. D. Fonseca, M. G. Vivas, E. Ishow, S. Canuto, C. R. Mendonca, and L. De Boni, "Experimental and theoretical investigation of the first-order hyperpolarizability of a class of triarylamine derivatives," *J. Chem. Phys*, vol. 142, no. 6, p. 064 312, 2015.

[178] M. G. Vivas, D. L. Silva, R. D. Rodriguez, S. Canuto, J. Malinge, E. Ishow, C. R. Mendonca, and L. De Boni, "Interpreting the first-order electronic hyperpolarizability for a series of octupolar push–pull triarylamine molecules containing trifluoromethyl," *J. Phys. Chem. C*, vol. 119, no. 22, pp. 12 589–12 597, 2015.

[179] A. S. Dutra, M. A. Castro, T. L. Fonseca, E. E. Fileti, and S. Canuto, "Hyperpolarizabilities of the methanol molecule: A ccsd calculation including vibrational corrections," *J. Chem. Phys*, vol. 132, no. 3, p. 034 307, 2010.

[180] P. Salek, T. Helgaker, O. Vahtras, H. Ågren, D. Jonsson, and J. Gauss, "A comparison of density-functional-theory and coupled-cluster frequency-dependent polarizabilities and hyperpolarizabilities," *Mol. Phys.*, vol. 103, no. 2-3, pp. 439–450, 2005.

[181] F. Pawłowski, P. Jørgensen, and C. Hättig, "The hyperpolarizability of the ne atom in the approximate coupled cluster triples model cc3," *Chem. Phys. Lett.*, vol. 391, no. 1, pp. 27–32, 2004.

[182] A. Rizzo, S. Coriani, B. Fernández, and O. Christiansen, "A coupled cluster response study of the electric dipole polarizability, first and second hyperpolarizabilities of hcl," *Phys. Chem. Chem. Phys.*, vol. 4, no. 13, pp. 2884–2890, 2002.

[183] L. Jensen, P. T. van Duijnen, and J. G. Snijders, "A discrete solvent reaction field model for calculating frequency-dependent hyperpolarizabilities of molecules in solution," *J. Chem. Phys*, vol. 119, no. 24, pp. 12 998–13 006, 2003.

[184] P. Norman, D. M. Bishop, H. J. A. Jensen, and J. Oddershede, "Nonlinear response theory with relaxation: The first-order hyperpolarizability," *The Journal of chemical physics*, vol. 123, no. 19, p. 194 103, 2005.

[185] Y. Luo, P. Norman, P. Macak, and H. Ågren, "Solvent effects on the polarizabilities and hyperpolarizabilities of conjugated polymers," *J. Chem. Phys.*, vol. 111, no. 21, pp. 9853–9858, 1999.

[186] L. Ferrighi, L. Frediani, C. Cappelli, P. Sałek, H. Ågren, T. Helgaker, and K. Ruud, "Density-functional-theory study of the electric-field-induced second harmonic generation (efishg) of push–pull phenylpolyenes in solution," *Chem. Phys. Lett.*, vol. 425, no. 4, pp. 267–272, 2006.

[187] C. Cappelli, B. Mennucci, and S. Monti, "Environmental effects on the spectroscopic properties of gallic acid: A combined classical and quantum mechanical study," *J. Phys. Chem. A*, vol. 109, no. 9, pp. 1933–1943, 2005.

[188] C. Cappelli and B. Mennucci, "Modeling the solvation of peptides. the case of (s)-n-acetylproline amide in liquid water," *J. Phys. Chem. B*, vol. 112, no. 11, pp. 3441–3450, 2008.

[189] R. Cammi, *Molecular Response Functions for the Polarizable Continuum Model: Physical basis and quantum mechanical formalism*. Springer Science & Business Media, 2013.

[190] R. Cammi, "Coupled-cluster theory for the polarizable continuum model. iii. a response theory for molecules in solution," *Int. J. Quantum Chem.*, vol. 112, no. 13, pp. 2547–2560, 2012.

[191] G. Del Frate, F. Bellina, G. Mancini, G. Marianetti, P. Minei, A. Pucci, and V. Barone, "Tuning of dye optical properties by environmental effects: A qm/mm and experimental study," *Phys. Chem. Chem. Phys.*, vol. 18, no. 14, pp. 9724–9733, 2016.

[192] J. Kongsted, A. Osted, K. V. Mikkelsen, and O. Christiansen, "Second harmonic generation second hyperpolarizability of water calculated using the combined coupled cluster dielectric continuum or different molecular mechanics methods," *J. Chem. Phys*, vol. 120, no. 8, pp. 3787–3798, 2004.

[193] L. Jensen and P. T. van Duijnen, "The first hyperpolarizability of p-nitroaniline in 1, 4-dioxane: A quantum mechanical/molecular mechanics study," *J. Chem. Phys*, vol. 123, no. 7, p. 074 307, 2005.

[194] M. Hidalgo Cardenuto and B. Champagne, "Qm/mm investigation of the concentration effects on the second-order nonlinear optical responses of solutions," *J. Chem. Phys.*, vol. 141, no. 23, p. 234 104, 2014.

[195] M. H. Cardenuto and B. Champagne, "The first hyperpolarizability of nitrobenzene in benzene solutions: Investigation of the effects of electron correlation within the sequential qm/mm approach," *Phys. Chem. Chem. Phys.*, vol. 17, no. 36, pp. 23 634–23 642, 2015.

[196] S. W. Rick, S. J. Stuart, J. S. Bader, and B. Berne, "Fluctuating charge force fields for aqueous solutions," *J. Mol. Liq.*, vol. 65, pp. 31–40, 1995.

[197] M. S. Gordon, L. Slipchenko, H. Li, and J. H. Jensen, "The effective fragment potential: A general method for predicting intermolecular interactions," *Annu. Rep. Comput. Chem.*, vol. 3, pp. 177–193, 2007.

[198] I. Adamovic and M. S. Gordon, "Dynamic polarizability, dispersion coefficient c6 and dispersion energy in the effective fragment potential method," *Mol. Phys.*, vol. 103, no. 2-3, pp. 379–387, 2005.

[199] K. Ohno, "Some remarks on the Pariser-Parr-Pople method," *Theor. Chim. Acta*, vol. 2, no. 3, pp. 219–227, 1964.

[200] F. Lipparini, C. Cappelli, and V. Barone, "Linear response theory and electronic transition energies for a fully polarizable QM/classical hamiltonian," *J. Chem. Theory Comput.*, vol. 8, no. 11, pp. 4153–4165, 2012.

[201] F. Lipparini, C. Cappelli, G. Scalmani, N. De Mitri, and V. Barone, "Analytical first and second derivatives for a fully polarizable QM/classical hamiltonian," *J. Chem. Theory Comput.*, vol. 8, no. 11, pp. 4270–4278, 2012.

[202] R. McWeeny, *Methods of molecular quantum mechanics*. Academic press, 1992.

[203] R. Wortmann, P. Krämer, C. Glania, S. Lebus, and N. Detzer, "Deviations from kleinman symmetry of the second-order polarizability tensor in molecules with low-lying perpendicular electronic bands," *Chem. Phys.*, vol. 173, no. 1, pp. 99–108, 1993.

[204] J. E. Rice, R. D. Amos, S. M. Colwell, N. C. Handy, and J. Sanz, "Frequency dependent hyperpolarizabilities with application to formaldehyde and methyl fluoride," *J. Chem. Phys.*, vol. 93, no. 12, pp. 8828–8839, 1990.

[205] J. E. Rice and N. C. Handy, "The calculation of frequency-dependent hyperpolarizabilities including electron correlation effects," *Int. J. Quantum Chem.*, vol. 43, no. 1, pp. 91–118, 1992.

[206] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, "Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions," *J. Phys. Chem. B*, vol. 113, no. 18, pp. 6378–6396, 2009.

[207] M. Cossi, G. Scalmani, N. Rega, and V. Barone, "New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution," *J. Chem. Phys.*, vol. 117, no. 1, pp. 43–54, 2002.

[208] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, "Improved side-chain torsion potentials for the amber ff99sb protein force field," *Proteins: Struct., Funct., Bioinf.*, vol. 78, no. 8, pp. 1950–1958, 2010.

[209] M. Macchiagodena, G. Mancini, M. Pagliai, G. Del Frate, and V. Barone, "Fine-tuning of atomic point charges: Classical simulations of pyridine in different environments," *Chem. Phys. Lett.*, vol. 677, pp. 120–126, 2017.

[210] C. I. Bayly, P. Cieplak, W. Cornell, and P. A. Kollman, "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The resp model," *J. Phys. Chem.*, vol. 97, no. 40, pp. 10 269–10 280, 1993. DOI: 10.1021/j100142a004.

[211] M. Macchiagodena, G. Mancini, M. Pagliai, and V. Barone, "Accurate prediction of bulk properties in hydrogen bonded liquids: Amides as case studies," *Phys. Chem. Chem. Phys.*, vol. 18, pp. 25 342–25 354, 36 2016. DOI: 10.1039/C6CP04666E. [Online]. Available: http://dx.doi.org/10.1039/C6CP04666E.

[212] M. Macchiagodena, G. Mancini, M. Pagliai, G. Cardini, and V. Barone, "New atomistic model of pyrrole with improved liquid state properties and structure.," *Int. J. Quantum Chem.*, DOI:10.1002/qua.25554, 2017.

[213] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.*, vol. 79, no. 2, pp. 926–935, 1983.

[214] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of chemical physics*, vol. 126, no. 1, p. 014 101, 2007.

[215] T. Giovannini, M. Ambrosetti, and C. Cappelli, "A polarizable embedding approach to second harmonic generation (shg) of molecular systems in aqueous solutions," *Theoretical Chemistry Accounts*, vol. 137, no. 6, p. 74, 2018.

[216] P. C. Ray, P. K. Das, and S. Ramasesha, "A comparative study of first hyperpolarizabilities of the acidic and basic forms of weak organic acids in water," *The Journal of chemical physics*, vol. 105, no. 21, pp. 9633–9639, 1996.

[217] K. Clays and A. Persoons, "Hyper-rayleigh scattering in solution," *Phys. Rev. Lett.*, vol. 66, no. 23, p. 2980, 1991.

[218] K. Clays and A. Persoons, "Hyper-rayleigh scattering in solution," *Rev. Sci. Instrum.*, vol. 63, no. 6, pp. 3285–3289, 1992.

[219] E. Hendrickx, K. Clays, and A. Persoons, "Hyper-rayleigh scattering in isotropic solution," *Acc. Chem. Res.*, vol. 31, no. 10, pp. 675–683, 1998.

[220] E. Botek, M. Spassova, B. Champagne, I. Asselberghs, A. Persoons, and K. Clays, "Hyper-rayleigh scattering of neutral and charged helicenes," *Chem. Phys. Lett.*, vol. 412, no. 4, pp. 274–279, 2005.

[221] R. Bersohn, Y.-H. Pao, and H. Frisch, "Double-quantum light scattering by molecules," *J. Chem. Phys.*, vol. 45, no. 9, pp. 3184–3198, 1966.

[222] A. Plaquet, M. Guillaume, B. Champagne, F. Castet, L. Ducasse, J.-L. Pozzo, and V. Rodriguez, "In silico optimization of merocyanine-spiropyran compounds as second-order nonlinear optical molecular switches," *Phys. Chem. Chem. Phys.*, vol. 10, no. 41, pp. 6223–6232, 2008.

[223] P. A. Limacher, K. V. Mikkelsen, and H. P. Lüthi, "On the accurate calculation of polarizabilities and second hyperpolarizabilities of polyacetylene oligomer chains using the cam-b3lyp density functional," *J. Chem. Phys.*, vol. 130, no. 19, p. 194 114, 2009.

[224] T. Pluta and A. J. Sadlej, "Hypol basis sets for high-level-correlated calculations of electric dipole hyperpolarizabilities," *Chem. Phys. Lett.*, vol. 297, no. 5, pp. 391–401, 1998.

[225] Y. Luo, H. Ågren, O. Vahtras, P. Jørgensen, V. Spirko, and H. Hettema, "Frequency-dependent polarizabilities and first hyperpolarizabilities of h2o," *The Journal of chemical physics*, vol. 98, no. 9, pp. 7159–7164, 1993.

[226] S. L. Murov, I. Carmichael, and G. L. Hug, *Handbook of photochemistry*. CRC Press, 1993.

[227] C. Reichardt and T. Welton, *Solvents and solvent effects in organic chemistry*. John Wiley & Sons, 2011.

[228] A. Marini, A. Muñoz-Losa, A. Biancardi, and B. Mennucci, "What is solvatochromism?" *J. Phys. Chem. B*, vol. 114, no. 51, pp. 17 128–17 135, 2010.

[229] M. J. Kamlet and R. Taft, "The solvatochromic comparison method. i. the. beta.-scale of solvent hydrogen-bond acceptor (hba) basicities," *J. Am. Chem. Soc.*, vol. 98, no. 2, pp. 377–383, 1976.

[230] R. Taft and M. J. Kamlet, "The solvatochromic comparison method. 2. the. alpha.-scale of solvent hydrogen-bond donor (hbd) acidities," *J. Am. Chem. Soc.*, vol. 98, no. 10, pp. 2886–2894, 1976.

[231] M. J. Kamlet, J. L. Abboud, and R. Taft, "The solvatochromic comparison method. 6. the. pi.*scale of solvent polarities," *J. Am. Chem. Soc.*, vol. 99, no. 18, pp. 6027–6038, 1977.

[232] M. J. Kamlet, J. L. M. Abboud, M. H. Abraham, and R. Taft, "Linear solvation energy relationships. 23. a comprehensive collection of the solvatochromic parameters,. pi.*,. alpha., and. beta., and some methods for simplifying the generalized solvatochromic equation," *J. Org. Chem.*, vol. 48, no. 17, pp. 2877–2887, 1983.

[233] Y. Marcus, "The properties of organic liquids that are relevant to their use as solvating solvents," *Chem. Soc. Rev.*, vol. 22, no. 6, pp. 409–416, 1993.

[234] C. Reichardt, "Solvatochromic dyes as solvent polarity indicators," *Chem. Rev.*, vol. 94, no. 8, pp. 2319–2358, 1994.

[235] S. Nigam and S. Rutan, "Principles and applications of solvatochromism," *Appl. Spectrosc.*, vol. 55, no. 11, 362A–370A, 2001.

[236] J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," *Chemical reviews*, vol. 105, no. 8, pp. 2999–3094, 2005.

[237] B. Mennucci, "Polarizable continuum model," *WIREs Comput. Mol. Sci.*, vol. 2, pp. 386–404, 2012.

[238] B. Mennucci and S. Corni, "Multiscale modelling of photoinduced processes in composite systems," *Nat. Rev. Chem.*, vol. 3, pp. 315–330, 2019.

[239] Q. Sun and G. K.-L. Chan, "Quantum embedding theories," *Acc. Chem. Res.*, vol. 49, no. 12, pp. 2705–2712, 2016.

[240] F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller III, "A simple, exact density-functional-theory embedding scheme," *J. Chem. Theory Comput.*, vol. 8, no. 8, pp. 2564–2568, 2012.

[241] T. Giovannini, F. Egidi, and C. Cappelli, "Molecular spectroscopy of aqueous solutions: A theoretical perspective," *Chem. Soc. Rev.*, vol. 49, no. 16, pp. 5664–5677, 2020.

[242] C. R. Jacob, J. Neugebauer, L. Jensen, and L. Visscher, "Comparison of frozen-density embedding and discrete reaction field solvent models for molecular properties," *Phys. Chem. Chem. Phys.*, vol. 8, no. 20, pp. 2349–2359, 2006.

[243] L. Goletto, T. Giovannini, S. D. Folkestad, and H. Koch, "Combining multilevel hartree–fock and multilevel coupled cluster approaches with molecular mechanics: A study of electronic excitations in solutions," *Phys. Chem. Chem. Phys.*, vol. 23, no. 7, pp. 4413–4425, 2021.

[244] F. Egidi, S. Angelico, P. Lafiosca, T. Giovannini, and C. Cappelli, "A polarizable three-layer frozen density embedding/molecular mechanics approach," *J. Chem. Phys.*, vol. 154, no. 16, p. 164 107, 2021.

[245] T. A. Wesolowski and A. Warshel, "Frozen density functional approach for ab initio calculations of solvated molecules," *J. Phys. Chem.*, vol. 97, no. 30, pp. 8050–8053, 1993.

[246] T. Giovannini, P. Lafiosca, and C. Cappelli, "A general route to include pauli repulsion and quantum dispersion effects in qm/mm approaches," *J. Chem. Theory Comput.*, vol. 13, no. 10, pp. 4854–4870, 2017.

[247] L. V. Slipchenko, "Effective fragment potential method," in *Many-Body Effects and Electrostatics in Biomolecules*, Pan Stanford, 2016, pp. 147–187.

[248] M. Bondanza, M. Nottoli, L. Cupellini, F. Lipparini, and B. Mennucci, "Polarizable embedding qm/mm: The future gold standard for complex (bio) systems?" *Phys. Chem. Chem. Phys.*, vol. 22, no. 26, pp. 14 433–14 448, 2020.

[249] T. Giovannini, F. Egidi, and C. Cappelli, "Theory and algorithms for chiroptical properties and spectroscopies of aqueous systems," *Phys. Chem. Chem. Phys.*, vol. 22, pp. 22 864–22 879, 2020.

[250] J. M. Olsen, K. Aidas, and J. Kongsted, "Excited states in solution through polarizable embedding," *J. Chem. Theory Comput.*, vol. 6, no. 12, pp. 3721–3734, 2010.

[251] J. M. H. Olsen, C. Steinmann, K. Ruud, and J. Kongsted, "Polarizable density embedding: A new qm/qm/mm-based computational strategy," *J. Phys. Chem. A*, vol. 119, no. 21, pp. 5344–5355, 2015.

[252] A. Marini, A. Muñoz-Losa, A. Biancardi, and B. Mennucci, "What is solvatochromism?" *J. Phys. Chem. B*, vol. 114, no. 51, pp. 17 128–17 135, 2010.

[253] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio Jr, *et al.*, "Current status of the amoeba polarizable force field," *The journal of physical chemistry B*, vol. 114, no. 8, pp. 2549–2564, 2010.

[254] E. Boulanger and W. Thiel, "Toward qm/mm simulation of enzymatic reactions with the drude oscillator polarizable force field," *J. Chem. Theory Comput.*, vol. 10, no. 4, pp. 1795–1809, 2014.

[255] T. Giovannini, A. Puglisi, M. Ambrosetti, and C. Cappelli, "Polarizable qm/mm approach with fluctuating charges and fluctuating dipoles: The qm/fqfμ model," *Journal of chemical theory and computation*, vol. 15, no. 4, pp. 2233–2245, 2019.

[256] T. Giovannini, R. R. Riso, M. Ambrosetti, A. Puglisi, and C. Cappelli, "Electronic transitions for a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles: Linear and corrected linear response regimes," *The Journal of chemical physics*, vol. 151, no. 17, p. 174 104, 2019.

[257] D. Loco, F. Buda, J. Lugtenburg, and B. Mennucci, "The dynamic origin of color tuning in proteins revealed by a carotenoid pigment," *J. Phys. Chem. Lett.*, vol. 9, no. 9, pp. 2404–2410, 2018.

[258] D. Loco, L. Lagardère, O. Adjoua, and J.-P. Piquemal, "Atomistic polarizable embeddings: Energy, dynamics, spectroscopy, and reactivity," *Acc. Chem. Res.*, pp. 3654–3661, 2021.

[259] T. Giovannini, P. Lafiosca, B. Chandramouli, V. Barone, and C. Cappelli, "Effective yet reliable computation of hyperfine coupling constants in solution by a qm/mm approach: Interplay between electrostatics and non-electrostatic effects," *J. Chem. Phys.*, vol. 150, p. 124 102, 2019.

[260] T. Giovannini, M. Macchiagodena, M. Ambrosetti, A. Puglisi, P. Lafiosca, G. Lo Gerfo, F. Egidi, and C. Cappelli, "Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling," *International Journal of Quantum Chemistry*, vol. 119, no. 1, e25684, 2019.

[261] S. Gómez, T. Giovannini, and C. Cappelli, "Absorption spectra of xanthines in aqueous solution: A computational study," *Phys. Chem. Chem. Phys.*, vol. 22, no. 10, pp. 5929–5941, 2020.

[262] S. D. Folkestad, E. F. Kjønstad, R. H. Myhre, J. H. Andersen, A. Balbi, S. Coriani, T. Giovannini, L. Goletto, T. S. Haugland, A. Hutcheson, *et al.*, "E t 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods," *J. Chem. Phys.*, vol. 152, no. 18, p. 184 103, 2020.

[263] H. A. Stern, G. A. Kaminski, J. L. Banks, R. Zhou, B. Berne, and R. A. Friesner, "Fluctuating charge, polarizable dipole, and combined models: Parameterization from ab initio quantum chemistry," *The Journal of Physical Chemistry B*, vol. 103, no. 22, pp. 4730–4737, 1999.

[264] S. Naserifar, D. J. Brooks, W. A. Goddard III, and V. Cvicek, "Polarizable charge equilibration model for predicting accurate electrostatic interactions in molecules and solids," *J. Chem. Phys.*, vol. 146, no. 12, p. 124 117, 2017.

[265] D. E. Goldberg and J. H. Holland, "Genetic algorithms and machine learning," 1988.

[266] O. Kramer, "Genetic algorithms," in *Genetic algorithm essentials*, Springer, 2017, pp. 11–19.

[267] A. Garrett, "Inspyred (version 1.0. 1) inspired intelligence," 2012.

[268] A. Tonda, "Inspyred: Bio-inspired algorithms in python," *Genetic Programming and Evolvable Machines*, vol. 21, no. 1, pp. 269–272, 2020.

[269] J. Chen and T. J. Martıénez, "Charge conservation in electronegativity equalization and its implications for the electrostatic properties of fluctuating-charge models," *The Journal of chemical physics*, vol. 131, no. 4, p. 044 114, 2009.

[270] M. J. Abrahama, T. Murtola, R. Schulz, S. Pálla, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1-2, pp. 19–25, 2015.

[271] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," *Journal of computational chemistry*, vol. 25, no. 9, pp. 1157–1174, 2004.

[272] J. Wang, W. Wang, P. A. Kollman, and D. A. Case, "Automatic atom type and bond type perception in molecular mechanical calculations," *Journal of molecular graphics and modelling*, vol. 25, no. 2, pp. 247–260, 2006.

[273] J. Wang, W. Wang, P. A. Kollman, and D. A. Case, "Antechamber: An accessory software package for molecular mechanical calculations," *J. Am. Chem. Soc*, vol. 222, U403, 2001.

[274] A. W. S. Da Silva and W. F. Vranken, "Acpype-antechamber python parser interface," *BMC research notes*, vol. 5, no. 1, pp. 1–8, 2012.

[275] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.*, vol. 79, no. 2, pp. 926–935, 1983.

[276] C. I. Bayly, P. Cieplak, W. Cornell, and P. A. Kollman, "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The resp model," *J. Phys. Chem.*, vol. 97, no. 40, pp. 10 269–10 280, 1993. DOI: 10.1021/j100142a004.

[277] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh Ewald method," *J. Chem. Phys.*, vol. 103, no. 19, pp. 8577–8593, 1995.

[278] H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of chemical physics*, vol. 81, no. 8, pp. 3684–3690, 1984.

[279] S. Corni, R. Cammi, B. Mennucci, and J. Tomasi, "Electronic excitation energies of molecules in solution within continuum solvation models: Investigating the discrepancy between state-specific and linear-response methods," *J. Chem. Phys.*, vol. 123, no. 13, p. 134 512, 2005.

[280] I. Duchemin, C. A. Guido, D. Jacquemin, and X. Blase, "The bethe–salpeter formalism with polarisable continuum embedding: Reconciling linear-response and state-specific features," *Chem. Sci.*, vol. 9, no. 19, pp. 4430–4443, 2018.

[281] C. A. Guido, A. Chrayteh, G. Scalmani, B. Mennucci, and D. Jacquemin, "Simple protocol for capturing both linear-response and state-specific effects in excited-state calculations with continuum solvation models," *J. Chem. Theory Comput.*, DOI: 10.1021/acs.jctc.1c00490, 2021.

[282] D. Kosenkov and L. V. Slipchenko, "Solvent effects on the electronic transitions of p-nitroaniline: A qm/efp study," *J. Phys. Chem. A*, vol. 115, no. 4, pp. 392–401, 2011.

[283] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, "Electronic absorption spectra and solvatochromic shifts by the vertical excitation model: Solvated clusters and molecular dynamics sampling," *J. Phys. Chem. B*, vol. 119, no. 3, pp. 958–967, 2015.

[284] G. Scalmani, M. J. Frisch, B. Mennucci, J. Tomasi, R. Cammi, and V. Barone, "Geometries and properties of excited states in the gas phase and in solution: Theory and application of a time-dependent density functional theory polarizable continuum model," *J. Chem. Phys.*, vol. 124, no. 9, p. 094 107, 2006.

[285] M. Stähelin, D. Burland, and J. Rice, "Solvent dependence of the second order hyperpolarizability in p-nitroaniline," *Chem. Phys. Lett.*, vol. 191, no. 3-4, pp. 245–250, 1992.

[286] C.-K. Wang, Y.-H. Wang, Y. Su, and Y. Luo, "Solvent dependence of solvatochromic shifts and the first hyperpolarizability of para-nitroaniline: A nonmonotonic behavior," *J. Chem. Phys.*, vol. 119, no. 8, pp. 4409–4412, 2003.

[287] C. Thomsen, J. Thøgersen, and S. Keiding, "Ultrafast charge-transfer dynamics: Studies of p-nitroaniline in water and dioxane," *J. Phys. Chem. A*, vol. 102, no. 7, pp. 1062–1067, 1998.

[288] S. Kovalenko, R. Schanz, V. Farztdinov, H. Hennig, and N. Ernsting, "Femtosecond relaxation of photoexcited para-nitroaniline: Solvation, charge transfer, internal conversion and cooling," *Chem. Phys. Lett.*, vol. 323, no. 3-4, pp. 312–322, 2000.

[289] J. J. Eriksen, S. P. Sauer, K. V. Mikkelsen, O. Christiansen, H. J. A. Jensen, and J. Kongsted, "Failures of tddft in describing the lowest intramolecular charge-transfer excitation in para-nitroaniline," *Mol. Phys.*, vol. 111, no. 9-11, pp. 1235–1248, 2013.

[290] T. Giovannini, M. Ambrosetti, and C. Cappelli, "Quantum confinement effects on solvatochromic shifts of molecular solutes," *The journal of physical chemistry letters*, vol. 10, no. 19, pp. 5823–5829, 2019.

[291] F. Santoro and D. Jacquemin, "Going beyond the vertical approximation with time-dependent density functional theory," *WIREs Comput. Mol. Sci.*, vol. 6, no. 5, pp. 460–486, 2016.

[292] L. P. Novaki and O. A. El Seoud, "Solvatochromism in pure solvents: Effects of the molecular structure of the probe," *Berichte der Bunsengesellschaft für physikalische Chemie*, vol. 100, no. 5, pp. 648–655, 1996.

[293] L. Brooker, G. Keyes, R. Sprague, R. VanDyke, E. VanLare, G. VanZandt, F. White, H. Cressman, and S. Dent Jr, "Color and constitution. x. 1 absorption of the merocyanines2," *J. Am. Chem. Soc.*, vol. 73, no. 11, pp. 5332–5350, 1951.

[294] D. C. da Silva, I. Ricken, M. A. d. R. Silva, and V. G. Machado, "Solute–solvent and solvent–solvent interactions in the preferential solvation of brooker's merocyanine in binary solvent mixtures," *J. Phys. Org. Chem.*, vol. 15, no. 7, pp. 420–427, 2002.

[295] J. O. Morley, R. M. Morley, and A. L. Fitton, "Spectroscopic studies on brooker's merocyanine," *J. Am. Chem. Soc.*, vol. 120, no. 44, pp. 11 479–11 488, 1998.

[296] V. Cavalli, D. C. da Silva, C. Machado, V. G. Machado, and V. Soldi, "The fluorosolvatochromism of brooker's merocyanine in pure and in mixed solvents," *J. Fluoresc.*, vol. 16, no. 1, pp. 77–86, 2006.

[297] S. T. Abdel-Halim and M. K. Awad, "Absorption, fluorescence, and semiempirical ased-mo studies on a typical brooker's merocyanine dye," *J. Mol. Struct.*, vol. 754, no. 1-3, pp. 16–24, 2005.

[298] P. Jacques, "On the relative contributions of nonspecific and specific interactions to the unusual solvtochromism of a typical merocyanine dye," *J. Phys. Chem.*, vol. 90, no. 22, pp. 5535–5539, 1986.

[299] S. T. Abdel-Halim, M. H. Abdel-Kader, and U. E. Steiner, "Thermal cis-trans isomerization of solvatochromic merocyanines: Linear correlations between solvent polarity and adiabatic and diabatic transition energies," *J. Phys. Chem.*, vol. 92, no. 15, pp. 4324–4328, 1988.

[300] C. Reichardt, "Pyridinium n-phenolate betaine dyes as empirical indicators of solvent polarity: Some new findings," *Pure Appl. Chem.*, vol. 76, no. 10, pp. 1903–1919, 2004.

[301] C. Reichardt, "Polarity of ionic liquids determined empirically by means of solvatochromic pyridinium n-phenolate betaine dyes," *Green Chem.*, vol. 7, no. 5, pp. 339–351, 2005.

[302] F. Santoro, C. Cappelli, and V. Barone, "Effective time-independent calculations of vibrational resonance raman spectra of isolated and solvated molecules including duschinsky and herzberg–teller effects," *J. Chem. Theory Comput.*, vol. 7, no. 6, pp. 1824–1839, 2011.

[303] J. J. Oppenheim, S. Naserifar, and W. A. Goddard III, "Extension of the polarizable charge equilibration model to higher oxidation states with applications to ge, as, se, br, sn, sb, te, i, pb, bi, po, and at elements," *J. Phys. Chem. A*, vol. 122, no. 2, pp. 639–645, 2018.

[304] A. Mayer, "Formulation in terms of normalized propagators of a charge-dipole model enabling the calculation of the polarization properties of fullerenes and carbon nanotubes," *Phys. Rev. B*, vol. 75, no. 4, p. 045 407, 2007.

[305] L. L. Jensen and L. Jensen, "Atomistic electrodynamics model for optical properties of silver nanoclusters," *J. Phys. Chem. C*, vol. 113, no. 34, pp. 15 182–15 190, 2009.

[306] Z. Rinkevicius, X. Li, J. A. Sandberg, K. V. Mikkelsen, and H. Ågren, "A hybrid density functional theory/molecular mechanics approach for linear response properties in heterogeneous environments," *J. Chem. Theory Comput.*, vol. 10, no. 3, pp. 989–1003, 2014.

[307] S. Jakobsen and F. Jensen, "Systematic improvement of potential-derived atomic multipoles and redundancy of the electrostatic parameter space," *Journal of chemical theory and computation*, vol. 10, no. 12, pp. 5493–5504, 2014.

[308] S. Jakobsen and F. Jensen, "Searching the force field electrostatic multipole parameter space," *J. Chem. Theory Comput.*, vol. 12, no. 4, pp. 1824–1832, 2016.

[309] S. M. Morton and L. Jensen, "A discrete interaction model/quantum mechanical method for describing response properties of molecules adsorbed on metal nanoparticles," *J. Chem. Phys*, vol. 133, no. 7, p. 074 103, 2010.

[310] T. Löytynoja, X. Li, K. Jänkälä, Z. Rinkevicius, and H. Ågren, "Quantum mechanics capacitance molecular mechanics modeling of core-electron binding energies of methanol and methyl nitrite on ag (111) surface," *J. Chem. Phys.*, vol. 145, no. 2, p. 024 703, 2016.

[311] Z. Rinkevicius, J. A. Sandberg, X. Li, M. Linares, P. Norman, and H. Ågren, "Hybrid complex polarization propagator/molecular mechanics method for heterogeneous environments," *J. Chem. Theory Comput.*, vol. 12, no. 6, pp. 2661–2667, 2016.

[312] X. Li, V. Carravetta, C. Li, S. Monti, Z. Rinkevicius, and H. Ågren, "Optical properties of gold nanoclusters functionalized with a small organic compound: Modeling by an integrated quantum-classical approach," *J. Chem. Theory Comput.*, vol. 12, no. 7, pp. 3325–3339, 2016.

[313] C.-K. Lim, X. Li, Y. Li, K. L. Drew, J. P. Palafox-Hernandez, Z. Tang, A. Baev, A. N. Kuzmin, M. R. Knecht, T. R. Walsh, M. T. Swihart, H. Ågren, and P. N. Prasad, "Plasmon-enhanced two-photon-induced isomerization for highly-localized light-based actuation of inorganic/organic interfaces," *Nanoscale*, vol. 8, no. 7, pp. 4194–4202, 2016.

[314] A. K. Rappe and W. A. Goddard III, "Charge equilibration for molecular dynamics simulations," *J. Phys. Chem*, vol. 95, no. 8, pp. 3358–3363, 1991.

[315] J. A. Lemkul, J. Huang, B. Roux, and A. D. MacKerell Jr, "An empirical polarizable force field based on the classical drude oscillator model: Development history and recent applications," *Chem. Rev.*, vol. 116, no. 9, pp. 4983–5013, 2016.

[316] D. P. Geerke, S. Thiel, W. Thiel, and W. F. van Gunsteren, "Combined qm/mm molecular dynamics study on a condensed-phase sn2 reaction at nitrogen: The effect of explicitly including solvent polarization," *J. Chem. Theory Comput.*, vol. 3, no. 4, pp. 1499–1509, 2007.

[317] Z. Lu and Y. Zhang, "Interfacing ab initio quantum mechanical method with classical drude osillator polarizable model for molecular dynamics simulation of chemical reactions," *J. Chem. Theory Comput.*, vol. 4, no. 8, pp. 1237–1248, 2008.

[318] S. Caprasecca, S. Jurinovich, L. Viani, C. Curutchet, and B. Mennucci, "Geometry optimization in polarizable qm/mm models: The induced dipole formulation," *J. Chem. Theory Comput.*, vol. 10, no. 4, pp. 1588–1598, 2014.

[319] S. Caprasecca, L. Cupellini, S. Jurinovich, D. Loco, F. Lipparini, and B. Mennucci, "A polarizable qm/mm description of environment effects on nmr shieldings: From solvated molecules to pigment–protein complexes," *Theor. Chem. Acc.*, vol. 137, no. 6, p. 84, 2018.

[320] J. Huang, A. C. Simmonett, F. C. Pickard IV, A. D. MacKerell Jr, and B. R. Brooks, "Mapping the drude polarizable force field onto a multipole and induced dipole model," *J. Chem. Phys.*, vol. 147, no. 16, p. 161 702, 2017.

[321] T. Giovannini, P. Lafiosca, B. Chandramouli, V. Barone, and C. Cappelli, *submitted*,

[322] H. Berendsen, D. van der Spoel, and R. van Drunen, "Gromacs: A message-passing parallel molecular dynamics implementation," *Comp. Phys. Comm.*, vol. 91, no. 1–3, pp. 43–56, 1995, ISSN: 0010-4655. DOI: 10.1016/0010-4655(95)00042-E. [Online]. Available: http://www.sciencedirect.com/science/article/pii/001046559500042E.

[323] E. Lindahl, B. Hess, and D. van der Spoel, "Gromacs 3.0: A package for molecular simulation and trajectory analysis," *J. Mol. Model.*, vol. 7, pp. 306–317, 8 2001, ISSN: 1610-2940. [Online]. Available: http://dx.doi.org/10.1007/s008940100045.

[324] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, "Gromacs: Fast, flexible, and free," *J. Comput. Chem.*, vol. 26, no. 16, pp. 1701–1718, 2005, ISSN: 1096-987X. DOI: 10.1002/jcc.20291. [Online]. Available: http://dx.doi.org/10.1002/jcc.20291.

[325] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, "Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation," *J. Chem. Theory Comput.*, vol. 4, no. 3, pp. 435–447, 2008. DOI: 10.1021/ct700301q. eprint: http://pubs.acs.org/doi/pdf/10.1021/ct700301q. [Online]. Available: http://pubs.acs.org/doi/abs/10.1021/ct700301q.

[326] K. Kitaura and K. Morokuma, "A new energy decomposition scheme for molecular interactions within the hartree-fock approximation," *Int. J. Quantum Chem.*, vol. 10, no. 2, pp. 325–340, 1976.

[327] K. Morokuma and K. Kitaura, "Energy decomposition analysis of molecular interactions," in *Chemical applications of atomic and molecular electrostatic potentials*, Springer, 1981, pp. 215–242.

[328] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, "General atomic and molecular electronic structure system," *J. Comput. Chem.*, vol. 14, no. 11, pp. 1347–1363, 1993.

[329] M. S. Gordon and M. W. Schmidt, "Advances in electronic structure theory: Gamess a decade later," in *Theory and Applications of Computational Chemistry: the first forty years*. Elsevier, 2005, pp. 1167–1189, ISBN: 9780470399545. DOI: 10.1002/9780470399545.ch1. [Online]. Available: http://dx.doi.org/10.1002/9780470399545.ch1.

[330] K. Szalewicz, "Symmetry-adapted perturbation theory of intermolecular forces," *WIRES: Comput. Mol. Sci.*, vol. 2, pp. 254–272, 2012.

[331] E. G. Hohenstein and C. D. Sherrill, "Wavefunction methods for noncovalent interactions," *WIRES: Comput. Mol. Sci.*, vol. 2, pp. 304–326, 2012.

[332] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, "Psi4 1.1: An open-source electronic structure program emphasizing automation, advanced libraries, and interoperability," *J. Chem. Theory Comput.*, vol. 13, no. 7, pp. 3185–3197, 2017. DOI: 10.1021/acs.jctc.7b00174. eprint: http://dx.doi.org/10.1021/acs.jctc.7b00174. [Online]. Available: http://dx.doi.org/10.1021/acs.jctc.7b00174.

[333] E. G. Kratz, A. R. Walker, L. Lagardère, F. Lipparini, J.-P. Piquemal, and G. Andrés Cisneros, "Lichem: A qm/mm program for simulations with multipolar and polarizable force fields," *J. Comput. Chem.*, vol. 37, pp. 1019–1029, 2016.

[334] A. Tkatchenko and M. Scheffler, "Accurate molecular van der waals interactions from ground-state electron density and free-atom reference data," *Phys. Rev. Lett.*, vol. 102, no. 7, p. 073 005, 2009.

[335] A. Tkatchenko, L. Romaner, O. T. Hofmann, E. Zojer, C. Ambrosch-Draxl, and M. Scheffler, "Van der waals interactions between organic adsorbates and at organic/inorganic interfaces," *MRS bulletin*, vol. 35, no. 06, pp. 435–442, 2010.

[336] A. Tkatchenko, R. A. DiStasio Jr, R. Car, and M. Scheffler, "Accurate and efficient method for many-body van der waals interactions," *Phys. Rev. Lett.*, vol. 108, no. 23, p. 236 402, 2012.

[337] J. Hermann, R. A. DiStasio, and A. Tkatchenko, "First-principles models for van der waals interactions in molecules and materials: Concepts, theory, and applications," *Chem. Rev.*, vol. 117, no. 6, pp. 4714–4758, 2017.

[338] D. Jacquemin, B. Moore, A. Planchat, C. Adamo, and J. Autschbach, "Performance of an optimally tuned range-separated hybrid functional for 0–0 electronic excitation energies," *J. Chem. Theory Comput.*, vol. 10, no. 4, pp. 1677–1685, 2014.

[339] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.*, vol. 78, pp. 1396–1396, 7 Feb. 1997. DOI: 10.1103/PhysRevLett.78.1396. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.78.1396.

[340] A. D. Becke, "Density-functional thermochemistry. v. systematic optimization of exchange-correlation functionals," *J. Chem. Phys.*, vol. 107, no. 20, pp. 8554–8560, 1997.

[341] H. L. Schmider and A. D. Becke, "Optimized density functionals from the extended g2 test set," *J. Chem. Phys.*, vol. 108, no. 23, pp. 9624–9631, 1998.

[342] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, L. A. Constantin, and J. Sun, "Workhorse semilocal density functional for condensed matter physics and quantum chemistry," *Phys. Rev. Lett.*, vol. 103, p. 026 403, 2 Jul. 2009. DOI: 10.1103/PhysRevLett.103.026403. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.103.026403.

[343] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Phys. Rev. A*, vol. 38, pp. 3098–3100, 6 Sep. 1988. DOI: 10.1103/PhysRevA.38.3098. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.38.3098.

[344] Y. Zhao and D. G. Truhlar, "The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four m06-class functionals and 12 other functionals," *Theor. Chem. Acc.: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, vol. 120, no. 1, pp. 215–241, 2008.

[345] C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: The pbe0 model," *J. Chem. Phys.*, vol. 110, no. 13, pp. 6158–6170, 1999.

[346] A. D. Becke, "Density-functional thermochemistry. iii. the role of exact exchange," *J. Chem. Phys.*, vol. 98, no. 7, pp. 5648–5652, 1993.

[347] R. Peverati and D. G. Truhlar, "Communication: A global hybrid generalized gradient approximation to the exchange-correlation functional that satisfies the second-order density-gradient constraint and has broad applicability in chemistry," *J. Chem. Phys.*, vol. 135, p. 191 102, 2011.

[348] C. Adamo and V. Barone, "Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mpw and mpw1pw models," *J. Chem. Phys.*, vol. 108, no. 2, pp. 664–675, 1998.

[349] T. Yanai, D. P. Tew, and N. C. Handy, "A new hybrid exchange–correlation functional using the coulomb-attenuating method (cam-b3lyp)," *Chemical physics letters*, vol. 393, no. 1-3, pp. 51–57, 2004.

[350] T. H. Dunning Jr, "Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen," *J. Chem. Phys.*, vol. 90, no. 2, pp. 1007–1023, 1989.

[351] Y. Mao, Y. Shao, J. Dziedzic, C.-K. Skylaris, T. Head-Gordon, and M. Head-Gordon, "Performance of the amoeba water model in the vicinity of qm solutes: A diagnosis using energy decomposition analysis," *J. Chem. Theory Comput.*, vol. 13, no. 5, pp. 1963–1979, 2017.

[352] R. M. Parrish, T. M. Parker, and C. D. Sherrill, "Chemical assignment of symmetry-adapted perturbation theory interaction energy components: The functional-group sapt partition," *J. Chem. Theory Comput.*, vol. 10, no. 10, pp. 4417–4431, 2014.

[353] R. M. Parrish, K. C. Thompson, and T. J. Martıénez, "Large-scale functional group symmetry-adapted perturbation theory on graphical processing units," *J. Chem. Theory Comput.*, vol. 14, no. 3, pp. 1737–1753, 2018.

[354] C. A. Guido, P. Cortona, B. Mennucci, and C. Adamo, "On the metric of charge transfer molecular excitations: A simple chemical descriptor," *J. Chem. Theory Comput.*, vol. 9, no. 7, pp. 3118–3126, 2013.

[355] M. Savarese, C. A. Guido, E. Bremond, I. Ciofini, and C. Adamo, "Metrics for molecular electronic excitations: A comparison between orbital-and density-based descriptors," *J. Phys. Chem. A*, vol. 121, no. 40, pp. 7543–7549, 2017.

[356] A. Baiardi, J. Bloino, and V. Barone, "General formulation of vibronic spectroscopy in internal coordinates," *J. Chem. Phys.*, vol. 144, no. 8, p. 084 114, 2016.

[357] J. Bloino, A. Baiardi, and M. Biczysko, "Aiming at an accurate prediction of vibrational and electronic spectra for medium-to-large molecules: An overview," *Int. J. Quantum Chem.*, vol. 116, no. 21, pp. 1543–1574, 2016.

[358] R. Improta, F. Santoro, V. Barone, and A. Lami, "Vibronic model for the quantum dynamical study of the competition between bright and charge-transfer excited states in single-strand polynucleotides: The adenine dimer case," *J. Phys. Chem. A*, vol. 113, no. 52, pp. 15 346–15 354, 2009.

[359] M. Barbatti, "Nonadiabatic dynamics with trajectory surface hopping method," *WIREs Comput. Mol. Sci.*, vol. 1, no. 4, pp. 620–633, 2011.

[360] M. Barbatti, A. J. Aquino, J. J. Szymczak, D. Nachtigallová, P. Hobza, and H. Lischka, "Relaxation mechanisms of uv-photoexcited dna and rna nucleobases," *Proc. Natl. Acad. Sci. USA*, vol. 107, no. 50, pp. 21 453–21 458, 2010.

[361] L. González, D. Escudero, and L. Serrano-Andrés, "Progress and challenges in the calculation of electronic excited states," *ChemPhysChem*, vol. 13, no. 1, pp. 28–51, 2012.

[362] O. Cannelli, T. Giovannini, A. Baiardi, B. Carlotti, F. Elisei, and C. Cappelli, "Understanding the interplay between the solvent and nuclear rearrangements in the negative solvatochromism of a push–pull flexible quinolinium cation," *PCCP*, vol. 19, no. 48, pp. 32 544–32 555, 2017.

[363] B. Carlotti, A. Cesaretti, O. Cannelli, T. Giovannini, C. Cappelli, C. Bonaccorso, C. G. Fortuna, F. Elisei, and A. Spalletti, "Evaluation of hyperpolarizability from the solvatochromic method: Thiophene containing push- pull cationic dyes as a case study," *J. Phys. Chem. C*, 2018.

[364] V. Barone, A. Baiardi, M. Biczysko, J. Bloino, C. Cappelli, and F. Lipparini, "Implementation and validation of a multi-purpose virtual spectrometer for large systems in complex environments," *Phys. Chem. Chem. Phys.*, vol. 14, no. 36, pp. 12 404–12 422, 2012.

[365] B. Mennucci, R. Cammi, and J. Tomasi, "Excited states and solvatochromic shifts within a nonequilibrium solvation approach: A new formulation of the integral equation formalism method at the self-consistent field, configuration interaction, and multiconfiguration self-consistent field level," *J. Chem. Phys.*, vol. 109, no. 7, pp. 2798–2807, 1998.

[366] D. Jacquemin, B. Mennucci, and C. Adamo, "Excited-state calculations with td-dft: From benchmarks to simulations in complex environments," *Phys. Chem. Chem. Phys.*, vol. 13, no. 38, pp. 16 987–16 998, 2011.

[367] L. Cupellini, M. Corbella, B. Mennucci, and C. Curutchet, "Electronic energy transfer in biomacromolecules," *WIREs Comput. Mol. Sci.*, vol. 9, no. 2, e1392, 2019.

[368] A. Puglisi, T. Giovannini, L. Antonov, and C. Cappelli, "Interplay between conformational and solvent effects in uv-visible absorption spectra: Curcumin tautomers as a case study," *Phys. Chem. Chem. Phys.*, vol. 21, pp. 15 504–15 514, 28 2019. DOI: 10.1039/C9CP00907H. [Online]. Available: http://dx.doi.org/10.1039/C9CP00907H.

[369] T. Giovannini, R. Di Remigio, L. Frediani, and C. Cappelli, *in preparation*,

[370] M. E. Casida, "Time-dependent density functional response theory for molecules," in *Recent Advances in Density Functional Methods Part I*, D. P. Chong, Ed., World Scientific, Singapore, 1995, pp. 155–192.

[371] L. Jensen, P. T. Van Duijnen, and J. G. Snijders, "A discrete solvent reaction field model for calculating molecular linear response properties in solution," *J. Chem. Phys.*, vol. 119, no. 7, pp. 3800–3809, 2003.

[372] J. Kongsted, A. Osted, K. V. Mikkelsen, and O. Christiansen, "The qm/mm approach for wavefunctions, energies and response functions within self-consistent field and coupled cluster theories," *Mol. Phys.*, vol. 100, no. 11, pp. 1813–1828, 2002.

[373] M. Caricato, B. Mennucci, J. Tomasi, F. Ingrosso, R. Cammi, S. Corni, and G. Scalmani, "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.*, vol. 124, no. 12, p. 124 520, 2006.

[374] A. V. Marenich, C. J. Cramer, D. G. Truhlar, C. A. Guido, B. Mennucci, G. Scalmani, and M. J. Frisch, "Practical computation of electronic excitation in solution: Vertical excitation model," *Chem. Sci.*, vol. 2, no. 11, pp. 2143–2161, 2011.

[375] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, "Electronic absorption spectra and solvatochromic shifts by the vertical excitation model: Solvated clusters and molecular dynamics sampling," *J. Phys. Chem. B*, vol. 119, no. 3, pp. 958–967, 2014.

[376] S. Budzak, M. Medved, B. Mennucci, and D. Jacquemin, "Unveiling solvents effect on excited-state polarizabilities with the corrected linear-response model," *J. Phys. Chem. A*, vol. 118, no. 30, pp. 5652–5656, 2014.

[377] R. Improta, V. Barone, G. Scalmani, and M. J. Frisch, "A state-specific polarizable continuum model time dependent density functional theory method for excited state calculations in solution," *J. Chem. Phys.*, vol. 125, no. 5, p. 054 103, 2006.

[378] M. Caricato, "A corrected-linear response formalism for the calculation of electronic excitation energies of solvated molecules with the ccsd-pcm method," *Comput. Theor. Chem.*, vol. 1040, pp. 99–105, 2014.

[379] C. A. Guido, B. Mennucci, G. Scalmani, and D. Jacquemin, "Excited state dipole moments in solution: Comparison between state-specific and linear-response td-dft values," *J. Chem. Theory Comput.*, vol. 14, no. 3, pp. 1544–1553, 2018.

[380] H. Schr oder and T. Schwabe, "Corrected polarizable embedding: Improving the induction contribution to perichromism for linear response theory," *J. Chem. Theory Comput.*, vol. 14, no. 2, pp. 833–842, 2018.

[381] C. A. Guido and S. Caprasecca, "On the description of the environment polarization response to electronic transitions," *Int. J. Quantum Chem.*, vol. 119, no. 1, e25711, 2019.

[382] C. A. Guido, G. Scalmani, B. Mennucci, and D. Jacquemin, "Excited state gradients for a state-specific continuum solvation approach: The vertical excitation model within a lagrangian tddft formulation," *J. Chem. Phys.*, vol. 146, no. 20, p. 204 106, 2017.

[383] C. A. Guido, D. Jacquemin, C. Adamo, and B. Mennucci, "Electronic excitations in solution: The interplay between state specific approaches and a time-dependent density functional theory description," *J. Chem. Theory Comput.*, vol. 11, no. 12, pp. 5782–5790, 2015.

[384] Q. Zeng and W. Liang, "Analytic energy gradient of excited electronic state within tddft/mmpol framework: Benchmark tests and parallel implementation," *J. Chem. Phys.*, vol. 143, no. 13, p. 134 104, 2015.

[385] S. Sok, S. Y. Willow, F. Zahariev, and M. S. Gordon, "Solvent-induced shift of the lowest singlet $\pi \rightarrow \pi^*$charge-transfer excited state of p-nitroaniline in water: An application of the tddft/efp1 method," *J. Phys. Chem. A*, vol. 115, no. 35, pp. 9801–9809, 2011.

[386] D. Kosenkov and L. V. Slipchenko, "Solvent effects on the electronic transitions of p-nitroaniline: A qm/efp study," *J. Phys. Chem. A*, vol. 115, no. 4, pp. 392–401, 2010.

[387] K. Sneskov, T. Schwabe, O. Christiansen, and J. Kongsted, "Scrutinizing the effects of polarization in qm/mm excited state calculations," *Phys. Chem. Chem. Phys.*, vol. 13, no. 41, pp. 18 551–18 560, 2011.

[388] J. J. Eriksen, S. P. Sauer, K. V. Mikkelsen, H. J. A. Jensen, and J. Kongsted, "On the importance of excited state dynamic response electron correlation in polarizable embedding methods," *J. Comput. Chem.*, vol. 33, no. 25, pp. 2012–2022, 2012.

[389] S. Frutos-Puerto, M. A. Aguilar, and I. Fdez. Galv'an, "Theoretical study of the preferential solvation effect on the solvatochromic shifts of para-nitroaniline," *J. Phys. Chem. B*, vol. 117, no. 8, pp. 2466–2474, 2013.

[390] A. DeFusco, N. Minezawa, L. V. Slipchenko, F. Zahariev, and M. S. Gordon, "Modeling solvent effects on electronic excited states," *J. Phys. Chem. Lett.*, vol. 2, no. 17, pp. 2184–2192, 2011.

[391] B. Mennucci, "Hydrogen bond versus polar effects: An ab initio analysis on n $\rightarrow$ π*absorption spectra and n nuclear shieldings of diazines in solution," *J. Am. Chem. Soc.*, vol. 124, no. 7, pp. 1506–1515, 2002.

[392] M. Pagliai, G. Mancini, I. Carnimeo, N. De Mitri, and V. Barone, "Electronic absorption spectra of pyridine and nicotine in aqueous solution with a combined molecular dynamics and polarizable qm/mm approach," *J. Comput. Chem.*, vol. 38, no. 6, pp. 319–335, 2017.

[393] M. Biczysko, J. Bloino, G. Brancato, I. Cacelli, C. Cappelli, A. Ferretti, A. Lami, S. Monti, A. Pedone, G. Prampolini, *et al.*, "Integrated computational approaches for spectroscopic studies of molecular systems in the gas phase and in solution: Pyrimidine as a test case," *Theoretical Chemistry Accounts*, vol. 131, no. 4, p. 1201, 2012.

[394] M. Cossi and V. Barone, "Time-dependent density functional theory for molecules in liquid solutions," *J. Chem. Phys.*, vol. 115, no. 10, pp. 4708–4717, 2001.

[395] S. Mason, "The electronic spectra of n-heteroaromatic systems. part i. the n$\rightarrow$ π transitions of monocyclic azines," *J. Chem. Soc.*, pp. 1240–1246, 1959.

[396] S. Millefiori, G. Favini, A. Millefiori, and D. Grasso, "Electronic spectra and structure of nitroanilines," *Spectrochim. Acta A*, vol. 33, no. 1, pp. 21–27, 1977.

[397] Z.-L. Cai and J. R. Reimers, "The low-lying excited states of pyridine," *J. Phys. Chem. A*, vol. 104, no. 36, pp. 8389–8408, 2000.

[398] M. Schreiber, M. R. Silva-Junior, S. P. Sauer, and W. Thiel, "Benchmarks for electronically excited states: Caspt2, cc2, ccsd, and cc3," *J. Chem. Phys.*, vol. 128, no. 13, p. 134 110, 2008.

[399] F. F. da Silva, D. Almeida, G. Martins, A. Milosavljević, B. Marinković, S. V. Hoffmann, N. Mason, Y. Nunes, G. Garcia, and P. Limao-Vieira, "The electronic states of pyrimidine studied by vuv photoabsorption and electron energy-loss spectroscopy," *Phys. Chem. Chem. Phys.*, vol. 12, no. 25, pp. 6717–6731, 2010.

[400] L. Prabhumirashi and S. Kunte, "Solvent effects on electronic absorption spectra of nitrochlorobenzenes, nitrophenols and nitroanilines–i. studies in nonpolar solvents," *Spectrochim. Acta A*, vol. 42, no. 4, pp. 435–441, 1986.

[401] K. J. de Almeida, K. Coutinho, W. B. De Almeida, W. R. Rocha, and S. Canuto, "A monte carlo–quantum mechanical study of the solvatochromism of pyrimidine in water and in carbon tetrachloride," *Phys. Chem. Chem. Phys.*, vol. 3, no. 9, pp. 1583–1587, 2001.

[402] M. R. Silva-Junior, M. Schreiber, S. P. Sauer, and W. Thiel, "Benchmarks of electronically excited states: Basis set effects on caspt2 results," *J. Chem. Phys.*, vol. 133, no. 17, p. 174 318, 2010.

[403] T. Giovannini, M. Ambrosetti, and C. Cappelli, "Quantum confinement effects on solvatochromic shifts of molecular solutes," *J. Phys. Chem. Lett.*, vol. 10, no. 19, pp. 5823–5829, 2019, PMID: 31518133. DOI: 10.1021/acs.jpclett.9b02318. eprint: https://doi.org/10.1021/acs.jpclett.9b02318. [Online]. Available: https://doi.org/10.1021/acs.jpclett.9b02318.

[404] H.-H. Perkampus Ed., *UV-VIS atlas of organic compounds*. Wiley, Weinheim, 1996.

[405] F. Maschietto, M. Campetella, M. J. Frisch, G. Scalmani, C. Adamo, and I. Ciofini, "How are the charge transfer descriptors affected by the quality of the underpinning electronic density?" *J. Comput. Chem.*, vol. 39, no. 12, pp. 735–742, 2018.

[406] J. Bloino and V. Barone, "A second-order perturbation theory route to vibrational averages and transition properties of molecules: General formulation and application to infrared and vibrational circular dichroism spectroscopies," *J. Chem. Phys.*, vol. 136, no. 12, p. 124 108, 2012.

[407] J. Bloino, M. Biczysko, and V. Barone, "Anharmonic effects on vibrational spectra intensities: Infrared, raman, vibrational circular dichroism, and raman optical activity," *J. Phys. Chem. A*, vol. 119, no. 49, pp. 11 862–11 874, 2015.

[408] B. Mennucci and J. M. Martıénez, "How to model solvation of peptides? insights from a quantum-mechanical and molecular dynamics study of n-methylacetamide. 1. geometries, infrared, and ultraviolet spectra in water," *J. Phys. Chem. B*, vol. 109, no. 19, pp. 9818–9829, 2005.

[409] B. Mennucci, R. Cammi, and J. Tomasi, "Analytical free energy second derivatives with respect to nuclear coordinates: Complete formulation for electrostatic continuum solvation models," *J. Chem. Phys.*, vol. 110, no. 14, pp. 6858–6870, 1999.

[410] M. Cossi and V. Barone, "Analytical second derivatives of the free energy in solution by polarizable continuum models," *J. Chem. Phys.*, vol. 109, no. 15, pp. 6246–6254, 1998.

[411] J. Huang, Y. Mei, G. König, A. C. Simmonett, F. C. Pickard IV, Q. Wu, L.-P. Wang, A. D. MacKerell Jr, B. R. Brooks, and Y. Shao, "An estimation of hybrid quantum mechanical molecular mechanical polarization energies for small molecules using polarizable force-field approaches," *J. Chem. Theory Comput.*, vol. 13, no. 2, pp. 679–695, 2017.

[412] T. Giovannini, R. R. Riso, M. Ambrosetti, A. Puglisi, and C. Cappelli, "Electronic Transitions for a Fully Polarizable QM/MM Approach Based on Fluctuating Charges and Fluctuating Dipoles: Linear and Corrected Linear Response Regimes," *arXiv e-prints*, arXiv:1906.03852, arXiv:1906.03852, Jun. 2019. arXiv: 1906.03852 [physics.chem-ph].

[413] Q. Cui and M. Karplus, "Molecular properties from combined qm/mm methods. i. analytical second derivative and vibrational calculations," *J. Chem. Phys.*, vol. 112, no. 3, pp. 1133–1149, 2000. DOI: 10.1063/1.480658. [Online]. Available: http://link.aip.org/link/?JCP/112/1133/1.

[414] H. Li and M. S. Gordon, "Polarization energy gradients in combined quantum mechanics, effective fragment potential, and polarizable continuum model calculations," *J. Chem. Phys*, vol. 126, no. 12, 124112, p. 124 112, 2007. DOI: 10.1063/1.2711199. [Online]. Available: http://link.aip.org/link/?JCP/126/124112/1.

[415] M. A. Thompson, "Qm/mmpol: A consistent model for solute/solvent polarization. application to the aqueous solvation and spectroscopy of formaldehyde, acetaldehyde, and acetone," *J. Phys. Chem.*, vol. 100, no. 34, pp. 14 492–14 507, 1996.

[416] D. Loco, L. Lagardère, S. Caprasecca, F. Lipparini, B. Mennucci, and J.-P. Piquemal, "Hybrid qm/mm molecular dynamics with amoeba polarizable embedding," *J. Chem. Theory Comput.*, vol. 13, no. 9, pp. 4025–4033, 2017.

[417] J. Dziedzic, Y. Mao, Y. Shao, J. Ponder, T. Head-Gordon, M. Head-Gordon, and C.-K. Skylaris, "Tinktep: A fully self-consistent, mutually polarizable qm/mm approach based on the amoeba force field," *J. Chem. Phys.*, vol. 145, no. 12, p. 124 106, 2016.

[418] G. Yang and Y. Xu, "Probing chiral solute-water hydrogen bonding networks by chirality transfer effects: A vibrational circular dichroism study of glycidol in water.," *J. Chem. Phys.*, vol. 130, no. 16, pp. 164 506–164 506, 2009.

[419] R. Kawiecki, F. Devlin, P. Stephens, and R. Amos, "Vibrational circular dichroism of propylene oxide," *J. Phys. Chem*, vol. 95, no. 24, pp. 9817–9831, 1991.

[420] S. Jin and J. D. Head, "Theoretical investigation of molecular water adsorption on the al (111) surface," *Surf. Science*, vol. 318, no. 1, pp. 204–216, 1994.

[421] M. D. Calvin, J. D. Head, and S. Jin, "Theoretically modelling the water bilayer on the al (111) surface using cluster calculations," *Surf. science*, vol. 345, no. 1, pp. 161–172, 1996.

[422] A. Biancardi, R. Cammi, C. Cappelli, B. Mennucci, and J. Tomasi, "Modelling vibrational coupling in DNA oligomers: A computational strategy combining QM and continuum solvation models," *Theor. Chem. Acc.*, vol. 131, no. 3, pp. 1–10, 2012.

[423] M. Frisch, M. Head-Gordon, and J. Pople, "Direct analytic SCF second derivatives and electric field properties," *Chem. Phys.*, vol. 141, no. 2, pp. 189–196, 1990.

[424] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," *J. Comput. Chem.*, vol. 25, no. 9, pp. 1157–1174, 2004, ISSN: 1096-987X. DOI: 10.1002/jcc.20035. [Online]. Available: http://dx.doi.org/10.1002/jcc.20035.

[425] J. Wang, W. Wang, P. A. Kollman, and D. A. Case, "Automatic atom type and bond type perception in molecular mechanical calculations," *J. Mol. Graph. Model.*, vol. 25, no. 2, pp. 247–260, 2006.

[426] A. V. Marenich, S. V. Jerome, C. J. Cramer, and D. G. Truhlar, "Charge model 5: An extension of hirshfeld population analysis for the accurate description of molecular interactions in gaseous and condensed phases," *J. Chem. Theory Comput.*, vol. 8, no. 2, pp. 527–541, 2012, PMID: 26596602. DOI: 10.1021/ct200866d. eprint: http://dx.doi.org/10.1021/ct200866d. [Online]. Available: http://dx.doi.org/10.1021/ct200866d.

[427] V. Barone, M. Biczysko, J. Bloino, and C. Puzzarini, "Accurate molecular structures and infrared spectra of trans-2, 3-dideuterooxirane, methyloxirane, and trans-2, 3-dimethyloxirane," *J. Chem. Phys.*, vol. 141, no. 3, p. 034 107, 2014.

[428] M. Hodecker, M. Biczysko, A. Dreuw, and V. Barone, "Simulation of vacuum uv absorption and electronic circular dichroism spectra of methyl oxirane: The role of vibrational effects," *J. Chem. Theory Comput.*, vol. 12, no. 6, pp. 2820–2833, 2016.

[429] K. Ruud and R. Zanasi, "The importance of molecular vibrations: The sign change of the optical rotation of methyloxirane," *Angew. Chem. Int. Ed.*, vol. 117, no. 23, pp. 3660–3662, 2005.

[430] M. C. Tam, N. J. Russ, and T. D. Crawford, "Coupled cluster calculations of optical rotatory dispersion of (s)-methyloxirane," *J. Chem. Phys.*, vol. 121, no. 8, pp. 3550–3557, 2004.

[431] P. Mukhopadhyay, G. Zuber, M.-R. Goldsmith, P. Wipf, and D. N. Beratan, "Solvent effect on optical rotation: A case study of methyloxirane in water," *Chem. Phys. Chem*, vol. 7, no. 12, pp. 2483–2486, 2006.

[432] M. Carnell, S. Peyerimhoff, A. Breest, K. Gödderz, P. Ochmann, and J. Hormes, "Experimental and quantum-theoretical investigation of the circular dichroism spectrum of r-methyloxirane," *Chem. Phys. Lett.*, vol. 180, no. 5, pp. 477–481, 1991.

[433] B. A. McGuire, P. B. Carroll, R. A. Loomis, I. A. Finneran, P. R. Jewell, A. J. Remijan, and G. A. Blake, "Discovery of the interstellar chiral molecule propylene oxide (ch3chch2o)," *Science*, vol. 352, no. 6292, pp. 1449–1452, 2016.

[434] W. Sun, J. Wu, B. Zheng, Y. Zhu, and C. Liu, "Dft study of vibrational circular dichroism spectra of (s)-glycidol–water complexes," *J. Mol. Struct.: THEOCHEM*, vol. 809, no. 1, pp. 161–169, 2007.

[435] T. Giovannini, L. Grazioli, M. Ambrosetti, and C. Cappelli, "Calculation of ir spectra with a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles," *Journal of chemical theory and computation*, vol. 15, no. 10, pp. 5495–5507, 2019.

[436] M. Losada, P. Nguyen, and Y. Xu, "Solvation of propylene oxide in water: Vibrational circular dichroism, optical rotation, and computer simulation studies," *J. Phys. Chem A*, vol. 112, no. 25, pp. 5621–5627, 2008.

[437] M. Losada, H. Tran, and Y. Xu, "Lactic acid in solution: Investigations of lactic acid self-aggregation and hydrogen bonding interactions with water and methanol using vibrational absorption and vibrational circular dichroism spectroscopies," *J. Chem. Phys.*, vol. 128, no. 1, p. 014 508, 2008.

[438] A. Edelmann and B. Lendl, "Toward the optical tongue: Flow-through sensing of tannin- protein interactions based on ftir spectroscopy," *J. Am. Chem. Soc.*, vol. 124, no. 49, pp. 14 741–14 747, 2002.

[439] C. Reichardt, "Solvatochromism, thermochromism, piezochromism, halochromism, and chiro-solvatochromism of pyridinium n-phenoxide betaine dyes," *Chem. Soc. Rev.*, vol. 21, no. 3, pp. 147–153, 1992.

[440] Y.-l. Lin and J. Gao, "Solvatochromic shifts of the n $\rightarrow$ π*transition of acetone from steam vapor to ambient aqueous solution: A combined configuration interaction qm/mm simulation study incorporating solvent polarization," *J. Chem. Theory Comput.*, vol. 3, no. 4, pp. 1484–1493, 2007.

[441] C. A. Guido, S. Knecht, J. Kongsted, and B. Mennucci, "Benchmarking time-dependent density functional theory for excited state geometries of organic molecules in gas-phase and in solution," *J. Chem. Theory Comput.*, vol. 9, no. 5, pp. 2209–2220, 2013.

[442] T. J. Zuehlsdorff and C. M. Isborn, "Modeling absorption spectra of molecules in solution," *Int. J. Quantum Chem.*, vol. 119, no. 1, e25719, 2019.

[443] A. Moskvin, O. Yablonskii, and L. Bondar, "An experimental investigation of the effect of alkyl substituents on the position of the k and r absorption bands in acrolein derivatives," *Theor. Exp. Chem.*, vol. 2, no. 5, pp. 469–472, 1966.

[444] K. Aidas, A. Møgelhøj, E. J. Nilsson, M. S. Johnson, K. V. Mikkelsen, O. Christiansen, P. Söderhjelm, and J. Kongsted, "On the performance of quantum chemical methods to predict solvatochromic effects: The case of acrolein in aqueous solution," *J. Chem. Phys.*, vol. 128, no. 19, p. 194 503, 2008.

[445] F. Aquilante, V. Barone, and B. O. Roos, "A theoretical investigation of valence and rydberg electronic states of acrolein," *J. Chem. Phys.*, vol. 119, no. 23, pp. 12 323–12 334, 2003.

[446] G. Brancato, N. Rega, and V. Barone, "A quantum mechanical/molecular dynamics/mean field study of acrolein in aqueous solution: Analysis of h bonding and bulk effects on spectroscopic properties," *J. Chem. Phys.*, vol. 125, no. 16, p. 164 515, 2006.

[447] C. Bistafa, L. Modesto-Costa, and S. Canuto, "A complete basis set study of the lowest n–π*and π–π*electronic transitions of acrolein in explicit water environment," *Theor. Chem. Acc.*, vol. 135, no. 5, p. 129, 2016.

[448] H. C. Georg, K. Coutinho, and S. Canuto, "A sequential monte carlo quantum mechanics study of the hydrogen-bond interaction and the solvatochromic shift of the n–π*transition of acrolein in water," *J. Chem. Phys.*, vol. 123, no. 12, p. 124 307, 2005.

[449] S. A. do Monte, T. Müller, M. Dallos, H. Lischka, M. Diedenhofen, and A. Klamt, "Solvent effects in electronically excited states using the continuum solvation model cosmo in combination with multireference configuration interaction with singles and doubles (mr-cisd)," *Theor. Chem. Acc.*, vol. 111, no. 2-6, pp. 78–89, 2004.

[450] M. E. Martıén, A. Muñoz Losa, I. Fdez.-Galván, and M. Aguilar, "A theoretical study of solvent effects on the 1 (n $\rightarrow \pi^*$) electron transition in acrolein," *J. Chem. Phys.*, vol. 121, no. 8, pp. 3710–3716, 2004.

[451] A. Muñoz Losa, I. Fdez. Galván, M. A. Aguilar, and M. E. Martıén, "A caspt2//casscf study of vertical and adiabatic electron transitions of acrolein in water solution," *J. Phys. Chem. B*, vol. 111, no. 33, pp. 9864–9870, 2007.

[452] R. Fukuda, M. Ehara, H. Nakatsuji, and R. Cammi, "Nonequilibrium solvation for vertical photoemission and photoabsorption processes using the symmetry-adapted cluster–configuration interaction method in the polarizable continuum model," *J. Chem. Phys.*, vol. 134, no. 10, p. 104 109, 2011.

[453] H. Ren, M. Ming, J. Zhu, J. Ma, and X. Li, "Solvent effect on uv/vis absorption and emission spectra in aqueous solution based on a modified form of solvent reorganization energy," *Chem. Phys. Lett.*, vol. 583, pp. 213–217, 2013.

[454] C. Amovilli and B. Mennucci, "Self-consistent-field calculation of pauli repulsion and dispersion contributions to the solvation free energy in the polarizable continuum model," *J. Phys. Chem. B*, vol. 101, no. 6, pp. 1051–1057, 1997.

[455] M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys, and W. J. Stevens, "The effective fragment potential method: A qm-based mm approach to modeling environmental effects in chemistry," *J. Phys. Chem. A*, vol. 105, no. 2, pp. 293–307, 2001.

[456] H. Gokcan, E. G. Kratz, T. A. Darden, J.-P. Piquemal, and G. A. Cisneros, "Qm/mm simulations with the gaussian electrostatic model, a density-based polarizable potential," *J. Phys. Chem. Lett.*, vol. 9, pp. 3062–3067, 2018.

[457] T. Giovannini, L. Grazioli, M. Ambrosetti, and C. Cappelli, "On the Calculation of IR Spectra with a Fully Polarizable QM/MM Approach Based on Fluctuating Charges and Fluctuating Dipoles," *J. Chem. Theory Comput.*, DOI: 10.1021/acs.jctc.9b00574, 2019.

[458] J. Chen and T. J. Martıénez, "Qtpie: Charge transfer with polarization current equalization. a fluctuating charge model with correct asymptotics," *Chem. Phys. Lett.*, vol. 438, no. 4, pp. 315–320, 2007.