Scuola Normale Superiore
Classe di Scienze
Corso di Perfezionamento in Methods and Models for Molecular Sciences

Development of Fully Atomistic Approaches to Model Response Properties of Complex Systems

Supervisor:
Prof. Chiara Cappelli

Tommaso Giovannini

To my family

Contents

Papers

Attached

1. Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution.
T. Giovannini, G. Del Frate, P. Lafiosca, C. Cappelli, *Phys. Chem. Chem. Phys.*, **2018**, *20*, 9181-9197

2. Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling
T. Giovannini, M. Macchiagodena, M. Ambrosetti, A. Puglisi, P. Lafiosca, G. Lo Gerfo, F. Egidi, C. Cappelli, *Int. J. Quantum Chem.*, **2018**, e25684

3. A polarizable embedding approach to second harmonic generation (SHG) of molecular systems in aqueous solutions
T. Giovannini, M. Ambrosetti, C. Cappelli, *Theor. Chem. Acc.*, **2018**, *137*, 74

4. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches
T. Giovannini, P. Lafiosca, C. Capelli, *J. Chem. Theory Comput.*, **2017**, *13*, 4854-4870

5. Effective yet Reliable Computation of EPR Spectra in Solution by a QM/MM Approach: Interplay Between Electrostatics and Non-electrostatic Effects
T. Giovannini, P. Lafiosca, B. Chandramouli, V. Barone, C. Cappelli, *submitted*

6. Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQFμ model
T. Giovannini, A. Puglisi, M. Ambrosetti, C. Cappelli, *submitted*

7. A Classical Picture of Subnanometer Junctions: an Atomistic Drude Approach to Nanoplasmonics
T. Giovannini, M. Rosa, S. Corni, C. Cappelli, *submitted*

Other papers during the Ph.D. internship

8. Effective fully polarizable QM/MM approach to model vibrational circular dichroism spectra of systems in aqueous solution
T. Giovannini, M. Olszowka, C. Cappelli, *J. Chem. Theory Comput.*, **2016**, *12*, 5483-5492

9. Can the resonance raman optical activity display sign alternation?
L. Vidal, **T. Giovannini**, C. Cappelli, *J. Phys. Chem. Lett.*, **2016**, *7*, 3585-3590

10. Polarizable Embedding Approach for the Analytical Calculation of Raman and Raman Optical Activity Spectra of Solvated Systems
T. Giovannini, M. Olszowka, F. Egidi, J. R. Cheeseman, G. Scalmani, C. Cappelli, *J. Chem. Theory Comput.*, **2017**, *13*, 4421-4435

Chapter 1

Introduction

This thesis aims at developing a theoretical framework, based on a multiscale QM/MM description, to describe the energy and response properties of complex molecular systems. Complexity in theoretical chemistry can arise from several aspects. First, complexity can be related to the complexity of the property under investigation, meaning the difficulty in defining effective algorithms to calculate it. Second, complexity can arise from the large number of atoms/molecules that need to be considered in order to have a reliable representation of the physico-chemical phenomenon. This is for instance the case of molecular systems in the condensed phase or embedded in a biological matrix, in which an interplay between the description of the solvent/embedding environment and an adequate sampling of all possible configurations is mandatory. Finally, complexity can arise from the strong interaction between light and environment, as in the case of nanoparticles/nanoaggregates. The peculiarity of such environments stands in the fact that they can enhance by orders of magnitude the resonant electric field near the surface. Even more complexity arises when the above mentioned phenomena are combined together, as for instance in the case of high order molecular properties of molecular systems embedded in a complex environment. Therefore, in order to obtain an appropriate physico-chemical description of the target phenomena, an effective theoretical model able to accurately treat all players is needed. This is exactly the aim of the present work.

The problem of describing the interaction between a molecule and its embedding environment is one of the pillars of Quantum Chemistry. The interplay between the molecule and the environment can in fact dramatically alter both the structure and the molecular response to external electromagnetic fields.

An effective theoretical modeling of molecular phenomena in external environments needs to resort to Quantum-Mechanical (QM) descriptions. If all the atoms/molecules were treated at the QM level, several issues would arise, due to the large number of degrees of freedom that need to be considered. For these reasons, to describe the energy and spectroscopic/response properties of molecules in the condensed phase at the same level of accuracy as isolated systems is a particularly challenging task. Due to the huge size of the whole system, any attempt to use the same QM approaches adopted for isolated system would be unrealistic. Even if the QM calculation on the

whole system were computationally affordable, the huge number of degrees of freedom of the environment would give rise to several configurations which would need to be taken into consideration to achieve a reliable sampling of the phase space. However, since the molecular properties are local properties of the molecule, which are modified but not determined by the presence of the environment, "brute force" approaches are usually unwarranted, because most of the computational effort would be devoted to the simulation of the properties of the environment itself, which are not those determining the system's signals.

The most successful answer to this problem has been found within the realm of multi-scale approaches and focused models:[1–6] 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 properties are local properties of the molecule, which are modified but not determined by the presence of the environment. The most renowned focused models belong to the family of QM/Classical approaches, in which the classical portion ranges from an an atomistic description (giving rise to QM/MM models[1,2,7–63] to a blurred continuum description.[3–6,64–82] These models have had great success in modern chemical research, because they can be effectively coupled with most QM descriptions, ranging from semi-empirical methods to Density Functional Theory (DFT) or Wavefunction Theory, without increasing the computational costs with respect the corresponding QM calculation of the isolated molecule. This is due to the fact that such models limit the number of degrees of freedom to be treated in the QM calculation to those of the the QM portion, without a substantial increase in the dimension of the QM problem. Also, by introducing the environment-related terms in the molecular Hamiltonian, the machinery of Quantum Chemistry can be exploited to obtain the desired spectral signals in the same way as they are calculated for isolated systems, with the addition of extra terms to be introduced in the formulation of energy derivatives and response equations.[56,64,67,83] The difference among the various types of focused models mainly lies in the specific structure of the environment-related extra terms. Most such models focus on reproducing the electrostatic component of the system-environment interactions, which is in many cases the most important term.

The Polarizable Continuum Model (PCM),[3–6] which belongs to the family of QM/-Continuum approches, has been particularly successful. In such an approach, the QM described target molecule is accommodated into a molecular cavity, whose shape depends on the molecular geometry.[69] The environment is instead described as an homogeneous, continuum dielectric with given dielectric properties, which characterize the environment. The QM electron density and the dielectric mutually polarize. In such a framework, the molecule-environment interaction is restrained to the electrostatic interaction, which is of course an approximation, although it is the main force in several systems. Methods to include non-electrostatic terms, such as repulsion and dispersion, both empirical and QM based approaches have been proposed.[84–88] QM/PCM model has been particularly successful because it has been extended to almost all the computational spectroscopies,[28,64,71,72,74,82,83,89–132] and it has been coupled to various QM descriptions.[119,133–141] Remarkably, in QM/PCM all possible environment

configurations are implicitly taken into consideration, and thus no statistical average is needed. This has a practical consequence: a QM/PCM study requires the same number of calculations as for the corresponding isolated system. On the other hand, since all the information about the structure of the environment is neglected, the specific molecule-environment interactions (e.g. Hydrogen Bonding), cannot be described.
In order to recover the atomistic description of the environment, QM/Molecular Mechanics is exploited,[1,2,7–62,142–145] where the target molecule is described at the QM level, whereas the environment is described by resorting to MM force fields. The fully atomistic description of QM/MM approaches is not gained for free: in fact, whereas QM/PCM implicitly includes the statistical average of the possible configurations of the environment, QM/MM approaches need an explicit sampling of the phase space. Such a sampling is usually carried out by firstly performing a Molecular Dynamics simulation, and then extracting some uncorrelated snapshots, whose number can vary depending on the property under consideration, until convergence is reached.[54,58,146–155]
In the past years, several hierarchical ways to couple the QM and MM portions have been proposed:

1. *Mechanical embedding*: the QM/MM coupling term is treated at the MM level only, therefore it is independent of the QM electronic density. Such a coupling is expressed in terms of dispersion-repulsion potentials (usually, the Lennard-Jones potential) and classical electrostatics. The contribution of the surrounding environment to the energy and properties of the QM portion is thus only indirect.

2. *Electrostatic embedding*: the MM atoms are endowed with fixed atomic charges, that produce an electric potential which polarizes the QM electronic density. From the point of view of the QM solute, the electrostatic embedding introduces a new term in the molecular Hamiltonian, i.e. the interaction between the potential generated by the MM fixed charges and the electron density:

$$\mathscr{H}_{int} = \sum_{j=1}^{N_q} q_i V_{QM}(\mathbf{r}_i)$$

   where the sum runs over the N_q MM charges q_i. $V_{QM}(\mathbf{r}_i)$ is the QM potential calculated at charge positions $\mathbf{r}_i$. Dispersion - repulsion contributions are usually considered by resorting to classical potentials. The actual quality of the final results crucially depends on a proper choice of values of the fixed charges representing the MM portion.

3. *Polarizable embedding (PE)*: the mutual polarization between MM and QM portions is explicitly taken into account. In fact, the MM force field contains a *response* term, which modifies the electrostatics as a result of the presence of the QM density. A polarization term is included in the QM Hamiltonian to represent the interaction of the electronic density with the MM portion. Dispersion - repulsion contributions are usually considered by means of classical potentials.
   Several QM/PE schemes have been proposed in the literature. In the Fluctuating Charge (FQ) force field,[156–159] each MM atom is endowed with an atomic charge,

which can vary according to the Electronegativity Equalization Principle (EEP), which states that at the equilibrium each atom must have the same electronegativity.[160] In the resulting QM/FQ approach,[54–56,146–153,161–163] (see Ref.[56] for a recent review) the atomic charges on each MM atom vary according to the differences in electronegativity and to the differences in the values of the QM potential calculated at charge positions. Thus, the QM Hamiltonian is modified as:

$$\mathscr{H}_{int} = \sum_{j=1}^{N_q} q_i V_{QM}(\mathbf{r}_i)$$

where, differently from electrostatic embedding, charge values are obtained by solving a linear equation, in which the right hand side depends on the QM density:

$$\mathbf{Dq} = \mathbf{C}(\rho)$$

where $\mathbf{D}$ is the response matrix.

In the induced point dipole model (ID, or Thole method)[10,35,37,39,155,164–169] mutual polarization effects are introduced in terms of induced point dipoles placed in the MM region. In this case the electric field produced by the QM density enters the equations that determine the induced dipoles. The latter, together with a set of fixed charges define the interaction term in the QM Hamiltonian.

$$\mathscr{H}_{int} = \sum_{j=1}^{N_q} q_i V_{QM}(\mathbf{r}_i) - \sum_{k=1}^{N_\mu} \boldsymbol{\mu}_k \cdot \mathbf{E}_{QM}(\mathbf{r}_k)$$

where, N_μ is the number of induced dipoles $\boldsymbol{\mu}$ in the MM region. $\mathbf{E}_{QM}(\mathbf{r}_i)$ is the QM electric field calculated at dipole positions $\mathbf{r}_k$. The induced dipoles $\boldsymbol{\mu}$ are obtained by solving a linear equation, in which the right hand side contains the electric field produced by the QM density:

$$\mathbf{T}\boldsymbol{\mu} = \mathbf{K}(\rho)$$

where $\mathbf{T}$ is the dipole response matrix, and $\mathbf{K}(\rho)$ is the QM density-dependent right hand side.

An alternative approach is represented by the Drude Oscillator (DO) model,[59,170–176] which represents the induced electric dipole at each polarizable MM site in terms of a couple of charges, of the same magnitude and opposite sign, linked by a harmonic spring. The first charge is located at the nucleus of the MM atom, while the second one is mobile, so that polarization arises from the competition between the forces acting on the charges, which are due to the harmonic potential, and the electrostatic interactions with the remaining environment. In this model the molecular Hamiltonian is modified as:

$$\mathscr{H}_{int} = \sum_{j=1}^{N_q} q_i V_{QM}(\mathbf{r}_i) + \sum_{k}^{N_Z} Z_k V_{QM}(\mathbf{r}_k) - \sum_{l}^{N_Z} Z_l V_{QM}(\mathbf{r}_l)$$

where Z_k and Z_l are the two fixed opposite charges connected by the harmonic spring. In such an equation, the positions of the negative mobile charges $(Z_l, \mathbf{r}_l)$ are found by imposing the total force acting on them to be set to zero:

$$\mathbf{F}_{\text{tot}} = \mathbf{F}_{\text{spring}} + \mathbf{F}_{\text{elec}} + \mathbf{F}(\rho_{QM}) = 0$$

where $\mathbf{F}_{\text{spring}}$ are the forces due to the springs, $\mathbf{F}_{\text{elec}}$ are electrostatic forces arising from other MM charges, and $\mathbf{F}(\rho_{\text{QM}})$ is the force due to the QM density. Differently from FQ and ID, QM/DO equations are by definition non-linear.[59,175–177] Remarkably, DO can be mapped onto ID, thus resulting in the same electrostatic description.[178]

Building upon the aforementioned models, the AMOEBA force field[58,154,179–184] is based on the inclusion of both permanent and induced multipolar terms in the electrostatic term. Molecular polarization is achieved via a Thole-style damped interactive induction model based upon distributed atomic polarizabilities. In particular, charges q, induced dipoles $\boldsymbol{\mu}$ and quadrupoles $\boldsymbol{\Theta}$ are considered, so that the QM Hamiltonian is modified as:

$$\mathscr{H}_{int} = \sum_{j=1}^{N_q} q_i V_{QM}(\mathbf{r}_i) - \sum_{k=1}^{N_\mu} \boldsymbol{\mu}_k \cdot \mathbf{E}_{QM}(\mathbf{r}_k) - \sum_{l=1}^{N_{quad}} \boldsymbol{\Theta}_l \cdot \nabla \mathbf{E}_{QM}(\mathbf{r}_l)$$

where $\nabla \mathbf{E}_{QM}(\mathbf{r}_l)$ is the electric field gradient calculated at quadrupole positions $\mathbf{r}_l$. Again, polarization is described by letting atomic dipoles $\boldsymbol{\mu}$ vary as a response of the electric field generated by the QM density:

$$\mathbf{T}\boldsymbol{\mu} = \mathbf{K}(\rho)$$

where $\mathbf{T}$ is the dipole response matrix, and $\mathbf{K}(\rho)$ is the QM density-dependent right hand side.

4. *QM-based embedding*: the MM force field is constructed on the basis of explicit QM calculations on the environment. In such approaches, all the interactions, i.e. electrostatics, induction, repulsion and dispersion, can be included in the QM Hamiltonian.

   The most used approach is the QM/Effective Fragment Potential (EFP) model,[185–190] which constructs the MM force field employing QM quantities calculated by using localized molecular orbitals of the MM portion. Such a model is usually used to calculate the energy of the system, however it has also been extended to gradients and UV-VIS spectroscopy.[188,190]

A similar QM-based approach, namely the Polarizable Density Embedding (PDE), has been recently proposed to calculate UV-VIS spectra.[60–62] Differently from QM/EFP, this model does not include QM/MM dispersion contributions.

Expanding the QM/FQ approach

Most of the aforementioned Polarizable Embedding (PE) approaches are designed and employed for reproducing structural properties and only few of them are able to treat electronic excitations, or magnetic perturbations.[10,43,58,165,167–169,191] The only currently QM/PE approach specifically designed for response properties and computational spectroscopy is the QM/Fluctuating Charge (FQ) approach, which has been developed and extended to the calculation of several spectroscopic and response properties in our group.[54–56,146–153,161–163] Such an extension follows from its variational formalism.[54] In this thesis QM/FQ is further developed and tested to the calculation of Vibrational Optical Activity (VOA) spectra of (L)-Methyl Lactate and (S)-Glycidol in aqueous solution (**Chapter 2**).[151] In **Chapter 3**, the performances of QM/PCM, the QM/FQ and hybrid approaches are analyzed with respect to the description of UV-VIS vertical excitation energies of selected chromophores in aqueous solution.[153] In **Chapter 4**, QM/FQ model is extended to the calculation of the first electric hyperpolarizability (β) and Second Harmonic Generation (SHG) and tested for a set of organic acids in aqueous solution.[152]
QM/FQ is based on the assumption that electrostatic energy terms dominate the QM/MM interaction. Although such an assumption can be in principle reasonable for aqueous solutions, it may fail for other environments. In addition, in FQ the MM portion is represented only in terms of a set of charges. This poses conceptual issues, because in this way only monopoles, i.e. the zeroth order of the electrostatic Taylor expansion, are considered. In the present thesis, these two approximations are overtaken. In **Chapters 5** and **6** a model to include repulsion/dispersion contributions in the calculation of QM/MM energies and Electron Paramagnetic Resonance (EPR) parameters is proposed. **Chapter 7** focuses on the development and testing of an alternative fully polarizable QM/MM model, in which polarization effects are modelled both in terms of fluctuating charges and fluctuating dipoles (FQFμ). Both the two novel approaches are based on the strong points of QM/FQ. In fact, they are both formulated in terms of the QM density, so that they can be further extended to properties/spectroscopies. Also, the variational formalism of QM/FQ is maintained in QM/FQFμ, thus foreseeing its extension to properties and spectroscopies.

QM/MM Non-electrostatic contributions Non-bonding intermolecular interactions find their physical origin in the forces between the charge densities of the molecules involved, resulting in inductive forces, and dispersion and repulsion interactions, the latter essentially due to the quantum nature of the electrons.[192,193] In particular, repulsion forces between molecules arise from the Pauli Exclusion principle, whereas dispersion forces are related to the long-range correlation between the electrons' motions of two molecules. Because of their quantum nature, both these forces are not well reproduced by simple approximations based on classical models. If the whole sys-

tem is described at the QM level the repulsion terms are automatically included in the calculations, while, in order to have a good account for dispersion interactions, it is compulsory to consider electron correlation, i.e. to resort to Post-Hartree-Fock or suitably parametrized DFT methods.[194] In QM/Classical approaches, electrostatic forces can be fully included in the model provided the environment be polarizable, but to account for dispersion and repulsion interactions requires to go beyond the basic models. The simplest way, which is actually the common strategy in Mechanical Embedding, Electrostatic Embedding and Polarizable Embedding, is by calculating the Lennard-Jones potential, which is parametrized on equilibrium distances and dissociation energies.[195,196] However, contrary to the electrostatic term which acts as an external potential in the Hamiltonian, the Lennard-Jones energy is only added to the total energy of the system, because it does not depend on the QM density but only on atom positions. Thus, it gives only indirect contributions to molecular properties, because it does not directly affect the Hamiltonian, its derivatives and response equations.
In **Chapter 5**, a theoretical approach,[197] which formulates repulsion as a function of an auxiliary density on the MM portion and extends the Tkatchenko-Scheffler (TS) approach to DFT[198–202] to treat QM/MM dispersion terms is presented. Such an approach can easily be coupled to any kind of QM/MM approach, because repulsion and dispersion are formulated in a way which is totally independent of the choice of the FF to model the electrostatics (i.e. fixed-charges or polarizable embedding). Remarkably, in our model repulsion and dispersion contributions are indeed dependent on the QM density. Thus, an explicit contribution to the QM Fock operator can be defined and the resulting calculated QM properties/spectra are modified by such interactions.
In **Chapter 5**,[197] the approach is tested against the reproduction of QM calculated non-electrostatic interactions. Then, it is challenged to the calculation of EPR hyperfine coupling constants (hcc) of two stable organic nitroxyl radicals (TEMPO and PROXYL) in aqueous solution in **Chapter 6**.[203]

Refining QM/FQ electrostatic coupling FQ does not explicitly take into account the intrinsic anisotropy of specific molecule-environment interactions, such as HB, because it only works in terms of charges. To overcome this problem, the electrostatic description of the FQ force field is refined in Chapter 7, where an additional source of polarization, i.e. fluctuating dipoles is included. A novel polarizable force field, the Fluctuating Charge Fluctuating Dipoles (FQFμ) is proposed, in which both monopoles (charges) and dipoles can vary as a response to the external Maxwell sources, i.e. electric potential/field. In **Chapter 7**, FQFμ is coupled to a QM SCF description, following the general structure of QM/MM approaches. Remarkably, the variational formalism of QM/FQ is maintained, so that the novel QM/FQFμ can be extended to molecular response/spectral properties by using the machinery of quantum chemistry.

FQ for the spectroscopy on plasmonic materials In the past years, FQ has only been applied to molecular systems in solution. However, recently much interest has been devoted to nanoplasmonic materials under the effects of external fields, which

have been reported to give huge (up to $10^{10} \sim 10^{12}$) enhancement of the spectral response of a molecule adsorbed on them, up to the point of allowing single molecule detection.[204–216] Such enhancement is primary due to the enhancement of the electric field in correspondence of the surface of the plasmonic material.[217]This has stimulated the growth of theoretical approaches to model such phenomena, giving rise to diverse methods, employing both Quantum Mechanical (QM) and classical descriptions. In this context, classical approaches are very promising because of their low computational cost, which allows the investigation of nanoparticles of size by far larger than what can be treated at the QM level.[107,113,218–234]

The last chapter of this thesis (**Chapter 8**) is dedicated to the formulation of a novel classical approach based on the FQ force field to describe the plasmonic response of metal nanoparticles/nanoaggregates to an external electric field. In the novel ωFQ approach, each atom of the nanoparticle is endowed with an electric charge, which can vary as a response to the external electric field. In particular, when nanoparticle dimers are considered, and a sub-nanometer gap emerges, purely QM effects become crucial and cannot be neglected.[225,230,235–242] Such effects are essentially related to electron tunneling, thus the description of the resulting Charge Transfer (CT) is crucial.[225,233,234,241] Due to the CT nature of plasmonic excitation, classical models based on the electronegativity equalization principle appear to be very promising, because, differently from other polarizable force fields, they are in principle able to describe charge transfer between the atoms constituting the system (although with some issues related to unphysical charge transfer at long distances[243–245]).

Chapter 2

Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution

Abstract We present a computational methodology, based on a polarizable Quantum Mechanical (QM)/Molecular Mechanics (MM) approach to accurately compute Vibrational Optical Activity (VOA) spectra of chiral systems. The approach is applied to the calculation of Infrared (IR), Vibrational Circular Dichroism (VCD), Raman and Raman Optical Activity (ROA) spectra of aqueous solutions of (L)-Methyl Lactate and (S)-Glycidol. Remarkable agreement between calculations and experiments is reported, showing the reliability and accuracy of the methodology, especially with respect to standard continuum solvation approaches.

2.1 Introduction

Chiroptical spectroscopic methods are nowadays the most useful tools to study chiral systems and assign their molecular absolute configuration. Chiroptical spectral signals arise from the differential response of the chiral system to polarized light, either in absorption/emission or scattering of the right and left components of the circularly polarized light. Basically, two different families of chiroptical molecular responses have been developed, focusing on the electronic or vibrational molecular degrees of freedom. The former, such as the Optical Rotation Dispersion (ORD) and the Electronic Circular Dichroism (ECD), provide relatively little molecular information in comparison with vibrational chiroptical spectroscopies, i.e. Vibrational Circular Dichroism (VCD) and Raman Optical Activity (ROA), because there are many more bands sensitive to the details of the molecular structure in the vibrational domain than for the different electronic states accessible to experimental investigations.[246–248] VCD and ROA, which are collectively named as Vibrational Optical Activity (VOA), have demonstrated high reliability and a wide range of applicability, as it has been amply documented by recent papers.[247,249–258]

VOA has been so far mostly employed to study biomolecules or natural biopolymers in solution,[246–248,259] however their horizons are broadening, as they have also applied to the study of structural patterns in synthetic chiral polymers,[260] fibrillar patterns,[261] ionic liquids[262] and the transfer of chirality from a chiral analyte to an achiral molecule in the vicinity of a plasmon resonance of an achiral metallic nanostructure.[263] The wide applicability of VOA is due to its sensitivity to the local environment experienced by the chiral system. However, if such a feature is beneficial for the experimental investigation, it is a severe issue to deal with in the development of reliable modeling strategies. In fact, the unambiguous assignment of VOA spectra to enantiomers and structural patterns is only possible by a subtle interplay of experiment and theory, which requires the availability of reliable and algorithms for predicting the spectroscopic signals in a computationally viable manner.[150,264] Such algorithms, cannot neglect the presence of the environment, because that can cause a totally wrong reproduction of the molecular chiroptical signal, up to impede a correct assignment of the absolute configuration.[56,94] For this reason, in the last years significant advances have been made in the coupling of reliable approaches to treat environmental effects and *ab initio* Quantum Mechanical (QM) methods to predict these signals.[91,94,97,111,112,120,122,246–248,256,264–277]

The most successful approaches in this field belong to the family of the so-called focused models, where the system is divided in two portions: a target molecule (e.g. the solute in case of solvated systems), which is responsible for the spectral signal and is treated accurately with *ab initio* QM methods, and the environment (e.g. the solvent), which is treated at a lower level, usually by resorting to classical physics. By resorting to such approaches, the molecule/environment interactions and their effects on the molecular structure and properties are accurately captured.[56,94] The most renowned focused models belong to the family of QM/Classical approaches, in which the classical portion can keep an atomistic description (giving rise to QM/Molecular Mechanics(MM) models)[1,2,12,13,278] or even be blurred to a continuum.[4–6,279]

Continuum solvation approaches, and the Polarizable Continuum Model (PCM) in par-

ticular, have been extended to VOA several years ago.[91,115] Nowadays, they represent the most used strategy to include the effects of the environment on calculated VOA spectra, due to their low computational cost. However, whenever the target system and the environment strongly interact, the mean-field continuum approach may fail and the use of discrete approaches, able to account for specific and directional interactions, is compulsory. This is for instance the case of a solute-solvent couple interacting thorugh hydrogen bonding.[10,148]

In such cases, the limitations of continuum strategies can be effectively overcome by using explicit QM/MM approaches,[1,273] and in particular those accounting for the mutual polarization between the QM and MM moieties of the system, i.e. the so-called polarizable QM/MM methods.[10,56,58,59,155,164,185,186,280,281]

Although several papers have addressed the prediction of electronic response properties in the polarizable QM/MM approach,[10,37,155,167] much less work has been done to extend such approaches to vibrational spectroscopies (IR, Raman) and especially VOA. To the best of our knowledge, the only polarizable QM/MM currently able to calculate such spectra is the QM/Flucuating Charges (FQ) model, based on the Polarizable Fluctuating Charges (FQ) Force Field,[156,157,282] which has recently been extended to several molecular properties and spectroscopies by some of the present authors.[54–56,146,148–150,161,162]

To illustrate our approach, and report simultaneously for the first time on the IR, Raman, VCD and ROA spectra calculated with our method, two chiral systems are here considered, i.e. aqueous solutions of (L)-Methyl Lactate (ML) and (S)-Glycidol (GL) (Figure 2.1). For both systems, experimental IR, Raman, VCR and ROA spectra have been reported in the literature, as they have been employed as test systems to get a deep understanding of their intermolecular interaction with water.[249,252,283–285] In fact, their vibrational spectra have been reported to be strongly modified by the presence of the surrounding aqueous solution, giving rise to several specific features due to hydrogen bonding interactions between the chiral systems and the nearby water molecules.[249–252,284,285]

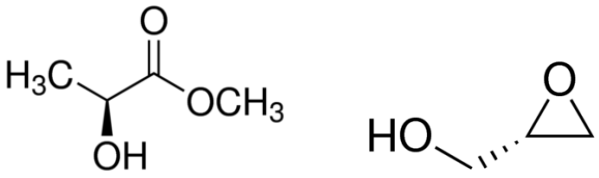


Figure 2.1. (L)-Methyl Lactate and (S)-Glycidol molecular structures

The paper is organized in the following way. In the next section the computational protocol for the calculation of vibrational spectra with the fully polarizable QM/FQ approach is recalled. Then, its application to IR, Raman, VCD and ROA spectra of (L)-Methyl Lactate and (S)-Glycidol is reported, with particular emphasis on the comparison between calculations and experiments. A section focusing on the conclusions of this work and its future perspectives end the manuscript.

2.2 Methodology

The aim of this section is to introduce the reader with the computational protocol for the evaluation of vibrational (IR, VCD, Raman and ROA) spectra within the fully polarizable QM/FQ approach. The theoretical fundamentals of the method are briefly recalled in the next section, as well as the definition of the computational protocol allowing the calculation of the spectra for a given chemical system.

2.2.1 QM/FQ Approach to Vibrational Spectra

The QM/FQ model is a multiscale approach defined in the framework of focused models. When it is applied to molecular systems in solution, this means treating the solute at QM level of theory, whereas the solvent is described by means of the polarizable Fluctuating Charge force field. In Figure 2.2, a schematic picture of the partitioning is shown.

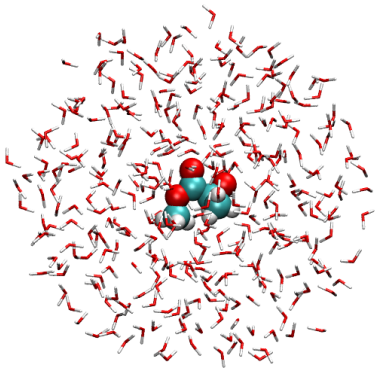


Figure 2.2. Schematic picture of the partition in the QM/FQ model. The van der Waals spheres represent the QM portion, whereas the water molecules the FQ one.

The FQ force field[56,159] represents each atom of the MM portion with a set of fluctuating charges. The polarization arises from the difference of electronegativities between each atom in the electronegativity equalization principle (EEP)[286,287] framework. The EEP states that, at equilibrium, the instantaneous electronegativity χ of each atom have the same value,[160,286] which give a minimization principle in a variational meaning of the term.
The FQs (q) can be defined as those minimizing the following functional[159]

$$\begin{aligned} F(\mathbf{q},\lambda) =& \sum_{\alpha,i} q_{\alpha i}\chi_{\alpha i} + \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j} q_{\alpha i}J_{\alpha i,\beta j}q_{\beta j} + \sum_{\alpha}\lambda_{\alpha}(\sum_{i} q_{\alpha i} - Q_{\alpha}) \\ =& \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} \end{aligned} \qquad \textbf{2.1}$$

where 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. **J** is the charge interaction kernel: there are several ways to treat this term[287–290] and in our implementation the Ohno kernel[291] is exploited.

By following the general philosophy of the so-called "focused" models, in the QM/FQ model, a classical electrostatic interaction between the FQs and the QM density is considered:[54]

$$E_{\mathrm{QM/FQ}} = \sum_{i=1}^{N_q} V_{\mathrm{QM}}[\rho](\mathbf{r}_i) q_i \qquad \mathbf{2.2}$$

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 some of the authors recently extended the present model to the inclusion of non-electrostatic interactions, although they are not considered in the present work.[197] If a Self Consistent Field (SCF) description of the QM portion is adopted, the global QM/MM energy functional reads:[54,55,192]

$$\mathcal{E}[\mathbf{P}, \mathbf{q}, \boldsymbol{\lambda}] = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{1}{2}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} + \mathbf{q}^{\dagger}\mathbf{V}(\mathbf{P}) \qquad \mathbf{2.3}$$

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

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q - \mathbf{V}(\mathbf{P}) \qquad \mathbf{2.4}$$

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

Once the basic QM/FQ approach is set up, the extension to spectroscopic and transition properties is obtained through the definition of analytical energy derivatives and response equations to electric and magnetic perturbations.[192] We refer the reader to refs.[54–56,161] for major details on the implementation and to ref.[149] for major details on VCD spectra. Extension of the QM/FQ model to the analytical evaluation of Raman and Raman Optical Activity has recently been presented by some of the present authors.[150]

The resorting to the physical framework of the so-called "focused models" implies that external perturbations (i.e., electric/magnetic fields and/or a nuclear displacement) only act on the QM portion of the system, whereas the environment is only indirectly affected through the perturbation on the QM density. In vibrational spectroscopy, the focusing on the QM portion of the system means that the geometric displacements of the MM molecules are not taken into account; this framework is well-defined within the Partial Hessian Vibrational Approach (PHVA).[89,292,293] The calculation of vibrational frequencies in the harmonic regime requires the evaluation of energy second derivatives. They can be obtained by differentiating twice Equation 2.1 with respect to x, y perturbations:[55]

$$\begin{aligned}\mathcal{E}^{xy} = &\sum_{\mu\nu}\left[h^{xy}_{\mu\nu} + \frac{1}{2}G^{(xy)}_{\mu\nu}(\mathbf{P}) + \mathbf{q}^{\dagger}\mathbf{V}^{xy}_{\mu\nu}\right]P_{\mu\nu} - \text{tr }\mathbf{W}\mathbf{S}^{xy} - \text{tr }\mathbf{W}^{y}\mathbf{S}^{x} \\ &+ \sum_{\mu\nu}\left[h^{x}_{\mu\nu} + G^{(x)}_{\mu\nu}(\mathbf{P}) + \mathbf{q}^{\dagger}\mathbf{V}^{x}_{\mu\nu}\right]P^{y}_{\mu\nu} + \sum_{\mu\nu}\mathbf{q}^{y\dagger}\mathbf{V}^{x}_{\mu\nu}P_{\mu\nu}\end{aligned} \tag{2.5}$$

where μ, ν are atomic basis functions, $\mathbf{W}$ is the energy-weighted density matrix, and $\mathbf{S}$ is the overlap matrix. Notice that Equation 2.5 requires the computation of the perturbed density matrix, which is accessible through a Coupled Perturbed Hartree-Fock or Kohn-Sham (CPHF/KS) procedure by solving a modified set of equations including FQ terms.[54,55] In the construction of the CPHF equations the Fock matrix derivative is used:

$$\begin{aligned}\tilde{\mathbf{F}}^{x} =&\mathbf{h}^{x} + \mathbf{G}^{(x)}(\mathbf{P}) + \mathbf{q}^{\dagger}\mathbf{V}^{x} + \mathbf{G}(\mathbf{P}^{x}) + \mathbf{V}^{\dagger}\mathbf{q}^{x} \\ =&\mathbf{F}^{(x)} + \mathbf{q}^{\dagger}\mathbf{V}^{x} + \mathbf{G}(\mathbf{P}^{x}) + \mathbf{V}^{\dagger}\mathbf{q}^{x}\end{aligned} \tag{2.6}$$

which is included in the $\tilde{\mathbf{Q}}_X$ and $\tilde{\mathbf{Q}}_Y$ terms of the Casida Equation:

$$\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^{*} & \tilde{\mathbf{A}}^{*} \end{pmatrix}\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{Q}}_X \\ \tilde{\mathbf{Q}}_Y \end{pmatrix} \tag{2.7}$$

The solution of this set of equations yields the density matrix derivatives, $P^{x}_{jb} = X_{jb}$ and $P^{x}_{bj} = Y_{jb}$. Eq.2.5 also allows the calculation of IR intensities, if x, y are specified as nuclear coordinate and electric field component. In case of magnetic perturbations, other terms need to be considered to assure gauge invariance in the computed results. We refer the interested reader to ref.[161] for further details. VCD intensities are proportional to Rotational Strengths (RS), i.e. the imaginary part of the product between the electric and the magnetic dipole moments. RS can be expressed in terms of two tensors, namely the Atomic Polar Tensor (APT) and the Atomic Axial Tensor (AAT), which are defined as:[248,268–271]

$$\begin{aligned}(\text{APT})^{\lambda}_{\alpha\beta} &= E^{\lambda}_{\alpha\beta} + N^{\lambda}_{\alpha\beta} = \\ &= 2\left\langle\left(\frac{\partial\Psi_G}{\partial X_{\lambda\alpha}}\right)_{R_0}\middle|(\mu^{e}_{el})_{\beta}\middle|\Psi^{0}_{G}\right\rangle + Z_{\lambda}e\delta_{\alpha\beta}\end{aligned} \tag{2.8}$$

$$\begin{aligned}(\text{AAT})^{\lambda}_{\alpha\beta} &= I^{\lambda}_{\alpha\beta} + J^{\lambda}_{\alpha\beta} = \\ &= \left\langle\left(\frac{\partial\Psi_G}{\partial X_{\lambda\alpha}}\right)_{R^0}\middle|\left(\frac{\partial\Psi_G}{\partial B_{\beta}}\right)_{B_{\beta}=0}\right\rangle + \frac{i}{4hc}\sum\varepsilon_{\alpha\beta\gamma}R^{0}_{\lambda\gamma}(Z_{\lambda}e)\end{aligned} \tag{2.9}$$

where μ^e_{el} is the electronic part of μ_{el} while $Z_\lambda e$ and R^0_λ are the charge and position of nucleus λ at the equilibrium geometry R^0. Ψ_G is the wave function of the ground electronic state while $(\partial\Psi_G/\partial X_{\lambda\alpha})$ and $(\partial\Psi_G/\partial B_\beta)$ are the derivatives of the wave function with respect to nuclear displacement and magnetic field, respectively. The former enters into a vibrational transition moment; the latter is appropriate to a magnetic dipole transition moment. FQ contributions affect the wavefunction and its derivatives. In particular, by solving the CPHF/CPKS equations (modified according to the Gauge Including Atomic Orbital–GIAO approach) taking into account FQ contributions, the APT and AAT in the QM/FQ framework are obtained.
By exploiting the Placzek approach within the double harmonic approximation, Raman and ROA intensities are obtained in terms of the geometric derivatives of the electric dipole-electric dipole polarizability α^x, electric dipole-electric quadrupole polarizability A^x and electric dipole-magnetic dipole polarizability G'^x. In particular, Raman intensities depend only on α^x, whereas ROA intensities depend on α^x, A^x and G'^x. In the following equations the QM/FQ contributions to these quantities are reported. We refer the interested reader to Ref.[150] for further details.

$$\alpha^x_{QM/FQ} = \sum_{\mu\nu}\left[\mathbf{q}^\dagger(\mathbf{P}^{e'}(\omega'))\mathbf{V}^x_{\mu\nu}P^e_{\mu\nu}(\omega)\right] + \sum_{\mu\nu}\left[\mathbf{q}^\dagger(\mathbf{P}^{e}(\omega))\mathbf{V}^x_{\mu\nu}P^{e'}_{\mu\nu}(\omega')\right] \tag{2.10}$$

$$A^x(\omega)_{QM/FQ} = \sum_{\mu\nu}\left[\mathbf{q}^\dagger(\mathbf{P}^{e}(\omega))\mathbf{V}^x_{\mu\nu}P^q_{\mu\nu}\right] + \sum_{\mu\nu}\left[\mathbf{q}^\dagger(\mathbf{P}^{q})\mathbf{V}^x_{\mu\nu}P^e_{\mu\nu}(\omega)\right] \tag{2.11}$$

$$\begin{aligned} G'^x_{QM/FQ} = &\sum_{\mu\nu}\left[\mathbf{q}^\dagger(\mathbf{P}^{m})\mathbf{V}^x_{\mu\nu}P^e_{\mu\nu}(\omega)\right] + \sum_{\mu\nu}\left[\mathbf{q}^\dagger(\mathbf{P}^{e}(\omega))\mathbf{V}^x_{\mu\nu}P^m_{\mu\nu}\right] + \\ &+ \sum_{\mu\nu}\left[\mathbf{q}^\dagger(\mathbf{V}^m(\mathbf{P}^e(\omega)))P^x_{\mu\nu}\mathbf{V}_{\mu\nu}\right] + \\ &+ \sum_{\mu\nu}\left[\mathbf{q}^\dagger\mathbf{V}^{m,x}_{\mu\nu}P^e_{\mu\nu}(\omega) + \mathbf{q}^{x\dagger}\mathbf{V}^m_{\mu\nu}P^e_{\mu\nu}(\omega)\right] \end{aligned} \tag{2.12}$$

2.2.2 QM/FQ Computational Protocol

Besides the development of the theoretical methodology to actually calculate vibrational spectral signals within the QM/FQ approach, another crucial point of the method is the definition of the QM/FQ model system being investigated. This first implies a sensible choice of the part of the system which is modelled at the QM level. Such a choice can be tricky in case of covalently-bound systems (e.g. in case of a receptor in a protein), however in case of solutions, the most natural choice is to reserve the QM treatment to the solute, and resort to classical physics for the solvent. This basic

choice may be possibly refined, so to include a (small) number of solvent molecules in the QM moiety, that based on the analysis of the solute-solvent interaction patterns (e.g. by analyzing the hydrogen-bonding patterns).
The definition of the QM/FQ partition is not the only issue in this kind of modelling. As a matter of fact, the measured spectrum of a given molecular system embedded in a surrounding environment arises from the spectral signals of all the possible configurations spanned by the system. Therefore, in order to get a reliable theoretical description, it is mandatory to reliably take into account both the flexibility of the target QM molecule and the spatial arrangement of the surrounding classical environment, which can evolve in time. The best strategy to do that is to resort to Molecular Dynamics (MD) simulations based on high level force fields (FF), which permit a dynamical description of both the solute flexibility and the solvation phenomenon at the same time.[294,295]
Based on the aforementioned considerations, the QM/FQ computational protocol for the calculation of spectral properties of solvated systems, involves a number of steps:

1. *Definition of the system:* the model systems is composed by the target 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/Raman/VCD/ROA 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/Raman/VCD/ROA spectra for the system are obtained as the arithmetic mean of the spectra for all the snapshots.

2.3 Computational details

Geometry optimization of (L)-Methyl Lactate and (S)-Glycidol was performed at the B3LYP/aug-cc-pVDZ level of theory, by using the conductor-like variant of PCM (C-

PCM) to reproduce water bulk solvent effects.[74,296] An analysis of the C-PCM Potential Energy Surfaces(PES) was performed in order to search for the different conformers. Starting from the most populated C-PCM conformer for both systems, MD simulation was performed to sample the configurational space. Because both molecular systems are flexible, the intramolecular FF was reparametrized. Stiff bonded FF terms (i.e., stretching, bending, rigid dihedral force constants and equilibrium values) were transferred from the OPLS FF.[297] OPLS Lennard-Jones parameters and CM5 point charges[298] computed on the global minimum were used for the FF nonbonded part. CM5 were chosen because of their tiny dependence on molecular conformations and on the level of theory adopted in the fitting procedure.[298,298] Flexible dihedrals were further re-parametrized by using the Joyce algorithm,[294] by fitting the FF energies of different twisted conformations to their QM counterparts calculated at the B3LYP/aug-cc-pVDZ level. Such conformations were obtained as a results of a relaxed energy scan around the torsion angles of interest, which were varied from 0 to 360 degrees in steps of 30 degrees.

The two molecules, in their lowest C-PCM free energy conformation, were placed in the centre of a cubic box with an edge of 50 Å containing roughly 5000 TIP3P[299] water molecules. Such systems were minimized using the conjugate gradient algorithm until an energy threshold of 0.5 kJ mol^{-1} was reached. Preliminary equilibration steps of 200 ps were performed in the NPT ensemble, using a time step of 0.2 fs. The systems were slowly heated from 150 to 300 K. Berendsen thermostat and barostat were used, with time constants of 0.1 and 1.0 ps, respectively. Periodic boundary conditions were applied in all directions. VdW forces were computed using a cutoff distance of 10 Å. Long-range electrostatic interactions were treated with the particle mesh Ewald (PME) method, using a real-space cutoff radius of 10 Å. Production runs were performed in the NVT ensemble at 300 K. The simulation time was set to 50 ns, and the time step was increased to 2 fs. The LINCS algorithm was used in order to freeze all chemical bonds. Systems coordinates were stored every ps of simulation. All molecular simulations were carried out with the Gromacs 4.6.5 package.[300]

200 uncorrelated snapshots were extracted from the last 40 ns of the MD simulation (one snapshot every 200 ps). For each snapshot a sphere centered at the solute's geometric center was cut, of radius equal to 13 Å. The partial optimization of the QM solute in each snapshot was performed by keeping fixed all water molecules. Finally, IR, VCD, Raman and ROA spectra were calculated with the QM/FQ model, at the B3LYP/aug-cc-pVDZ level for the solute and the SPC FQ parameters for the FQ portion.[156] All reported spectra were obtained by convoluting peak intensities with a Lorentzian function, with Full Width at Half Maximum (FWHM) of 4 cm^{-1}. Data were finally averaged to obtain the final spectra. All DFT and QM/FQ calculations were performed by using a development version of the Gaussian16 package.[301]

2.4 Results and Discussion

In this section, the potentialities of the methodology outlined in the previous sections are illustrated by taking as example the calculation of IR, VCD, Raman and ROA

spectra of aqueous solutions of (L)-Methyl Lactate (ML) and the (S)-Glycidol (GL). The spectra of the two systems are discussed and compared with experiments, taken from the recent literature.[249,252,284,285,302]
ML and GL are floppy molecules, showing several populated conformers in aqueous solution, possibly separated by low energy barriers.[249,303–308] Therefore, the first step of the simulation protocol is a reliable sampling of such conformations, which will be achieved by resorting to MD runs based on an high level, customized FF.
The subsections reporting on the two systems are structured in the following way: first, conformational analyses and hydration patterns are discussed. Then, vibrational spectra are presented and compared to experiments.

2.4.1 (L)-Methyl Lactate

Conformational Analysis

Conformers were first located by exploiting the implicit PCM model, at the B3LYP/aug-cc-pVDZ level of theory. Minima geometries were optimized and validated by means of frequency calculations. As expected on the basis of the previous literature,[249,306] three stable conformers were located. Their structures are depicted in Figure 2.3.

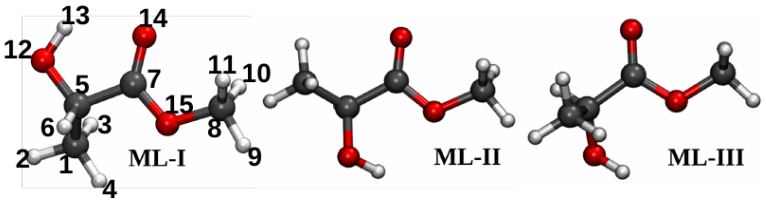


Figure 2.3. B3LYP/aug-cc-pVDZ/PCM most stable conformers of ML in aqueous solution. ML-I presents also the labeling used in the following.

All PCM minima energy structures are stabilized *via* intramolecular H-bonding interaction between the hydroxyl group and the oxygen of the carbonyl group (ML-I) or of the ester group (ML-II and ML-III). Calculated PCM Boltzmann populations at 298 K are summarized in Table 2.1, where also the corresponding values obtained in vacuo by Borho et al. at the MP2/aug-cc-pVDZ level of theory are reported.[306]

Conformer	Vacuum[306]	Water (PCM)
ML-I	91.8	82.8
ML-II	4.6	9.7
ML-III	3.6	7.5

Table 2.1. Calculated Boltzmann populations with ZPE corrections included in vacuum (B3LYP/6-311++G(d,p)) and water (B3LYP/aug-cc-pVDZ)

In order to describe the dynamical fluctuations of the solvent molecules and improve the description of the intermolecular solute-solvent interactions, MD simulations were

also performed by starting from the PCM conformers. To properly account for the flexibility of the molecule, a new FF for ML was generated by fitting the potential energy associated to flexible torsion angles with their QM counterpart. In the case of ML, two dihedral angles were considered: i) θ_1, which involves the α hydroxy ester group, and ii) θ_2, which is related to the methyl ester (see Figure 2.4).

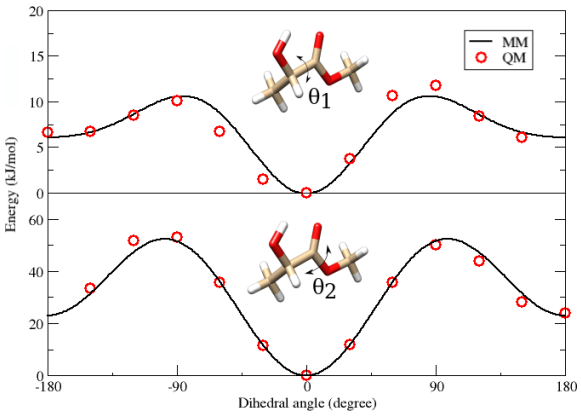


Figure 2.4. QM (red circles) and MM (black continuous line) potential energy curves along the two dihedral angles θ_1 and θ_2 of ML.

The results of the fitting procedure are reported in Figure 2.4. The energy profiles associated to θ_1 and θ_2 are very similar and in both cases, the global minimum is located at 0°, while a partial minimum is found at 180°. The minimum found at 0° along θ_1 is not surprising, because the hydroxil group interacts with the carbonyl oxygen *via* intramolecular H-bonding. Focusing on the θ_2 angle, the two minima are separated by a high energy barrier of more than 50 kJ·mol^{-1}, whereas the barrier is lower for θ_1 (about 10 kJ·mol^{-1}). Overall, a very good matching between the classical and QM descriptions is noticed, thus supporting the use of the developed FF in the forthcoming MD simulations. The whole FF parameter set is given in Section S1 of the electronic supplementary information (ESI).
Figure 2.5 shows the θ_1 and θ_2 distributions obtained in a 50 ns MD simulation performed with the refined FF. Clearly, the global minima are well populated with respect to the other possible conformations. Also, a small, but not negligible number of configurations, are generated from the partial minima at 180 degrees.

Hydration Patterns

ML hydration patterns were analyzed by calculating the radial distribution function *g(r)* of water hydrogen and oxygen atoms around the three ML oxygen atoms (O12, O14, O15). The data are reported in Figure 2.6 (see Figure 2.3 for the labeling of the atoms).
In the left panel of Figure 2.6, the radial distribution functions of water hydrogen atoms around ML oxygen atoms are reported. A well-defined peak at about 2 Å for both carbonyl (O14) and hydroxyl (O12) oxygen atoms is present, thus indicating a

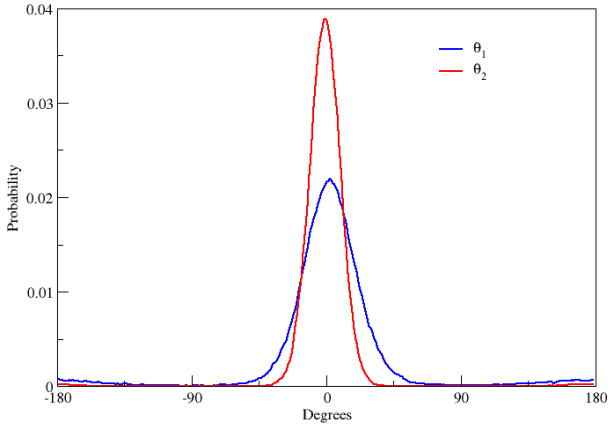


Figure 2.5. Calculated θ_1 and θ_2 distributions for ML in aqueous solution. The data were extracted from a 50 ns MD simulation exploiting the refined FF (see text).

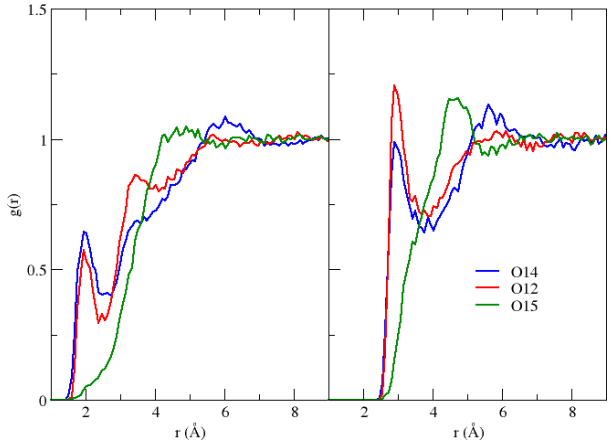


Figure 2.6. Radial distribution function ($g(r)$) of water hydrogen (left panel) and oxygen (right panel) atoms around the three ML oxygen atoms: O14 (blue line), O12 (red line) and O15 (green line).

strong hydrogen bonding (HB) interaction between ML and the surrounding water molecules. Coordination numbers of approximately 1.30 and 1.04 for O14 and O12 are found, respectively. Thus, a slight preferential HB interaction is reported for the ML carbonyl oxygen atom, if compared to the hydroxyl oxygen. On the contrary, the oxygen atom of the ML alkoxy group (O15, see Figure 2.6), does not interact with the hydrogen atoms of the surrounding water molecules.
The radial distribution functions related to the interaction of ML with water oxygen atoms are depicted in the right panel of Figure 2.6; an opposite behaviour with respect to what has been commented above is noticed. In fact, the hydroxyl oxygen atom (O12) can act both as HB donor and HB acceptor: therefore, the water oxygen atoms are placed, on average, preferentially around O12 than around the carbonyl site, O14. We note that our calculated intensities of the $g(r)$) peaks are lower if compared with the data reported in a previous paper,[309] however their relative positions are in good agreement. Such findings show that the here employed electrostatic description of the water solvent (especially the use of the CM5 charges) results in a weaker, albeit well defined, HB interaction of ML with the surrounding solvent.

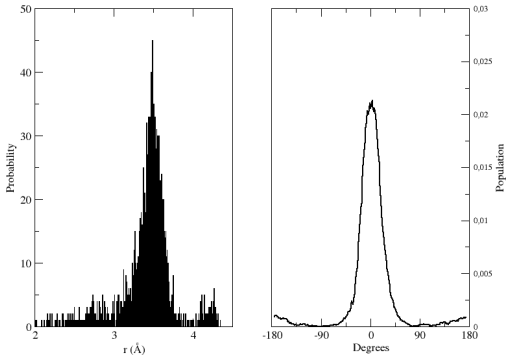


Figure 2.7. Left panel: H13-O14 distance distribution, sampled in the last 2 ns of simulation time. Right panel: θ_1 distribution, sampled in the last 2 ns of simulation time (right panel).

To end this section, let's focus on the intramolecular HB interaction, which can occur between the ML hydroxyl group and the carbonyl oxygen. Such an interaction is highly favored when the solvent is treated as a continuum dielectric medium (see Table2.1), in fact the PCM global minimum exhibits an intramolecular *HO-O* distance of 2.052 Å (see Table 2.1 and Figure 2.3). In order to check whether this preferential interaction is maintained by the explicit solvent description, the *HO-O* distance during the last 2ns of MD simulation was monitored: the results are plotted in the left panel of Figure 2.7. Only a small population of conformers exhibit the intramolecular HB interaction. In fact, most conformers show an *HO-O* distance of about 3.5 Å thus showing that ML-water intermolecular interactions are highly favoured wiht respect to the intramolecular HB by adopting the explicit solvation modeling. As a result, the conformational distribution predicted by exploiting the PCM is expected not to be preserved in the explicit modeling. We also note that our analysis is not biased towards the selection of irrelevant configurations: the θ_1 distribution (see the right

panel of Figure 2.7), which could directly affect the occurrence of intramolecular HB interactions, is in excellent agreement with the corresponding data reported in Figure 2.5.

IR Spectrum

To calculate the QM/FQ IR spectrum 200 uncorrelated snapshots were extracted from the MD simulation; such a number is enough to yield a converged spectrum, as already pointed out by some of the present authors.[149] The raw data extracted from the single QM/FQ calculations are reported as stick spectrum in Figure S1, given as ESI. Clearly, the overall shape of the final, averaged spectrum is already visible from the data reported in Figure S1 already depicts the shape of the spectrum, also giving insight into the spreading of the vibrational bands, both in wavenumbers and intensities. This is due to the fact that in the different snapshots the spatial distribution of water molecules around ML varies, as well as the conformation of ML.
In order to obtain the final, averaged spectrum, each transition in Figure S1 was convoluted with a Lorentzian function and averaged. The final results are given in (Figure 2.8), which also shows the experimental spectrum.[249] We first notice that the inhomogeneous band broadening is naturally obtained as a result of the averaging procedure on the stick spectrum. The computed spectrum is characterized by three main bands: 1150 cm^{-1} (bending of the O-CH3 group), 1200-1300 cm^{-1} (composite bending modes, see Figure S2 in the ESI) and 1750 cm^{-1} (CO stretching mode). All the normal modes corresponding to bands in the region 400-1800 cm^{-1} are depicted in Figure S2 in the ESI for a randomly chosen snapshot.

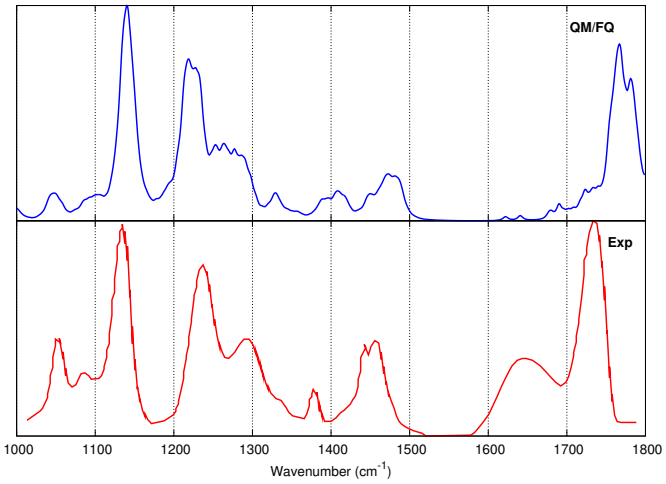


Figure 2.8. Convoluted QM/FQ IR spectrum of ML in aqueous solution (top). Experimental spectrum taken from Ref.[249] (bottom).

Overall, the calculated spectrum is in very good agreement with the experiment (Figure 2.8). In fact, almost all peaks relative intensities are correctly reproduced, as well as the band broadening. This is particularly evident for the composite band between

1200-1300 cm^{-1}, of which the structure is almost perfectly reproduced. Notice that, as it can be seen from the inspection of the normal models in Figure S2 in the ESI, the vibrational modes causing in this band involve the OH group, which experiences a strong hydrogen bonding interaction with the surrounding water molecules (see the $g(r)$ depicted in Figure 2.6). The relevance of a correct description of the HB interaction for a correct reproduction of this part of the spectrum is even more emphasized if the results obtained by exploiting a continuum PCM solvent description are considered (See Figure S3 in the ESI). In fact, the PCM completely fails at describing the intensity pattern of this band, due to the lack of the explicit HB interaction in this mean-field approach. Coming back to the the comparison between the QM/FQ and the experimental spectra (see Figure 2.8), some small discrepancies are indeed present. First, there are small deviations in peak's wavenumbers (especially in the higher energy region), which are due both to the QM level exploited in the calculation, and to the lack of anharmonicity in our calculations. Second, the broad band between 1600-1700 cm^{-1} is not reproduced by our model. As already reported by some of the present authors[149] and in Refs.,[250–252] this band is attributed to the OH bending mode of water molecules. Our modelling, which focuses on the QM portion of the system only, cannot reproduce such spectral features.

VCD Spectrum

The VCD spectrum of ML was calculated by following the same procedure adopted for the IR spectrum. The computed VCD stick spectrum is plotted in Figure S4, given as ESI. Clearly, the same vibrational normal mode can result in peaks intensities of opposite sign depending on the selected snapshot, i.e. on the particular spatial arrangement of the solvent around the QM molecule and on the specific conformation of the latter. This applies to almost all transitions in the studied region (1000-1800 cm^{-1}). These findings, which have already been pointed out by some of us in previous papers[146,149,150] are particularly remarkable, because in VCD, such as in other chiroptical spectroscopies, it is of crucial importance to correctly reproduce the sign of the single transitions. The data reported in Figure S4 in the ESI confirm the importance of coupling an explicit and dynamic description of the solvation phenomenon (reproduced through MD) with an accurate description of the aqueous solution.

The final sign of the bands is obtained as a result of the averaging procedure, of which the results are depicted in Figure 2.9, where also the experimental spectrum is reported.[249]

The calculated spectrum is characterized by a very intense pattern (+,-,-,-,+) in the region between 1200 and 1500 cm^{-1}. The normal modes involved in these transitions have been discussed before for the IR spectrum. However, it is important to remark that the high negative peak at about 1280 cm^{-1} and the band at 1380 cm^{-1} are due to the bending mode involving the OH group, which, as stated before, interacts with the solvent via hydrogen bonding interactions. Furthermore, the small negative peak between 1700-1800 cm^{-1} is due to the carbonyl stretching.

Figure 2.9 also reports the experimental spectrum reproduced from Ref.[249] All the signs of the bands and most of the bands relative intensities in the region between

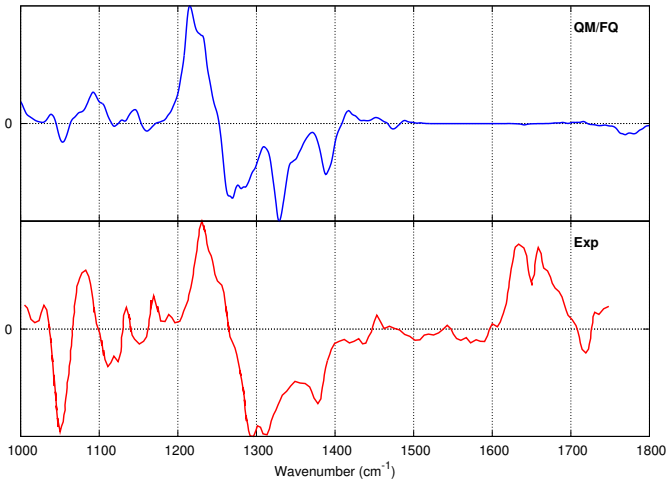


Figure 2.9. Convoluted QM/FQ VCD spectrum of (L)-Methyl Lactate in aqueous solution (top). Experimental spectrum taken from [249] (bottom).

1200-1500 cm^{-1} are correctly reproduced in the calculated spectrum. Also, the inhomogeneous band broadening is accurately calculated; this is especially evident for the most intense peak of the spectrum (1220 cm^{-1}), and the negative bands between 1300 and 1400 cm^{-1}. The only region where calculated and experimental data are not in a perfect agreement with the experiment is that below 1200 cm^{-1}; in fact, the computed intensities are too low. The normal modes involved in the regions do not involve any potential site for hydrogen bonding (See Figure S2 in ESI). Therefore, the not perfect reproduction of the experimental spectrum is probably due to the huge alternation of the sign of the bands for the single snapshots in this region, as depicted in Figure S4. Nevertheless, the band sign is correctly reproduced also in this region, as well as the band inhomogeneous broadening. The relevance of exploiting a dynamic explicit solvation modeling for the description of the VCD spectrum of ML is manifest if the results commented above are commented with what can be obtained by resorting to a purely continuum, static QM/PCM approach. In fact, (see Figure S5 given as ESI), the spectral features in the region 1200-1400 cm^{-1} are badly described by the continuum approach: as already discussed, such a range is dominated by vibrational modes directly involving the OH group, which strongly interacts with the surrounding water molecules as a result of HB (see Figure 2.6).

To end the discussion on the QM/FQ VCD spectrum of ML, it is worth pointing out that our approach, which focuses on the QM portion only of the multilevel system, cannot reproduce the intense band structure in the 1600-1700 cm^{-1} region; this is not surprising, because such spectral features have been atributed to the so-called "chirality transfer" from water molecules to ML, i.e. they are actually due to water molecules vibrational motions, of which the chiral signal is activated as a result of the interaciton with the chiral ML solute. [250–252]

Raman Scattering Spectrum

The QM/FQ Raman Scattering Spectrum of ML was calculated on the same 200 snapshots extracted from the MD simulation. Figure S6, given as ESI, reports the raw data in the region 400-1800 cm^{-1}. Generally, the same considerations already reported for IR and VCD spectra (Figures S1 and S4) also apply to Raman.

Figure 2.10 reports the convoluted QM/FQ Raman spectrum and the experimental spectrum.[284] The computed spectrum is characterized by two intense bands at about 830 cm^{-1} and 1500 cm^{-1}, the latter being associated to the bending mode of the O-CH3 moiety. Notice that this normal vibrations has almost zero intensity in the VCD spectrum (see Figure 2.9), and very low intensity in the IR spectrum (see Figure 2.8), thus confirming the complementarity of these vibrational spectroscopies in structural studies. The predicted Raman intensities in the region between 1200-1300 cm^{-1}, which correspond to normal modes of groups involved in intermolecular hydrogen bonding interactions, are very low. The same also applies to the most intense peak in the IR spectrum at 1130 cm^{-1}, which exhibits a very low intensity in the Raman spectrum.

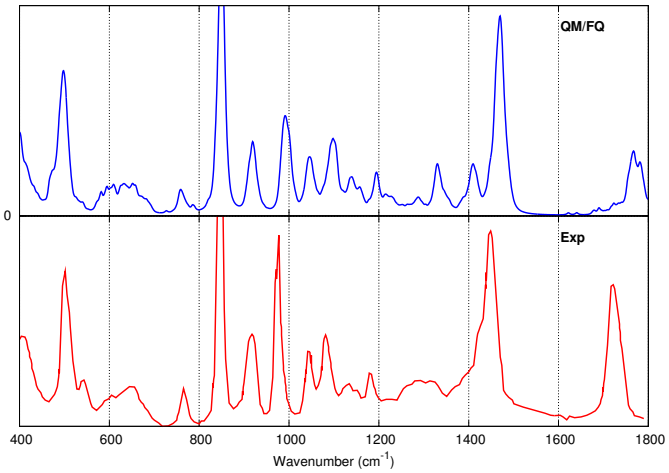


Figure 2.10. QM/FQ Raman spectrum of ML in aqueous solution (top) and experimental spectrum taken from Ref.[284] (neat liquid, bottom). The external excitation wavelenght is 488 nm.

Figure 2.10 also reports the experimental spectrum,[284] for the sake of comparison. Notice that the experimental spectrum of ML in aqueous solution has not been reported in the literature. However, the differences between the neat liquid spectrum (plotted in Figure 2.10) and the acqueous solution spectra have been discussed,[284] and the major differences have been ascribed to peaks' wavenumbers more than intensities. The comparison of the spectra in Figure 2.10 shows a very good agreement between calculated and experimental data. In particular, the inhomogeneous band broadening is adequately reproduced for all the bands and relative intensities are also well reproduced, except for the peak at 980 cm^{-1}. This is probably due to the fact that the Raman intensities are not much affected by the presence of the environment, as it is confirmed

by the QM/PCM spectrum (see Figure S7 in the ESI), where only small differences between the calculated and the experimental spectra can be appreciated. We finally notice that the reported spectrum is very similar to what was already reported from some of us for the same molecule studied by exploiting a three-layer QM/FQ/PCM approach:[150] clearly, the presence of the third PCM layer is not crucial for the description of the Raman spectral features, and the strong similarity is also a further proof of the little dependence of the calculated spectra on the presence of the external environment.

ROA

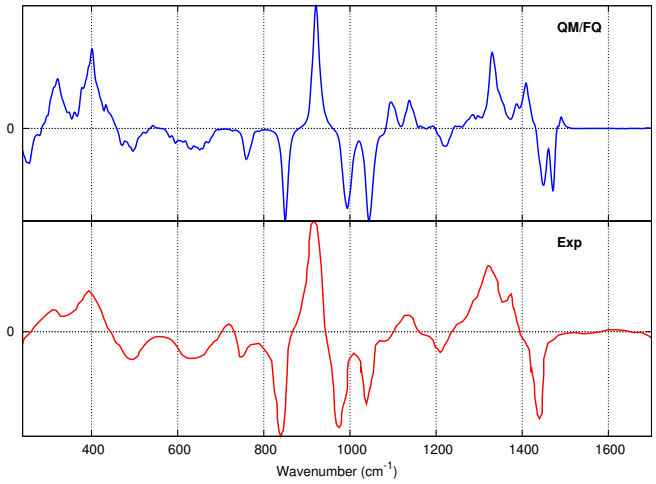


Figure 2.11. QM/FQ ROA spectrum of ML in aqueous solution (top) and experimental spectrum taken from Ref.[310] (neat liquid, bottom). The external excitation wavelength is 532 nm.

Similar to the previous spectra, also the QM/FQ ROA spectrum of ML was calculated on the same 200 snapshots extracted from the MD simulation. Figure S8, given as ESI, reports the raw data in the region 400-1800 cm^{-1}; similar to VCD, the same vibrational normal mode can result in ROA peaks intensities of opposite sign depending on the particular spatial arrangement of the solvent around the QM molecule and on its conformation, as sampled by the MD. This especially applies to the regions below 400 cm^{-1} and between 1100 and 1450 cm^{-1}.

The QM/FQ averaged convoluted spectrum is shown in Figure 2.11, as well as the experimental spectrum taken from Ref.[310] and measured for the neat liquid. Notice that, as already commented for Raman in the previous section the presence of a third PCM layer[150] is not crucial for the description of the Raman spectral features.

Remarkably, the signs of all peaks are correctly reproduced, as well as their relative intensities and the band broadening, with an accuracy similar to what has already been pointed out for VCD in a previous section. This is not the case of the application of the purely continuum PCM approach (see Figure S9 given in the ESI), which fails at correctly reproducing not ony the band broadening but remarkably the sign of some

bands.[150] Therefore, our dynamic computational protocol based on the polarizable QM/FQ approach confirms its remarkable potentialities and reliability for the study of vibrational optical activity spectra.

2.4.2 (S)-Glycidol

Conformational Analysis

The same protocol as used for ML was exploited to perform (S)-Glycidol (GL) conformational analysis. Thus, it was first performed by describing the aqueous solution by means of the continuum PCM approach, at B3LYP/aug-cc-pVDZ level of theory. GL geometry was optimized, followed by frequency calculations in order to validate the minima structures. As previously reported in the literature[252,307] eight stable conformers were located; their structures are depicted in Figure 2.12.

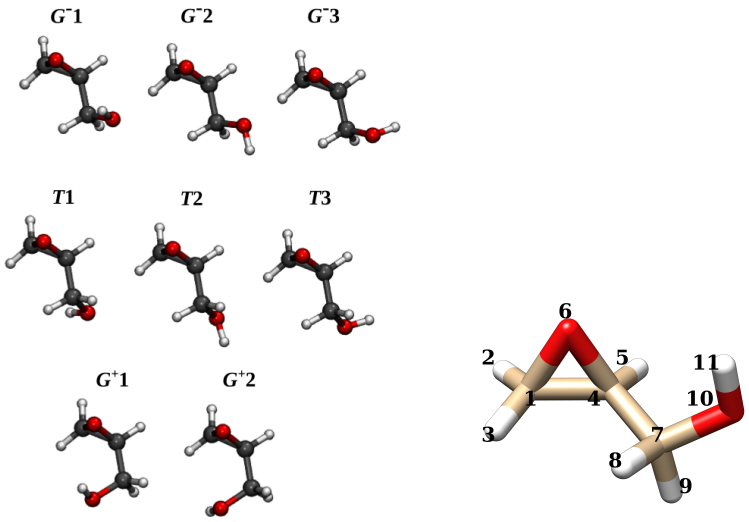


Figure 2.12. (S)-Glycidol most stable conformers in aqueous solution. In the right panel, the labeling used in the following is reported.

Calculated QM/PCM Boltzmann populations are summarized in Table 2.2, where also the data calculated in vacuum are reported for the sake of comparison.
We first note that, similarly to ML, the most stable conformers both in vacuo and in solution (G^-1 and G^+1) are stabilized *via* intramolecular H-bonding interaction. Also, solvent effects significantly change the relative populations of the conformers.
Remarkably, the population of G^+1 decreases and simultaneously the populations of G^-2 and G^-3 increase. These two conformers are potentially characterized by an intermolecular hydrogen bonding with the solvent molecules (see Figure 2.3), although this is not explicitly described by the continuum PCM model. In order to improve the conformational analysis by explicitly considering the solvent molecules in their spatial coordinates, MD simulations were performed. To this end, a customized FF for GL was generated by re-parametrizing the flexible torsional coordinate, i.e. the dihedral

Conformer	Vacuum	Water (PCM)
G^-1	40.2	31.0
G^-2	3.0	16.0
G^-3	1.8	15.3
$T1$	4.7	7.7
$T2$	3.8	4.2
$T3$	2.8	4.9
G^+1	43.2	17.2
G^+2	0.5	0.7

Table 2.2. Calculated B3LYP/aug-cc-pVDZ Boltzmann populations (%) in vacuo and aqueous solution (PCM). ZPE corrections are included.

angle defined by the hydroxyl group and the oxirane moiety. The results of the FF fitting procedure are reported in Figure 2.13.

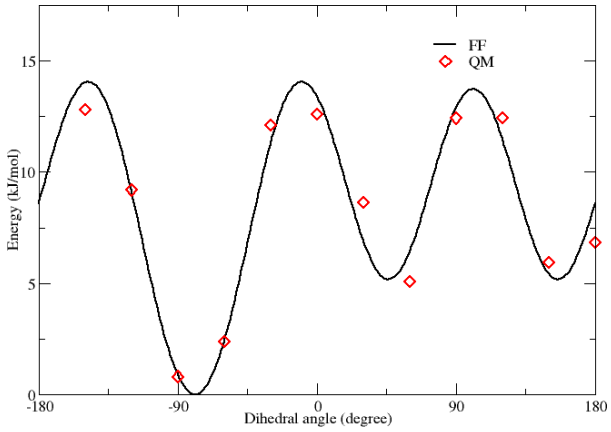


Figure 2.13. Calculated energy profile along GL flexible dihedral angle. QM (red circles) and fitted FF (black continuous line).

Figure 2.13 clearly shows that the description offered by the refined FF reproduces pretty well the energy profile computed at the B3LYP/aug-cc-pVDZ level. The global minimum is found at approximately -78 degrees, and the other two minima are located at 50 and 150 degrees and separated by an energy barrier of about 10 kJ/mol. The developed FF was further validated by monitoring the dihedral distributions during the 50 ns of MD production run in aqueous solution. As reported in Figure 2.14, the three different minima along the main dihedral coordinate are selectively populated during the MD simulation. Furthermore, the conformation at -78 degrees is confirmed to be the most populated due to its stability, as also suggested by the previous QM calculations. The whole GL FF parameter set is given in Section S2.1 of the ESI.

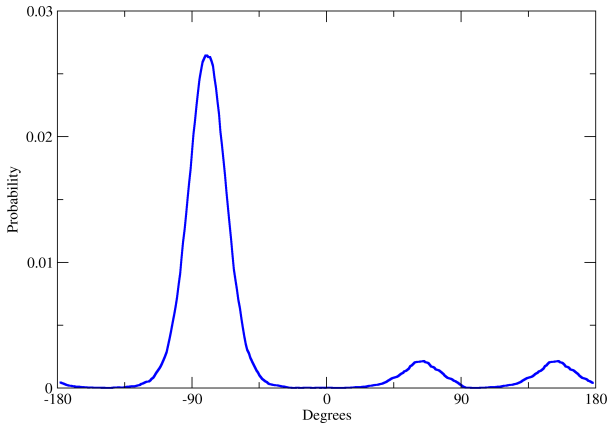


Figure 2.14. Dihedral distributions of GL in the 50 ns of the MD simulation.

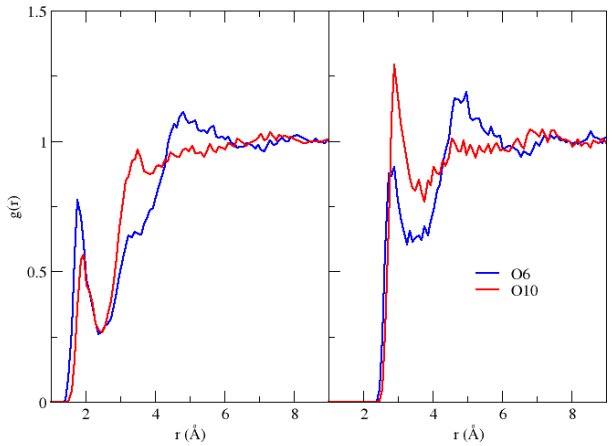


Figure 2.15. Radial distribution function ($g(r)$) of water hydrogen (left panel) and oxygen (right panel) atoms around the two GL oxygen atoms: O6 (blue line) and O10 (red line).

Hydration Patterns

HB patterns were computed for the two GL oxygen atoms O6 and O10 (see Figure 2.12 for the labeling); their relative $g(r)$ profiles are reported in Figure 2.15, in which the $g(r)$ distribution of water hydrogen (left) or oxygen (right) atoms around GL oxygen atoms are reported. Considering the left panel, in the case of O6 the H_w-O_{gly} distribution shows a peak at 1.8 Å, which is more intense then that of the hydroxyl O10, which is located at 1.9 Å. The coordination numbers are 1.3 and 1.08, respectively, thus a clear HB pattern is established between the solute and the solvent molecules. On the contrary, by looking to the right panel of Figure 2.15, the radial distribution related to of O_w-O_{GL} shows a first peak which is higher in the case of the hydroxyl oxygen O10, that due to this atoms is both HB donor and HB acceptor. The second peak in the O_w-$O6_{GL}$ can be assigned to the water molecules which interact with O10, when O6 and O10 are close to each other. Notice that the results here reported are similar to previuous findings of Xu and coworkers,[252] especially as concerns the position of the peaks. However, the intensity of the O_w-$O10_{GL}$ *g(r)* is remarkably underestimated. This can be once again attributed to the use of CM5 charges, whose absolute values are significantly lower if compared to QM-electrostatic potential derived charges.[311] To confirm such findings, a 50 ns MD with the RESP[312] charge scheme was performed: in this case, the computed $O10_w$-O_{GL} g(r) profile closely reproduces the available literature data (see Figure S10 given as ESI).[252]

IR Spectrum

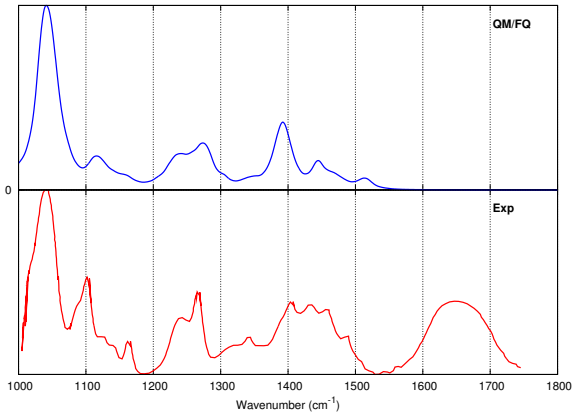


Figure 2.16. Convoluted QM/FQ IR spectrum of (S)-Glycidol in aqueous solution (top). Experimental spectrum taken from[252] (bottom).

Similar tyo ML, the QM/FQ IR spectrum of GL in aqueous solution was calculated on 200 snapshots extracted from the MD simulation.[149] The raw data extracted form the snapshots are reported in Figure S11 in the ESI. The case of GL is more complicated than ML, because several conformers are present in aqueous solution, as predicted by the MD (see Figure 2.14). Together with the dynamical description of the solvent molecules and the atomistic description of the QM/FQ model, this conformational

flexibility results in a large variability of the peaks of the single snapshots both in intensities and wavenumbers.
Each peak of Figure S11 was convoluted with a Lorentzian function and then averaged in order to obtain the average spectrum in Figure 2.16, where also the experimental spectrum[252] is reported. The calculated IR spectrum shows an intense peak at about 1050 cm^{-1}, which can be assigned to a diffuse stretching/bending normal mode, involving the hydroxyl group. The composite band between 1200 and 1300 cm^{-1} is mainly constituted by two peaks, the first (1230 cm^{-1}) due to the C-OH bending, the second (1270 cm^{-1}) due to the C-CH bending modes. The computed spectrum also presents a broad band between 1400 and 1500 cm^{-1}, due to C-OH bending (1395 cm^{-1}), a diffuse C-CH bending (1440 cm^{-1}) and a CH_2 bending (1465 cm^{-1}). The normal modes for a randomly chosen snapshot of the MD in the region 700-1800 cm^{-1} are depicted in Figure S12 in the ESI. It is also worth noticing that the variability in the wavenumbers in the stick spectrum (Figure S11) results in the inhomogeneous broadening of the bands in the convoluted spectrum.
The calculated and the experimental spectra are in very good agreement. In fact, most of the relative intensities and the band broadening are correctly reproduced. This is particularly evident for the structured peak between 1200 and 1300 cm^{-1}. Some discrepancies are instead present in the region 1400-1500 cm^{-1} which is characterized in the experimental spectrum by a very broad band. In the QM/FQ spectrum the broadening is not perfectly reproduced due to a not perfect description of the peaks relative intensities, however the three-band structure is correctly described. As for ML, a minor deviation in the vibrational energies due to the DFT level of theory and to the lack of anharmonicity is reported. Overall, the continuum PCM approach, is inadequate to correctly reproduce the experimental spectrum (see Figure S13 in the ESI), thus remarking once again the huge potentialities of our approach to model vibrational spectra of solutes strongly interacting with the aqueous environment.
To end the discussion on the IR spectrum, we notice that the same broad band between 1600-1700 cm^{-1} already observed for ML is reported also for GL in the experimental spectrum. Obviously, also in this case the focused model cannot reproduce this band.

VCD Spectrum

The QM/FQ VCD spectrum was calculated on the same 200 snapshots extracted from the MD. In Figure S14 the raw VCD data for the different snapshots are depicted. It is worth noticing that almost all spectral regions are characterized by a huge variability in peaks intensities and sign. This behaviour, which has already been discussed for ML, is in this case even more complicated. This is due to the dynamical description of the solvent around GL, but especially to the larger conformational flexibility of GL; the alternation of sign is caused both by the fluctuations of the solvent molecules in time and by the the interconversion between the several GL conformers.
The convoluted average QM/FQ spectrum is reported together with the experiment[252] in Figure 2.17. The calculated VCD spectrum is characterized by a (-,+,-,+) sign pattern in the region between 1100 and 1330 cm^{-1}, followed by two negative bands between 1400 and 1550 cm^{-1}. All signs result from the averaging of the sticks in

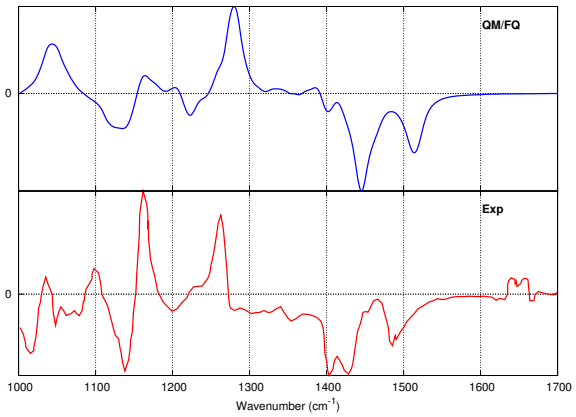


Figure 2.17. Convoluted QM/FQ (top) and experimental spectrum[252] (bottom) VCD spectrum of GL in aqueous solution.

Figure S14 and is it remarkable that such a sing alternation particularly affects the region 1100-1300 cm^{-1}.

Moving to the comparison with experimental data (Figure 2.17) it is worth noticing that most of the experimental signs are correctly reproduced by the QM/FQ model. However, some differences between the computed and the experimental spectra are present and deserve discussion. In fact, the experimentally most intense peak is located at about 1160 cm^{-1}, while the calculated rotational strength in this region is very low. The associated normal mode involves a diffuse mode dominated by the bending of the hydroxyl group, which is hugely affected by the atomistic description of the surrounding water molecules. Moving to the negative peak at about 1220 cm^{-1}, as a matter of fact, the VCD spectrum of GL in CCl_4 solution shows at this frequency a high and positive peak, which becomes negative and with low intensity in aqueous solution.[252] This is the main difference between the spectra measured in CCl_4 and the H_2O spectrum, thus meaning the this band in the whole VCD spectrum is the most modified by the nature of the surrounding environment. The correct sign reproduction of this peak is one of the most remarkable results achievable by resorting to our discrete model, in which the correct sign results from the averaging of negative and positive sticks in Figure S13. It is also to be noticed that the use of a continuum PCM solvent description (see Figure S15 in the ESI) totally fails to correctly reproduce the sign of this band.

Still on Figure 2.17, the calculated spectrum in the region between 1400 and 1500 cm^{-1} is in good agreement with the experiment, both sign and intensity. The first intense negative band in this region is experimentally composed by two distinct peaks, while our calculated spectrum reports only a broad band. Probably, the second peak is hidden by the first one (see Figure S14), thus resulting in a broad band instead of two separate peaks.

To end the discussion on the VCD spectrum, we note that the band between 1600-1700 cm^{-1} is again due to chiral imprinting in the solvent molecules, and cannot therefore reproduced with our focused model.

Raman Scattering Spectrum

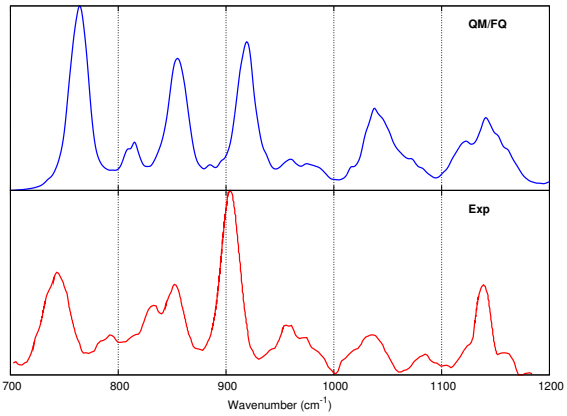


Figure 2.18. Convoluted QM/FQ (top) and experimental[285] (bottom) Raman spectrum of GL in aqueous solution. Excitation wavelength: 1064 nm

The Raman spectrum was calculated through analytical evaluation of energy third derivatives.[150] Notice that the same sampling as exploited for the other properties (i.e. 200 snapshots) was exploited, based on previous convergence studies.[150] Raw data calculated at 1064 nm are depicted in Figure S16. Similarly to IR and VCD, also Raman intensities vary as a function of the snapshot. Again, it is worth remarking that this results from the complexity of this molecule, which presents several conformers in aqueous solution and to the atomistic description of the solvent molecules obtained by exploiting the QM/FQ model.

The convoluted calculated spectrum together with experiment[285] is reported in Figure 2.18. The computed spectrum is characterized by 3 principal bands that can be associated to diffuse normal modes involving O6 and the OH group, and to a diffuse vibration involving all the atoms in the molecule. Furthermore, two broad bands are predicted in the region 1000-1200 cm^{-1}. Inhomogeneous broadening, particularly evident in the latter bands, occurs from the spreading of the single peaks depicted in Figure S10.

Figure 2.18 also reports the experimental spectrum, which was measured at 1064 nm.[285] Some differences between the computed and the experimental relative intensities are present. This is particularly evident for the bands at 750, 850 and 900 cm^{-1}, probably due to a not perfect description of the statistical distribution of the several conformations of GL in aqueous solution. However, bands inhomogeneous broadening is almost perfectly reproduced, especially for the small bands at about 800-820 cm^{-1} and above 950 cm^{-1}, which are in almost perfect agreement with the experiment. We note that also for Raman the continuum PCM solvent description (see Figure S17 in the ESI) is not able to correctly reproduce the spectral patterns.

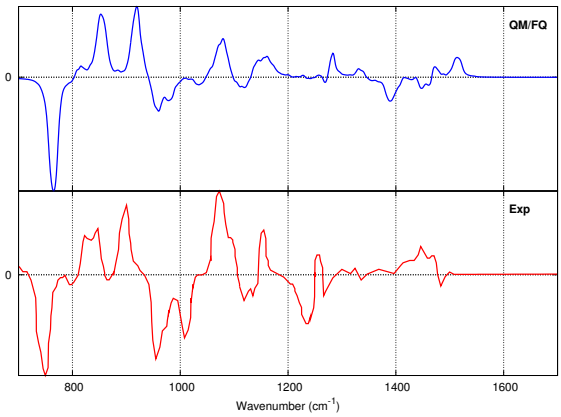


Figure 2.19. Convoluted QM/FQ (top) and experimental [302] (bottom, neat liquid) ROA spectrum of GL in aqueous solution. Excitation wavelength: 514 nm

ROA

ROA raw data calculated at 514 nm are depicted in Figure S18 in the ESI. Similarly to VCD, a great variability in peaks intensities and sign is reported.
The convoluted calculated spectrum is reported in Figure 2.19, together with the experimental spectrum measured for the neat liquid. [302] The computed spectrum is characterized by an intense (-,+,+) pattern that can be associated to diffuse normal modes involving O6 and the OH group.
Moving to the comparison between calculations and experimental spectra , we note that the relative intensities of the bands between 950 and 1200 cm^{-1} are not perfectly reproduced by our model, however the inhomogeneous band broadening is well described. Such differences are again due to a not perfect description of the statistical distribution of the several conformations of GL in aqueous solution. In addition, the experimental spectrum refers to GL as neat liquid, in which the intermolecular interactions between the GL-GL molecules are largely different from those modelled in our computational sample. Also in this case, the limitations of the continuum PCM approach are evident from the inspection of the spectra reported in Figure S19 in the ESI.

2.5 Conclusions and Future Perspectives

In this paper, the potentialities of a multi-scale focused approach based on the combination of classical MD simulations and a fully polarizable atomistic QM/FQ Hamiltonian are shown through its application to the calculation of VOA spectra of two chiral systems, able to strongly interact, via hydrogen bonding, with the surrounding aqueous solution. The fully atomistic character of the approach permits a reliable modeling of specific solute-solvent interactions, and the fully account of electrostatic solute-solvent mutual polarization effects yields a reliable description of the solvation phenomenon. Also, the coupling with classical MD simulations permits to account for the solva-

tion dynamics. Remarkably, the application of the method to the calculation of IR, VCD, Raman and ROA spectra of (L)-Methyl Lactate and (S)-Glycidol in aqueous solution shows an excellent agreement between computed and experimental spectra, and of higher quality if compared to standard continuum solvation approaches. Such discrepancies are not only due to the inaccurate description of HB interactions in the continuum approach, but also to a different sampling of the PES resulting from the static PCM or dynamic QM/FQ+MD approaches. Obviously, there is a price to pay for the greatest accuracy. As far as the single QM/FQ calculation on a single snapshot is concerned, the required computational cost is comparable to the corresponding PCM calculation. However, the single calculation needs to be repeated for the number of snapshots required to reach a good modeling of the configurational space, i.e. the final cost of the calculation strongly depends on the number of snapshots which have to be considered to reach the convergence of the desired property. Such a number cannot be defined *a priori*, and especially can wildly vary as a function of the property/spectroscopy to model, from hundreds[56,148–150,162] to thousands[146] of representative snapshots.

A limitation of our approach comes from a closer inspection of the IR/VCD spectra of the two systems in the region 1600-1700 cm^{-1}, which was assigned[249,250,250,252] to the bending mode of water molecules bound to the chiral solutes as a result of HB interactions. Interestingly, such bands are also active in the VCD spectra, giving rise to the so-called "chirality transfer" phenomenon, i.e. the chiral supramolecular arrangement of the achiral solvent molecules around a chiral solute. Such a phenomenon cannot be modelled by our approach, which focuses on the vibrational transitions of the solutes, and their modifications as a result of the interaction with the surroundings. A possible way to extend the model to such phenomena would be to extend the definition of "solute", i.e. including not only the chiral system but a few water molecules in the QM portion, similar to what is done for continuum solvation studies.[28,313,314] An alternative approach, which appears to be more adequate to our general framework, would be to resort to ab-initio MD[315] techniques coupled to the FQ approach.

As last conclusion, we note that only purely electrostatic solute-solvent interactions have been considered in the current approach. QM/FQ non-electrostatic interactions may play a role and should be included, similar to what has recently been proposed by some of the present authors but to date only limited to the evaluation of energetic properties.[197]

Electronic Supplementary Information

Electronic supplementary information (ESI) available online.
See DOI: 10.1039/C8CP00487K

Chapter 3

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

DOI: 10.1002/qua.25684

FULL PAPER

WILEY International Journal of QUANTUM CHEMISTRY

Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling

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

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

3.1 Introduction

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

With the gradual increase in the complexity of the systems under investigation, the correct interpretation of experimental data is increasingly reliant upon their calculated *ab-initio* counterparts. Many theoretical models based on quantum mechanics (QM), accompanied by their computational implementations, have been presented over the years offering different levels of compromise between the computational cost and the accuracy of the results.[316–318] At present, methods based on density functional theory (DFT) and its time-dependent counterpart (TD-DFT) have become the most popular choice for the simulation of absorption spectra of medium-large organic molecular systems thanks to their versatility stemming from the freedom of choice of density functional and basis set, as well as the favorable scaling with system size which allows their application to increasingly large systems.[316,319–321]

Many benchmarks studies have been presented elaborating upon the merits and limitations of TD-DFT for the simulation of UV-Vis spectroscopy, as well as on the most appropriate choice of functional and basis set combination for different types of system.[321–331] And though many computational studies are carried out on isolated systems, solvent effects should not be neglected for the presence of the solvation environment can significantly alter the electronic absorption properties of a system, both qualitatively and quantitatively.[102–105,332–343] For this reason, theoretical models have been developed to tackle this problem and then combined with DFT and TD-DFT to include solvent effects within the theoretical model.

The standard protocol for such cases requires the usage of Polarizable Continuum Model (PCM) to describe solvation.[6,70,104,105] This approach falls into the category of implicit solvent models where the environment is represented in a continuous way, while the solute molecule sits within a cavity and the surrounding continuum possesses dielectric properties that mimic the given solvent. Implicit methods prove to simulate correctly the properties of non-aqueous solutions lacking specific interactions between the solute and solvent molecules. Meanwhile, in water as well as in numerous other media the directional interactions, such as hydrogen bonds (HB), can play a crucial role. Hydrogen bonding can be introduced within continuum models by including explicit solvent molecules treated quantum mechanically in the system. This quantum treatment assures that both the directional nature of hydrogen and its covalent contribution are treated, however this still relies upon a static description of the system, whereas in reality the solvent moves about around the solute and a physically correct picture should not neglect the fact that the system is but an ensemble of many different configurations that may have varying spectroscopic properties. In fact, even though the positions of the explicit solvent molecules may be optimized to obtain a minimum-energy-structure, many such structures may be obtained in principle because of the high flexibility of the supramolecular system, but none of them taken singularly may be representative of the whole. In pure PCM, the converged quantum mechanical

density is considered an *implicit* average over the configurational space of the solvent, as the latter is viewed as smeared out, which justifies the continuum picture. A supermolecule approach which re-introduces some solvent molecules as explicit static entities therefore helps to include some crucial interaction into the picture[343] but carries its own problems.

In order to overcome the limitations of implicit solvent models, explicit solvation models have been developed in the past years. The widest used explicit approach is based on the Quantum Mechanics/Molecular Mechanics (QM/MM) multi-scale scheme. The system is divided into the portion directly responsible for a given property (e.g. a chromophore interacting with light) that is described at the QM level and the surrounding molecules described at the MM (classical) level through *ad-hoc* constructed Force Field (FF). In the most commonly used QM/MM approaches, only the electrostatic interaction between the two portion is considered. In particular, each atom is endowed with a fixed pre-parametrized charges, giving rise to the so called non-polarizable QM/MM models. Hovewer, to recover a better and more physical description of the electrostatic interaction between the QM and MM portions, several polarizable QM/MM models, in which the MM atoms can be polarized by the QM density, have been developed. Such models can be based on distributed multipoles,[58,185,186,281] induced dipoles,[10,155,164,166] Drude oscillators[59] or Fluctuating Charges (FQ).[56]

The FQ model was firstly developed into a 3-layer fully polarizable approach with non-periodic boundary conditions (QM/FQ/PCM). The method has subsequently been extended to allow calculations of numerous molecular properties by including features like analytical first and second derivatives,[55] response equations,[54] magnetic perturbations with Gauge Including Atomic Orbitals (GIAOs),[161] excitation energies (at the TD-DFT and equation-of-motion coupled cluster model with single and double substitutions levels of theory),[148,344] Vibrational Optical Activity,[149,150] excitation energy gradients.[162]

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

The article is organized as follows: in the next section we briefly recall the fundamentals of the QM/FQ model, by also focusing on its extension to calculate TD-DFT/FQ excitation energies. After a section dedicated to the details of the computations, the results for selected organic dyes in aqueous solution, previously studied with different approaches,[102,345] are discussed. In particular, the discussion focuses on Molecular Dynamics (MD) and TD-DFT results. Eventually, we come to conclusions and considerations for further perspectives.

3.2 Theoretical Model

The FQ model provides a computationally efficient and chemically consistent way of introducing polarization effects within both classical dynamics simulations and QM/MM

calculations.[56,346] In the FQ model, each atom is endowed with a charge that is allowed to fluctuate. Such fluctuation is ruled by the difference in atomic electronegativities. Thus, two set of parameters are needed to describe the FQ energy, namely atomic hardnesses and electronegativities. It is worth noticing that such terms can be rigorously defined in the "conceptual DFT" framework.[286,289] Through them, the atomic charges can be calculated by solving a linear system.[156–158] More in detail, the value of the fluctuating charge on each MM atom is related to the electronegativity[156–158] according to the Electronegativity Equalization Principle (EEP),[160,286] which states that, at equilibrium, the instantaneous electronegativity χ of each atom has the same value.[160,286] The FQs ($\mathbf{q}$) can be defined as those minimizing the following functional:[159]

$$\begin{aligned} F(\mathbf{q},\lambda) =& \sum_{\alpha,i} q_{\alpha i}\chi_{\alpha i} + \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j} q_{\alpha i}J_{\alpha i,\beta j}q_{\beta j} + \sum_{\alpha}\lambda_{\alpha}(\sum_{i} q_{\alpha i} - Q_{\alpha}) \\ =& \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} \end{aligned} \tag{3.1}$$

where $\mathbf{q}$ is a vector containing the FQs, the Greek indices α run over molecules and the Latin ones i over the atoms of each molecule. $\boldsymbol{\lambda}$ is a set of Lagrangian multipliers used to impose charge conservation constraints on each molecule. In this work, the charge interaction kernel $\mathbf{J}$ is the Ohno kernel.[291] Atomic units are used throughout the manuscript. The stationarity conditions of the functional in eq.3.1 are defined through the following equation:[159]

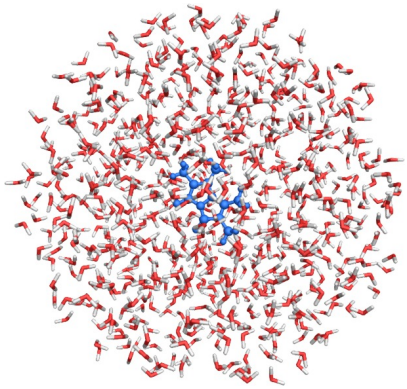
$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_{Q} \tag{3.2}$$

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

The FQ force field (FF) can be effectively coupled to QM methods. The resulting QM/FQ approach[56] has been shown to be especially suited to the modeling of response and spectral properties because, as it is shown below, its energy expression can be easily differentiated up to high orders. The QM/FQ describes also polarization effects: in this contest, the charges equilibrate to both the electrostatic potential generated by the QM moiety and their electronegativies, while the QM core feels the presence of the FQs through specific additional terms in the QM Hamiltonian, in a mutual polarization fashion.

The QM/FQ model system is usually partitioned in a QM core region placed at the center of a spherical region defining the environment (see Figure 3.1), which is described classically by exploiting the FQ FF. The size of this region is chosen to guarantee the convergence of the desired property/spectrum. Notice that the position of QM and MM atoms is obtained by a previous performed classical Molecular Dynamics (MD) allowing the exploration of the configuration space. This gives rise to a dynamic approach to the solvation phenomenon, which is instead neglected in purely continuum approaches.

Figure 3.1. Representation of the QM/FQ scheme. The QM portion is in blue.



The coupling between the QM and the MM portions is defined as the classical electrostatic interaction:[54]

$$E_{\mathrm{QM/FQ}} = \sum_{i=1}^{N_q} V_{\mathrm{QM}}[\rho](\mathbf{r}_i) q_i \qquad \textbf{3.3}$$

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,[197] 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:[54,55,192]

$$\mathcal{E}[\mathbf{P}, \mathbf{q}, \boldsymbol{\lambda}] = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{1}{2}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} + \mathbf{q}^{\dagger}\mathbf{V}(\mathbf{P}) \qquad \textbf{3.4}$$

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

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q - \mathbf{V}(\mathbf{P}) \qquad \textbf{3.5}$$

which includes the coupling term $\mathbf{V}(\mathbf{P})$ between the QM and MM moieties.
In case of the calculation of response/spectroscopic properties, such terms propagate to the solute's response equations, so that polarization effects are fully considered also in the computed final spectral data.[54,55,147,149,150,161]

3.2.1 Linear Response Theory in QM/FQ

In order to calculate excitation UV-VIS spectra, we briefly recall how linear response equations have to be changed to account for the presence of the FQ portion. For a more detailed discussion on this topic, we refer the reader to ref.[54] The following matrices,

depending on the FQ charges, are defined:

$$\tilde{A}_{ai,bj} = (\varepsilon_a - \varepsilon_i)\delta_{ab}\delta_{ij} + \langle aj||ib\rangle - \sum_{kl}^{N_q} V_{ia}^{\dagger} D^{-1} V_{jb} \tag{3.6}$$

$$\tilde{B}_{ai,bj} = \langle ab||ij\rangle - \sum_{kl}^{N_q} V_{ia}^{\dagger} D^{-1} V_{bj} \tag{3.7}$$

where i, j are occupied orbitals whereas a, b are virtual orbitals. ε are orbital energies. The sum runs over the molecules in the classical portion and V is the electrostatic potential. D is the FQ matrix introduced in the previous section (see Eq. 3.5). Then, excitation energies and transition amplitudes are obtained by solving the so-called Casida's equations for the QM/FQ linear response theory:[54]

$$\begin{pmatrix} \tilde{A} & \tilde{B} \\ \tilde{A}^* & \tilde{B}^* \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{3.8}$$

3.3 Computational Details

For this work we have selected seven molecules, depicted in Figure 3.2. These systems are all organic molecules for which experimental measurements of their UV-Vis absorption properties in aqueous solution exist. Furthermore, several of these system are capable of forming intermolecular hydrogen bonds, and their absorption spectra exhibit bright excitations with varying degree of charge-transfer character, and can therefore be affected by the presence of the highly polar solvation environment to a different extent. All QM calculations were performed using the Gaussian16 program,[301] and employed CAM-B3LYP density functional[347] and 6-311++G(d,p) basis set. Continuum solvation effects were included using the polarizable continuum model (PCM). The ground-state geometry for each molecule was optimized at the QM/PCM taking into account the possible presence of multiple conformers, then the first five excited states for each were converged at the TD-DFT/PCM level. Non-equilibrium solvation effects[6,105,106,121] were modeled by resorting to the Linear Response (LR) formalism. In order to estimate the effect of the covalent and directional components of hydrogen bond, both the geometry optimizations and spectra calculations were repeated after saturating every hydrogen bonding site with a water molecule (QM/QM$_w$/PCM results). The QM/MM calculations of excitation energies and intensities were performed by resorting to the following computational steps:

1. *Definition of the systems and calculation of atomic charges.* The solute molecules were surrounded by a number of water molecules large enough to represent all the solute-solvent interactions. The atomic charges of the solute were computed by using the Charge Model 5 (CM5).[298]

2. *Classical MD simulations in aqueous solution.* The MD simulations were performed in a cubic box reproduced periodically in every direction, satisfying the Periodic Boundary Conditions (PBC). A minimization step ensures that the several simulations were started from a minimum of the classical PES. From the MD runs, a set of snapshots was extracted to be used in the QM/MM and QM/FQ calculations.

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

4. *QM/MM or QM/FQ calculations and comparison with experimental data.* QM/MM or QM/FQ excitation energies calculations were performed on the set of structures obtained for the seven molecules in the previous step of the protocol. The results obtained for each spherical snapshot were extracted and averaged to produce the final value.

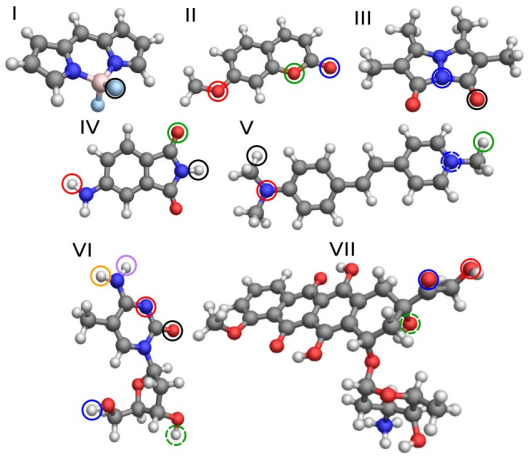


Figure 3.2. Representation of the studied molecules: **I** (bodipy); **II** (7-methoxycoumarin); **III** (bimane); **IV** (5-aminophtalimide) ; **V** (pyridinium dye); **VI** (5-methylcytidine); **VII** (doxorubucin).

The systems studied in this manuscript are depicted in Figure 3.2 and their simplified name are reported in Table 3.1. Notice that IUPAC names are reported in Table S1, given as Supporting Information (SI).
In step 1, the systems were optimized and CM5 charges were calculated at the B3LYP/6-31+G* level of theory including solvent effects by means of the PCM.[6,70] For sake of completeness, the vertical energies were also computed using the PCM model at the CAM-B3LYP/6-311++G** level of theory. Depending on the case of study, explicit solvent molecules were included (QM/QM$_w$/PCM).
The MD simulations were performed by using GROMACS,[348] with the GROMOS,[349] GAFF[350] and Amber11[351] force fields to describe intra-/inter-molecular interactions.

The force field used depend on the considered solute **I-VII** and are summarized in Table 3.1. CM5 charges were used to account for electrostatic interactions. The TIP3P-FB FF was used to describe the water molecules.[352] A single molecule was dissolved in a cubic box containing at least 3000 water molecules. The number of water molecules varies depending on the dimension of the considered molecules **I-VII** (see Table 3.1 for the exact number for each structure). For molecule **V**, a chloride ion has been included in the box to neutralize the system. The molecular systems were initially brought to 0 K with the steepest descent minimization procedure and then heated to 298.15 K in an NVT ensemble using the velocity-rescaling[353] method with an integration time step of 0.2 fs and a coupling constant of 0.1 ps for 200 ps. The time step and temperature coupling constant were then increased to 2.0 fs and 0.2 ps, respectively, and an NPT simulation (using the Parrinello-Rahman barostat and a coupling constant of 1.0 ps) for 1 ns was performed to obtain a uniform distribution of molecules in the box. 100 ns production runs in the NVT ensemble were then carried out, fixing the fastest internal degrees of freedom by means of the LINCS algorithm (δt=2.0 fs).[354] Electrostatic interactions are treated by using particle-mesh Ewald (PME)[355] method with a grid spacing of 1.2 Å and a spline interpolation of order 4. The cross interactions for Lennard-Jones terms are calculated using the Lorentz-Berthelot[356,357] mixing rules and we have excluded intramolecular interactions between atom pairs separated up to three bonds. A snapshot every 500 ps was extracted in order to obtain a total of 200 uncorrelated snapshots for each system. For each snapshot a solute-centered sphere with radius of at least 17 Å of explicit waters was cut. The radii used for each molecule **I-VII** are summarized in Table 3.1. Notice however that for molecule **V**, the chloride ion was not present in any of the extracted spherical snapshots. For each snapshot, the excitation energies were then calculated with two QM/MM models, treating the QM portion at the CAM-B3LYP/6-311++G** level. The water molecules were modeled by means of the non-polarizable TIP3P FF,[358] and the FQ SPC parametrization proposed by Rick *et al.*[156] Further calculations were performed by adding the closest water molecules in the QM portion, in a QM/QM$_w$/FQ framework. The water molecules to be included in the QM/QM$_w$/FQ calculations were chosen case by case by looking to the maximum/a in the Radial Distribution Function (RDF or $g(r)$) calculated from the MDs. The average number of water molecules included in the QM portion are reported in Table 3.1. All the QM(/QM$_w$)/MM(FQ) calculations were performed by using a locally modified version of Gaussian 16 package.[301]

3.4 Results

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

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

Table 3.1. Assigned number and simplified names for the studied molecules. Number of water molecules (N_{H_2O}) and Force Field used for molecules **I-VII** for MD simulations. The sphere radius for the QM/MM calculations and the average number of water molecules included in QM/QM_w/FQ calculations (N_{QM_w}) are also reported.

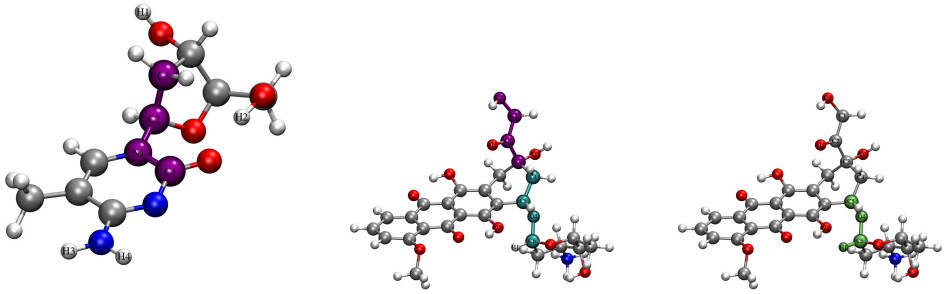
interaction can affect the predicted spectroscopic signature is provided for each system.

3.4.1 MD Analysis

The MD trajectories of molecules **I-VII** were analyzed to provide information about how the solvation environment affects the conformational space explored by the systems and about the intermolecular interaction through Hydrogen Bond (HB) with water molecules. This analysis was performed by using TRAVIS package.[359]

Conformational analysis based on MD simulations

Among the seven considered molecules, only **VI** and **VII** are flexible and present different minima in the potential energy landscape that can be optimized using QM techniques. In Figure 3.3 the molecular structures of these molecules are reported highlighting the dihedral angles that define the different conformations.



(3.3.1). VI. DO(purple)

(3.3.2). VII. D1 (purple) and D2 (cyan)

(3.3.3). VII. D3 (green)

Figure 3.3. (a) **VI** (5-methylcytidine, *syn* conformer),(b) and (c) **VII** (doxorubicin) structures. The dihedral angles studied in the conformation analysis are colored. Relevant atoms are also labeled.

In molecule **VI**, the flexibility is due to the rotation around the D0 dihedral angle (see Figure 3.3.1), which defines the relative orientation of the sugar and the pyrimidine ring. Two main conformers can be present, namely *syn* and *anti* (see section S2 in the Supporting Information). Notice that in Figure 3.3.1, only the *syn* conformer is depicted. Figure 3.4 reports the time evolution of the D0 dihedral angle highlighted in Figure 3.3.1 as the classical dynamics unfolds, together with its resulting distribution. The *syn* conformer (D0 ≈ -70 degrees) results to be by far the most abundant, however the *anti* conformer (D0 ≈ 130 degrees) is also present, albeit with a much lower and less sharp population.
The plot in figure 3.4 also points out to one of the problems that may be encountered when resorting to a continuum solvation model, as a QM/PCM geometry optimization yields one structure which may not be representative of the whole conformational space spanned by the solute around the relative minimum which is, in this case, quite wide.

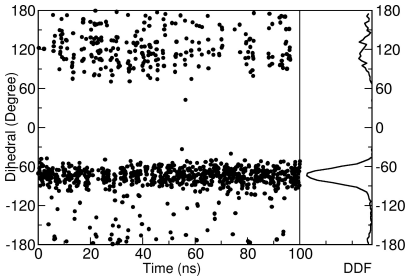


Figure 3.4. Time development and dihedral distribution of the dihedral angle D0 of molecule **VI**.

Concerning molecule **VII**, the main free rotation is given by the amino-sugar moiety with respect to the rest of the molecule. Three different dihedrals can be identified, namely D1, D2, D3 (see Figures 3.3.2 and 3.3.3). The distributions of such dihedrals during the MD are depicted in Figure 3.5. The D1 maximum value is placed at about -135 degree indicating that no intramolecular interaction between O3 and the hydrogen bounded to O1 can occur. In figure 3.5, the D2 distribution shows a three peaks profile (81, 110, 157 degrees), whereas two peaks are identified in the D3 profile (-34, 44 degrees). It is worth noticing that D2 and D3 profiles show a correlation. In fact, such dihedrals are responsible of intramolecular interaction O5$\cdots$H1, which occurs when D2 and D3 are at about 44 and 110/157 degrees, respectively, which is not a highly populated area of the conformational landscape. This is confirmed also by the Radial Distribution Function (RDF, or $g(r)$) reported in the inset of Figure 3.5. It shows that the least intense peak at about 2.5 Å is associated with the intramolecular interaction just discussed.

Hydration Patterns

Molecule **I**, whose solvation properties were previously investigated, albeit with a lack of polarization effects in the classical portion,[360] is characterized by two fluorine atoms that can be involved in Hydrogen Bonds (HB) with water hydrogen atoms. In Figure

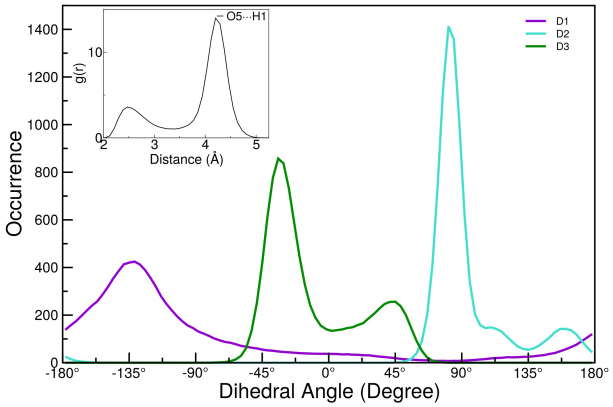


Figure 3.5. Population analysis of D1, D2 and D3 depicted in Figure 3.3.2 for molecule **VII**. In the inset, the radial distribution function between O5 and H1 is also reported (see Figure 3.3.2 for atom labeling).

3.6(a), the $g(r)$ between F(**I**) and H(water) atoms is reported. The first peak at about 1.95-2.00 Å is intense and broad and it clearly represents an HB pattern.

Concerning molecule **II**, the three chromophore oxygen atoms are potentially responsible for HB interactions with water molecules. The $g(r)$ of molecule **II** is depicted in Figure 3.6(b). The carbonyl oxygen atom is the only one that presents values of O(**II**)$\cdots$H(water) distances characteristic of hydrogen bond. To further analyze this interaction, the Combined Distribution Function (CDF), i.e. $g(r)$ as function of the O(**II**)$\cdots$H(water) distance and the angle O(chromophore)$\cdots$HO(water), was calculated. The plot is depicted in Figure S3 given as SI. This analysis confirms that the carbonyl oxygen atom is involved in an HB with the surrounding water molecules.

Molecule **III** is characterized by two carbonyl oxygen atoms which are potentially bonded to water molecules through an HB interaction. The $g(r)$ is plotted in Figure 3.6(c). The O(**III**)-H(Water) rdfs are equal due to the molecular symmetry and they present a peak at about 1.8 Å which is an indicator of HB interaction. Notice that the integral value of the $g(r)$ shows that molecule **III** interacts on average with two water molecules at the same time.

Molecule **IV** presents two carbonyl oxygen atoms which may act as HB acceptors as well as an amine and an imide group which instead can act as HB donors. Molecule **IV** $g(r)$ is depicted in Figure 3.6(d). The HB donor character is the predominant one and the imide hydrogen atom is the most involved in the HB interaction. This is also confirmed by the Spatial Distribution Function (SDF) reported in Figure S4, given as SI.

Molecule **V** is characterized by the presence of two nitrogen atoms (aminic and pyridinic) whose ibridization character does not allow for the formation of hydrogen bonds with the solvent. The $g(r)$ presented in Figure 3.6(e) confirms this, with distances around 4.5-5.0 Å.

Molecule **VI** is characterized by several potential HB sites. The carbonyl and ether oxygen atoms, together with the iminic atom are potentially HB acceptors, whereas

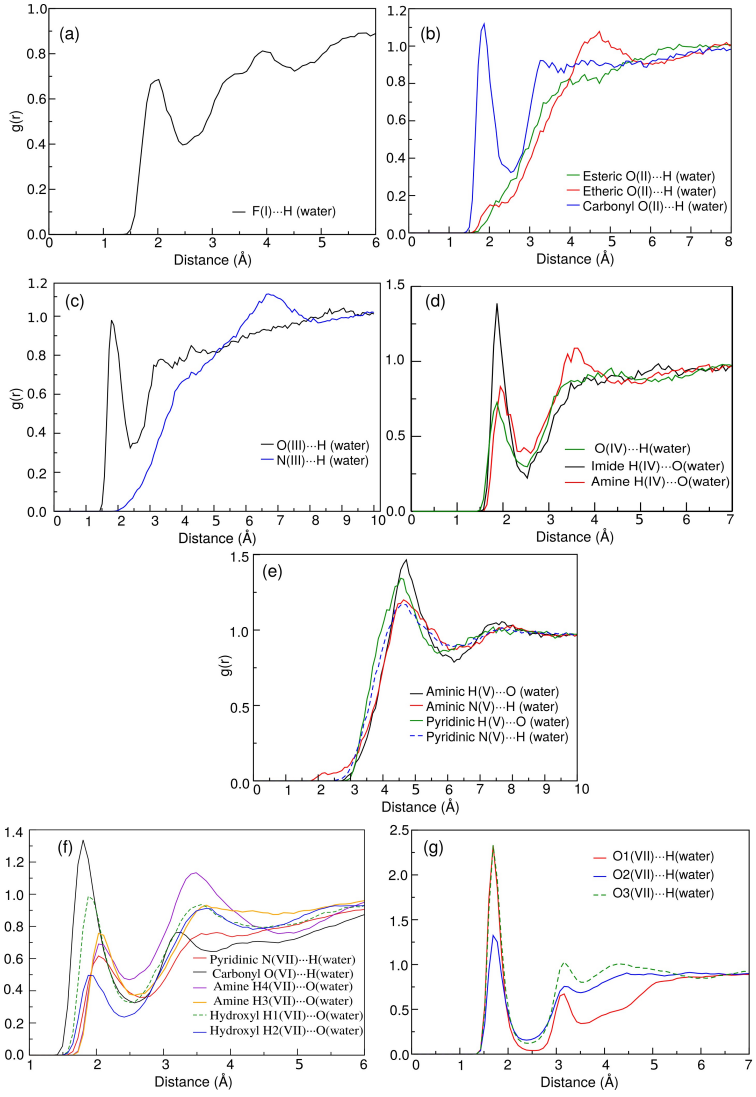


Figure 3.6. Radial distribution function between selected sites of all the analyzed chromophores and water molecules: (a) **I**, (b) **II**, (c) **III**, (d) **IV**, (e) **V**, (f) **VI**, (g) **VII**. Sites are highlighted in Figure 3.2.

the oxygen and nitrogen atoms which are bound to an hydrogen atoms are potential HB donors. The $g(r)$ is depicted in Figure 3.6(f). The most intense peak of the several $g(r)$ is shown by the carbonyl oxygen atom, which means that molecule **VI** behaves mostly as HB acceptor. It is worth noticing that $g(r)$ of the hydrogen atoms (H3 and H4 in Figure 3.3.1) of the amine group are not equivalent. In fact, at about 3.6 Å, the $g(r)(\mathrm{H4}\cdots\mathrm{O}_{H_2O})$ presents a broad peak which is instead absent in case of H3. This result suggests that the rotation around the C-N bond is blocked, and also that this bond has a partially double character. This can be chemically explained through the resonance of the lone electron pair on nitrogen with the π electrons of the aromatic ring. This conclusion is also supported by the CDF between the intermolecular distance of the pyridinic nitrogen and the water hydrogen atoms and the $g(r)$ of the amine $\mathrm{H4}\cdots\mathrm{O}_{H_2O}$ (see Figure S5 given as SI). The CDF shows that when a water molecule interacts with the pyridinic nitrogen atom the relative distance H4···O(water) is exactly around 3.6 Å explaining the second broad peak in the g(r)$\mathrm{H4}\cdots\mathrm{O}$(water).
Molecule **VII** hydration pattern has already been studied extensively in a previous paper by some of the present authors and we direct an interested reader to this publication for more details. In this manuscript, the main intermolecular interactions between the chromophore and the water molecules are reported (see Figure 3.6(g)). Such interactions involve the oxygen atoms O1, O2 and O3 (see Figure 3.3.2, for atom labeling) and the hydrogen atoms of water. The three $g(r)$ are characterized by a peak placed at about 1.8 Å , and the most intense ones are related to the hydroxyl oxygen O1 and O3.

3.4.2 Excitation Energies

We now move on to present the results obtained by exploiting the continuum and explicit (with or without polarization effects) approach to the calculation of the absorption spectra of each of the seven molecules. In general, five models of increasing complexity will be considered: (1) a purely continuum QM/PCM model, (2) a QM/QM_w/PCM model where explicit solvent molecules are included in the QM part to saturate hydrogen bond sites, (3) a non-polarizable QM/TIP3P hybrid quantum-classical model where the solvent is non-polarizable but rather treated using fixed charges, (4) the polarizable QM/FQ model to expose the role of solvent polarization in generating the spectroscopic response, and finally (5) a QM/QM_w/FQ model which treats some water molecules close to the solute hydrogen bonding site(s) at the DFT level to model any covalent effects that might be of importance. As aforementioned, to calculate the QM/FQ spectra 200 uncorrelated snapshots were extracted from the MD simulations; such a number is enough to yield a converged spectrum as already pointed out by some of the present authors.[149–152] The QM/FQ convoluted and averaged spectra for molecules **I-VII** are reported in Figure 3.7. Therein the QM/PCM, QM/QM_w/PCM, QM/TIP3P, QM/QM_w/FQ and experimental data are also plotted. For QM/QM_w/PCM (i.e. with some water molecules explicitly introduced in the QM portion) further details are given in Figure S6 and in Table S2, given as SI. Notice that for molecule **V** and **VII** no explicit water are included in QM/PCM calculations: for molecule **V** this is due to the fact that no specific interactions are present

in aqueous solution, as confirmed by Figure 3.6. Molecule **VII** presents three distinct conformers and several hydrogen bonding sites. Because of this, converging the geometries for all structures proved difficult, therefore we omitted this molecule from the analysis, though the inclusion of explicit QM solvent molecules was still done for the QM/QM$_w$/FQ model, which showed no significant change compared to the QM/FQ results (vide infra). For all molecules we considered the first bright excitation. Before comparing the different solvation models, the character of each excitation was investigated by looking at the molecular orbitals (MO) involved in the transition , which are depicted in Figure S7, given as SI. The MOs show that most of the excitations can be classified as charge-transfer states, however the exact degree to which the electron density is displaced during the excitations should be evaluated in order to provide more quantitative results. To this end we employ a simple and intuitive numerical index recently developed[361] which considers the baricenters of the positive and negative difference density. To analyze the Charge Transfer (CT) nature of the first electronic transition, the extension of the length of the electron transfer, we used a simple and intuitive index, denoted as D_{CT}, that was recently developed.[361] The barycenters of the positive and negative density distributions are calculated by the difference of the Ground State (GS) and Excited State (ES) densities. The CT length (D_{CT}) is defined as the distance between the two barycenters. In table 3.3, the D_{CT} for molecule **I-VII** are reported. Notice that also another quantitative index (Δr) proposed by Guido *et al.* based on MOs was employed in the analysis of the CT nature.[101] The values obtained by exploiting this alternative index are reported in Table S3, given as SI.

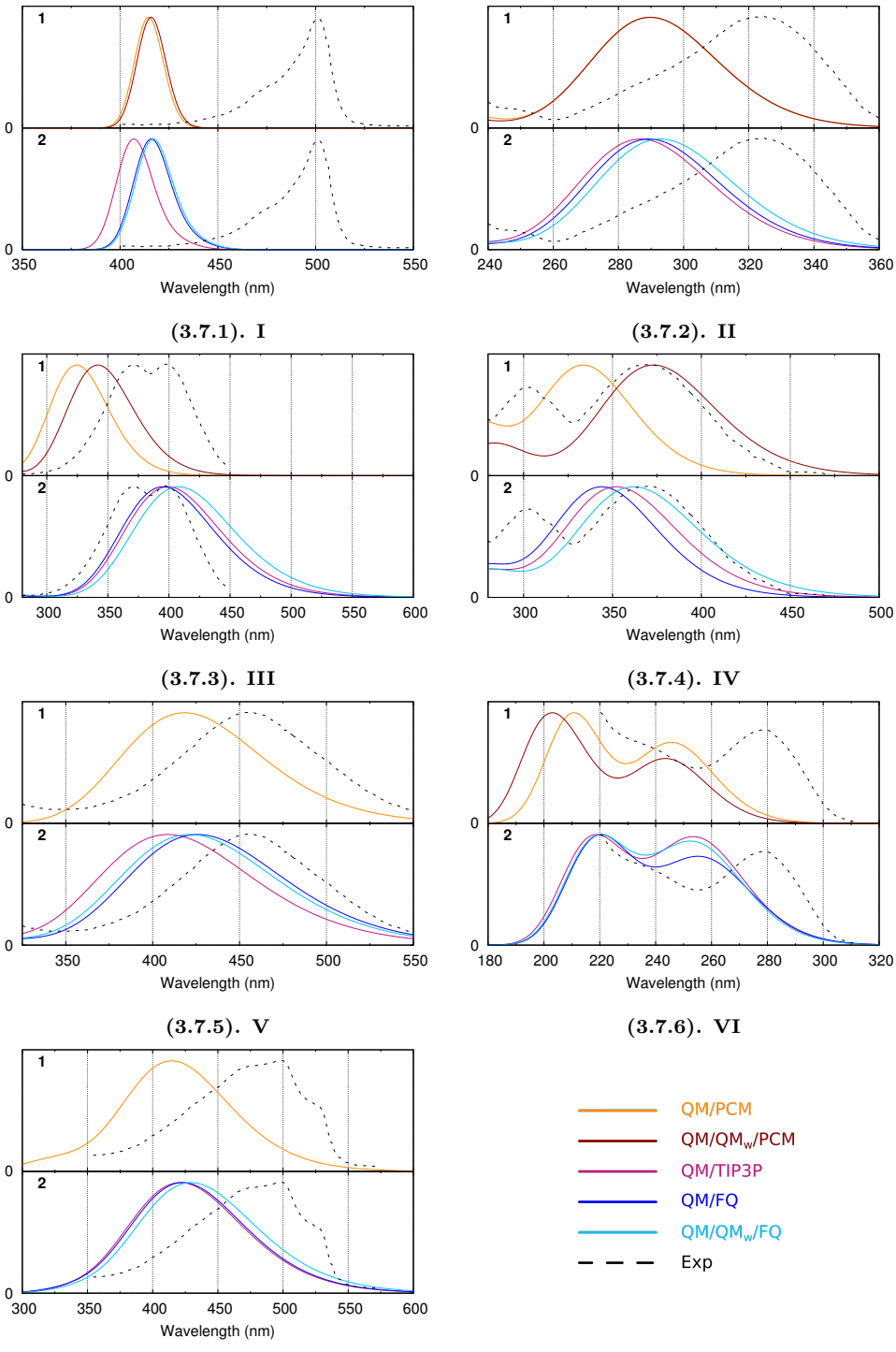
We start the discussion on excitation energies by focusing on molecule **I**. The experimental spectrum in Figure 3.7.1 is characterized by a band at about 500 nm. To guarantee a direct comparison with the experimental spectrum, for this molecule all the data have been convoluted with a Gaussian function with a Full Width Half Maximum (FWHM) of 0.13 eV. The calculated D_{CT} reported in Table 3.3 clearly show that the first excitation has no CT character. This is also confirmed by the MO involved in the transition depicted in Figure S7, given as SI. From Figure 3.7.1 and Table 3.2, it is clear that the several solvation models used in this work result in very similar excitation energies which differ from the experiment of about 20%, except for the non polarizable QM/TIP3P which is the worst method employed. Because the inclusion of polarization effects into the solvation model, while moving the results in right direction does not lead to a quantitative agreement, and considering the rigidity of the molecular structure, the observed difference between theory and experiment is most likely due to poor description afforded by chosen DFT functional. In fact, it has been shown that bodipy dyes do require higher level QM theory models that adequately describe differential electron correlation to produce accurate results.[362,363] On the other hand, all the polarizable models predict the same energy, meaning that polarization effects play indeed a role in capturing solvation effects for this molecule. Also, the inclusion of some water molecules in the QM portion, both in the case of QM/PCM and QM/FQ, does not give a relevant improvement in the comparison with experimental data.

Concerning molecule **II**, the experimental spectrum depicted in Figure 3.7.2 presents one main broad band with a maximum placed at about 325 nm. The calculated D_{CT} reported in Table 3.3 show that the first excitation has a low CT character, which is

enhanced by the explicit solvation models. This is also confirmed by the MO involved in the transition depicted in Figure S7, given as SI. The general trend discussed for molecule **I** is valid also in this case. In fact, QM/PCM and QM/QM$_w$/PCM give very similar results, with a discrepancy of about 0.5 eV from the experimental value. The QM/TIP3P model is again the most inaccurate one, resulting in a shift of 0.05 eV with respect all the other methods. Notice however that the inclusion of water molecule in the QM portion in QM/QM$_w$/FQ calculations results in a shift of the excitation energy towards the experiment. This can be rationalized by considering that the a strong and directional HB interaction is present in the MD (see previous section). Also in this case, it is worth noticing that the several methods employed give similar excitation energies, with an error that is almost constant (12% on average). Thus, such a discrepancy is again probably due to the computational level of theory.
The experimental spectrum of molecule **III** presents a double peak profile in the region 350-400 nm, which is probably due to vibronic effects (see Figure 3.7.3). The calculated D$_{CT}$ data reported in Table 3.3 and the associated MO depicted in Figure S7 in the SI show that the first excitation has no CT character. The most intense is the peak at about 400 nm. The QM/PCM approach predicts an excitation energy which is 0.62 eV blue shifted with respect to the experimental value. The inclusion of three explicit water molecules redshift the energy of about 0.2 eV, however resulting in a discrepancy with respect the experiment of about 0.4 eV. The explicit solvation models coupled with the dynamic approach of the MD results are in fair agreement with the experimental data, also at the QM/TIP3P level of theory. This is particularly interesting and it can be due to the fact the in this case polarization effects are not crucial in the description of the excitation energy.
Molecule **IV** experimental UV-VIS spectrum has a maximum at about 375 nm (see Figure 3.7.4). In this case, MO (see Figure S7, given as SI) and D$_{CT}$ values show for the first transition a CT character, although it is generally small. Similarly to the previous cases, QM/PCM predicts an excitation energy lower of about 0.4 eV with respect the experiment. The inclusion of explicit water molecules, however, gives an almost perfect agreement with only a 1% error. The errors obtained by using the explicit solvation models are lower with respect the QM/PCM model, with an error of 2% in the case of QM/QM$_w$/FQ approach. Notice that in this case the purely QM/FQ model gives a discrepancy of about 0.25 eV with respect the experimental value. Considering that the inclusion of explicit water molecules both in QM/QM$_w$/PCM and QM/QM$_W$/FQ models, are crucial in the reproduction of the excitation energies, some non-electrostatic effects can play a relevant role in this case.
Concerning molecule **V**, the experimental spectrum is characterized by a broad band placed at about 450 nm (see Figure 3.7.5). As resulting from the D$_{CT}$ calculations (see Table 3.3 and from the MO involved (see Figure S7 in the SI), the first transition has a CT character. As pointed out before, in this case no explicit water molecules were included in the QM/PCM calculations, thus no results for the QM/QM$_w$/PCM are discussed. This was due to the fact that no specific solute-solvent interactions were identified from the RDF depicted in Figure 3.6, panel (e). This explains also why the results obtained by using an implicit or an explicit model are very similar and in general in fair agreement. Notice that the QM/TIP3P error is the highest, and again

this can be explained by the fact that polarization effects in the solvation model may be crucial in this case.
The experimental excitation spectrum of molecule **VI** presents a main band at about 280 nm (see Figure 3.7.6). The calculated D_{CT} reported in Table 3.3 clearly show that the first excitation has not a CT character. This is also confirmed by the MO involved in the transition depicted in Figure S7, given as SI. QM/PCM and QM/QM_w/PCM were calculated by weighting the spectra of the two conformers by their Boltzmann population (see Section S2 and Figure S6 given as SI). The excitation energies predicted by exploiting such methods differ of an average 14% with respect the experimental value. In particular, the discrepancy is of 0.6 eV on average. A shifting towards the experiment is recovered by using the explicit solvation models: the QM/FQ gives the best agreement (0.4 eV, 8%), however no great differences between the three exploited models is reported.
The last molecule (**VII**)'s experimental spectrum presents a huge vibronic band with a maximum at about 500 nm (see Figure 3.7.7). Similarly to molecule **V**, also in this case no explicit water molecules were added to the QM region in QM/QM_w/PCM calculations. Again, also in this case, the first excitation has no CT character (see Table 3.3 and Figure S7 in the SI). The four solvation approaches give very similar results, with the maximum shifting giving passing through the QM/PCM to the QM/QM_w/FQ model (0.11 eV). However, as pointed out before for molecule **I**, the similarity in the results may be due to the computational level adopted in this work. Probably, vibronic contributions influence the position and the shape of the absorption band, thus resulting in a shifting and a better agreement with the experiment.

To conclude the discussion on calculated excitation energies, in Table 3.2 the Mean Relative Deviation (MRD), Root Mean Square Deviation (RMSD) and Maximum Absolute Deviation (MAD) obtained for all the considered models are reported. As it was expected the less MRD is shown by the QM/QM_w/FQ, however the inclusion of explicit water molecules is not crucial to reproduce the excitation energy of the variety of studied molecules. In fact, QM/FQ model results in a MRD of about 10%. RMSD confirms this trend. In fact, QM(/QM_w)/PCM reports an RMSD 0.13 eV greater than the QM/FQ model. The same conclusions can be also extracted by the MAD values, which are reported for molecule **I** in every adopted model.
To end the discussion, we show how the spectra are obtained in the case of the QM/MM methods from raw data. The data extracted from the single QM/FQ calculations on each snapshot for molecule **VI** in aqueous solution are reported as a stick spectrum in Figure 3.8. The same spectra for all of the investigated molecules **I-VII** are reported in Figure S8, given as SI.
Clearly, the overall shape of the final, averaged spectrum is already visible from the data reported in Figure 3.8, which also gives insight into the spreading of the transition bands, both in wavelengths and intensities. This is due to the fact that in the different snapshots the spatial distribution of water molecules around the molecule varies and also the conformational freedom of the molecule is sampled. The same also applies to the other investigated systems (see SI). In order to obtain the final, averaged spectrum, each transition in Fig. 3.8 was convoluted with a Gaussian function and averaged. It



(3.7.1). I **(3.7.2). II**

(3.7.3). III **(3.7.4). IV**

(3.7.5). V **(3.7.6). VI**

(3.7.7). VII **(3.7.8). Key**

Figure 3.7. QM/PCM, QM/QM$_w$/PCM, QM/TIP3P, QM/QM$_w$/FQ and experimental UV-VIS spectra for molecules **I-VII**. The key is shown in panel h.

Chromophore	QM/PCM	QM/QM$_w$/PCM	QM/TIP3P	QM/FQ	QM/QM$_w$/FQ	Exp
I	2.99 (21%)	2.98 (20%)	3.04 (23%)	2.97 (20%)	2.97 (20%)	2.48 [364]
II	4.28 (12%)	4.28 (12%)	4.32 (13%)	4.28 (12%)	4.23 (11%)	3.81 [102]
III	3.82 (20%)	3.63 (14%)	3.11 (3%)	3.14 (2%)	3.04 (5%)	3.20 [365]
IV	3.72 (11%)	3.32 (1%)	3.52 (5%)	3.60 (7%)	3.42 (2%)	3.36 [366]
V	2.96 (6%)	-	3.04 (9%)	2.91 (5%)	2.94 (6%)	2.78 [102]
VI	5.05 (13%)	5.09 (14%)	4.87 (9%)	4.83 (8%)	4.88 (9%)	4.46 [345]
VII	2.99 (16%)	-	2.94 (14%)	2.93 (13%)	2.88 (12%)	2.58 [367]
MRD	14%	12%	11%	10%	9%	
RMSD	0.47	0.46	0.37	0.34	0.32	
MAD	0.63	0.63	0.56	0.49	0.49	

Table 3.2. QM/PCM, QM/QM$_w$/PCM, QM/TIP3P, QM/QM$_w$/FQ excitation energies. Experimental data taken from the indicated references are reported in the last column. Relative deviations with respect the experiments are given in parentheses. Mean Relative Deviation (MRD), Root Mean Square Deviation (RMSD) and Maximum Absolute Deviation (MAD) are also reported. All data are in eV.

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

Table 3.3. QM/PCM, QM/QM$_w$/PCM, QM/TIP3P, QM/QM$_w$/FQ calculated D_{CT}(Å) index for the first excitation.

should be clear that the final spectrum emerges as an ensemble average of many different snapshots which may have widely varying spectroscopic response properties. This is in stark contrast with results obtained using methods like standard QM/PCM which most commonly only takes minimum-energy-structures as representative of the whole, and whose results are then usually convoluted with wide empirical lineshape functions meant to represent the spread of conformational and solvation degrees of freedom, and whose true distribution may be far from being represented by a simple Gaussian or Lorentzian lineshape function. Notice however that other more sophisticated approaches have been proposed in the literature to overcome this problem.[343]

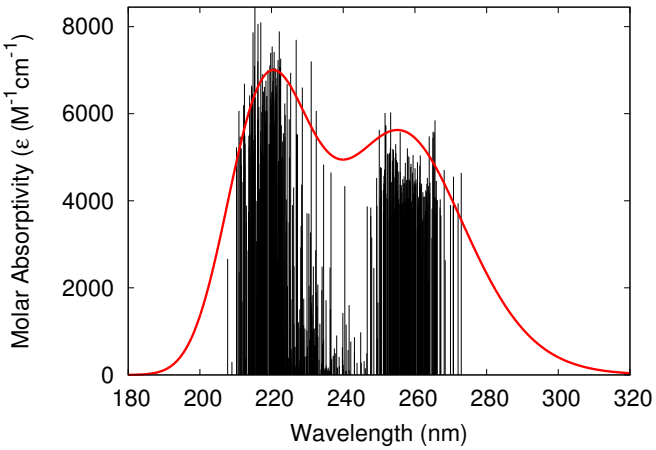


Figure 3.8. Molecule **VI** QM/FQ calculated data reported as stick spectrum and convoluted with a Gaussian band shape (FWHM=0.5 eV)

3.5 Summary and Conclusions

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

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

These results should still be seen as preliminary as still much more work remains to be done in benchmarking polarizable QM/MM methods for the purpose of calculating spectra of systems in solution. UV-Vis absorption spectroscopy is but the simplest type of observable that may be studied, and more complex spectroscopic observables might be considered for future benchmarks, including mixed electric-magnetic properties such as circular dichroism intensities or higher order spectroscopies such as Raman or Raman optical activity. In this work we have only considered a set of medium-sized organic molecules, however more complex systems, with larger CT effects, may be of interest. Biological molecules such as peptides or nucleic acids, for which aqueous solution is the natural environment, are often both highly flexible and able to form multiple hydrogen bonds with the solvent, and are therefore the ideal systems for the application of the method. The study of the spectroscopic properties of solvated inorganic systems such as transition metal complexes through polarizable QM/MM methods is also a largely unexplored field and may present its own peculiar challenges. Finally, this method can be extended to non-aqueous solvents both polar and non-polar. For apolar solvents the effect of polarization may be much less important compared to water, in that case significant contributions to the solute-solvent interactions may instead come from non-electrostatic forces such as dispersion and repulsion. Methods for the inclusion of such effects in a QM/MM picture have been recently presented[197] however their extension to the computation of molecular spectra is still lacking and will be the topic of future investigations.

Acknowledgements

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

Supporting Information

Supporting Information avaiable online. See DOI: 10.1002/qua.25684

Chapter 4

A Polarizable Embedding Approach to Second Harmonic Generation (SHG) of Molecular Systems in Aqueous Solutions

Theoretical Chemistry Accounts (2018) 137:74
https://doi.org/10.1007/s00214-018-2247-7

REGULAR ARTICLE

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

Tommaso Giovannini[1] · Matteo Ambrosetti[1] · Chiara Cappelli[1]

Received: 18 January 2018 / Accepted: 23 April 2018

Abstract 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.

4.1 Introduction

Non Linear Optics (NLO)[368–371] has been reported to impact many different research fields, such as material science, communications, medicine and the developoment of electronic devices.[372–374] 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.[375] Through the variety of NLO properties,[98,99,376] 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.[123,377–381] 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.[382–385] Notice that this is not generally true for static properties, where vibrational contributions can be even larger than electronic terms.[386–388]

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.[389–391] However, whenever electron correlation plays a crucial role, Coupled Cluster (CC) approaches are best suited,[392–397] 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.[28,100,114,398] The standard approach to include such effects is to resort to continuum solvation approaches, such as the Polarizable Continuum Model (PCM).[6] This model has been successfully applied to the modeling of several properties and spectroscopies, and also to SHG of solvated systems.[83,114,389,399] 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.[149,150,155,400] In all such cases, the best strategy has demonstrated to be to resort to the so called QM/Molecular Mechanics (MM) methods.[1,8,12,13] Such approaches have already been employed for the description of SHG.[31,401–404]

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,[58,185,186,281] induced dipoles,[10,155,164] Drude oscillators[59] or Fluctuating Charges (FQ).[156,157] 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.[56,149,150] 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.[61,187,189,197] Notice also that polarizable embedding approaches based on induced dipoles have been applied to the modeling of SHG with promising results.[396,405]
The paper is organized as follows: after a brief section recalling the fundamentals of the QM/FQ approach, the analytical evaluation of SHG within such an approach is discussed. Test applications to aqueous solutions of organic acids is then discussed, for which the computed data are compared with experimental values taken from the literature. Finally, the main outcomes of the study are summarized and some future perspectives are discussed.

4.2 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".[286,289] Through these parameters, atomic charges can be computed based on the difference of electronegativities between the atoms.[156–158] More in detail, the FQ FF describes polarization effects by endowing each MM atom with a fluctuating charge whose value depends on the electronegativity[156–158] according to the Electronegativity Equalization Principle (EEP)[160,286] which states that, at equilibrium, the instantaneous electronegativity χ of each atom has the same value.[160,286] The FQs ($\mathbf{q}$) can be defined as those minimizing the following functional[159]

$$\begin{aligned} F(\mathbf{q},\lambda) =& \sum_{\alpha,i} q_{\alpha i}\chi_{\alpha i} + \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j} q_{\alpha i}J_{\alpha i,\beta j}q_{\beta j} + \sum_{\alpha}\lambda_{\alpha}(\sum_{i} q_{\alpha i} - Q_{\alpha}) \\ =& \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} \end{aligned} \qquad \textbf{4.1}$$

where $\mathbf{q}$ is a vector containing the FQs, the Greek indices α run over molecules and the Latin ones i over the atoms of each molecule. $\boldsymbol{\lambda}$ is a set of Lagrangian multipliers used to impose charge conservation constraints on each molecule. The charge interaction kernel $\mathbf{J}$ is, in our implementation, the Ohno kernel.[291] Atomic units are used throughout the article. The stationarity conditions of the functional in eq.4.1 are defined through the following equations[159]

$$\begin{cases} \sum_{\beta,j} J_{\alpha i,\beta j}q_{\beta j} + \lambda_{\alpha} = -\chi_{\alpha i} \\ \\ \sum_{i} q_{\alpha_i} = Q_{\alpha} \end{cases} \qquad \textbf{4.2}$$

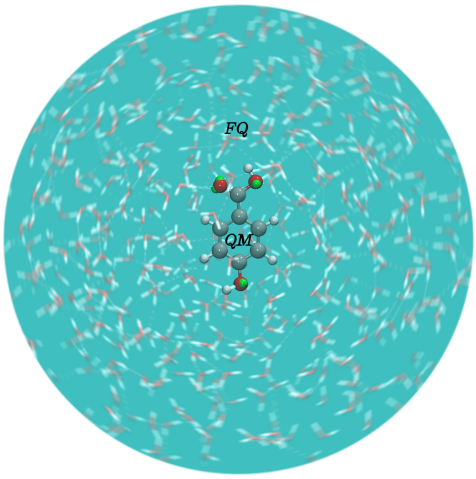


Figure 4.1. Schematic representation of the QM/FQ model.

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

$$\mathbf{D} = \left(\begin{array}{c|c} \mathbf{J} & \mathbf{1_\lambda} \\ \hline \mathbf{1_\lambda^\dagger} & \mathbf{0} \end{array}\right)$$

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

$$\mathbf{D q_\lambda} = -\mathbf{C} \qquad \textbf{4.3}$$

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[56,159] has been shown to be especially suited to the modeling of response and spectral properties because, as it is shown below, its energy expression can be easily differentiated up to high orders. The QM/FQ 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.[54,55,147,149,150,161]

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 4.1), 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:[54]

$$E_{\mathrm{QM/FQ}} = \sum_{i=1}^{N_q} V_{\mathrm{QM}}[\rho](\mathbf{r}_i) q_i \tag{4.4}$$

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,[197] 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:[54,55,192]

$$\mathcal{E}[\mathbf{P}, \mathbf{q}, \boldsymbol{\lambda}] = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{1}{2}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \mathbf{q}^\dagger\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^\dagger\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q} + \mathbf{q}^\dagger\mathbf{V}(\mathbf{P}) \tag{4.5}$$

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

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q - \mathbf{V}(\mathbf{P}) \tag{4.6}$$

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

4.2.1 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 $\boldsymbol{\mu}(t)$:

$$\boldsymbol{\mu}(t) = \boldsymbol{\mu}^0 + \boldsymbol{\mu}^\omega \cos(\omega t) + \boldsymbol{\mu}^{2\omega} \cos(2\omega t) + ... \tag{4.7}$$

Each Fourier amplitude in Eq. 4.7, can be rewritten as a Taylor expansion with respect to the external electric field.[370] In particular, SHG, i.e. the generation of a photon at 2ω as a result of the interaction with an incident ω photon reads:

$$\boldsymbol{\mu}^{2\omega} = \frac{1}{4}\boldsymbol{\beta}(-2\omega;\omega,\omega) : \mathbf{E}^\omega\mathbf{E}^\omega \tag{4.8}$$

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. $\beta_{\alpha\beta\beta} = \beta_{\beta\alpha\beta} = \beta_{\beta\beta\alpha}$.[406]
By exploiting the response theory formalism, the first-order hyperpolarizability $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ can be calculated as:[407,408]

$$\boldsymbol{\beta}(-2\omega;\omega,\omega) = 2\,\mathrm{tr}\,\boldsymbol{\mu}\mathbf{P}^{(2)} \tag{4.9}$$

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.[407–409] 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:

$$(\tilde{\mathbf{A}}+\tilde{\mathbf{B}})(\mathbf{X}+\mathbf{Y})+2\mathbf{Q}=0 \qquad \textbf{4.10}$$

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:[54]

$$\begin{pmatrix} \tilde{\mathbf{Q}}_X \\ \tilde{\mathbf{Q}}_Y \end{pmatrix} + \begin{pmatrix} \tilde{\mathbf{A}}-\omega\mathbb{I} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}} & \tilde{\mathbf{A}}+\omega\mathbb{I} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = 0 \qquad \textbf{4.11}$$

4.3 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 4.2 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.

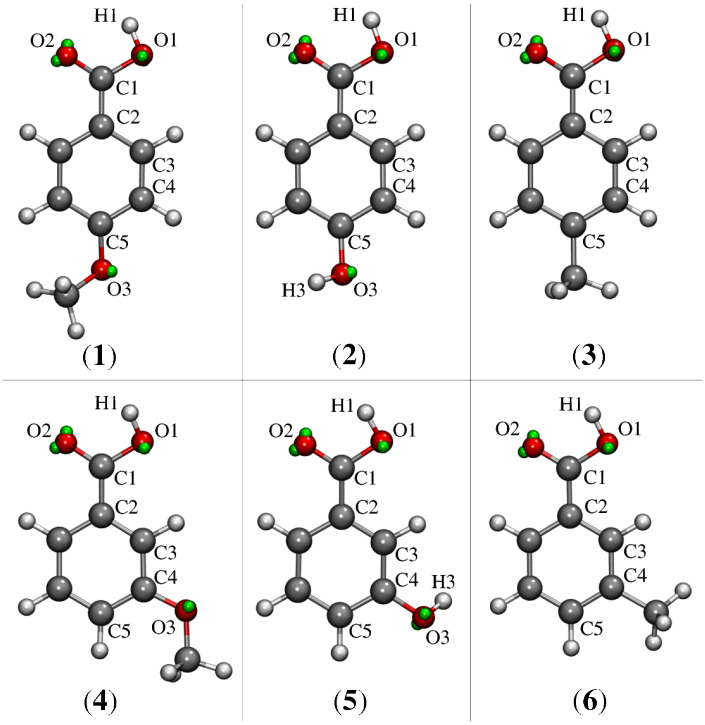


Figure 4.2. Structure of the molecules studied in this work. The green spheres depicted on each Oxygen atom represent the Virtual Sites (VS).

Molecular geometries of the molecules depicted in Figure 4.2 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.[6,74] MD simulations were carried out by using GROMACS[348] with Amber 99ffSB-ILDN force field to describe intramolecular and intermolecular potentials.[410] RESP charges were used for electrostatic interactions.[312,411] VS were possibly placed on the centroids of Boys orbitals.[412,413] In particular a couple of VS was assigned to each carboxylic and hydroxylic Oxygen atom (see Figure 4.2). 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 (3385-3872) modeled using

the TIP3P parameter set, depending on the considered molecule **1**-**6**.[414] Electrostatic interactions were taken into account by means of the particle mesh Ewald method[355] 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.[415] The LINCS algorithm was used in order to constrain all bonds of the solute molecules.[354] 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 Electronic Supplementary Material (ESM)). 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[156] and the parametrization proposed by some of the present authors.[162] The TIP3P[358] 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.[301] Finally, the calculated values were compared with experimental Hyper Rayleigh Scattering (HRS) data taken from ref.[416] Experimental HRS values were divided by the Lorentz local-field factor in order to be directly compared with our calculated data.[417–419]

In Ref.[416] a comparison between computed and experimental data was done by referring to the following quantity:

$$\left|\vec{\beta}\right| = \sqrt{\beta_x^2 + \beta_y^2 + \beta_z^2} \qquad \textbf{4.12}$$

where

$$\beta_i = \sum_{k=1}^{3} \beta_{ikk} + \beta_{kik} + \beta_{kki} \qquad \textbf{4.13}$$

Therefore, our calculated data refer to Eq.4.12. We note, however, that alternative definitions for HRS values, giving computed results directly comparable with experimental data, can be found in the literature.[418,420–422]

4.4 Numerical Results

4.4.1 MD Analysis

The analysis of MD trajectories was performed by using the TRAVIS package.[359] Three different results are presented and discussed: the radial distribution function $g(r)$ (RDF), the spatial distribution function (SDF) and the dihedral distribution function

(DDF). It is worth noticing that from a structural point of view, the molecules depicted in Figure 4.2 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 4.3), whereas the data for the other molecules are given as ESM (see Figure S2). The left and right panels of Figure 4.3 report the $g(r)$ obtained from the two MD runs performed with or without virtual sites (MD_{VS} and MD_{noVS}, respectively).

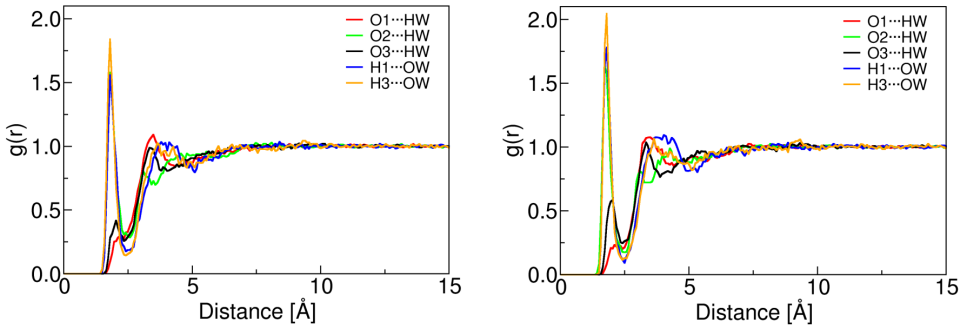


Figure 4.3. RDF of system **2**, obtained from the analysis of the MD run without (left panel) or with (right panel) virtual sites.

Focusing on both panels of Figure 4.3, 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 4.1 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 4.1).

		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 4.1. Number of Hydrogen Bonds for the different molecules depicted in Figure 4.2. 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 4.4 for all the molecules depicted in Figure 4.2. SDF calculated from MD_{noVS} are reported in Figure S3, as ESM. Calculated SDF isodensity values are equal to 70 and 100 nm^{-3} for water Hydrogen and Oxygen atoms, respectively. Figure 4.4 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 4.1. **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 4.5 for molecule **2**. In fact, the distribution of the C4C5O3H3 dihedral angle shows two maxima at 0 and 180 degrees, thus confirming two most probable configurations. In Figure 4.5, the distribution related to the dihedral angle of the donor hydroxyl group (C2C1O1H1) is plotted. Notice that in this case a single maximum occurs at 0 degrees, thus confirming the SDF sampling.

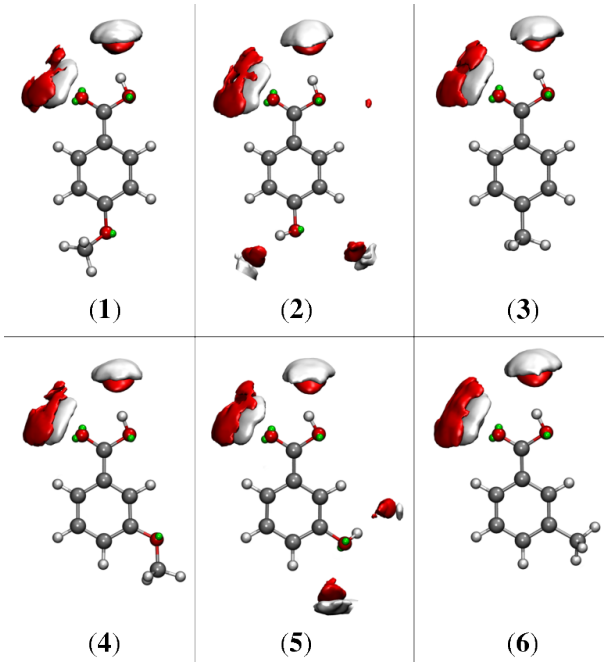


Figure 4.4. Spatial distribution functions extracted from MD_{VS}. Calculated SDF isodensity values are given in nm^{-3} and are equal to 70 and 100 nm^{-3} for water hydrogen and oxygen atoms, respectively.

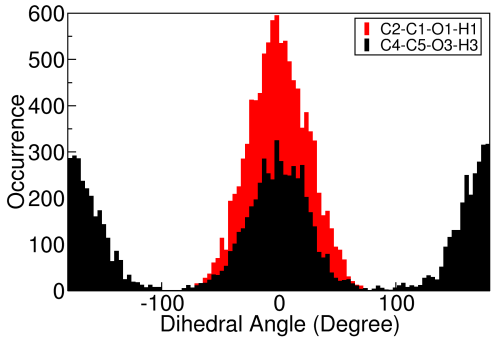
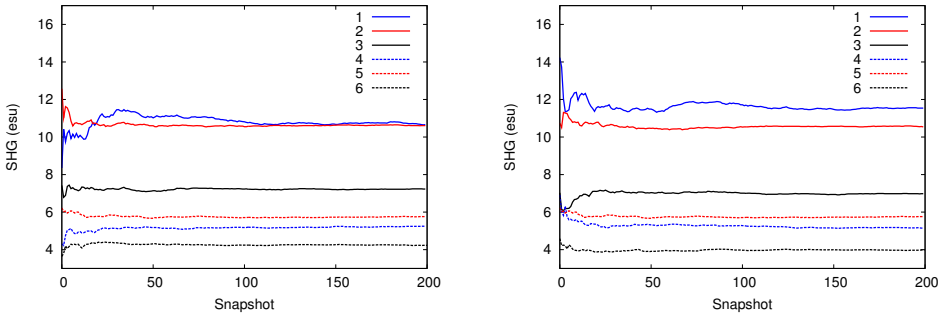


Figure 4.5. Dihedral distribution function of molecule **2** obtained with MD_{VS}.

4.4.2 $\beta(-2\omega;\omega,\omega)$ QM/MM Results

Convergence with respect to sampling extracted from the MD

Before analyzing the QM/FQ $\beta(-2\omega;\omega,\omega)$ 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 4.6, the average values of $\beta(-2\omega;\omega,\omega)$ as a function of the number of snapshots for the six selected systems are plotted. Notice that panel **a** refers to the snapshots extracted 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.



(4.6.1). a **(4.6.2). b**

Figure 4.6. Average values of $\beta(-2\omega;\omega,\omega)$ as a function of the number of snapshots extracted from the MD for all the molecules depicted in Figure 4.2. **a.** Snapshots extracted from MD_{noVS}. **b.** Snapshots extracted from MD_{VS}.

Dependence on DFT Functional and Basis Set

The dependence of the $\beta(-2\omega;\omega,\omega)$ 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 $\beta(-2\omega;\omega,\omega)$ results are reported in Table 4.2.

Table 4.2 shows some general trends. First, the calculated $\beta(-2\omega;\omega,\omega)$ 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 $\beta(-2\omega;\omega,\omega)$ values have already been reported to be extremely sensitive to the inclusion of diffuse functions in the basis set.[423–425] 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.

		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 4.2. 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.

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,[114] 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 4.7 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 4.7 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 4.7 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 ESM (see Figures S4-S8).
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 4.3. Experimental data, taken from Ref.[416] 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 4.3), we note that both MD runs, with or without VS,

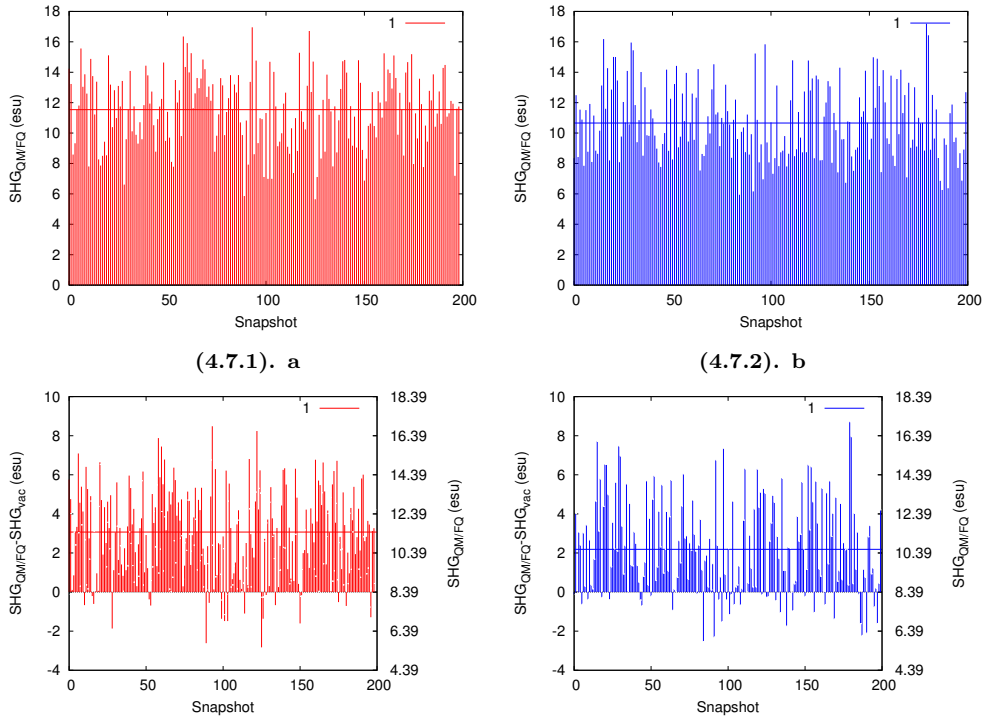


Figure 4.7. QM/FQ $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ values for molecule **1** calculated for the different snapshots extracted from the MD runs without (**a**)-(**c**) or with (**b**)-(**d**) VS. Panels **a** and **b** report $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ raw data, panels **c** and **d** report difference between QM/FQ and vacuum $\boldsymbol{\beta}(-2\omega;\omega,\omega)$ data. All values are given in esu.

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 4.3, follow indeed a specific trend. In fact, solvent 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.[156] and by Carnimeo et al.[162] are compared (see columns $\mathrm{QM/FQ}^a$ and $\mathrm{QM/FQ}^b$ in Table 4.3) . The most relevant discrepancy between the two parametrizations is the difference in electronegativities between water and oxygen and hydrogen atoms, which in case b is greater than case a. FQ charges are calculated by solving Eq. 4.5, 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 4.3). 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 $\mathrm{MD}_{\mathrm{noVS}}$ and $\mathrm{MD}_{\mathrm{VS}}$ respectively, with the minimal error (20%) being reported for $\mathrm{MD}_{\mathrm{noVS}}$ and the 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.

4.5 Summary and Conclusions

We have reported on the extension of the atomistic fully polarizable QM/FQ method to the calculation of SHG of systems in aqueous solution. The modeling which has been proposed focuses in achieving an adequate modeling of solvent effects both in

		Vacuum	QM/TIP3P	QM/FQ[a]	QM/FQ[b]	Exp.[416]
MD_{noVS}	**1**	8.47	9.63±0.15 (14%)	11.54±0.17 (36%)	12.11±0.18 (43%)	13.06
	2	7.10	8.61±0.07 (21%)	10.54±0.08 (48%)	10.96±0.09 (54%)	10.93
	3	4.15	6.00±0.08 (44%)	6.98 ±0.07 (68%)	7.48 ±0.09 (80%)	10.28
	4	3.71	4.44±0.06 (20%)	5.15 ±0.07 (39%)	5.45 ±0.07 (47%)	8.91
	5	3.63	4.83±0.05 (33%)	5.76 ±0.05 (58%)	5.98 ±0.06 (65%)	6.78
	6	2.35	3.43±0.05 (46%)	3.99 ±0.05 (70%)	4.28 ±0.06 (82%)	6.57
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 4.3. CAM-B3LYP/6-311++G** vacuum, QM/TIP3P and QM/FQ $\beta(-2\omega;\omega,\omega)$ (± 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.[156] [b] FQ parametrization proposed by Carnimeo et al.[162]

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 systems, 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 HRS experiments. The good performance of our approach is especially due to the inclusion of polarization effects, as it is shown by the comparison of our data with non-polarizable QM/TIP3P results. The disagreement in the final reported absolute values may be due to the only partial account of electron correlation effects in DFT approaches, and in the neglecting of vibrational corrections, which may in principle play a relevant role in the computation of the final property. Such corrections have not been included, due to their high computational cost, especially when combined with the need of repeating the computation for hundreds of representative snapshots. Also, non-electrostatic (repulsion/dispersion) interactions have not been considered, however they can contribute to reach a numerical agreement between calculations and experimental values.

Acknowledgment

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

Electronic Supplementary Material

Electronic Supplementary Material avaiable online. See DOI: 10.1007/s00214-018-2247-7

Chapter 5

A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches

JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches

Tommaso Giovannini, Piero Lafiosca, and Chiara Cappelli*

Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

Abstract A methodology to account for non-electrostatic interactions in Quantum Mechanical (QM)/Molecular Mechanics(MM) approaches is developed. Formulations for Pauli repulsion and dispersion energy, explicitly depending on the QM density are derived. Such expressions are based on the definition of an auxiliary density on the MM portion and the Tkatchenko-Scheffler (TS) approach, respectively. The developed method is general enough to be applied to any QM/MM method and partition, provided an accurate tuning of a small number of parameters is obtained. The coupling of the method with both non-polarizable and the fully polarizable QM/Fluctuating Charges(FQ) approaches is reported and applied. A suitable parametrization for the aqueous solution, so that its most representative features are well reproduced, is outlined. Then, the obtained parametrization and method are applied to calculate the non-electrostatic (repulsion and dispersion) interaction energy of nicotine in aqueous solution.

5.1 Introduction

Multiscale computational approaches rooted in the so-called hybrid quantum mechanics (QM) molecular mechanics (MM) methods (QM/MM)[1,12,13,23,278,426–429] have nowadays been amply and successfully applied to a variety of chemical systems and their physico-chemical properties.[7,82,146,162,187,430–435]

The idea behind those approaches is to treat accurately, by QM methods, a small but critical part of the overall system, while resorting to much cheaper and less accurate MM methods for the remaining portion of the whole system. Such a partition is sometimes naturally applicable (such as in solvation phenomena and non-covalent interactions), however in some cases (i.e. covalently bound systems) the QM and MM portions are more difficult to define. In any case, a specific choice of the QM/MM partition introduces some assumptions on the system, which in the worst cases can negatively affect the quality of the final computed results. The quality of the results which can be obtained with QM/MM models do not only depend on the definition of the two moieties, but also (and crucially) on the approach exploited to model the interaction between the two portions.[12,429]

Different choices are possible in this context, however the model for the QM/MM coupling must be capable of treating both bonded and non-bonded interactions (electrostatic and non-electrostatic). The way of treating the electrostatic interaction is generally a key element of any QM/MM approach, largely affecting the quality of the computed results.[149,155,429,436,437] Two groups of methods exist, the so called mechanical embedding schemes and the electrostatic embedding methods.[29] The latter may o may not include mutual polarization effects between the QM and MM portions: in the first case, a set of atomic-centered partial point charges are used for calculating the electrostatic interaction at the MM level, which also enters in definition of the effective QM Hamiltonian. Polarization effects can be included by using either Fluctuating Charges (FQ),[156,159] distributed multipoles,[185,186,281] induced dipoles,[10,155,164] or Drude oscillators.[59]

Limiting QM/MM interactions to electrostatic-only terms may yield an unphysical description of the systems. Non electrostatic interactions,[193] also called London interactions, play a crucial role in many chemical processes. For instance, most of DNA and RNA functionalities, as well as the adsorption of a molecule on a surface are regulated by repulsion/dispersion interactions. Moreover, these interactions can also play an important role in solvation phenomena.

Although their paramount importance, in most QM/MM approaches non-electrostatic interactions between the QM and MM moieties are only retained at the MM level and treated by means of Lennard-Jones or similar parametrized analytical functions.[195]

This approach, if computationally inexpensive, introduces a rough approximation in the computational modeling. In fact, non-electrostatic interactions are primarily due to the Pauli repulsion principle, which cannot be postulated in a classical framework, and to long-range electron correlation effects, which are again not defined in the classical realm. In purely QM approaches, such interactions are modeled by resorting to correlated expensive QM methods, such as coupled cluster with single, double, and perturbative triple excitations - CCSD(T) coupled to large atomic basis sets in order to reduce the

Basis Set Superposition Error (BSSE).[438,439]

The formulation of QM/MM approaches able to account for QM effects affecting dispersion/repulsion interactions between the QM and MM portions has received so far only little attention in the literature. To the best of our knowledge, the only approach which has been proposed is the so-called Effective Fragment Potential (EFP) method.[185–190] In this approach, empirical force-fields are not exploited, but the force-field (FF) for the "MM" portion is obtained from electronic structure calculations of the single fragments. In this way, the FF is defined in terms of point charges, multipoles, static and dynamic polarizabilities, localized molecular orbitals and related QM quantities.

Due to the nature of the EFP method, the inclusion of dispersion and Pauli repulsion terms can be formulated in terms of QM quantities calculated for the fragments. Therefore, such an approach cannot be straightforwardly extended to generic QM/MM methods based on empirical potentials.

The account for QM-based non-electrostatic interactions (explicitly depending on the QM density) in QM/MM calculations will permit not only a more reliable description of the interaction between the QM and the MM moieties, but also to include them in the QM Hamiltonian and to propagate such terms also to molecular properties and spectra. The common approaches, based on Lennard-Jones and similar potentials, which do not bear any explicit dependence on QM quantities, do not give any contribution to the QM Hamiltonian; therefore they only result in a correction to the QM/MM energy. As it will be detailed in the following sections, the development of a model with the aforementioned features is the goal of this paper.

Notice that we do not aim to propose a way of decomposing the intermolecular energy terms. Such kind of calculations can be performed by exploiting other approaches, for instance the general effective fragment potential (EFP2)[187] or the Symmetry Adapted Perturbation Theory (SAPT)[440,441] approach.

The manuscript is organized as follows: first, a general formulation of Pauli repulsion and dispersion energy in a QM/MM framework is presented. The formulation that is reported is based on the definition of an auxiliary density on the MM portion and the Tkatchenko-Scheffler (TS) approach[198–201] for the repulsion and dispersion terms, respectively. Next, the inclusion of such terms in the QM/MM Hamiltonian is derived, with specific emphasis in the coupling with the polarizable QM/MM approach which is developed in our group.[54–56,146,149,159] The derived repulsion/dispersion terms depend on some parameters. A parametrization to treat aqueous solution is then proposed, which allows the application of the methodology to treat non-electrostatic interaction energies of solvated systems. To this end, aqueous solutions of (L)-Methyl Lactate (MLAT) and (R)-Methyloxirane (MOXY) are considered, as well as the more complicated case of Nicotine in aqueous solution, where the focus is on the influence of non electrostatic interactions on conformational populations and on the electric dipole. Summary, conclusions and future perspectives end the presentation.

5.2 Theory

The total energy of a system composed by two interacting moieties A and B can be expressed as:[192,442]

$$E_{AB} = E_{AB}^{ele} + E_{AB}^{pol} + E_{AB}^{pen} + E_{AB}^{ex} + E_{AB}^{dis} \tag{5.1}$$

where, E_{AB}^{ele} arises from electrostatic interactions and E_{AB}^{pol} is the polarization contribution. E_{AB}^{pen} is the so-called penetration term, E_{AB}^{ex} is the exchange contribution and E_{AB}^{dis} arises from dispersion interactions. In the context of QM/MM approaches, A can represent the QM portion of the system, while B the MM one. E_{AB}^{ele} and E_{AB}^{pol} are the energy terms considered within electrostatic embedding schemes and in particular in polarizable QM/MM approaches[1,10,12,59,155,159,164]

5.2.1 Pauli Repulsion Energy

The Pauli Repulsion energy, E_{AB}^{rep}, also known as Exchange-Repulsion energy, is formally the sum between the Penetration (E_{AB}^{pen}) and the Exchange (E_{AB}^{ex}) contributions in Eq. 5.1 above. The penetration term is considered to be twice the exchange term in the van der Waals region, thus giving rise to the following expression:[86,192]

$$E_{AB}^{rep} = \frac{1}{2} \int \frac{\mathrm{d}\mathbf{r_1}\, \mathrm{d}\mathbf{r_2}}{r_{12}} \rho_A(\mathbf{r_1}, \mathbf{r_2}) \rho_B(\mathbf{r_2}, \mathbf{r_1}) \tag{5.2}$$

where ρ_x is the density matrix of the A or B moieties, placed at distance r_{12}.

The extension of Eq. 5.2 to QM/MM partitions, is obstructed by the the difficulty to define the density matrix ρ_B of the classical region B (no electrons are present in this region).
In the following derivation we will work out an approximate expression for Pauli repulsion in the framework of the so-called focused models,[6] namely models in which the main component, that essentially bearing the property, is described at a higher level than the remainder, which plays a complementary, but not negligible role. Within such an approach, terms related to two fragments of the MM portion are not considered.

The starting point for the derivation of the equations is the formulation of quantum repulsion effects for the Polarizable Continuum Model.[86,87,95,96] In particular, each fictitious valence electron pair of the MM molecules is localized in bond and lone pair (if they are present) regions and represented by an s-gaussian-type function. Due to the different physical nature of the two (bond or lone pair) regions, the two are discerned by using two different sets of parameters, so that the expression for ρ_B becomes:[86]

$$\rho_{\mathrm{MM}}(\mathbf{r_1}, \mathbf{r_2}) = \sum_{\mathbf{R}} \xi_{\mathbf{R}}^2 e^{-\beta_{\mathbf{R}}(\mathbf{r_1}-\mathbf{R})^2} \cdot e^{-\beta_{\mathbf{R}}(\mathbf{r_2}-\mathbf{R})^2} \tag{5.3}$$

where, $\mathbf{R}$ collects the centers of the gaussian functions used to represents the fictitious MM electrons. The β and ξ parameters are generally different for lone-pairs or bond-pairs: their values are adjusted to the specific kind of environment (MM portion) to be modeled (vide infra). By substituting Eq. 5.3 in Eq. 5.2, the QM/MM repulsion energy reads:

$$E^{rep}_{\mathrm{QM/MM}} = \frac{1}{2}\sum_{\mathbf{R}} \int \frac{\mathrm{d}\mathbf{r_1}\,\mathrm{d}\mathbf{r_2}}{r_{12}} \rho_{\mathrm{QM}}(\mathbf{r_1},\mathbf{r_2}) \left[\xi^2_{\mathbf{R}} e^{-\beta_{\mathbf{R}}(\mathbf{r_1}-\mathbf{R})^2} \cdot e^{-\beta_{\mathbf{R}}(\mathbf{r_2}-\mathbf{R})^2}\right] \qquad \mathbf{5.4}$$

In this formalism, the QM/MM Pauli Repulsion energy is calculated as a two-electron integral. Interestingly, this differs from the formulation of the same quantity in the PCM model,[86] where this term is a pure one-electron term.
Eq. 5.4 holds for every kind of MM environment, independent from its nature, i.e. the formalism not only holds for solvents, but can be extended to other substrates (proteins, surfaces) surrounding the QM core region. The specification for the different external environments is simply done by defining the number of different electron-pair types and the relative β and ξ parameters in Eq. 5.3. Also, due to its simplicity, this formalism is retained also in case of polarizable QM/MM approaches, such as our polarizable QM/Fluctuating Charge(FQ) approach[56] (vide infra), by only refining the parameters if necessary. To end the discussion, it is worth noticing that in the present work ρ_{MM} will only defined in terms of spherical gaussian-type functions. Extension to p/d-type functions is possible, and mainly implies the definition of additional parameters in Eq. 5.3. Such an extension (and the related parametrization work) will be the topic of future communications.

Practical Formulation of $E^{rep}_{\mathbf{QM/MM}}$

As pointed out above, Eq. 5.4 requires the calculation of a two-electron integral. Such an integral is formally similar to the exchange integral (with opposite sign), where one of the densities in Eq. 5.2 has been replaced with an explicit function of $\mathbf{r_1}$ and $\mathbf{r_2}$. A similar approach is sometimes used within the framework of Density Functional Theory (DFT), in the development of Hybrid Density Functionals with a non-local contribution to the energy. Similar to the definition of exchange term, in DFT an exchange-repulsion energy density can be defined:[443]

$$\varepsilon^{\mathrm{rep}}(\mathbf{r_1}) = \frac{1}{2}\int \mathrm{d}\mathbf{u}\ \frac{\rho_{\mathrm{QM}}(\mathbf{r}_1,\mathbf{r_1}+\mathbf{u})\rho_{\mathrm{MM}}(\mathbf{r}_1,\mathbf{r_1}+\mathbf{u})}{u} \qquad \mathbf{5.5}$$

where $\mathbf{u} = \mathbf{r_2} - \mathbf{r_1}$ has been introduced (u is its module). In this formalism, ρ_{MM} acts as "semi-local density" in the DFT framework.[443–445,445–447] By inserting ρ_{MM} in Eq. 5.5, we obtain:

$$\varepsilon^{\mathrm{rep}}(\mathbf{r_1}) = \frac{1}{2}\int \mathrm{d}\mathbf{u}\ \frac{\rho_{\mathrm{QM}}(\mathbf{r_1},\mathbf{r_1}+\mathbf{u})}{u}\ \left(\sum_{\mathbf{R}} \xi^2_{\mathbf{R}} e^{-\beta_{\mathbf{R}}|\mathbf{r_1}-\mathbf{R}|^2} e^{-\beta_{\mathbf{R}}|\mathbf{r_1}+\mathbf{u}-\mathbf{R}|^2}\right) \qquad \mathbf{5.6}$$

By exploiting the standard approach of expanding noninteracting reference system's Kohn-Sham orbitals in a finite basis set of real, nonorthogonal, Gaussian-type atomic orbitals $\{\chi_\mu\}$, the non-local one-particle density becomes:

$$\rho_{\mathrm{QM}}(\mathbf{r_1}, \mathbf{r_1}+\mathbf{u}) = \sum_{\mu\nu} P_{\mu\nu}\chi_\mu(\mathbf{r_1})\chi_\nu(\mathbf{r_1}+\mathbf{u}) \tag{5.7}$$

where $P_{\mu\nu}$ is the μ,ν element of the density matrix P. By substituting Eq. 5.7 into Eq. 5.6, and symmetrizing, to recover the notation of Rung 3.5 functionals, $\varepsilon^{\mathrm{rep}}(\mathbf{r_1})$ becomes:

$$\varepsilon^{\mathrm{rep}}(\mathbf{r_1}) = \frac{1}{2}\sum_{\mu\nu} P_{\mu\nu}\left[\frac{\chi_\mu(\mathbf{r_1})A_\nu(\mathbf{r_1}) + A_\mu(\mathbf{r_1})\chi_\nu(\mathbf{r_1})}{2}\right] \tag{5.8}$$

where, similarly to what is done in the context of the definition of the so-called "hybrid Rung 3.5 density functionals" [447–450] the A_μ function is introduced:

$$A_\mu(\mathbf{r_1}) = \int \mathrm{d}\mathbf{u}\ \frac{\chi_\mu(\mathbf{r_1}+\mathbf{u})}{u}\left(\sum_{\mathbf{R}} \xi_{\mathbf{R}}^2 e^{-\beta_{\mathbf{R}}|\mathbf{r_1}-\mathbf{R}|^2} e^{-\beta_{\mathbf{R}}|\mathbf{r_1}+\mathbf{u}-\mathbf{R}|^2}\right) \tag{5.9}$$

Eq. 5.9 has the form of an electrostatic potential integral, yielding the potential at point $\mathbf{r_1}$ due to the product of a basis function centered at $\mathbf{R}_\mu$ and the sum of the Gaussian functions representing ρ_{MM}, centered at $\mathbf{R}$. Such an integral can be calculated analytically, for instance by specifying the Obara-Saika algorithm [451–453] to the evaluation of Eq. 5.9. The details on the formulation and implementation of this algorithm in the context of the present work are given as Supporting Information (SI) (Section S1). Notice however, that a straightforward adaption of the current implementations of Rung 3.5 density functionals [447–450] to the evaluation of Eq. 5.9 is impossible. In addition, the definition of ρ_{MM} (Eq. 5.3) does not allow the use of the auxiliary basis sets exploited in Rung 3.5 functionals, because the gaussian functions which we are using (Eq. 5.3) are centered in the MM grid. Since they are by definition non-symmetric functions, the Müntz theorem [454] can not be applied to our case.
By re-writing Eq. 5.9 as a function of $\mathbf{r_2}$:

$$A_\mu(\mathbf{r_1}) = \sum_{\mathbf{R}} \xi_{\mathbf{R}}^2 e^{-\beta_{\mathbf{R}}|\mathbf{r_1}-\mathbf{R}|^2} \int \mathrm{d}\mathbf{r_2} \frac{\chi_\mu(\mathbf{r_2}) e^{-\beta_{\mathbf{R}}|\mathbf{r_2}-\mathbf{R}|^2}}{|\mathbf{r_2}-\mathbf{r_1}|} \tag{5.10}$$

the exchange-repulsion energy, $E^{rep}_{\mathrm{QM/MM}}$, can be calculated by numerical integration of the energy density in Eq. 5.8:

$$E^{rep}_{\mathrm{QM/MM}} = \frac{1}{2}\sum_{\mu\nu} P_{\mu\nu}\int \mathrm{d}\mathbf{r_1}\ \left[\frac{\chi_\mu(\mathbf{r_1})A_\nu(\mathbf{r_1}) + A_\mu(\mathbf{r_1})\chi_\nu(\mathbf{r_1})}{2}\right] \tag{5.11}$$

and the matrix element of the Pauli-Repulsion potential, to be added to the QM Fock matrix, reads:

$$F_{\mu\nu}^{rep} = \frac{\partial E^{rep}}{\partial P_{\mu\nu}} = \frac{1}{2}\int \mathrm{d}\mathbf{r_1}\ \left[\frac{\chi_\mu(\mathbf{r_1})A_\nu(\mathbf{r_1}) + A_\mu(\mathbf{r_1})\chi_\nu(\mathbf{r_1})}{2}\right] \quad \textbf{5.12}$$

It is worth remarking that the resorting to the DFT formalisms allows to transform the two-electron integral in Eq. 5.4 into a one-electron integral, which can be evaluated by integration over grid points defined in the DFT formalism.
Also, Eq. 5.11 depends explicitly both on $P_{\mu\nu}$ and the atomic basis $\{\chi_\mu\}$. This introduces an explicit contribution to the QM Hamiltonian, which propagates to the calculation of molecular properties and spectra, through the definition of suitable analytical procedures. Such an extension will be the topic of further investigations.

5.2.2 Quantum Dispersion Energy

The exact quantum-mechanical definition of the dispersion interaction originally proposed by McWeeny, results in a computational expensive approach, depending on transition densities of the QM portion.[192,455] A popular remedy to this issue, widely used in the case of dispersion corrected density functionals,[194,198,199,456–466] consists of adding a pairwise interatomic C_6R^{-6} term to the DFT energy:

$$E^{dis} = -\frac{1}{2}\sum_{A,B} f_{damp}(R_{AB}, R_A^0, R_B^0)C_{6AB}R_{AB}^{-6} \quad \textbf{5.13}$$

where, R_{AB} is the distance between atoms (portions) A and B, C_{6AB} is the corresponding C_6 coefficient, R_A^0 and R_B^0 are the van der Waals (vdW) radii. The R_{AB}^{-6} singularity at small distances is eliminated by the short-range damping function $f_{damp}(R_{AB}, R_A^0, R_B^0)$.
Among the several corrections proposed in the literature, we found the approach by Tkatchenko and Scheffler (TS) the most suitable for our purposes, due to its mathematical formulation and its performances.[198] Starting from the Casimir-Polder equation and the Padé series, the C_{6AB} coefficients are defined by only using homonuclear parameters, i.e. C_{6AA}, C_{6BB}, α_A^0, α_B^0 (the latter being static polarizabilities of the A and B moieties):

$$C_{6AB} = \frac{2C_{6AA}C_{6BB}}{\frac{\alpha_B^0}{\alpha_A^0}C_{6AA} + \frac{\alpha_A^0}{\alpha_B^0}C_{6BB}} \quad \textbf{5.14}$$

Similarly to repulsion, also for dispersion terms only the interaction between QM (A) and MM (B) atoms will be considered. The TS model resorts to an Atom in Molecules[467] approach and adopts the Hirshfeld[468] partition of the density to define effective homonuclear coefficients C_{6AA}^{eff} of the A atom in the molecule:

$$C_{6AA}^{eff} = \left(\frac{V_A^{eff}}{V_A^{free}}\right)^2 C_{6AA}^{free} = \eta_A^2 C_{6AA}^f \quad \textbf{5.15}$$

where (V_A^{eff}) is the effective volume of the A atom in the molecule, (V_a^{free}) is the free volume of the same atom, and C_{6AA}^{free} are the free homonuclear coefficients C_{6AA}^{free}. η_A can be written in terms of the electron density of the system by employing the Hirshfeld partitioning of the density:[468]

$$\eta_A = \frac{\int \mathrm{d}\mathbf{r}\ r^3 w_A(\mathbf{r})\rho(\mathbf{r})}{\int \mathrm{d}\mathbf{r}\ r^3 \rho_A^{free}(\mathbf{r})} \tag{5.16}$$

$$w_A(\mathbf{r}) = \frac{\rho_A^{free}(\mathbf{r})}{\sum_J \rho_J^{free}(\mathbf{r})} \tag{5.17}$$

where, $w_A(\mathbf{r})$ is the Hirshfeld atomic partitioning weight for the atom A, r is the distance from the nucleus , $\rho(\mathbf{r})$ is the total electron density, $\rho_A^{free}(\mathbf{r})$ is the electron density of the free atom A, and the summation runs over all atoms J in the system.
By following this approach, the QM/MM dispersion energy, $E_{QM/MM}^{dis}$, becomes:

$$E_{\mathrm{QM/MM}}^{dis} = -\frac{1}{2}\sum_{A\in\mathrm{QM}}\sum_{B\in\mathrm{MM}} f_{damp}(R_{AB}, R_A^0, R_B^0)\frac{\eta_A^2 C_{AA}^{free} C_{6BB}^{eff}}{\frac{\alpha_B^0}{\alpha_A^0}\eta_A^2 C_{AA}^{free} + \frac{\alpha_A^0}{\alpha_B^0} C_{6BB}^{eff}} R_{AB}^{-6} \tag{5.18}$$

The C_{6BB}^{eff} are the effective homonuclear coefficients of the B (MM) atoms. Due to the difficulty to express them through Eq. 5.15, in this work their values are parametrized in an atom-type fashion with respect to QM calculations based on the Hirshfeld partitioning proposed in TS.[469,470] α_A^0 and α_B^0 are parametrized with respect to high-level QM calculations (vide infra).
$f_{damp}(R_{AB}, R_A^0, R_B^0)$ in Eq. 5.18 is a Fermi-type damping function, which is specified by following the standard approaches exploited to define dispersion corrected density functionals:[194,198,466]

$$f_{damp}(R_{AB}, R_A^0, R_B^0) = \frac{1}{1+\exp\left[-d\left(\frac{R_{AB}}{s_R R_{AB}^0} - 1\right)\right]} \tag{5.19}$$

where, $R_{AB}^0 = R_A^0 + R_B^0$, and d, s_R are free parameters (see Section 5.5.1).
The dispersion contribution to the QM/MM Fock matrix is:

$$F_{\mu\nu}^{dis} = \frac{\partial E_{dis}}{\partial \rho}\frac{\partial \rho}{\partial P_{\mu\nu}} = -\frac{1}{2}\frac{\partial \rho}{\partial P_{\mu\nu}}\sum_{A\in\mathrm{QM}}\sum_{B\in\mathrm{MM}} f_{damp}(R_{AB})\frac{\partial C_{6AB}}{\partial \rho} R_{AB}^{-6} \tag{5.20}$$

By considering that the free atomic related quantities are independent of the density matrix, and that the same obviously applies to MM-related quantities, the terms in Eq. 5.20 can be written as follows:

$$\frac{\partial C_{6AB}^{eff}}{\partial \rho} = \frac{2\frac{\alpha_A^0}{\alpha_B^0} C_{6BB}^2}{\left(\frac{\alpha_B^0}{\alpha_A^0} C_{6AA}^{eff} + \frac{\alpha_A^0}{\alpha_B^0} C_{6BB}\right)^2}\frac{\partial C_{6AA}^{eff}}{\partial \rho} \tag{5.21}$$

$$\frac{\partial C_{6AA}^{eff}}{\partial \rho} = C_{6AA}^{free} 2\eta_A \frac{\partial \eta_A}{\partial \rho} \tag{5.22}$$

$$\frac{\partial \eta_A}{\partial \rho} = \eta_A^{\rho} = \frac{\int \mathrm{d}\mathbf{r}\ r^3 w_A(\mathbf{r})\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})}{\int \mathrm{d}\mathbf{r}\ r^3 \rho_A^{free}(\mathbf{r})} \tag{5.23}$$

where the eff superscript in the C_{6BB} term is omitted for the sake of readability of the equations. In Eq. 5.23 the term due to the partial derivative of the density with respect the density matrix is accounted for. By recollecting all the terms in of the above equations, the quantum dispersion contribution to the Fock matrix becomes:

$$F_{\mu\nu}^{dis} = -\frac{1}{2} \sum_{A \in \mathrm{QM}} \sum_{B \in \mathrm{MM}} f_{damp}(R_{AB}) \frac{2\frac{\alpha_A^0}{\alpha_B^0} C_{6BB}^2 C_{6AA}^{free} 2\eta_A}{\left(\frac{\alpha_B^0}{\alpha_A^0} C_{6AA}^{eff} + \frac{\alpha_A^0}{\alpha_B^0} C_{6BB}\right)^2} \eta_A^{\rho} R_{AB}^{-6} \tag{5.24}$$

Similarly to the Pauli repulsion term, Eq. 5.24 introduces an explicit contribution to the QM Hamiltonian, which propagates to the calculation of molecular properties and spectra, which will be considered in future communications.

5.3 Coupling dispersion/repulsion to non-polarizable QM/MM approaches

In electrostatic embedding QM/MM models the MM atoms are endowed with fixed atomic charges, that produce an electric field which polarizes the electron density. The electrostatic embedding introduces a new term in the molecular Hamiltonian, that is, the interaction between the potential generated by the MM charges and the electron density:

$$H_{\mathrm{QM/MM}}^{ele} = \sum_{j=1}^{N_{\mathrm{MM}}} \int_{\mathbf{R}^3} \frac{\rho_{\mathrm{QM}}(\mathbf{r}) q_j}{|\mathbf{r} - \mathbf{r}_j|} \mathrm{d}\mathbf{r} \tag{5.25}$$

In Eq. 5.25 the summation runs over the j MM charges. Notice how the MM charges, that are parameters of the employed force field, are a fundamental quantity: their quality is crucial as they provide a representation, albeit crude, of the electron density of the environment. Quantum Pauli repulsion and quantum dispersion act as additive

contributions to Eq. 5.25:

$$\begin{aligned} H_{\mathrm{QM/MM}} &= H^{ele}_{\mathrm{QM/MM}} + H^{rep}_{\mathrm{QM/MM}} + H^{dis}_{\mathrm{QM/MM}} = \\ &= \sum_{j=1}^{N_{\mathrm{MM}}} \int_{\mathbf{R}^3} \frac{\rho_{\mathrm{QM}}(\mathbf{r}) q_j}{|\mathbf{r}-\mathbf{r}_j|}\,\mathrm{d}\mathbf{r} + \\ &+ \frac{1}{2}\sum_{\mathbf{R}} \int \frac{\mathrm{d}\mathbf{r_1}\,\mathrm{d}\mathbf{r_2}}{r_{12}} \rho_{\mathrm{QM}}(\mathbf{r_1},\mathbf{r_2}) \left[\beta^2_{\mathbf{R}} e^{-\xi_{\mathbf{R}}(\mathbf{r_1}-\mathbf{R})^2} \cdot e^{-\xi_{\mathbf{R}}(\mathbf{r_2}-\mathbf{R})^2}\right] + \\ &- \frac{1}{2}\sum_{A\in\mathrm{QM}}\sum_{B\in\mathrm{MM}} f_{damp}(R_{AB}, R^0_A, R^0_B) \frac{\eta^2_A C^{free}_{AA} C^{eff}_{6BB}}{\frac{\alpha^0_B}{\alpha^0_A}\eta^2_A C^{free}_{AA} + \frac{\alpha^0_A}{\alpha^0_B} C^{eff}_{6BB}} R^{-6}_{AB} \end{aligned} \qquad \textbf{5.26}$$

By expanding ρ_{QM} in a finite basis set $\{\chi_\mu\}$ and taking the derivative with respect to the density matrix $P_{\mu\nu}$, it is possible to define the contribution to the Fock matrix:

$$\begin{aligned} F_{\mu\nu} &= h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{V}^{\dagger}_{\mu\nu}\mathbf{q} + \frac{1}{2}\int \mathrm{d}\mathbf{r_1} \left[\frac{\chi_\mu(\mathbf{r_1})A_\nu(\mathbf{r_1}) + A_\mu(\mathbf{r_1})\chi_\nu(\mathbf{r_1})}{2}\right] + \\ &- \frac{1}{2}\sum_{A\in\mathrm{QM}}\sum_{B\in\mathrm{MM}} f_{damp}(R_{AB}) \frac{2\frac{\alpha^0_A}{\alpha^0_B} C^2_{6BB} C^{free}_{6AA} 2\eta_A}{\left(\frac{\alpha^0_B}{\alpha^0_A} C^{eff}_{6AA} + \frac{\alpha^0_A}{\alpha^0_B} C_{6BB}\right)^2} \eta^\rho_A R^{-6}_{AB} \end{aligned} \qquad \textbf{5.27}$$

5.3.1 Coupling dispersion/repulsion to the polarizable QM/FQ model

In polarizable embedding QM/MM models, the mutual polarization of the MM and QM portions is explicitly taken into account. The MM force field contains a *response* term, which modifies the electrostatics as a reaction to the presence of the QM density. In a symmetric fashion, a polarization term is included in the core's Hamiltonian to represent the interaction of the electronic density with the MM electrostatics. If the polarizability of the MM region is introduced by means of induced point dipoles,[10,164,165] the electric field produced by the QM density appears in the equations that determine the dipoles and the dipoles appears in an interaction term in the Hamiltonian multiplied by a field operator. If instead a fluctuating charge (FQ) description[56,156–159] is adopted to make the force field polarizable, the electrostatic potential produced by the QM density gives rise to a charge flow in the MM region; the MM fluctuating charges in turn interact with the QM density. Therefore, the expression for the interaction between the QM and MM portions is the same as Eq. 5.25, but the charges are in this case calculated by solving the following response equation:[54]

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q - \mathbf{V}(\mathbf{P}_{\mathrm{QM}}) \qquad \textbf{5.28}$$

where $\mathbf{D}$ is the response matrix whose diagonal terms are the atomic electronegativities, $\mathbf{q}$ is a vector containing the FQs and the Lagrangian multipliers, $\mathbf{C}$ is a vector containing the atomic electronegativities and the constraints to ensure that each MM molecule has fixed charge, and $\mathbf{V}(\mathbf{P})$ is the potential due to the QM density matrix $\mathbf{P}$. The Pauli repulsion and quantum dispersion terms developed in the previous pages can be added to the QM/FQ Hamiltonian in the same fashion as in Eq. 5.26. The resulting expression is the same,
however the q_j charges in Eq. 5.26 and Eq. 5.27 this time are the FQs calculated through Eq. 5.28 at each step of the SCF procedure.

5.4 Computational Details

The equations presented in the previous section were implemented in the Gaussian16 computational package.[301] Notice that the current implementation of Eq. 5.9 is restricted to uncontracted basis sets of s-, p- and cartesian d-type primitive gaussian functions. In all QM/FQ and QM/non-polarizable MM calculations were performed by treating the QM portion at the DFT level of theory, combined with selected Pople-type basis sets. The parameters to treat the electrostatic component in FQ calculations were taken from Rick et al.[156] The TIP3P[358] force-field was exploited in non-polarizable MM calculations. All the classical Molecular Dynamics (MD) simulations were performed by using the Gromacs package,[471–474] with the same settings as previously reported by some of the present authors.[146,148] The Kitaura-Morokuma Energy Decomposition Analysis (KM-EDA)[475,476] was performed by using the GAMESS package.[477,478] Symmetry Adapted Perturbation Theory (SAPT)[440,441] calculations were performed by using Psi4 1.1.[479]

5.5 Numerical Results

In this section the methodology explained in Sections 2-3 is applied to test cases. In particular, the model is first parametrized to treat the aqueous solution, so to reliably reproduce some of the most relevant properties of bulk water. Then, the method is applied to the calculation of the non-electrostatic interaction energy of molecular systems in aqueous solution.

5.5.1 Parametrization Strategy: Aqueous Solutions

The methodology stretched in the previous sections is general enough to be applied to different polarizable and non-polarizable QM/MM approaches and to model any kind of external environment, pending an appropriate parametrization of the quantities entering Eqs. 5.11, 5.12, 5.18 and 5.24. Such a parametrization is a crucial step towards the routinely application of the method to real cases. The development of accurate parametrizations for various kinds of environments is beyond the scope of this

paper. Here, we will present the strategy that is followed to parametrize the method coupled with our polarizable QM/MM model based on FQs (QM/FQ), and specifically tailored to aqueous solutions.[54,146,149,150]

In the specific case of water, Eq. 5.4 requires the definition and the numerical setting of 4 parameters: the exponents of the lone pairs and bond pairs $\beta_{\mathbf{R}}$ and the coefficients $\xi_{\mathbf{R}}$. In fact, the use of the same parameters for bond and lone pairs would not been justified. Also, the actual positions of the gaussian-type functions (Eq. 5.3) has to be set, and this introduces a further degree of freedom in the parametrization procedure. In case of water, eq. 5.3 reads:

$$\rho_{\mathrm{H_2O}}(\mathbf{r_1},\mathbf{r_2}) = \sum_{j=1}^{N_{\mathrm{H_2O}}} \sum_{i=1}^{4} \xi_i^2 e^{-\beta_i(\mathbf{r_1}-\mathbf{R}_i^{(j)})^2} \cdot e^{-\beta_i(\mathbf{r_2}-\mathbf{R}_i^{(j)})^2} \qquad \textbf{5.29}$$

where j runs over the water molecules of the MM portion, while i runs over the electron pairs of a single water molecule. Two sets of indices β_i and ξ_i are set, and again they differ if a lone-pair or a bond-pair is considered. $\mathbf{R}_i^{(j)}$ collects the points where the gaussian-type functions are centered, which are chosen in analogy with what is done in the TIP4P force field.[480] In particular, the $\mathbf{R}_i^{(j)}$ centers are set as the charge centroids of the localized molecular orbitals, as defined according to the Boys method[481] (see Figure 5.1), which in the present case were calculated at the B3LYP/6-311++G** level of theory (see Section S2 in SI).

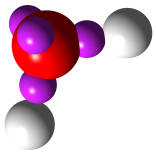


Figure 5.1. Calculated B3LYP/6-311++G** Boys localized orbitals centroids (purple spheres) for a single water molecule.

Notice that the way the ρ_{MM} is constructed, permits to extend this approach to solvents/environments other than water. Also, in the present work ρ_{MM} is defined in terms of spherical gaussian-type functions only.

By further substituting Eq. 5.29 in Eq. 5.2 we obtain:

$$E^{rep}_{\mathrm{QM/H_2O}} = \frac{1}{2} \sum_{j=1}^{N_{\mathrm{H_2O}}} \sum_{i=1}^{4} \int \frac{\mathrm{d}\mathbf{r_1}\,\mathrm{d}\mathbf{r_2}}{r_{12}} \rho_{\mathrm{QM}}(\mathbf{r_1},\mathbf{r_2}) \left[\xi_i^2 e^{-\beta_i(\mathbf{r_1}-\mathbf{R}_i^{(j)})^2} \cdot e^{-\beta_i(\mathbf{r_2}-\mathbf{R}_i^{(j)})^2}\right] \qquad \textbf{5.30}$$

In order to set the parameters entering Eq. 5.30, selected water clusters, chosen by following Refs.[482,483] (see Figure 5.2), were exploited. In particular, the reference full QM data for the Pauli repulsion energy of such clusters were calculated by performing a full QM calculation on each structure in Figure 5.2 at the Hartree-Fock (HF) level in combination with selected Pople-type basis sets, also including diffuse functions (6-31G, 6-31+G*, 6-311G, 6-311+G*). Then, the repulsion contribution to the energy

was extracted by resorting to the KM-EDA,[475,476] by following what has already been proposed in the literature.[482,483]

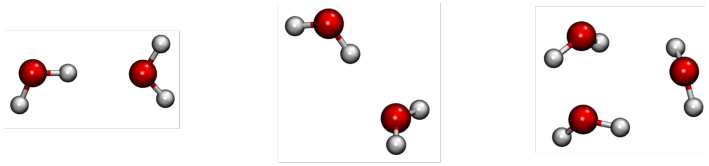


Figure 5.2. Water clusters used for the parametrization of $E^{rep}_{\mathrm{QM/H_2O}}$, see text.

Then, the Pauli Repulsion on the same water clusters was calculated with our method (Eq. 5.30). This has been done by treating only one water molecule at the QM level (B3LYP functional combined with the same selection of basis sets) and the other one (or two, in case of the trimer in Figure 5.2) at the FQ level. For each dimer structure, we performed two calculations, by exchanging the QM and FQ water molecules in order to average among hydrogen bond donor and acceptor moieties. Three calculations were performed for the trimer, by exchanging each time the QM molecule with one of the two FQ molecules.
Eq. 5.30 depends on 4 parameters (the exponents and the coefficients of each gaussian function): their best values were defined by performing a least square roots fitting on full QM data obtained with the KM-EDA approach, without setting any constraint on the parameters. The best fitted values are reported in Section S2 in SI. Notice that such values give repulsion energies not perfectly fitting the KM-EDA data (see Table S1 in SI); this is probably due to the absence of the contributions due to p-type gaussian functions in the MM moiety. Such functions can possibly be added by extending the formalism in a straightforward way, by only making the computations more cumbersome.
Moving to quantum dispersion (Eq. 5.18), its expression depends on several parameters, which were set according to the following scheme:

- α_0 are static atomic polarizabilities. They were calculated at the CCSD(T)/aug-cc-pVTZ level of theory, or taken from the literature.[484] The used values are reported in Table S2, in SI.

- The homonuclear C_6^{free} coefficients were taken from Chu and Dalgarno,[485,486] and are reported in Table S3 in SI.

- The homunuclear C_6^{eff} coefficients of the MM atoms cannot be intuitively defined, because they actually depend on the effective volume of a given atom in a molecule. Since our target environment is water, the effective volumes of oxygen and hydrogen atoms in a water molecule optimized at the B3LYP/6-311++G** were calculated. From these values, the C^{eff}_{6OO} and and C^{eff}_{6HH} were calculated, being 14.8 Hartree·bohr^6 and 2.8 Hartree·bohr^6, respectively. Notice that these data are in agreement with those proposed by TS.[198]

- The d coefficient of Eq. 5.19 was set to 20, according to the TS approach.[198] An extensive testing was however performed, showing that similar E_{dis} values are obtained for any choice of d between 15 and 40.
- The van der Waals radii R^0 were set to the Bondi reference values.[487]
- As previously noticed in the literature,[198] the s_R coefficient in Eq. 5.19 is actually the only empirical parameter. s_R was chosen in such a way E_{dis} of a water dimer as a function of the intermolecular distance O-O (calculated at the B3LYP-D3 level of theory) is accurately reproduced. The s_R coefficient was therefore set to 0.92, which gives an average error of about 5% in the region of hydrogen bonding (d_{O-O} 2.5-3.0 Å) (see below for more details). Notice that the calculated E_{dis} at the equilibrium distance between two neighbour water molecules (d_{O-O} = 2.9 Å)[488] is very close to the value reported recently by Guidez and Gordon.[489]

5.5.2 Dependence of E_{rep} and E_{dis} on the water-water intermolecular distance

In this section, the dependence of E_{rep} and E_{dis} on the water-water intermolecular distance is studied. To this end, the water dimer depicted in Figure 5.3 has been exploited, and the distance d between the oxygen atoms has been taken as reference. Notice that this distance has been chosen as reference because it is generally reported in the experimental/computational evaluation of radial distribution functions.[488]

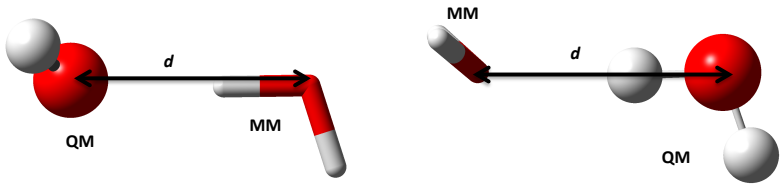


Figure 5.3. Structure of the water dimer used to study the dependence of E_{Qnel} on the water-water intermolecular distance.

In Figure 5.4, E_{rep} is reported as a function of d. The plot was constructed by performing 80 calculations increasing the O-O distance from 2.54 Å to 6.49 Å by a step of 0.05 Å. E_{rep} was calculated both with the QM/FQ and non-polarizable QM/MM(TIP3P) methods, by exploiting the B3LYP/6-31+G* level to treat the QM moiety. Also in this case, the QM and MM moieties were interchanged, and the average values were taken. In Figure 5.4, these data are compared with the repulsion energy obtained at the full QM level by means of the KM-EDA approach.
An almost perfect superposition of QM/FQ and QM/MM results is observed. For $d < 3.5$ Å, for which E_{rep} is large, the QM/FQ method shows an average percentage deviation from the full QM KM-EDA of around 10%, similarly to the non-polarizable QM/TIP3P approach. Notice that the QM/FQ results are in very good agreement with KM-EDA results in the region around $d = 3$ Å. This is not unexpected, because

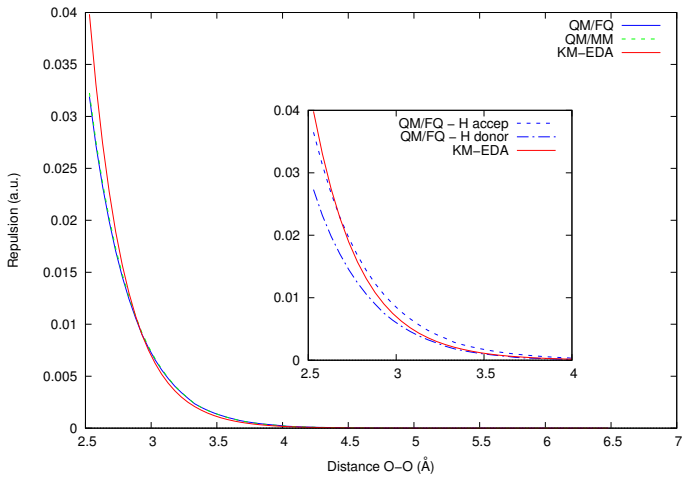


Figure 5.4. Plot of E_{rep} as a function of the O-O intermolecular distance in the water dimer depicted in Figure 5.3. QM/FQ and non polarizable QM/MM(TIP3P) values (B3LYP/6-31+G* for the QM moiety) are compared to KM-EDA (HF/6-31+G*) calculations. The inset shows E_{rep} calculated values with the QM/FQ approach, in case the QM water molecule acts as H-bond donor or acceptor.

the water-water structures that were exploited to perform the parametrization of this contribution (see Figure 5.2) were characterized by a similar intermolecular distance (2.04 Å).

The inset in Figure 5.4 shows in more detail the difference between the calculated QM/FQ and KM-EDA values in the region between 2.5 and 4 Å, the two curves obtained for the QM water molecule acting as H-bond donor and acceptor are given. We notice that E_{rep} is larger when the QM water molecule acts as H-bond acceptor. This is due to the fact that the Gaussian function in eq. 5.29 on the MM O-H water bond is larger than that related to the fictitious MM oxygen atom lone pair (i.e. the exponent of the function placed in the middle of the O-H distance is smaller than the exponent of the function placed at the position of the fictitious O lone pair). Therefore, the overlap between the Gaussian functions and the QM density is larger when the MM water molecule acts as H-bond acceptor, and this corresponds to a greater value of E_{rep}. The inset in Figure 5.4 also shows that the KM-EDA values lie almost always in between the two QM/FQ curves. This supports the averaging of the two values in the parametrization procedure (see the previous section). At small intermolecular distances the repulsion contribution is underestimated. This is probably related again to the absence of *p*-type Gaussian functions on the MM moiety, which would guarantee a greater overlap of the QM and MM densities.

Let us pass to discuss the dependence of the E_{dis} as a function of the intermolecular O-O distance. The data are plotted in Figure 5.5, which also reports the curve obtained with the B3LYP-D3 functional.[466]

Notice that also in this case the QM and MM moieties were interchanged (see inset

in Figure 5.5). Different from E_{rep}, E_{dis} is larger (in absolute value) when the QM water molecule acts as H-bond donor. This can be explained by considering the values obtained for the effective C_6^{eff} coefficients for the MM molecule (see previous section). In fact, when the QM molecule acts as H-bond donor, the oxygen atom of the MM water molecule is close to the QM portion:because C_{6OO}^{eff} is greater than C_{6HH}^{eff} (14.8 vs. 2.8 Hartree·bohr^6), E_{dis} increases.

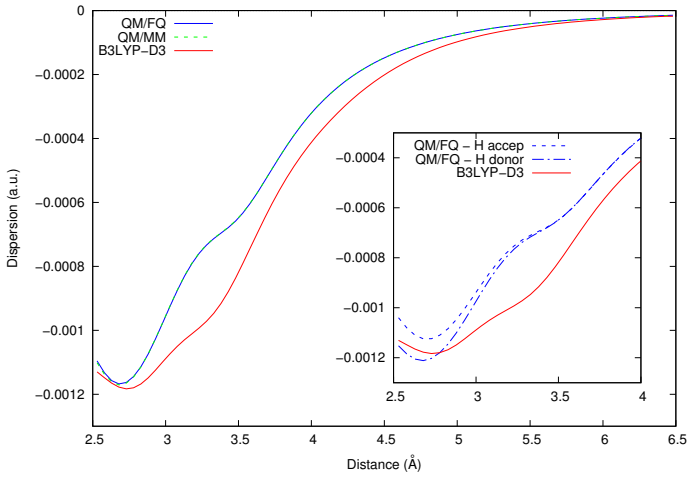


Figure 5.5. QM/FQ and non polarizable QM/TIP3P (B3LYP/6-31+G* for the QM moieties) quantum dispersion energy, E_{dis}, of the water dimer in Figure 5.3 as a function of the O-O distance. B3LYP-D3/6-31+G* data are also reported. The inset shows E_{dis} calculated values with the QM/FQ approach, in case the QM water molecule acts as H-bond donor or acceptor.

The behaviour of the total quantum non-electrostatic interaction energy E_{Qnel}, i.e. the sum of E_{rep} and E_{dis}, as a function of d is plotted in Figure 5.6, which also reports the SAPT2+3(CCD)/aug-cc-pVDZ curve. Also in this case the QM and MM moieties were interchanged (see inset in Figure 5.6). Comparison of Figure 5.5 and Figure 5.4 shows that the repulsion term is generally larger than the dispersion contribution: this clearly emerges from the trend reported in Figure 5.6, which closely resembles Figure 5.4. It is also worth pointing out that QM/FQ and non-polarizable QM/TIP3P give similar E_{Qnel} values. This is a further evidence of the stability of our parametrization, which gives similar results as changing the force field used to represent the MM portion.

To end this discussion, the total interaction energy as a function of d is plotted in Figure 5.7 and compared with CCSD(T)/aug-cc-pVTZ data (counterpoise corrections are included). The differences in the two curves can be attributed to the electrostatic contributions and the lack in our model of charge transfer effects and multipole terms. The equilibrium distance is a bit shifted in our model (3.14 Å vs. 2.99 Å), but the interaction energy at the equilibrium distance is accurately reproduced with an error of only 7% with respect to the CCSD(T) data and comparable with similar data reported in the literature.[346,489,490]

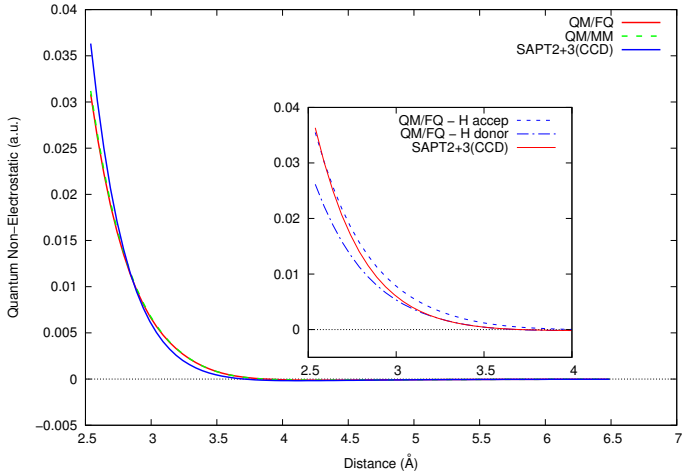


Figure 5.6. QM/FQ and non polarizable QM/TIP3P (B3LYP/6-31+G* for the QM moieties) quantum non electrostatic interaction energy, E_{Qnel}, of the water dimer in Figure 5.3 as a function of the O-O distance. SAPT2+3(CCD)/aug-cc-pVDZ data are also reported. The inset shows E_{Qnel} calculated values with the QM/FQ approach, in case the QM water molecule acts as H-bond donor or acceptor.

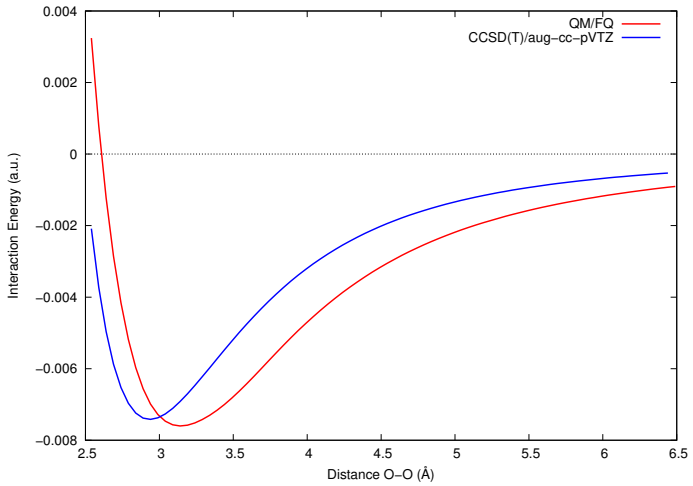


Figure 5.7. QM/FQ (B3LYP/6-31+G* for the QM moieties) and CCSD(T)/aug-cc-pVDZ total interaction energy for the water dimer in Figure 5.3 as a function of the O-O distance.

5.5.3 Testing on water dimers

In order to test the quality of the parametrization presented in the previous sections, the methodology was applied to ten water dimer structures, taken from Kratz and coworkers (see Figure 5.8).[491]

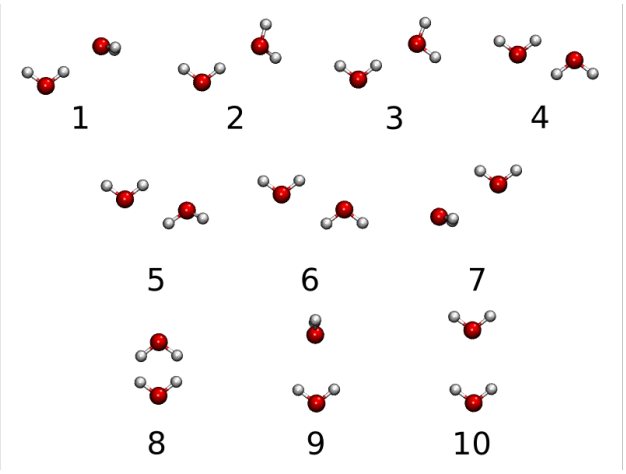


Figure 5.8. Selected water dimer structures, taken from Kratz et al.[491]

All the calculations were performed with both the QM/FQ and QM/MM(TIP3P) methods, coupled with the B3LYP/6-31+G* level for the QM moiety. Again, each time the QM and MM portions were exchanged, and the two values averaged to get the final results (see previous section). E_{rep} values were compared to KM-EDA data (see Figure 5.9 and Table 5.1).
QM/FQ and QM/TIP3P values are very similar, although the parametrization has been performed only with the QM/FQ method, thus confirming once again the stability of our approach. The largest deviation from KM-EDA results is given by structures 4-6, which show the smallest oxygen-oxygen distance, probably due to the absence of *p*-type gaussian functions placed at lone pair positions.
The Δ values in Table 5.1, i.e. the half-difference of the two calculations performed by exchanging the QM and MM moieties, are larger for structures characterized by a strong H-O intermolecular interaction. Furthermore, except for structure 7, the QM/FQ Δ values are greater than the corresponding QM/MM ones. This difference is linked to the different physical description of the MM portion in the polarizable and non-polarizable models. In fact, the polarizable QM/FQ approach emphasizes the H-bond acceptor or donor characters of the MM molecule, thus resulting in a different interaction with the QM densities.
The total calculated quantum non-electrostatic contribution (E_{Qnel},i.e. the sum of the Pauli repulsion and quantum dispersion) is reported in Table 5.2. Only QM/FQ data are shown, due to the similarity of the calculated results by exploiting the QM/FQ and QM/TIP3P approaches. According to what is expected for aqueous solutions, which are dominated by electrostatic interaction, E_{Qnel} is always smaller than the

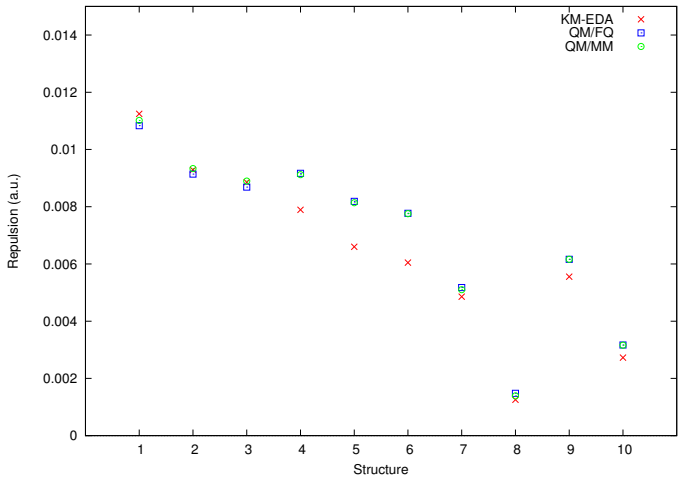


Figure 5.9. Calculated QM/FQ and QM/TIP3P (B3LYP/6-31+G*) E_{rep} values for the water dimer structures in Figure 5.8. KM-EDA (HF/6-31+G*) data are also reported for comparison.

	KM-EDA	QM/FQ	Δ	Err.%	QM/TIP3P	Δ	Err.%
1	11.245	10.831	1.517	3.68	11.013	0.550	2.06
2	9.284	9.138	1.203	1.57	9.334	0.320	0.54
3	8.856	8.687	1.154	1.91	8.897	0.285	0.46
4	7.892	9.167	-	16.16	9.118	-	15.54
5	6.600	8.185	0.003	24.02	8.139	0.001	23.32
6	6.046	7.769	-	28.51	7.756	-	28.28
7	4.858	5.174	0.039	6.51	5.100	0.516	4.99
8	1.255	1.478	-	17.75	1.391	-	10.81
9	5.556	6.161	1.506	10.89	6.164	1.159	10.94
10	2.727	3.168	0.738	16.17	3.162	0.533	15.95
	Average Error			12.72			11.29

Table 5.1. Calculated E_{rep} (10^{-3} Hartree) for the ten water dimers. Δ shows the deviation obtained by exchanging the QM and MM moieties. The percentage error from the KM-EDA values is also reported.

electrostatic term (E_{FQ}). However, the London contribution is not negligible, being as large as 35% of the total interaction energy.

	E_{rep}	E_{dis}	E_{Qnel}	E_{FQ}	E_{tot}
1	10.831	-1.123	9.708 (34%)	-18.483 (66%)	-8.775
2	9.138	-1.005	8.133 (33%)	-16.823 (67%)	-8.690
3	8.687	-0.977	7.710 (31%)	-17.014 (69%)	-9.304
4	9.167	-1.005	8.162 (35%)	-15.190 (65%)	-7.028
5	8.185	-0.988	7.197 (33%)	-14.516 (67%)	-7.319
6	7.769	-0.986	6.783 (31%)	-14.852 (69%)	-8.069
7	5.174	-1.687	3.487 (24%)	-11.068 (76%)	-7.581
8	1.478	-1.046	0.432 (12%)	-3.169 (88%)	-2.737
9	6.161	-1.174	4.987 (29%)	-12.498 (71%)	-7.511
10	3.168	-1.119	2.049 (18%)	-9.409 (82%)	-7.360

Table 5.2. E_{rep}, E_{dis}, E_{Qnel}, electrostatic energies (E_{FQs}) and total interaction energies, E_{tot}, for the ten water dimers. The values in parentheses give the percentage of the corresponding contribution with respect to E_{tot}. All energy values are given in 10^{-3} Hartree.

A closer look at Table 5.2 shows that E_{Qnel} is generally dominated by the Pauli repulsion interaction, which is always larger than E_{dis}. This behavior confirms the results reported above, where the dependence of such contributions on the intermolecular distance was outlined. Furthermore, the two terms are very similar for some structures, such as 8 and 10, thus demonstrating that the inclusion of both terms is compulsory to get a reliable description of E_{Qnel}.

To end the discussion on these water dimers, in Table 5.3 the total Quantum non electrostatic calculated by our approach is compared with SAPT2+3(CCD)/aug-cc-pVDZ values.

As already pointed out, the largest deviation from SAPT values is given by structures 4-6, which show the smallest oxygen-oxygen distance. Notice that the largest error is shown by dimer 8, for which, however, E_{Qnel} is very small and the deviation in absolute value is even smaller than for the other structures.

To further testing the quality of our approach, the model is applied to the water dimer in Figure 5.10, previously studied by Guidez and Gordon [489] by exploiting the EFP2($E_6 + E_7$) model.

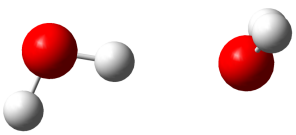


Figure 5.10. Structure of the water dimer optimized at the MP2/aug-cc-pVDZ level of theory [492] previously studied by Guidez and Gordon. [489]

In table 5.4, the three terms entering in the definition of the interaction energy are

	E_{Qnel}(QM/FQ)	E_{Qnel}(SAPT)
1	9.708 (1 %)	9.814
2	8.133 (3 %)	7.882
3	7.710 (3 %)	7.451
4	8.162 (19%)	6.863
5	7.197 (35%)	5.348
6	6.783 (44%)	4.707
7	3.487 (20%)	4.381
8	0.432 (50%)	0.866
9	4.987 (6 %)	4.693
10	2.049 (2 %)	2.000

Table 5.3. E_{Qnel} for the ten water dimers calculated by using our model and the SAPT2+3(CCD)/aug-cc-pVDZ. The values in parentheses give the percentage of the corresponding contribution with respect to the last column. All energy values are given in 10^{-3} Hartree.

reported as calculated by our model, the EFP2($E_6 + E_7$) and the Energy Decomposition Analysis performed at CCSD(T)/aug-cc-pVQZ level of theory, with the further inclusion of counterpoise corrections.

level of theory	QM/FQ	EPF2($E_6 + E_7$)[489]	EDA[490]
Electrostatic[a]	-11.01 (2 %)	-9.32 (14%)	-10.79
Exchange repulsion	6.57 (8 %)	5.59 (22%)	7.16
Dispersion	-0.70 (47%)	-0.51 (62%)	-1.33
Charge Transfer	N/A	-0.47	N/A
Total interaction energy	-5.15(4 %)	-4.71 (5%)	-4.95

Table 5.4. Electrostatic, exchange-Repulsion, dispersion and total interaction energy for the water dimer depicted in Figure 5.10 calculated by using our model, EFP2($E_6 + E_7$)[489] and Energy Decomposition Analysis EDA (CCSD(T)/aug-cc-pVQZ//MP2/aug-cc-pVQZ).[490] The values in parentheses give percentages with respect to the values in the last column. All energy values are given in kcal/mol. [a]: for EFP2 and EDA the electrostatic term is the sum of Coulomb and polarization contributions.

In the QM/FQ approach, electrostatic and polarization contribution cannot be separated, as previously reported by some of the present authors.[54] Therefore a single term is reported in Table 5.4. The largest deviation with respect to the EDA is shown by the dispersion term. However, as pointed out by Guidez and Gordon,[489] EDA overestimates the dispersion interaction due to the fact that it is computed as the difference between CCSD(T) and HF interaction energies. Overall, the agreement between our data and EDA is satisfactory, being the errors for the single terms generally small and the total interaction energy similar to the EDA value.

5.5.4 Dependence of E_{rep} and E_{dis} on the QM description

In this section, the dependence of calculated E_{rep} and E_{dis} values on the level used to model the QM moiety is studied. To this end, the water dimer depicted in Figure 5.3 with $d = 2.64$ Å is exploited. Eight different DFT functionals were selected, by following the recent literature,[147,493] ranging from pure (B97D[494,495]), to different classes of hybrid functionals (B3LYP,[496] B3PW91,[497] M062X,[498] PBE0,[499] SOGGA11-X[500]), also including long-range (CAM-B3LYP[347]) and dispersion corrections (ωB97xD[501]). Each functional was coupled to several Pople-type basis sets (see Figure 5.11 and Figure 5.12), in order to separate the contribution arising from polarization and diffuse functions.

Figure 5.11 reports schematically the trends obtained by computing E_{rep} with the different DFT functionals and the different basis sets. Numerical values are given in Table S4 in the SI. All DFT functionals predict very similar E_{rep} values as varying the basis set, with CAM-B3LYP always showing the highest values for a given basis set (on average, the CAM-B3LYP values are about 2% higher than the average value of the other functionals). This is not surprising, if the tendency of the CAM-B3LYP of spreading out the QM density is considered. Thus, the overlap between the QM density and the MM one is enlarged, resulting in an increase of E_{rep}.

The reported dependence on the choice of the basis set is also not surprising. In fact, the addition of functions on the hydrogen atoms (e.g. from 6-31G to 6-311G) increases the QM-MM overlap, and E_{rep} increases of about 6% on average. Such an increase is reduced when diffuse functions are included. The effect of polarization functions is usually negligible, however the addition of such functions generally results in the decreasing of E_{rep}, especially moving from single to double polarization functions. The addition of diffuse functions causes instead an increase of E_{rep}, due to an enlarged overlap between the QM and MM densities.

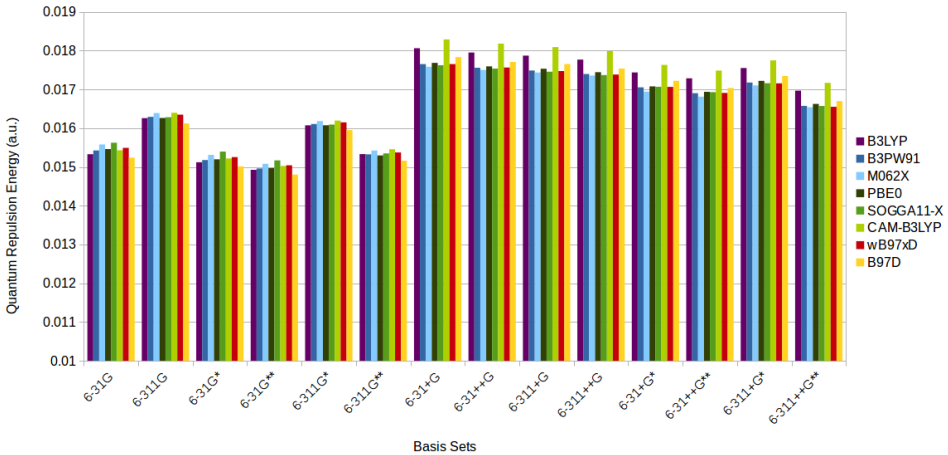


Figure 5.11. Dependence of E_{rep} on the basis set and DFT functional for the water dimer depicted in Figure 5.3.

Figure 5.12 reports the same analysis applied to E_{dis} (raw data are given in Table S5

in the SI).
The results reported in Figure 5.12 show some general trends, that are very similar to what already observed for E_{rep}. Firstly, CAM-B3LYP predicts the highest E_{dis}, and again this is probably due to the peculiarities of this functional. The observed trend as varying the basis set is similar to what has already been commented for E_{rep}. In fact, the addition of functions on hydrogen atoms increases E_{dis} of about 8%, and this increment is reduced when diffuse functions are considered. Once again, the effect of polarization functions is negligible, and finally results in the decreasing of E_{dis}, especially when double polarization functions are included. An opposite effect is observed when diffuse functions are added: an increase of E_{dis} is noticed, due to larger effective volumes of QM atoms and C_6^{eff} coefficients (see Eq. 5.15).

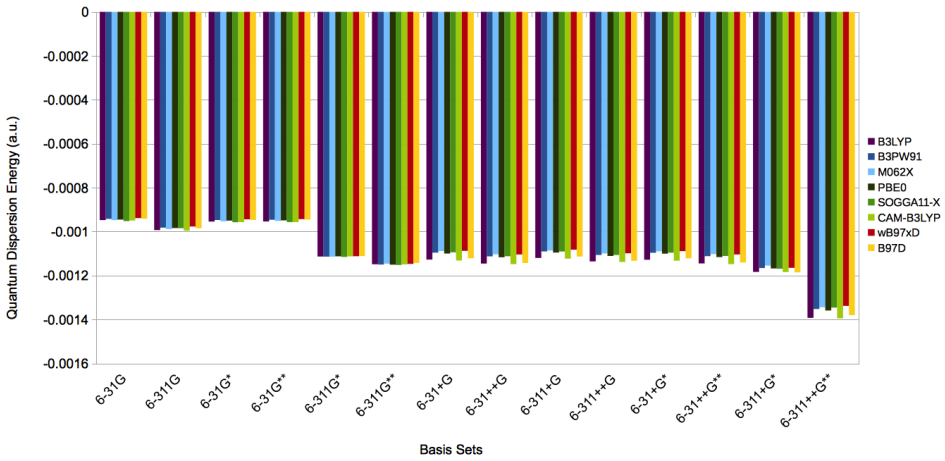


Figure 5.12. Dependence of E_{dis} on the basis set and DFT functional for the water dimer depicted in Figure 5.3.

The global effect of the choice of the DFT functional and basis set on E_{Qnel} is reported in Figure S3 in the SI (raw data are given in Table S6 in the SI). To end this discussion, it is worth pointing out that the results of the model here proposed are very stable as the functional and basis set vary. Furthermore, stable values of E_{rep} and E_{dis} are obtained by adding diffuse functions, so that their inclusion appears mandatory. For this reason, in the following section the 6-31+G* basis set, which adequately reproduces the total E_{Qnel}, is exploited.

5.6 Molecular Systems in aqueous solution described with the QM/FQ approach

We have shown in the previous section that our model permits a correct reproduction of the properties of the aqueous solution. In this section we will focus on the calculation of the non-electrostatic contribution to the energetic properties of molecular systems in bulk aqueous solution, as modeled with the QM/FQ approach. We first notice that

only the closest FQ solvent molecules will reasonably give a not negligible contribution to the solute-solvent interaction energy, due to the short-range character of E_{rep} and E_{dis}. This feature can help at reducing the computational cost of the calculation. In fact, suitable thresholds can be set. For quantum dispersion (Eq. 5.18) a cut-off is set, so that this term vanishes at intermolecular distances larger than 10 Å. Notice however that the further consideration of larger solvent shells would not increase much the computational demand.

On the contrary, the computational cost of the quantum repulsion term (Eq. 5.4) strongly depends on the number of water molecules around the solute, because each of them bears the gaussian functions used to represent ρ_{MM}; increasing the number of gaussian functions, makes the calculation of the the two-electron integral in Eq. 5.4 more and more cumbersome. Thus, the setting of a threshold appears beneficial.

To this end, only those FQ water molecules having a geometric center closer to at least a QM atom than a given geometric parameter R are included in the calculation of E_{rep}. In the practice, this requires to build up a cavity made of the union of identical spheres centered on each QM atom: only the MM molecules lying inside this cavity are considered in the evaluation of E_{rep}. Notice that in the present implementation the same radius R is used for different QM atom types: this may be possibly refined.

In order to validate this approach and to set a reasonable value of R, we took as test cases two random snapshots taken from a MD simulation of (L)-Methyl Lactate (MLAT) and (R)-Methyloxirane (MOXY) in aqueous solution. MOXY is a small rigid almost spherical molecule, whereas MLAT develops in the plane of the sp^2 carbon atom. On such snapshots, E_{rep} was calculated as a function of R by using the B3LYP/6-31+G* level to treat the QM solute. The results of such calculations are reported in Figure 5.13 for the resulting systems depicted in Figure 5.14 and 5.15. Notice that such figures only show the FQ water molecules relevant for the evaluation of E_{rep}, i.e. those within a range of variation of R between 1.5 Å and 5 Å with a step of 0.5. The other FQ water molecules are indeed present, but only contribute to the electrostatic interaction. The numbers of relevant water molecules associated at each radius for the structures depicted in Figure 5.14 and 5.15 are reported in Table 5.5.

Structure	Radius	N_{wat}	Structure	Radius	N_{wat}
ML-1	2.5	2	MO-1	3.0	5
ML-2	3.0	9	MO-2	3.5	13
ML-3	3.5	16	MO-3	4.0	19
ML-4	4.0	25	MO-4	4.5	27
ML-5	4.5	32	MO-5	5.0	32
ML-6	5.0	39			

Table 5.5. Radii (Å) of the spheres centered in each QM atom, and the total number of relevant waters for the structures depicted in Figures 5.14 and 5.15.

Figure 5.13 clearly shows that the trend of E_{rep} as a function of R strictly depends on the studied system. Also, the structure of the resulting clusters for a given R differs for

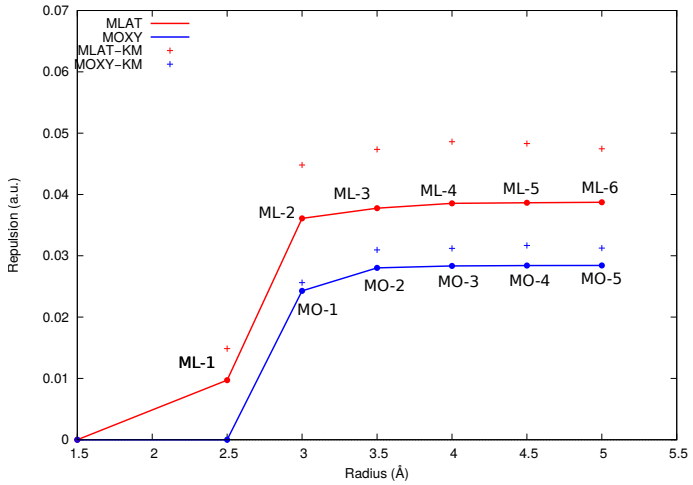


Figure 5.13. E_{rep} of (L)-Methyl Lactate and (R)-Methyloxirane in aqueous solution as a function of R. B3LYP/6-31+G* is used to treat the QM solute. E_{rep} calculated by exploiting the KM-EDA approach, at the HF/6-31+G* level of theory, is also reported.

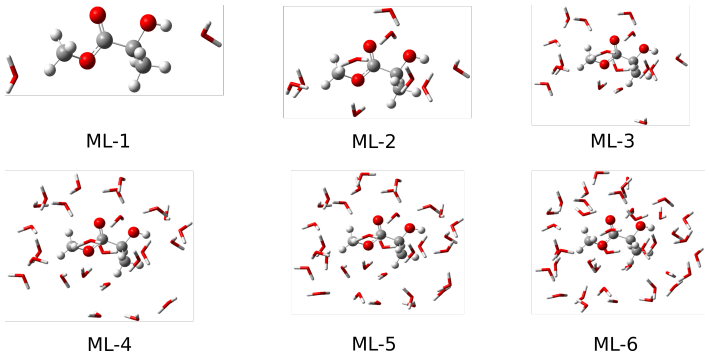


Figure 5.14. (L)-Methyl Lactate - water clusters arising from different choices of R.

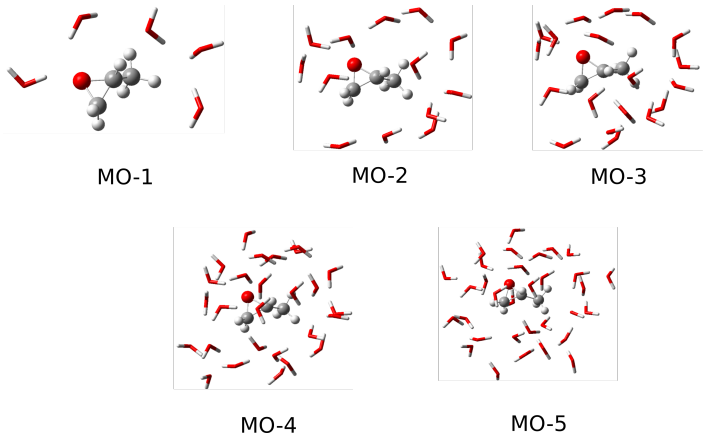


Figure 5.15. (R)-Methyloxirane - water cluster arising from different choices of R.

the two systems. MLAT has a greater surface area and a larger number of H-bond sites with respect to MOXY. This implies a greater exchange repulsion energy contribution for MLAT than for MOXY. Convergence in the repulsion energy value is reached at different values of the atomic radius R. In particular, imposing $R = 3.5$ Å is sufficient to describe the repulsion contribution for MOXY, whereas for MLAT a slightly larger value ($R = 4$ Å) is required. This is connected to the relative atomic positions: in case of MOXY, which is almost spherical, the majority of the relevant FQ molecules are shared by more than a single QM atom. MLAT has a more extended structure: therefore, increasing the R value causes new independent relevant FQ molecules to be included in the calculation.
Figure 5.13 also reports E_{rep} values obtained by using the KM-EDA approach. The error between our values and the reference KM-EDA data is about 9 % for MOXY, and about 20 % for MLAT. These findings confirm the applicability of our procedure to molecular systems in aqueous solution, in fact the calculated errors are of the same magnitude as what has been previously reported for water dimers (see Figure 5.8). Table S7 in the SI gives a more detailed comparison between our calculated values, KM-EDA data, and what can be obtained by exploiting the EFP2 approach.

5.6.1 Nicotine in aqueous solution

To end the section on the numerical testing of the developed procedure, the approach reported in this paper is applied to nicotine in aqueous solution (Figure 5.16, panel a)).
Starting from the MD performed previously by some of the present authors,[148] 300 snapshots were selected (more details on the MD protocol and the procedure for the extraction of the snapshots are given in the SI, section S5.1). The QM portion of the system was then described at the CAM-B3LYP/6-31+G* level of theory, according to previous studies on this molecule.[148,318] Figure 5.16, panel b) depicts a randomly

selected snapshot taken from the MD of Nicotine in aqueous solution.

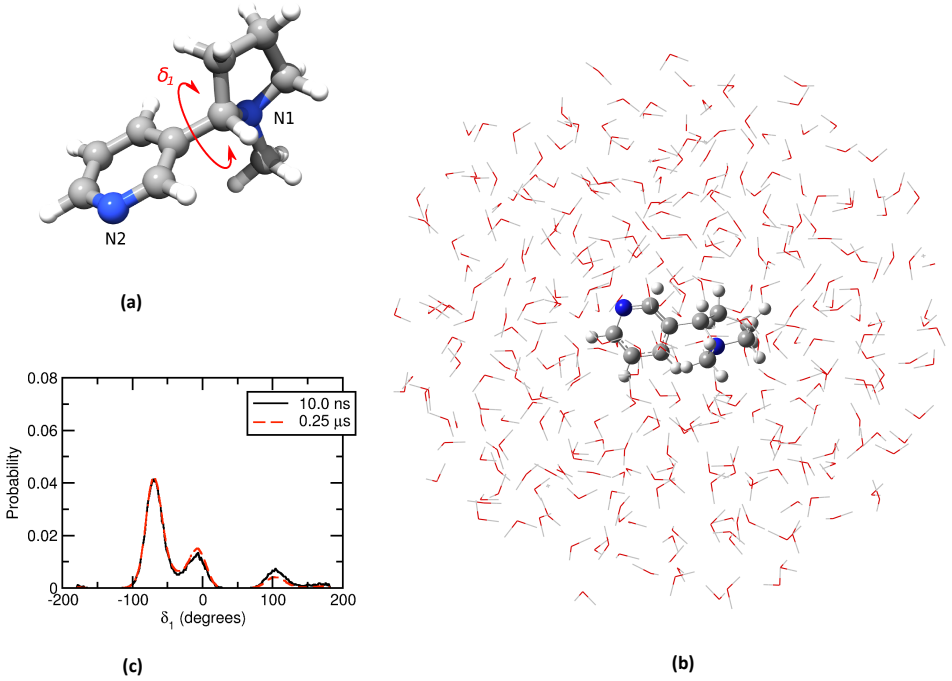


Figure 5.16. (a) Nicotine structure and definition of the δ_1 dihedral angl, defining the conformers; (b) a random snapshot selected from the MD simulation of nicotine in aqueous solution; c) conformational analysis extracted from MD.[148]

The analysis of the MD trajectory[148] shows that nicotine exhibits 3 different conformers in aqueous solution: the A conformer, having an average value of $\delta_1 = 106.4$ degrees, the B conformer ($\delta_1 = -65$ degrees) and the 0 conformer, where $\delta_1 = 0$ degrees. Figure 5.16, panel (c) shows the distribution of each conformer obtained from the MD simulation.[318] The most populated conformers belong to the B family, followed by the 0 and A families. Also, the analysis of the MD trajectory[148] shows that at least two water molecules are bound to nicotine nitrogen atoms through hydrogen-bonding interactions.

The distribution in panel (c) of Figure 5.16 is maintained in the 300 snapshots selected in this study, for which E_{rep}, E_{dis} and consequently E_{Qnel} were calculated. The resulting values of such energies are reported in the Figure 5.17 as a function of the snapshot. On the basis of the values depicted in Figure 5.13, E_{rep} was calculated by imposing the R= 5 Å, which in the present case implies on average 52 water molecules to be considered in the evaluation of this term.

Figure 5.17 clearly shows that, as previously reported by some of the present authors for other molecular properties,[56,146,149] E_{rep}, E_{dis} and E_{Qnel} may differ as a function of the snapshot. E_{Qnel} is always positive, showing that the attractive, negative, E_{dis} terms is always smaller than the repulsive, positive E_{rep} contribution. The average

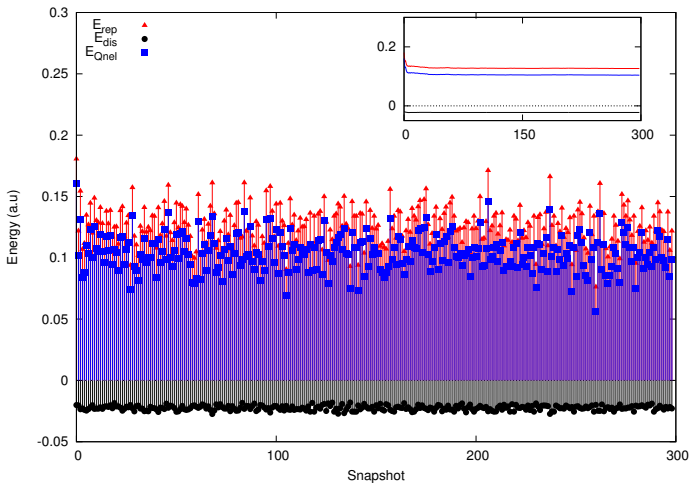


Figure 5.17. Calculated E_{rep}, E_{dis} and E_{Qnel} as a function of the snapshot for nicotine in aqueous solution. The trends in average values as a function of the snapshot are given as inset.

value of the three terms as a function of the snapshot is depicted as inset in Figure 5.17. Clearly, convergence is reached very quickly, when only 80 snapshots are considered, that different from other investigated properties of this and other molecular systems, requiring hundreds or thousands of snapshots to get a fully converged value for molecular spectral properties.[56,146,149]

To refine the analysis, the 300 snapshots were assigned to the three different conformers, that in order to dissect the role of the different E_{rep} and E_{dis} terms in each subclass of structures. To this end, a snapshot was considered to belong to the A family of conformers if $70 < \delta_1 < 180$, B conformers if $-180 < \delta_1 < -40$ and 0 conformers otherwise. This partitioning allows the snapshots to be divided in subclasses of structures and the contributions for each class to be calculated. The results of this analysis are summarized in Table 5.6, where also standard deviations are reported.

Conformer	E_{rep}	E_{dis}	E_{Qnel}	ΔE_{Qnel}
A	12.8 (±1.5)	-2.2 (±0.1)	10.6(±1.5)	0.4 (2.4 kcal/mol)
B	12.6 (±1.7)	-2.2 (±0.2)	10.4(±1.6)	0.2 (1.2 kcal/mol)
0	12.4 (±1.1)	-2.2 (±0.1)	10.2(±1.1)	0.0

Table 5.6. Calculated E_{rep}, E_{dis} and E_{Qnel} for nicotine in aqueous solution. ΔE_{Qnel} is the non-electrostatic energy difference between the various conformers and the most stabilized one (0). All data are given in 10^{-2} Hartree unless differently stated and refer to 300 selected snapshots. Standard deviations are reported in parentheses.

Table 5.6 clearly shows that the calculated E_{Qnel} values are similar for the three conformers, and they do not differ statistically. In particular, the lowest E_{Qnel} is exhibited

by conformer 0, due to the fact that this conformer shows the lowest value of E_{rep} in combination with a rather small absolute value of E_{dis}. As it has been noticed by some of us in previous works,[148] the analysis of the MD shows that the 0 conformer is characterized by a short lifetime (10 ps), and this is due to the weak hydrogen bonding pattern exhibited by this conformer with respect to the other two. Therefore, the nearest water molecules are placed farther than in the two other cases, thus giving rise to lower E_{rep} values.
The total non-electrostatic energy difference of the various conformers with respect to 0 is reported in the last column of Table 5.6. The values are small, but not negligible, and in particular their magnitude is such to potentially affect the predicted conformational weights in aqueous solutions.

$\mu_{\text{QM/TIP3P}}$	$\mu_{\text{QM/FQ}}$	$\mu_{\text{QM/FQ + Qnel}}$
3.1	5.9	4.2

Table 5.7. Calculated dipole moment of nicotine in aqueous solution, obtained by exploiting non polarizable (QM/TIP3P), polarizable (QM/FQ) and polarizable+non electrostatic (QM/FQ + Qnel) approaches. All data are given in Debye and refer to 300 selected snapshots.

In Table 5.7, the average dipole moments calculated by exploiting the three QM/TIP3P, QM/FQ and QM/FQ + Quantum non-electrostatic interactions approaches are reported. The difference between the dipole obtained by including polarization effects is huge, as expected (an increase of about 50% is observed). The third column in Table 5.7 shows the dipole moment calculated by using the QM/FQ approach coupled with our description of repulsive and dispersive interactions. By referring again to Table 5.6 and Figure 5.17, we note again that the repulsion term is dominant if compared to the dispersion. This results in a confinement of the molecular density, which causes the decrease of the molecular dipole, as expected.

5.7 Summary, Conclusions and Future Perspectives

In this paper, a general route to calculate quantum repulsion and quantum dispersion effects in polarizable and non-polarizable QM/MM approaches has been formulated. A remarkable peculiarity of the proposed approach is that repulsion/dispersion contributions are explicitly introduced in the QM Hamiltonian. Therefore, such terms not only enter the evaluation of the energetic properties of the systems but, remarkably, can be propagated to the calculation of molecular properties and spectra. Due to the specific form of the contributions, a reliable yet extensive application of the methodology requires a compulsory parametrization for different MM substrates, however the number of parameters entering the definition of our method is remarkably low. In this paper, a parametrization for the aqueous solution, which is the natural environment for most biomolecules, is proposed. Such a parametrization is able to reproduce the most important features of the aqueous solution, for which the reported data are in good agreement with reference data. The application of the obtained parametrization to the

calculation of the non-electrostatic interaction energy of aqueous nicotine shows that the Pauli repulsion contribution is larger than the dispersion term for all the representative snapshots extracted from the MD. This feature can potentially impact QM/MM geometry optimization of molecular systems in aqueous solution, which are currently performed by only resorting to the electrostatic term.[55] The results of this study pave the way to similar studies, aimed at extending the parametrization to environments other that water, in which non-electrostatic terms can compare (or even overcome) with the generally dominating electrostatic component of the molecule-environment interaction. Such studies, and the related parametrizations, together with the extension of our method to the calculation of molecular properties and spectroscopies, will be the topic of future communications.

5.8 Acknowledgment

The authors thank Giovanni Scalmani for useful discussion on implementation issues.

5.9 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.7b00776. Obara-Saika algorithm for the evaluation of A_μ in Eq. 5.9. Results of the parametrization of Eq. 5.29. Quantum dispersion parameters. Raw data for the dependence of E_{rep}, E_{dis} and E_{Qnel} on the QM description. Detailed comparison between our calculated values, KM-EDA and EFP2 approach for MLAT and MOXY in aqueous solution. Details on the MD protocol used for Nicotine in aqueous solution.

5.10 Appendix

Energy First Derivatives

Energy first derivatives (within the chain rule):

$$\mathcal{E}^x = \frac{\partial \mathcal{E}}{\partial x} + \frac{\partial \mathcal{E}}{\partial \mathbf{P}} \frac{\partial \mathbf{P}}{\partial x} \tag{5.31}$$

The second term is calculated by considering the idempotency condition, which gives rise to the usual energy-weighted density matrix contribution:

$$-\mathbf{PFPS}^x_{oo} = -\mathbf{WS}^x_{oo} \tag{5.32}$$

The first term is composed of three contributions: one arising from the QM, one from the electrostatic interaction between QM and MM portions, and the last one from the non-electrostatic interactions.
The first two elements have already been published and they are:[55]

$$\frac{\partial \mathcal{E}}{\partial x} = \mathrm{tr}\, \mathbf{h}^x \mathbf{P} + \frac{1}{2} \mathrm{tr}\, \mathbf{P} \mathbf{G}^{(x)}(\mathbf{P}) + \mathbf{q}^\dagger \mathbf{V}^{(x)}(\mathbf{P}) \tag{5.33}$$

Repulsion Term

Now, let us consider first the repulsion contribution (see Eq. 5.11):

$$E^x_{rep} = \frac{1}{2}\sum_{\mu\nu} P_{\mu\nu}\left[\int \mathrm{d}\mathbf{r_1}\ \left(\frac{\chi^x_\mu(\mathbf{r_1})A_\nu(\mathbf{r_1}) + A_\mu(\mathbf{r_1})\chi^x_\nu(\mathbf{r_1})}{2}\right) + \right.$$
$$\left. + \int \mathrm{d}\mathbf{r_1}\ \left(\frac{\chi_\mu(\mathbf{r_1})A^x_\nu(\mathbf{r_1}) + A^x_\mu(\mathbf{r_1})\chi_\nu(\mathbf{r_1})}{2}\right)\right] \qquad \mathbf{5.34}$$

where χ^x_i are the derivatives with respect to the QM coordinates of the i-th atomic orbital. A^x_i can be viewed instead as the electric field at point $\mathbf{r_1}$ due to the product of a basis function centered at $\mathbf{R}_\mu$ and the sum of the Gaussian functions representing ρ_{MM}, centered at $\mathbf{R}$. $A^x_\mu(\mathbf{r_1})$ can be calculated through:

$$A^x_\mu(\mathbf{r_1}) = \int \mathrm{d}\mathbf{u}\ \frac{\chi^x_\mu(\mathbf{r_1}+\mathbf{u})}{u}\left(\sum_{\mathbf{R}} \xi^2_{\mathbf{R}} e^{-\beta_{\mathbf{R}}|\mathbf{r_1}-\mathbf{R}|^2} e^{-\beta_{\mathbf{R}}|\mathbf{r_1}+\mathbf{u}-\mathbf{R}|^2}\right) \qquad \mathbf{5.35}$$

where the perturbation is only acting on the QM portion.

Dispersion Term

Let us consider the dispersion contribution to the interaction energy:

$$E_{dis} = -\frac{1}{2}\sum_{A\in\mathrm{QM}}\sum_{B\in\mathrm{MM}} f_{damp}(R_{AB}, R^0_A, R^0_B)\left(\frac{\eta^2_A C^{free}_{AA} C^{eff}_{6BB}}{\frac{\alpha^0_B}{\alpha^0_A}\eta^2_A C^{free}_{AA} + \frac{\alpha^0_A}{\alpha^0_B} C^{eff}_{6BB}}\right) R^{-6}_{AB}$$
$$= -\frac{1}{2}\sum_{A\in\mathrm{QM}}\sum_{B\in\mathrm{MM}} f_1 \cdot f_2 \cdot f_3 \qquad \mathbf{5.36}$$

Thus, energy first derivative reads:

$$E_{dis} = -\frac{1}{2}\sum_{A\in\mathrm{QM}}\sum_{B\in\mathrm{MM}} f^x_1 \cdot f_2 \cdot f_3 + f_1 \cdot f^x_2 \cdot f_3 + f_1 \cdot f_2 \cdot f^x_3 \qquad \mathbf{5.37}$$

First, let us consider f^x_1, which is the derivative of the damping function f_{damp}. This can be easily obtained from R_A:

$$f^x_1 = \frac{d(\mathbf{R}_A - \mathbf{R}_B)}{s_R \cdot R^0_{AB}\,|R_A - R_B|} \cdot \exp\left[-d\left(\frac{R_{AB}}{s_R R^0_{AB}} - 1\right)\right] \cdot (f_1)^2 \qquad \mathbf{5.38}$$

The second term f^x_2 is:

$$f^x_2 = \frac{\partial C^{eff}_{6AB}}{\partial \mathbf{R}_A} = \frac{2\frac{\alpha^0_A}{\alpha^0_B} C^2_{6BB}}{\left(\frac{\alpha^0_B}{\alpha^0_A} C^{eff}_{6AA} + \frac{\alpha^0_A}{\alpha^0_B} C_{6BB}\right)^2} \frac{\partial C^{eff}_{6AA}}{\partial \mathbf{R}_A} \qquad \mathbf{5.39}$$

$$\frac{\partial C^{eff}_{6AA}}{\partial \mathbf{R}_A} = C^{free}_{6AA} 2\eta_A \frac{\partial \eta_A}{\partial \mathbf{R}_A} \qquad \mathbf{5.40}$$

$$\frac{\partial \eta_A}{\partial \mathbf{R}_A} = \eta_A^x = \frac{\int \mathrm{d}\mathbf{r}\ r^3 w_A(\mathbf{r}) \left[\chi_\mu^x(\mathbf{r})\chi_\nu(\mathbf{r}) + \chi_\mu(\mathbf{r})\chi_\nu^x(\mathbf{r})\right]}{\int \mathrm{d}\mathbf{r}\ r^3 \rho_A^{free}(\mathbf{r})} + \\ - \frac{\int \mathrm{d}\mathbf{r}\ 3(\mathbf{r}-\mathbf{R}_A)\left|\mathbf{r}-\mathbf{R}_A\right| w_A(\mathbf{r})\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})}{\int \mathrm{d}\mathbf{r}\ r^3 \rho_A^{free}(\mathbf{r})} \qquad \textbf{5.41}$$

where the eff superscript in the C_{6BB} term is omitted for the sake of readability of the equations. In Eq. 5.41 the term due to the partial derivative of the density with respect the density matrix is accounted for.
The third term is the derivative of the R_{AB}^{-6} :

$$f_3^x = -6\frac{\mathbf{R}_{AB}}{|R_{AB}|^8} \qquad \textbf{5.42}$$

Chapter 6

Effective yet Reliable Computation of EPR Spectra in Solution by a QM/MM Approach: Interplay Between Electrostatics and Non-electrostatic Effects

Tommaso Giovannini*,†, Piero Lafiosca†, Balasubramanian Chandramouli†‡, Vincenzo Barone†, Chiara Cappelli*,†

†*Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy.*
‡*Compunet, Istituto Italiano di Tecnologia (IIT), Via Morego 30, 16163 Genova, Italy.*

E-mail: tommaso.giovannini@sns.it; chiara.cappelli@sns.it

Abstract In this paper, we have extended to the calculation of hyperfine coupling constants, the model recently proposed by some of the present authors (*J. Chem. Theory Comput.*, **2017**, *13*, 4854–4870) to include Pauli repulsion and dispersion effects in QM/MM approaches. The peculiarity of the proposed approach stands in the fact that repulsion/dispersion contributions are explicitly introduced in the QM Hamiltonian. Therefore, such terms not only enter the evaluation of energetic properties but, remarkably, propagate to molecular properties and spectra. The account of such contributions has permitted a quantitative analysis of QM/MM interaction energies, and this has also required a novel parametrization of the Fluctuating Charges force field, which has been then tested against the prediction of EPR parameters of prototypical nitroxide radicals in aqueous solutions.

6.1 Introduction

In the last decades multiscale models have been widely used for the study of molecular properties and spectra.[6,12,56,57,82,120,123,168,273,502,503] In this context, the most successful approaches fall within the class of "focused models", which aim at accurately modeling both the physico-chemical properties of the target and its interactions with the surrounding environment. The effect of the latter is seen as a perturbation on the target molecule, and is treated at a lower computational level of theory, e.g. by resorting to classical physics, whereas the target molecule is described accurately, generally at the Quantum Mechanical (QM) level. Due to such a partitioning, the computational cost of a QM/classical computation is comparable to that of the corresponding QM isolated system. Such a feature has strongly contributed to the increasing popularity of these models.

QM/Molecular Mechanics (MM) models are among the most renowned classes of QM/-classical approaches,[1,12,13,23,278,428,429,437] which have been formalized within different physical frameworks. Beyond the basic mechanical QM/MM embedding, in the last years much effort has been spent to define electrostatic QM/MM embedding approaches, in which a set of fixed charges is placed on the MM moiety (generally on MM atoms) and the interaction between QM and MM portions is modeled by resorting to the Coulomb law. Clearly, in such approaches the QM and MM moieties do not mutually polarize. 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, which can be based on distributed multipoles,[58,154,185,186,281] induced dipoles,[10,155,164] Drude oscillators[59] or Fluctuating Charges (FQ).[56,156,158]

The description of the molecular properties/spectra of embedded systems which is obtained by resorting to polarizable embedding is generally quite accurate.[10,58,146,150,155,183] However, such models are deeply based on the assumption that electrostatic energy terms dominate the target/environment interactions. Non-electrostatic (Pauli Repulsion and Dispersion) contributions between the QM and MM portions are roughly modeled by using parametrized functions, e. g. the Lennard-Jones potential,[195,196] which are however completely independent of the QM density. As a result, they are not taken into account in the QM operators, so that the calculated spectroscopic/response properties are not affected by such interactions. The reasons why such contributions are generally discarded are connected to the presumption of a numerically dominating effect of electrostatic terms. However, non-electrostatic contributions are crucial to get a physically consistent description of any embedded system, also in the case of target/environment interactions dominated by electrostatics.[504,505]

A way to include non-electrostatic energy terms is to resort to the Effective Fragment Potential (EFP).[185–190] The high accuracy of this method is essentially due to the explicit QM calculation of the molecular orbitals of the environment, drifting apart from the concept at the basis of MM Force Field (FF). A similar QM-based approach, namely the Polarizable Density Embedding (PDE), has been recently proposed to only include repulsion effects.[60,62]

A substantially different way of including non-electrostatic interactions in QM/MM approaches consists of exploiting a model recently developed by some of the present

authors,[197] which formulates repulsion as a function of an auxiliary density on the MM portion and extends the Tkatchenko-Scheffler (TS) approach to DFT[198–202] to treat QM/MM dispersion terms. Notice that the formulation of repulsion contributions is terms of gaussian functions placed in the MM region has also been proposed in the so-called Gaussian Electrostatic Model (GEM).[505–508] However, in both the aforementioned PDE and GEM models, repulsion interaction is modeled as a overlap one-electron integral. Our approach instead defines repulsion contributions in terms of a two-electron exchange integral, thus physically representing the Pauli repulsion. Moreover, differently from the stand-alone approaches discussed above (EFP, PDE, GEM), our approach can be easily coupled to any kind of QM/MM approach, because repulsion and dispersion are formulated in a way which is totally independent of the choice of the FF to model the electrostatics (i.e. fixed-charges or polarizable embedding). Remarkably, in our model repulsion and dispersion contributions are indeed dependent on the QM density. Thus, an explicit contribution to the QM Fock operator exists and the resulting calculated QM properties/spectra are modified by such interactions.

Our model for non-electrostatics in QM/MM has been so far only challenged on reproducing full QM non-electrostatic interaction energies, for which very good results have been obtained.[197] In this paper we start with the extension of of the model to spectroscopy. To this end we report the formulation of non-electrostatic QM/MM terms for EPR, for which environmental effects substantially contribute to the overall observable.[509–511] Environmental (solvent) effects on EPR are usually described by means of continuum models,[512–515] and only in few cases by adopting electrostatic QM/MM embedding coupled with a classical Molecular Dynamics (MD) to take into account the fluctuations of both the solute conformations and the solvent molecules.[11,313,314,431,516–518]

Nitroxide radicals are among the most thoroughly studied radicals from both experimental and computational points of view due to their remarkable stability coupled to strong sensitivity to the polarity of the surrounding and to the piramidality of the nitrogen atom. Given their importance, several nitroxide radicals have been synthesized to be either used as spin probes (when dispersed in an environment) or as spin labels (when chemically attached to a biological molecule, e.g. a protein).[519–521] High-field EPR spectroscopy provides quite rich information consisting essentially of the nitrogen hyperfine and gyromagnetic tensors.[519] However, interpretation of these experiments in structural terms strongly benefits from quantum mechanical calculations able to dissect the overall observables in terms of the interplay of several subtle effects.[313,314,522–528] This situation has prompted us to perform a comprehensive study of prototypical nitroxide radicals in aqueous solution coupling density functional and coupled cluster quantum mechanical computations to molecular dynamics simulations and average of properties for a sufficient number of snapshots including electrostatic, induction, repulsion and dispersion interactions with the surrounding evaluated by effective quantum mechanical approximations.

To the best of our knowledge, this work presents the first formulation and application of a QM/MM approach accounting at the same time for polarization and non-electrostatic interactions on EPR Hyperfine Coupling Constant (hcc).

The paper is organized as follows: firstly, the theoretical model is presented. Then, the computational approach is applied to the calculation of hcc_N of two nitroxyl radicals (PROXYL and TEMPO) in aqueous solution. Such compounds are characterized by the presence of the N–O group, which has been most widely used as "spin probe" and "spin label" for the study of structure and dynamics of macromolecular systems.[519–521] Summary and Conclusions end the manuscript.

6.2 Theoretical Model

The total energy of a system composed by two interacting moieties, one described at the QM level and the other at the MM level can be expressed as:[192,442]

$$E_{QM/MM} = E^{ele}_{QM/MM} + E^{pol}_{QM/MM} + E^{ex-rep}_{QM/MM} + E^{dis}_{QM/MM} \quad \textbf{6.1}$$

where, $E^{ele}_{QM/MM}$ accounts for electrostatic interactions and $E^{pol}_{QM/MM}$ is the polarization contribution. Such energy terms are those modeled in the electrostatic embedding approach, and in particular in polarizable QM/MM methods.[1,10,12,59,155,159,164] $E^{ex-rep}_{QM/MM}$ is the exchange-repulsion contribution and $E^{dis}_{QM/MM}$ arises from dispersion interactions.

In this work electrostatic and the polarization terms are modelled by exploiting the Fluctuating Charge (FQ) force field,[55,56,149–151,153,159] whereas non-electrostatic interactions (i.e. the sum of $E^{ex-rep}_{QM/MM}$ and $E^{dis}_{QM/MM}$) are modeled by using the model described in Ref.[197] In the next paragraphs, the mathematical formulation of the different energy contributions are discussed.

6.2.1 Electrostatic interactions

In order to model electrostatic and polarization terms (see Eq. 6.1), a polarizable QM/MM embedding needs to be adopted. In such a model, the MM force field adapts to the external field/potential originating from the QM density and electrostatic/polarization terms are included in the QM Hamiltonian, so to describe the mutual interaction between the QM density and the environment.

In this work we will resort to the FQ force field.[56] In the resulting QM/FQ model, the electrostatic potential due to the QM density together with the differences in electronegativities between different atoms in the MM region, give rise to a charge fluctuation in the MM region, up to the point that the differences in electrochemical potential between the MM atoms vanish. From a mathematical point of view, this results in the following linear equation:[54]

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q - \mathbf{V}(\mathbf{P}_{\mathrm{QM}}) \quad \textbf{6.2}$$

where $\mathbf{D}$ is a response matrix, whose diagonal terms are atomic chemical hardnesses, $\mathbf{q}$ is a vector containing the FQs and Lagrangian multipliers. $\mathbf{C}$ contains atomic electronegativies and those constraints which are needed to ensure each MM molecule to

have a fixed charge. $\mathbf{V}(\mathbf{P})$ is the potential due to the QM density matrix $\mathbf{P}$ calculated at MM charges positions. We refer the reader to Ref. [152] for further details.
The interaction between FQ charges and the QM density obeys the Coulomb law:

$$E^{ele}_{\mathrm{QM/MM}} + E^{pol}_{\mathrm{QM/MM}} = \sum_{j=1}^{N_{\mathrm{FQs}}} \int_{\mathbf{R}^3} \frac{\rho_{\mathrm{QM}}(\mathbf{r}) q_j}{|\mathbf{r} - \mathbf{r}_j|} \, \mathrm{d}\mathbf{r} \tag{6.3}$$

By deriving Eq. 6.3 with respect to the density matrix, $P_{\mu\nu}$, the contribution to the Fock matrix is obtained:[54])

$$F_{\mu\nu} = \frac{\partial E}{\partial P_{\mu\nu}} = \mathbf{V}^{\dagger}_{\mu\nu}\mathbf{q} \tag{6.4}$$

The Fock matrix defined in this way can enter a SCF procedure, so as to finally give a QM density mutually equilibrated with the FQs.

6.2.2 Pauli Repulsion Energy

The Exchange-Repulsion energy, $E^{ex-rep}_{QM/MM}$, also known as Pauli Repulsion energy, is formally due the Pauli principle, i.e. wavefunction antisimmetry. From a mathematical point of view, it can can be formulated as the opposite of an exchange integral: [86,192]

$$E^{ex-rep}_{QM/MM} = \frac{1}{2} \int \frac{\mathrm{d}\mathbf{r_1} \, \mathrm{d}\mathbf{r_2}}{r_{12}} \rho_{QM}(\mathbf{r_1}, \mathbf{r_2}) \rho_{MM}(\mathbf{r_2}, \mathbf{r_1}) \tag{6.5}$$

In order to define the density matrix ρ_{MM} we localize fictitious valence electron pairs for MM molecules in bond and lone pair regions and represent them by s-gaussian-type functions. The expression for ρ_{MM} becomes:

$$\rho_{\mathrm{MM}}(\mathbf{r_1}, \mathbf{r_2}) = \sum_{\mathbf{R}} \xi^2_{\mathbf{R}} e^{-\beta_{\mathbf{R}}(\mathbf{r_1} - \mathbf{R})^2} \cdot e^{-\beta_{\mathbf{R}}(\mathbf{r_2} - \mathbf{R})^2} \tag{6.6}$$

where, $\mathbf{R}$ collects the centers of the gaussian functions used to represent the fictitious MM electrons. The β and ξ parameters are generally different for lone-pairs or bond-pairs, their values being adjusted to the specific kind of environment (MM portion) to be modeled. By substituting Eq. 6.6 in Eq. 6.5, the QM/MM repulsion energy reads:

$$E^{rep}_{\mathrm{QM/MM}} = \frac{1}{2} \sum_{\mathbf{R}} \int \frac{\mathrm{d}\mathbf{r_1} \, \mathrm{d}\mathbf{r_2}}{r_{12}} \rho_{\mathrm{QM}}(\mathbf{r_1}, \mathbf{r_2}) \left[\xi^2_{\mathbf{R}} e^{-\beta_{\mathbf{R}}(\mathbf{r_1} - \mathbf{R})^2} \cdot e^{-\beta_{\mathbf{R}}(\mathbf{r_2} - \mathbf{R})^2} \right] \tag{6.7}$$

It is worth noticing that in this formalism, QM/MM Pauli Repulsion energy is calculated as a two-electron integral. Eq. 6.7 is general enough to hold for any kind of MM environment (solvents, proteins, surfaces ecc.). The nature of the external environments is specified by defining the number of different electron-pair types and the

corresponding β and ξ parameters in Eq. 6.6. Also, the formalism is general, so that it can be coupled to any kind of QM/MM approach.
By deriving Eq. 6.7 with respect to the density matrix, the corresponding contribution to the Fock matrix is obtained:

$$F_{\mu\nu}^{rep} = \frac{\partial E^{rep}}{\partial P_{\mu\nu}} = \frac{1}{2}\int \mathrm{d}\mathbf{r_1} \left[\frac{\chi_\mu(\mathbf{r_1})A_\nu(\mathbf{r_1}) + A_\mu(\mathbf{r_1})\chi_\nu(\mathbf{r_1})}{2}\right] \tag{6.8}$$

where χ_μ are atomic basis functions and A_μ are calculated as detailed in Ref.[197]

6.2.3 Quantum Dispersion Energy

To formulate dispersion interactions we start from the Tkatchenko and Scheffler (TS) DFT functional. In this model, the dispersion energy can be written as:

$$E^{dis} = -\frac{1}{2}\sum_{A,B} f_{damp}(R_{AB}, R_A^0, R_B^0) C_{6AB} R_{AB}^{-6} \tag{6.9}$$

where, R_{AB} is the distance between atoms A and B in a given system, C_{6AB} is the corresponding C_6 coefficient, R_A^0 and R_B^0 are their van der Waals (vdW) radii. The R_{AB}^{-6} singularity at small distances is eliminated by the short-range damping function $f_{damp}(R_{AB}, R_A^0, R_B^0)$.[198]
C_{6AB} coefficients can be expressed in terms of homonuclear parameters C_{6AA}, C_{6BB}, which in turn can be obtained through an Hirshfeld[468] partition of the density.[198]
Such an approach can be reformulated within a QM/MM formalism,[197,202] yielding:

$$E_{\mathrm{QM/MM}}^{dis} = -\frac{1}{2}\sum_{A\in\mathrm{QM}}\sum_{B\in\mathrm{MM}} f_{damp}(R_{AB}, R_A^0, R_B^0)\frac{\eta_A^2 C_{AA}^{free} C_{6BB}^{eff}}{\frac{\alpha_B^0}{\alpha_A^0}\eta_A^2 C_{AA}^{free} + \frac{\alpha_A^0}{\alpha_B^0} C_{6BB}^{eff}} R_{AB}^{-6} \tag{6.10}$$

where C_{6BB}^{eff} are effective homonuclear coefficients of B (MM) atoms and C_{6AA}^{free} are free homonuclear coefficients of A QM atoms. α_A^0 and α_B^0 are static dipole polarizabilities, whereas η_A is a function converting C_{6AA}^{free} into C_{6AA}^{eff}. Further details can be found in Refs.[197,198,202]
$f_{damp}(R_{AB}, R_A^0, R_B^0)$ in Eq. 6.10 is a Fermi-type damping function::[194,198,466]

$$f_{damp}(R_{AB}, R_A^0, R_B^0) = \frac{1}{1 + \exp\left[-d\left(\frac{R_{AB}}{s_R R_{AB}^0} - 1\right)\right]} \tag{6.11}$$

where, $R_{AB}^0 = R_A^0 + R_B^0$, and d, s_R are free parameters.
Similarly to what already done for electrostatic and repulsion contributions, by deriving Eq. 6.10 with respect to the QM density matrix the dispersion contribution to the Fock matrix is obtained:[197]

$$F_{\mu\nu}^{dis} = -\frac{1}{2}\sum_{A\in \mathrm{QM}}\sum_{B\in \mathrm{MM}} f_{damp}(R_{AB})\frac{2\frac{\alpha_A^0}{\alpha_B^0}C_{6BB}^2 C_{6AA}^{free} 2\eta_A}{\left(\frac{\alpha_B^0}{\alpha_A^0}C_{6AA}^{eff}+\frac{\alpha_A^0}{\alpha_B^0}C_{6BB}\right)^2}\eta_{A,\mu\nu}^{\rho} R_{AB}^{-6} \qquad \textbf{6.12}$$

The complete derivation and definition of $\eta_{A,\mu\nu}^{\rho}$ can be found in Ref.[197]

6.2.4 Hyperfine Coupling Constant

The spin Hamiltonian describing the interaction between the electron spin (S) of a free radical containing a magnetic nucleus of spin I and an external magnetic field (B) can be written as:

$$H_S = \mu_B \vec{S}\cdot\mathbf{g}\cdot\vec{B} + \frac{1}{\hbar\gamma_I}\vec{S}\cdot\mathbf{A}\cdot\vec{\mu}_I \qquad \textbf{6.13}$$

where the first term is the Zeeman interaction between the electron spin and the external magnetic field through the Bohr magneton μ_B and $\mathbf{g} = g_e\mathbf{1}_3 + \Delta g_{corr}$. Δg_{corr} accounts for the correction to the free electron value ($g_e = 2.0022319$) due to several terms including the relativistic mass ($\Delta\mathbf{g}_{RM}$), the gauge first-order corrections ($\Delta\mathbf{g}_C$) and a term arising from the coupling of the orbital Zeeman (OZ) and the spin–orbit coupling (SOC) operator.[529,530] The second term on the rhs of Eq. 6.13 describes the hyperfine interaction between **S** and the nuclear spin **I** through the hyperfine coupling tensor **A**. The latter, which is defined for each nucleus X, can be decomposed into two terms:

$$\mathbf{A}(X) = \mathrm{A_X}\mathbf{1}_3 + \mathbf{A}_{\mathrm{dip}}(\mathrm{X}) \qquad \textbf{6.14}$$

The dipolar term $\mathbf{A}_{\mathrm{dip}}(\mathrm{X})$ is a zero-trace tensor, whose contribution vanishes in isotropic media (e.g. solutions). The first term $\mathrm{A_X}$ (Fermi-contact interaction), which is an isotropic contribution, is also known as hyperfine coupling constant (hcc). It is related to the spin density (ρ_X) at nucleus X by the following relation:

$$\mathrm{A_X} = \frac{4\pi}{3}\mu_B\mu_X g_e g_X \langle S_Z\rangle^{-1}\rho_X^{\alpha-\beta} \qquad \textbf{6.15}$$

where $\rho_X^{\alpha-\beta}$ can be obtained as:

$$\rho_X^{\alpha-\beta} = \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta}\langle\chi_\mu(\mathbf{r})|\delta(\mathbf{r}-\mathbf{r}_X)|\chi_\nu(\mathbf{r})\rangle \qquad \textbf{6.16}$$

$P^{\alpha-\beta}$ is the difference between α and β density matrices. Because in our approach both electrostatic and non-electrostatic dispersion/repulsion interactions enter the definition of the QM Fock operators (see Eqs. 6.4, 6.8 and 6.12), $P^{\alpha-\beta}$ is modified. Therefore, hyperfine coupling constants with the account of electrostatic, polarization, dispersion and repulsion QM/MM interactions are obtained.

6.3 Computational Details

Molecular geometries of PROXYL and TEMPO radicals (Fig. 6.1) were optimized in vacuo by combining B3LYP and PBE0 hybrid density functionals with both aug-cc-pVDZ and 6-311++G(3df,2pd) basis sets. For all optimized structures the hyperfine coupling constant of Nitrogen atom was calculated by exploiting both B3LYP and PBE0 and the N07D basis set.[531,532] For the sake of comparison, on the reduced structures depicted in Figure 6.1, which are obtained by removing ring atoms for both TEMPO and PROXYL but keeping fixed all the geometrical parameters, additional CCSD/EPR-II calculations were performed.[533]

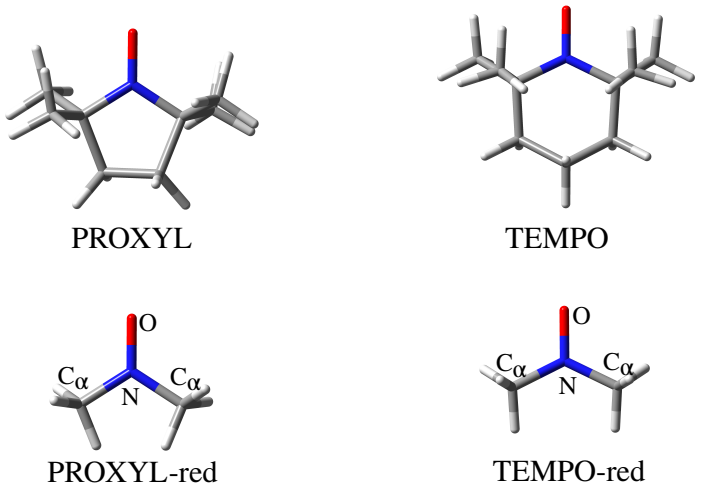


Figure 6.1. Top: PROXYL and TEMPO structures. Bottom: reduced structures used for CCSD/EPR-II calculations.

Clusters made of TEMPO and PROXYL radicals with two explicit water molecules (see Fig. 6.3) were optimized at the PBE0/6-311++G(3df,2pd) level, according to previous studies.[514] For those structures, the interaction energy between the radicals and the two water molecules was computed by exploiting SAPT0/aug-cc-pVTZ, or jun-ccp-pVDZ or N07D (as implemented in Psi4 1.1[479]) and CCSD(T)/aug-cc-pVTZ, jun-cc-pVDZ and N07D. Counter-Poise corrections were included in CCSD(T) calculations. QM/MM energy calculations were also performed at the PBE0/aug-cc-pVTZ, jun-cc-pVDZ and PBE0/N07D level, by including dispersion and repulsion energies obtained by exploiting our model.[197] The QM portion was restricted to the radical, whereas the two water molecules were treated at the MM level. The MM region was described by means of a non-polarizable force field (TIP3P[358]) and the polarizable FQ approach[56,159] by exploiting two literature parametrizations,[156,162] and a new parametrization proposed in this work. The parameters used for modeling dispersion and repulsion interactions were taken from Ref.[197] On the same structures, full QM and QM/MM nitrogen hyperfine coupling constants were calculated by exploiting the PBE0/N07D level of theory for treating the QM portion. For the sake of comparison, on the reduced cluster structures depicted in Figure 6.3, which are obtained by remov-

ing ring atoms for both TEMPO and PROXYL but keeping fixed all the geometrical parameters, additional CCSD/EPR-II hcc_N calculations were performed.[533]
Classical MD simulations were performed with the Amber software (v.12) using the ff99SB force field.[534,535] Parameters for nitroxides were obtained from a previous study by one of the present authors.[314] The nitroxides were embedded in a cubic box of TIP3P water molecules, which extended to 30 Å from the solute surface. The starting systems were equilibrated following a multistep protocol: o) minimization of the whole system for 10000 steps, ii) heating of the system from 103 to 303 K in 100 ps with a mild restraint of 0.5 kcal/mol Å^2 on the solute, iii) equilibration in NPT ensemble at a pressure of 1 bar and 303 K for 100 ps. The production phase was then initiated in NVT ensemble and continued for 10 ns. The simulation conditions involve PBS, a 1 gs time step for numerical integration, using SHAKE for constraining bonds involving hydrogens,[536] a 10 Å cut-off for non-bonded interactions, PME for evaluating the long-range electrostatics,[355] temperature regulation with Langevin coupling using a collision frequency of 1.0 ps^{-1}, snapshots collection in the trajectory at 1 ps interval. A total of 200 uncorrelated snapshots were extracted from the MDs (one snapshot every 50 ps). For each snapshot a 13 Å sphere centered at the solute's geometric center was cut. All hyperfine coupling constants were calculated within the QM/FQ or QM/TIP3P framework at the PBE0/N07D level. The FQ water molecules were modeled both with the SPC FQ parameters,[156] the parametrization proposed by some of the present authors[162] and the parameters proposed in this work. The convergence of the hcc_N values as increasing the number of representative snapshots was checked for both radicals. Dispersion and repulsion contributions to hcc_N were included by exploiting what has been explained in the previous section. All QM/FQ calculations were performed by using a locally modified version of Gaussian 16.[301] Finally, the calculated values were compared with experimental data taken from Refs.[537,538]

6.4 Numerical Results

In this section we will report the results issuing from the application of the developed methodology to the calculation of the nitrogen hyperfine coupling constant (hcc_N) of PROXYL and TEMPO radicals in aqueous solution. In order to evaluate the role of the different terms (electrostatic/polarization/dispersion/repulsion) concurring to overall solvent effect, we will present the results obtained by exploiting a hierarchy of different approaches, starting from a simple cluster model (isolated radical plus two water molecules) to averaging over a set of representative structures extracted form MD runs, with or without the inclusion of polarization/dispersion/repulsion solvent contributions. In addition, to allow a direct comparison with experimental hcc_N, reference values for the isolated radicals are discussed.

6.4.1 hcc_N of isolated radicals

PROXYL and TEMPO geometries (see Figure 6.1) were optimized in vacuo at different levels of theory. In particular, B3LYP and PBE0 functionals in combination with aug-cc-pVDZ (BS1) or 6-311++G(3df,2pd) (BS2) basis sets were employed . Selected

geometrical parameters are reported in Table 6.1. In particular, the N-O distance, the $C_\alpha NC_\alpha$ angle and the improper dihedral angle $C_\alpha NOC_\alpha$ were taken into consideration (see Figure 6.1 for atom labeling). Additional data obtained with B3LYP-D3 and PBE0-D3 functionals[194] can be found in Table 6.7 given as Supporting Information (SI). Geometries were also optimized by exploiting the B2PLYP double hybrid functional combined with the maug-cc-pVTZ-d(H) basis set (BS3), which has been reported to reliably describe molecular geometries.[539] The values reported in Table 6.1 clearly show that B3LYP/aug-cc-pVDZ and B2PLYP/maug-cc-pVTZ-d(H) perform in a similar way. However, all the considered combinations of functional and basis set do not differ much from the best calculated structure of both radicals. It is worth pointing out that the most relevant difference between PROXYL and TEMPO stands in the value of the improper dihedral angle $C_\alpha NOC_\alpha$, which is related to the Nitrogen atom pyramidalization. In fact, the angle is almost zero fro PROXYL and about -21 degrees for TEMPO.

		PBE0		B3LYP		B2PLYP
	Parameter	BS1	BS2	BS1	BS2	BS3
PROXYL	N–O	1.262	1.257	1.274	1.268	1.273
	$\angle C_\alpha NC_\alpha$	115.271	115.167	115.370	115.246	115.211
	$C_\alpha NOC_\alpha$	±0.026	±0.025	±0.015	±0.014	±0.014
TEMPO	N–O	1.271	1.266	1.283	1.278	1.282
	$\angle C_\alpha NC_\alpha$	124.141	124.241	124.452	124.584	124.381
	$C_\alpha NOC_\alpha$	±21.935	±21.316	±21.199	±20.632	±21.581

Table 6.1. Selected geometrical parameters of PROXYL and TEMPO radicals at the different levels of theory. BS1: aug-cc-pVDZ; BS2: 6-311++G(3dp,2pd); BS3: maug-cc-pVTZ-d(H).

For all the optimized structures obtained with PBE0 and B3LYP functionals in conjunction with BS1 and BS2, hcc_N were calculated by exploiting either PBE0 or B3LYP and the N07D basis sets purposely parametrized for both functionals (see Refs.[531,532] for more details). For the sake of comparison, additional hcc_N calculations were performed at the CCSD/EPR-II[533] level on the reduced structures depicted at the bottom of Figure 6.1. All results are reported in Table 6.2.

hcc_N for the two radicals differ of about 3 Gauss at all levels. Such differences are essentially due to the different pyramidalization of the nitroxyl group. The small discrepancies which are reported for the various optimized structures are due to small fluctuations in the improper dihedral angle (see Table 6.1). Notice that all calculated DFT hcc_N are underestimated with respect to CCSD/EPR-II values.

To further investigate on the role of nitrogen pyramidalization on hcc_N, PBE0/N07D hcc_N values for the reduced TEMPO structure as a function of $C_\alpha NOC_\alpha$ were calculated. The data are graphically reported in Fig. 6.2.

As it can be noticed, the value computed for PROXYL and TEMPO radicals are almost recovered at zero and ± 20 degrees, respectively. For larger $C_\alpha NOC_\alpha$ values, computed hcc_N values increase up to the maximum value (22 Gauss) at about ± 40 degrees. Such

Radical	Optimized structure	PBE0/N07D	B3LYP/N07D	CCSD/EPR-II
PROXYL	B3LYP/BS1	11.8	11.4	12.7
	B3LYP/BS2	12.0	11.3	12.6
	PBE0/BS1	11.8	11.1	12.4
	PBE0/BS2	11.7	11.0	12.3
TEMPO	B3LYP/BS1	15.0	14.4	15.9
	B3LYP/BS2	14.8	14.2	15.7
	PBE0/BS1	14.9	14.3	15.7
	PBE0/BS2	14.7	14.0	15.4

Table 6.2. Calculated hcc_N values (Gauss). BS1: aug-cc-pVDZ; BS2: 6-311++G(3dp,2pd).

a trend confirms what has already been reported by one of the present authors.[313]

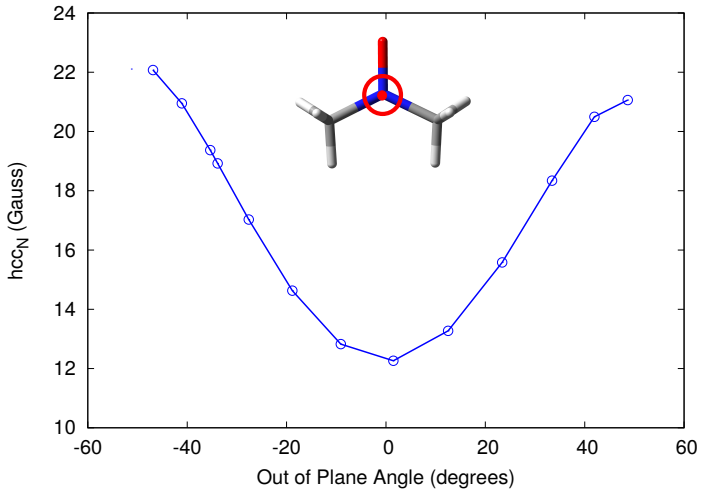


Figure 6.2. PBE0/N07D hcc_N values (Gauss) on the reduced TEMPO structure as a function of the out of plane $C_\alpha NOC_\alpha$ angle.

6.4.2 hcc_N of PROXYL/TEMPO+water clusters

The most basic method to describe hydrated radicals is to resort to a cluster approach. In particular, due to the presence of the oxygen atom, a natural choice consists of saturating oxygen doublets with two water molecules (see Figure 6.3).[313,514] According to what has already been proposed in previous studies, all structures were optimized at the PBE0/6-311++G(3df,2pd) level.[314,514]

To quantify the different contributions to the radical/water interaction energy, Energy Decomposition Analysis (EDA) as formulated in the Symmetry-Adapted Perturbation Theory (SAPT0),[440,441] was performed by exploiting the aug-cc-pVTZ basis set on the reduced structure of PROXYL cluster (see Figure 6.3). Additional SAPT0 calculations were performed by exploiting both the jun-cc-pVDZ or N07D basis sets (see Table 6.8

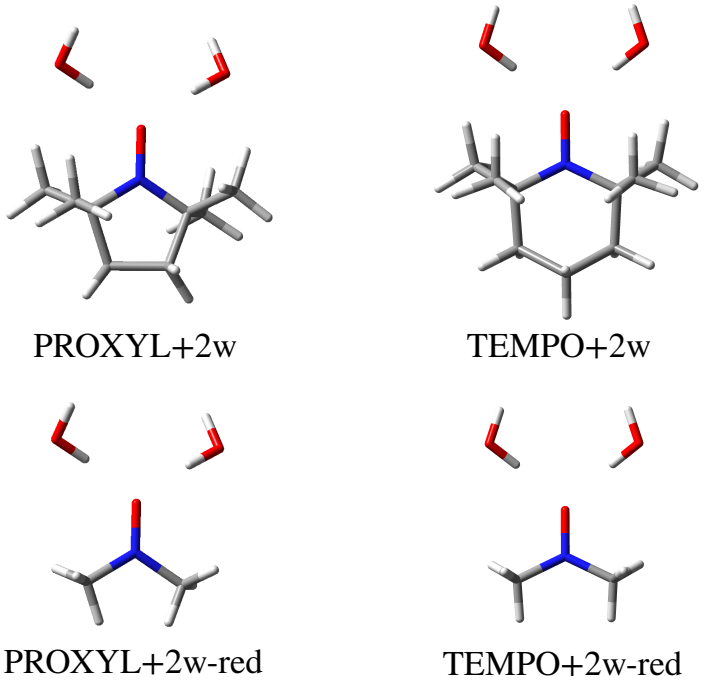


Figure 6.3. PBE0/6-311++G(3df,2pd) optimized structures of clusters of PROXYL (left) and TEMPO (right) with two water molecules.

given as SI). Such additional sets were selected because jun-cc-pVDZ has been reported to provide good results for closed shell systems,[540] whereas N07D is exploited in this study to calculate hcc_N.

SAPT0/aug-cc-pVTZ results are reported in Table 6.3, together with the corresponding values obtained by treating the radical at the QM level (PBE0/aug-cc-pVTZ) and the two water molecules at MM level. QM/MM electrostatic interactions were described by using the FQ approach with three different parametrizations (see Table 6.9 in SI), whereas QM/MM repulsion and dispersion contributions were modeled as reported above. Additional CCSD(T)/aug-cc-pVTZ calculations including Counter-Poise[541] corrections were also performed to quantify the accuracy of SAPT0 interaction energies.

SAPT0 values show that electrostatic interactions (i.e. the sum of electrostatic and induction terms) give larger contributions with respect to non-electrostatic (repulsion+dispersion). However, non-electrostatic interactions and in particular repulsion cannot be neglected, as it is commonly done in standard QM/MM models.

Moving to QM/FQ, we first notice that the available parametrizations (FQ^a and FQ^b in Table 6.3) focus on modeling electrostatic interactions, however they can indeed be inadequate whenever non-electrostatic terms are taken into consideration. This is confirmed by our results (Table 6.3): FQ^a and FQ^b electrostatic energies give a qualitatively correct description of SAPT0 or CCSD(T) total interaction energies. On the contrary, FQ^a and FQ^b total interaction energies are unsatisfactory; therefore, a

novel FQ parametrization is required (labeled FQ^c in Table 6.3). Differently from FQ^a and FQ^b, which were obtained to reproduce water bulk properties (FQ^a, ref.[156]) or QM atomic charges (FQ^b, ref.[162]), FQ^c is tuned to the total interaction energy at the CCSD(T) level (with an error of less than 1 kcal/mol). FQ^c yields an accurate description of SAPT0 electrostatic interactions. Notice that similar findings are given by both jun-cc-pVDZ and N07D basis sets (see Table 6.8 in SI). To end the discussion on interaction energies, it is worth noticing that the analysis reported above is only allowed when non-electrostatic interactions are included in QM/MM calculations, i.e. is not achievable by exploiting common purely electrostatic approaches.

	FQ^a	FQ^b	FQ^c	SAPT0	CCSD(T)
Electrostatic	-20.60	-26.80	-47.06	-31.35	–
Induction	–	–	–	-11.45	–
Repulsion	27.78	28.58	30.99	28.34	–
Dispersion	-3.28	-3.28	-3.28	-9.43	–
Total	3.90	-1.50	-19.35	-23.89	-20.62

Table 6.3. PROXYL+2w EDA obtained by expoiting PBE0/FQ with different parametrizations and SAPT0. CCSD(T) calculations include Counter-Poise corrections. All data are reported in mHartree and were obtained by using the aug-cc-pVTZ basis set.
[a] FQ parametrization taken from Ref.[156]
[b] FQ parametrization taken from Ref.[162]
[c] FQ parametrization proposed in this work

Calculated hcc$_N$ of the PROXYL/TEMPO+2w clusters are reported in Table 6.4. QM/MM calculations were performed by exploiting both the non-polarizable TIP3P[358] force field and FQ (with different parametrizations) to describe electrostatic interactions. Two set of QM/MM calculations were performed. The first employs TIP3P or FQ embedding and do not include non-electrostatic interactions. The corresponding results are reported in the first four columns of Table 6.4. In the second set of calculations, non-electrostatic interactions, as obtained with our model, are included. All results are also compared with full QM calculations, i.e. both the radicals and the two water molecules are described at the QM level (see column 9 in Table 6.4).
The reported data clearly show that the non-polarizable TIP3P approach gives large errors with respect to full QM calculations; remarkably, the inclusion of non-electrostatic terms does not improve the results. A different picture results from polarizable QM/FQ values. In fact, when only the electrostatic interactions are considered, the FQ^b parametrization gives values which are in fair agreement with the reference full QM data. However, the inclusion of non-electrostatic interactions shifts hcc$_N$ values in the wrong direction, thus increasing the absolute difference with respect to reference values. This is not surprising, because EDA analysis (see Table 6.3) already showed underestimated electrostatic interactions. The same considerations are also valid for FQ^a, whereas the novel FQ^c parametrization overestimates hcc$_N$ values if only electrostatic interactions are considered. Remarkably, the inclusion of non-electrostatic interactions shifts FQ^c values in the right direction, and the agreement with full QM

reference data is almost perfect (0.2 Gauss).
Furthermore, additional CCSD/EPR-II calculations were performed on the reduced structures depicted in Figure 6.3 (see Table 6.4). Full QM DFT calculations underestimate CCSD/EPR-II hcc_N valus by 0.9 and 1.2 Gauss for PROXYL and TEMPO, respectively. Notice that calculated CCSD/EPR-II hcc_N are still not comparable with experimental values, especially for PROXYL. This confirms that the cluster approach is inadequate to physically describe the solvation phenomenon, which is intrinsically a dynamical process.

	PBE0/N07D									CCSD/EPR-II	$\Delta_{CC/PBE0}$	Exp
	Elect.				Elect. + Dis/Rep							
	TIP3P	FQ[a]	FQ[b]	FQ[c]	TIP3P	FQ[a]	FQ[b]	FQ[c]	Full-QM	Full-QM		
PROXYL	13.4	13.1	13.4	14.3	13.1	12.8	13.1	13.9	13.7	14.6	0.9	16.4
TEMPO	17.9	15.3	15.7	16.7	16.9	14.8	15.1	16.1	15.9	17.1	1.2	17.3

Table 6.4. hcc_N of PROXYL/TEMPO+2w clusters obtained at different level of theory. All data are reported in Gauss.
[a] FQ parametrization taken from Ref. [156]
[b] FQ parametrization taken from Ref. [162]
[c] FQ parametrization proposed in this work

6.4.3 hcc_N of PROXYL/TEMPO from MD runs

An alternative and more accurate way of modeling solvation is to combine our approach with classical MD. Table 6.5 reports selected geometrical parameters (and their standard deviation) obtained by averaging 200 representative snapshots extracted from MD runs performed on PROXYL and TEMPO in aqueous solution. The improper dihedral angle $C_\alpha NOC_\alpha$, which as stated before plays a crucial role in determining EPR parameters, is drastically different with respect to what has been reported for the isolated radicals, especially for TEMPO. Furthermore, due to the dynamical picture given by the MD, the geometrical parameters are accompanied by standard deviations (in brackets), which are large in the case of this angle.

	<PROXYL>	<TEMPO>
N–O	1.27 (0.03)	1.27 (0.03)
$\angle$ $C_\alpha NC_\alpha$	115.3 (2.5)	123.6 (2.7)
$C_\alpha NOC_\alpha$	$\pm$ 0.4 (17.8)	$\pm$5.0 (20.1)

Table 6.5. Mean values and standard deviations (in brackets) of selected geometrical parameters of PROXYL and TEMPO structures extracted from MD runs.

In order to show how the variability in the improper dihedral affects calculated hcc_N values, two different set of calculations were performed. First, all solvent molecules in all snapshots were removed and hcc_N were calculated on the resulting structures. Second, all solvent molecules were indeed included and treated at the FQ level, with

the sole inclusion of electrostatic effects (c parametrization). In Figures 6.4 and 6.5 the resulting hcc_N values are reported as a function of the out-of-plane $\mathrm{C}_\alpha\mathrm{NOC}_\alpha$ angle.

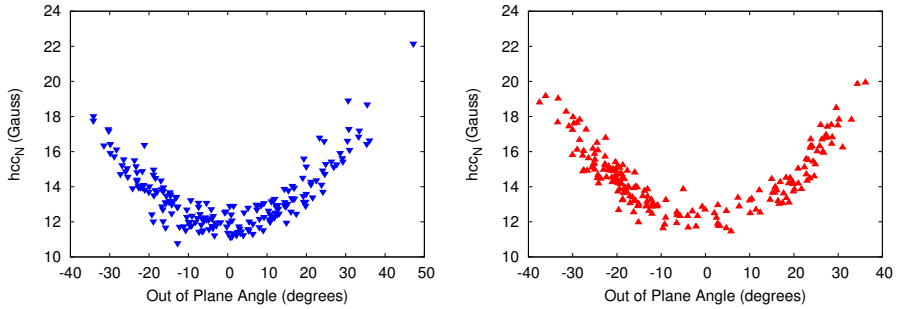


Figure 6.4. PBE0/N07D calculated hcc_N (Gauss) on the solute-only structures extracted from MD runs as a function of the out of plane $\mathrm{C}_\alpha\mathrm{NOC}_\alpha$ angle. (Left: PROXYL; Right: TEMPO).

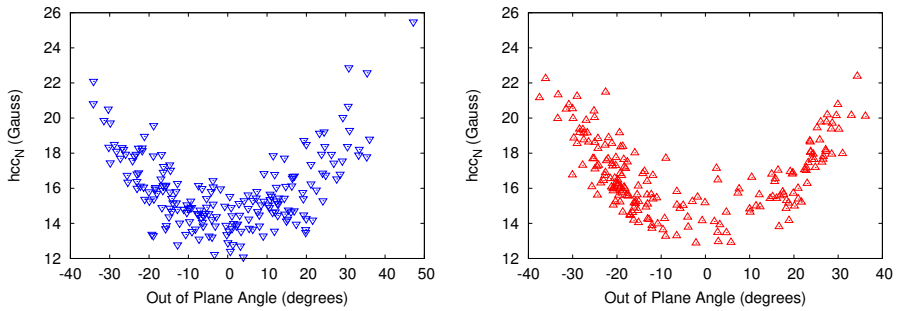


Figure 6.5. PBE0/N07D QM/FQ calculated hcc_N (Gauss) on the entire snapshots extracted from MD runs as a function of the out of plane $\mathrm{C}_\alpha\mathrm{NOC}_\alpha$ angle. (Left: PROXYL; Right: TEMPO).

As expected, the same picture as already reported for the isolated radicals emerges. Due to the large variability of hcc_N values as a function of the out of plane angle, the convergence of average values needs to be carefully checked. In Figure 6.6 QM/FQ hcc_N average values as a function of the number of snapshots are depicted for the two radicals. Clearly, hcc_N is well converged by using 200 snapshots.

Let us now compare our computed data with their experimental counterparts. Table 6.6 collects hcc_N values computed with different approaches. QM indicates calculations performed on the solute-only structures extracted from MD (see above). QM/FQ data were obtained by using the purely-electrostatic polarizable FQ with the c parametrization (the results obtained by exploiting the a, b parametrizations are reported in Table 6.10 in the SI). The contribution to hcc_N due to repulsion interactions is denoted as Δ_{rep}, whereas the contribution to hcc_N of both repulsion and dispersion interactions is denoted as $\Delta_{\mathrm{dis\text{-}rep}}$.

We first notice that, due to the different structural sampling given by the MD, QM data in Table 6.6 differ from what was reported for the isolated radicals (see Table 6.2). The

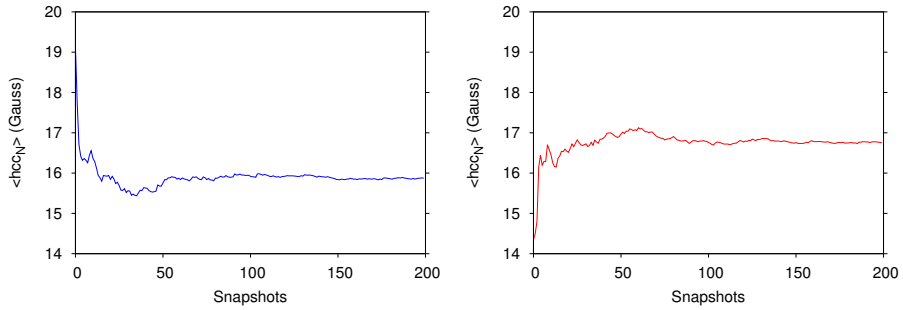


Figure 6.6. QM/FQ hcc_N mean value as a function of the number of snapshots extracted from MD runs. (Left: PROXYL; Right: TEMPO) All data are reported in Gauss.

dynamical sampling increases PROXYL and TEMPO hcc_N values by about 2.4 and 2.2 Gauss, respectively. As a result, the difference between hcc_N values of the two radicals (1.1 Gauss) is in good agreement with experimental data (0.9 Gauss).[537,538] When full solvent effects are included at the purely electrostatic FQ level (2nd column), hcc_N values are increased by about 2.3 Gauss on average for both radicals. This means that attractive interactions increase the computed property. As a result, the inclusion of repulsive interaction terms is expected to decrease computed values, and this is indeed confirmed by the values reported in the third column. In particular, for both radicals hcc_N decreases by 0.4 and 0.5 Gauss, respectively, i.e. of about 17 % and 23 % of the whole solvent effect. The further inclusion of dispersion terms does not affect the difference with FQ average values.

In order to best compare the results of our approach with experimental findings, DFT values were also corrected to account for some intrinsic deficiency. To this end, the difference between full DFT and full CCSD data obtained for clusters ($\Delta_{CC/PBE0}$, see Table 6.4) was added to the calculated QM/MM value. The resulting values are labeled "Best QM/MM" in Table 6.6. Remarkably, our best computed values are in excellent agreement with experimental data for both radicals, thus confirming the accuracy and reliability of our approach.

	QM	FQ	Δ_{rep}	$\Delta_{\text{dis-rep}}$	Best QM/MM FQ+$\Delta_{\text{dis-rep}}$+$\Delta_{CC/PBE0}$	Exp.
PROXYL	13.5	15.9	-0.4	-0.4	16.4 ± 0.1	16.4[537]
TEMPO	14.6	16.8	-0.5	-0.5	17.5 ± 0.1	17.3[538]

Table 6.6. PBE0/N07D hcc_N mean values calculated on 200 snapshots extracted from MD runs. QM indicates the calculation performed on solute-only structures. FQ refers to the purely electrostatic QM/FQ with c parametrization. Δ_{rep} and $\Delta_{\text{dis-rep}}$ are differences between FQ and hcc_N data obtained with our method. Best QM/MM data are obtained by summing FQ, $\Delta_{\text{dis-rep}}$ and $\Delta_{CC/PBE0}$ (see Table 6.4). All values are reported in Gauss.

To get further insight into solvent effects on hcc_N values, differences between FQ and QM values are reported as a function of the snapshot in Figure 6.7. As it can be noticed, for both PROXYL (left) and TEMPO (right) the electrostatic solvent contribution to hcc_N is always positive (only in one case a small negative contribution is reported for TEMPO). Notice that this is different from what has been reported for electric properties of higher order.[152]

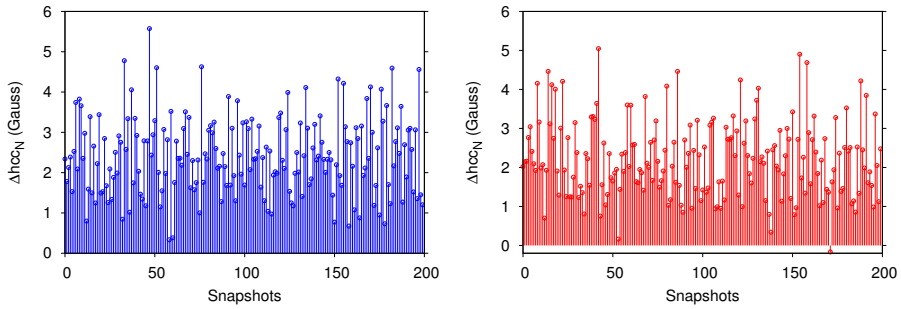


Figure 6.7. Calculated solvent effects (see text) on hcc_N as a function of the snapshot extracted from MD runs (Left: PROXYL; Right: TEMPO). All data are reported in Gauss.

In Figure 6.8 the difference between calculated solvent effects on hcc_N as obtained with the purely electrostatic FQ approach or with the further inclusion of the repulsion contribution is reported. Remarkably, repulsion contributions increase or decrease hcc_N value depending on the selected snapshot, thus showing that cluster approaches do not guarantee an adequate modeling of solvent effects. In fact, although repulsion effects give on average a negative contribution to hcc_N, by taking a random snapshot (cluster), a completely different picture could emerge.

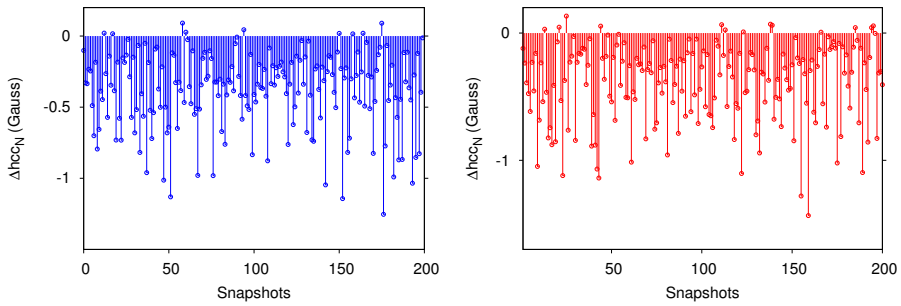


Figure 6.8. Difference between FQ and QM/FQ+repulsion solvent effects as a function of the snapshot extracted from the MD (Left: PROXYL; Right: TEMPO). All data are reported in Gauss.

6.5 Summary and Conclusions

In this paper, we have extended to the calculation of hyperfine coupling constants, the model proposed in Ref.[197] to include Pauli repulsion and dispersion effects in QM/MM approaches. The peculiarity of the proposed approach stands in the fact that repulsion/dispersion contributions are explicitly introduced in the QM Hamiltonian. Therefore, such terms not only enter the evaluation of energetic properties but, remarkably, propagate to molecular properties and spectra. The account of such contributions has permitted a quantitative analysis of QM/MM interaction energies, and this has also required a novel parametrization of the FQ force field, which has been then tested against the prediction of EPR hcc_N of PROXYL and TEMPO in aqueous solutions.
Numerical applications to the two radicals in vacuo, solvated within the so-called cluster approach or as modeled through MD, confirm the well known relevance of solvent effects and a proper account of their dynamical aspects. The further inclusion of dispersion and especially repulsion solute-solvent interactions gives, remarkably, an almost perfect agreement between calculated and experimental values. Therefore, although electrostatic effects have been invoked as dominating the solvation phenomenon in aqueous solution, we found that non-electrostatic effects are indeed relevant, contributing to 17 % and 23 % of the entire solvent effects on hcc_N for PROXYL and TEMPO, respectively. Remarkably, dispersion interactions seem not to play a crucial role.
To end the discussion, we remark that our model is general enough to be applied to any kind of solvent/environment, pending a reliable parametrization of both electrostatic and non-electrostatic interactions. Also, due to the inclusion of all terms in the molecular Hamiltonian, our approach can be extended to any kind of molecular properties and spectroscopies; this will be the topic of future communications.

Acknowledgments

We are thankful for the computer resources provided by the high performance computer facilities of the SMART Laboratory (http://smart.sns.it/). CC gratefully acknowledges the support of H2020-MSCA-ITN-2017 European Training Network “Computational Spectroscopy In Natural sciences and Engineering” (COSINE), grant number 765739.

Supporting Information

Geometries in vacuo

		PBE0 + D3		B3LYP + D3	
	Parameter	BS1	BS2	BS1	BS2
PROXYL	$N\cdots O$	1.262	1.257	1.274	1.267
	$C_\alpha NOC_\alpha$	-0.029	-0.030	-0.028	-0.026
	$\angle C_\alpha NC_\alpha$	115.178	115.064	115.194	115.074
TEMPO	$N\cdots O$	1.271	1.265	1.283	1.277
	$C_\alpha NOC_\alpha$	-22.668	-22.063	-22.664	-22.139
	$\angle C_\alpha NC_\alpha$	123.926	124.018	124.003	124.118

Table 6.7. Selected geometrical parameters of PROXYL and TEMPO radicals at the different levels of theory with the inclusion of Grimme empirical dispersion D3. BS1: aug-cc-pVDZ; BS2: 6-311++G(3dp,2pd).

Energy Decomposition Analysis

	jun-cc-pVDZ					N07D				
	FQ[a]	FQ[b]	FQ[c]	SAPT0	CCSD(T)	FQ[a]	FQ[b]	FQ[c]	SAPT0	CCSD(T)
Electrostatic	-21.00	-26.88	-45.22	-32.71	–	-20.25	-26.24	-45.77	-31.85	–
Induction	–	–	–	-11.33	–	–	–	–	-11.24	–
Repulsion	28.53	29.11	30.83	28.49	–	28.15	28.86	31.05	28.44	–
Dispersion	-3.28	-3.28	-3.28	-6.36	–	-3.29	-3.29	-3.29	-7.82	–
Total	4.25	-1.05	-17.67	-21.90	-17.72	4.61	-0.67	-18.01	-22.48	-18.57

Table 6.8. PROXYL+2w EDA obtained by expoiting PBE0/FQ with different parametrizations and SAPT0. CCSD(T) calculations include Counter-Poise corrections. All data are reported in mHartree.
[a] FQ parametrization taken from Ref. [156]
[b] FQ parametrization taken from Ref. [162]
[c] FQ parametrization proposed in this work

Electrostatic FQ parametrization

	FQ^a	FQ^b	FQ^c
η_O	0.584852	0.623700	0.523700
η_H	0.625010	0.637512	0.537512
χ_O	0.116859	0.189194	0.189194
χ_H	0.000000	0.012767	0.012767
$\Delta\chi$	0.116859	0.176427	0.176427

Table 6.9. O and H parameters for FQ calculations. η and χ are the chemical hardnesses and atomic electronegativities, respectively. All data are reported in a.u.
[a] FQ parametrization taken from Ref. [156]
[b] FQ parametrization taken from Ref. [162]
[c] FQ parametrization proposed in this work

hcc$_N$ of PROXYL/TEMPO from MD runs

	TIP3P	FQ^a	FQ^b	Δ_{rep}	$\Delta_{\text{dis-rep}}$
PROXYL	14.9	14.7	15.0	-0.3	-0.3
TEMPO	15.9	15.7	16.0	-0.3	-0.3

Table 6.10. PBE0/N07D hcc$_N$ mean values calculated on 200 snapshots extracted from MD runs. TIP3P refers to the purely electrostatic QM/MM embedding, where the water molecules were described by means of the non-polarizable TIP3P force field. FQ refers to the purely electrostatic QM/FQ with a, b parametrization. Δ_{rep} and $\Delta_{\text{dis-rep}}$ are differences between FQ^a and hcc$_N$ data obtained with our method. All values are reported in Gauss.
[a] FQ parametrization taken from Ref. [156]
[b] FQ parametrization taken from Ref. [162]

Chapter 7

Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQFμ model

Tommaso Giovannini*,†, Alessandra Puglisi†, Matteo Ambrosetti†, Chiara Cappelli*,†

†*Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy.*

E-mail: tommaso.giovannini@sns.it; chiara.cappelli@sns.it

Abstract The novel polarizable FQFμ force field is proposed and coupled to a QM SCF Hamiltonian. The peculiarity of the resulting QM/FQFμ approach stands in the fact the polarization effects are modeled in terms of both fluctuating charges and dipoles, which are not fixed but can vary in response to the external electric field/potential. The capabilities of the model to reproduce full QM reference electrostatic energies of molecular systems in aqueous solution are tested, showing a remarkable accuracy of the new approach.

7.1 Introduction

The problem of describing the interaction between a molecule and its embedding environment is one of the pillars of 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:[1–6] 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 intrisic 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).[12,13] 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,[58,154,185,186,281] induced dipoles,[10,155,164] Drude oscillators[59] or Fluctuating Charges (FQ).[56,156,158] 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,[542] Drude Oscillators (Polarizable Charge Equilibration PQEq),[543,544] or Gaussian-like induced atomic dipoles (Q+P iso $[R, \alpha_{iso}]$ model,[545] Discrete Interaction Model DIM,[546] Capacitance Polarization Model CMM[547]). 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.[545,546]

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.[545] 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.[54–56,151,152,161]
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.[548,549]
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.

7.2 Theoretical Model

7.2.1 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:

$$\begin{aligned}\rho_{q_i}(\mathbf{r}) &= \frac{q_i}{\pi^{\frac{3}{2}} R_{q_i}^3} \exp\left(-\frac{|\mathbf{r}-\mathbf{r}_i|^2}{R_q^2}\right) \\ \rho_{\boldsymbol{\mu}_i}(\mathbf{r}) &= \frac{|\boldsymbol{\mu}_i|}{\pi^{\frac{3}{2}} R_{\boldsymbol{\mu}_i}^3} \hat{\mathbf{n}}_i \cdot \boldsymbol{\nabla}\left[\exp\left(-\frac{|\mathbf{r}-\mathbf{r}_i|^2}{R_{\boldsymbol{\mu}_i}^2}\right)\right]\end{aligned} \qquad \textbf{7.1}$$

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:[545]

$$\begin{aligned}E(\mathbf{q},\boldsymbol{\mu}) = &\sum_i q_i\chi_i + \frac{1}{2}\sum_i q_i\eta_i q_i + \frac{1}{2}\sum_i\sum_{j\neq i} q_i \mathrm{T}_{ij}^{qq} q_j + \sum_i\sum_{j\neq i} q_i \mathbf{T}_{ij}^{q\mu}\boldsymbol{\mu}_j + \\ &- \frac{1}{2}\sum_i\sum_{j\neq i}\boldsymbol{\mu}_i^\dagger \mathbf{T}_{ij}^{\mu\mu}\boldsymbol{\mu}_j + \frac{1}{2}\sum_i \boldsymbol{\mu}_i^\dagger \alpha_i^{-1}\boldsymbol{\mu}_i\end{aligned} \qquad \textbf{7.2}$$

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 Eq. 7.1 are adopted, the functional form of the interaction kernels provided by Mayer[545] can be exploited. T_{ij}^{qq} term reads:

$$\mathrm{T}_{ij}^{qq} = \frac{1}{|\mathbf{r}_{ij}|}\mathrm{erf}\left(\frac{|\mathbf{r}_{ij}|}{R_{q_i-q_j}}\right) \tag{7.3}$$

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"):[545,546]

$$\lim_{\mathbf{r}_{ij}\to 0} \mathrm{T}_{ij}^{qq} = \mathrm{T}_{ii}^{qq} = \frac{2}{\sqrt{\pi}}\frac{1}{R_{q_i-q_i}} \tag{7.4}$$

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:

$$\mathrm{T}_{ii}^{qq} = \eta_i \Rightarrow R_{q_i} = \sqrt{\frac{2}{\pi}}\frac{1}{\eta_i} \tag{7.5}$$

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 Eq. 7.3:[545]

$$\mathbf{T}_{ij}^{q\mu} = -\boldsymbol{\nabla}_{r_i}\mathrm{T}_{ij}^{qq} = -\frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^3}\left[\mathrm{erf}\left(\frac{|\mathbf{r}_{ij}|}{R_{q_i-\mu_j}}\right) - \frac{2|\mathbf{r}_{ij}|}{\sqrt{\pi}R_{q_i-\mu_j}}\exp\left(\frac{-|\mathbf{r}_{ij}|}{R_{q_i-\mu_j}}\right)^2\right] \tag{7.6}$$

$$\mathbf{T}_{ij}^{\mu\mu} = \boldsymbol{\nabla}_{r_j}\mathbf{T}_{ij}^{q\mu} = \frac{3\mathbf{r}_{i,j}\otimes\mathbf{r}_{i,j} - |\mathbf{r}_{i,j}|^2\mathbf{I}}{|\mathbf{r}_{i,j}|^5}\left[\mathrm{erf}\left(\frac{|\mathbf{r}_{ij}|}{R_{\mu_i-\mu_j}}\right) - \frac{2}{\sqrt{\pi}}\frac{|\mathbf{r}_{ij}|}{R_{\mu_i-\mu_j}}\exp\left(-\frac{\mathbf{r}_{ij}}{R_{\mu_i-\mu_j}}\right)^2\right] +$$
$$- \frac{4}{\sqrt{\pi}R_{\mu_i-\mu_j}^3}\frac{\mathbf{r}_{ij}\otimes\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^2}\exp\left(-\frac{|\mathbf{r}_{ij}|}{R_{\mu_i-\mu_j}}\right)^2 \tag{7.7}$$

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:

$$\lim_{\mathbf{r}_{ij}\to 0} \mathbf{T}_{ij}^{q\mu} = \mathbf{T}_{ii}^{q\mu} = 0 \tag{7.8}$$

$$\lim_{\mathbf{r}_{ij}\to 0} \mathbf{T}_{ij}^{\mu\mu} = \mathbf{T}_{ii}^{\mu\mu} = -\sqrt{\frac{2}{\pi}}\frac{\mathbf{I}}{3R_{\mu_i}^3} \tag{7.9}$$

From Eq.7.9, R_{μ_i} can be defined in terms of the atomic polarizability α_i:

$$\alpha_i^{-1} = \sqrt{\frac{2}{\pi}}\frac{1}{3}\frac{1}{R_{\mu_i}^3} \Rightarrow R_{\mu_i} = \left(\sqrt{\frac{2}{\pi}}\frac{1}{3}\alpha_i\right)^{\frac{1}{3}} \tag{7.10}$$

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, Eq. 7.2 can be formally rewritten as:

$$E(\mathbf{q},\boldsymbol{\mu}) = \frac{1}{2}\sum_i\sum_j q_i \mathrm{T}_{ij}^{qq} q_j + \frac{1}{2}\sum_i\sum_j \boldsymbol{\mu}_i^\dagger \mathbf{T}_{ij}^{\mu\mu}\boldsymbol{\mu}_j + \sum_i\sum_j q_i \mathbf{T}_{ij}^{q\mu}\boldsymbol{\mu}_j^\dagger + \sum_i q_i\chi_i = \\ = \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} \qquad \textbf{7.11}$$

where a matrix notation has been adopted.
In Eq. 7.11, 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 intermolecular 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,[54,159] similar equations can be derived under the second assumption : they are given in Section 7.6.1 of the Supporting Information (SI). Notice that our implementation is general and can treat both cases. The energy functional F can be written by exploiting the Lagrangian multiplier (λ):

$$F(\mathbf{q},\boldsymbol{\mu},\lambda) = E(\mathbf{r},\mathbf{q},\boldsymbol{\mu}) + \lambda\left[\sum_i (q_i) - Q_{\mathrm{tot}}\right] = \\ = \frac{1}{2}\sum_i\sum_j q_i \mathrm{T}_{ij}^{qq} q_j + \frac{1}{2}\sum_i\sum_j \boldsymbol{\mu}_i^\dagger \mathbf{T}_{ij}^{\mu\mu}\boldsymbol{\mu}_j + \sum_i\sum_j q_i \mathbf{T}_{ij}^{q\mu}\boldsymbol{\mu}_j^\dagger + \sum_i q_i\chi_i \\ + \lambda\left[\sum_i (q_i) - Q_{\mathrm{tot}}\right] = \\ = \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \lambda\mathbf{q} \qquad \textbf{7.12}$$

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:

$$\begin{cases} \sum_j \mathbf{T}^{qq}_{i,j} q_j + \lambda + \sum_j \mathbf{T}^{q\mu}_{i,j} \boldsymbol{\mu}_j = -\chi_i \\ \sum_j \mathbf{T}^{\mu\mu}_{i,j} \boldsymbol{\mu}_j + \sum_j \mathbf{T}^{q\mu}_{i,j} q_j = 0 \\ \sum_i q_i = Q_{\text{tot}} \end{cases} \tag{7.13}$$

The whole system can be recast in a more compact form as:[542,547]

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_\lambda & \mathbf{T}^{q\mu} \\ \mathbf{1}_\lambda^\dagger & 0 & \mathbf{0} \\ -\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \lambda \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ Q_{\text{tot}} \\ \mathbf{0} \end{pmatrix} \Rightarrow \mathbf{D}\mathbf{L}_\lambda = -\mathbf{C}_Q \tag{7.14}$$

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 $\mathbf{L}_\lambda$ is a vector containing charges, dipoles and the Lagrange multiplier.

7.2.2 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:

$$\mathcal{E} = E_{QM} + E_{MM} + E_{QM/MM} \tag{7.15}$$

where E_{MM} is defined in Eq. 7.12. 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:

$$E_{QM/MM} = \sum_i V[\rho_{QM}](\mathbf{r}_i) q_i - \boldsymbol{\mu}_i^\dagger \mathbf{E}[\rho_{QM}](\mathbf{r}_i) \tag{7.16}$$

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:

$$V[\rho_{QM}](\mathbf{r}_i) = V_i[\mathbf{P}] = V_i^N(\mathbf{P}) + V_i^e(\mathbf{P}) = \sum_\zeta \frac{Z_\zeta}{|\mathbf{r}_i - \mathbf{R}_\zeta|} + \int_{\mathbb{R}^3} \frac{\rho_{el}(\mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|} \, \mathrm{d}\mathbf{r} \tag{7.17}$$

$$\tag{7.18}$$

$$\mathbf{E}[\rho_{QM}](\mathbf{r}_i) = \mathbf{E}_i(\mathbf{P}) = \mathbf{E}_i^N(\mathbf{P}) + \mathbf{E}_i^e(\mathbf{P}) = \sum_\zeta \frac{Z_\zeta(\mathbf{R}_\zeta - \mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{R}_\zeta|^3} - \int_{\mathbb{R}^3} \frac{\rho_{el}(\mathbf{r})(\mathbf{r}_i - \mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|^3} \, \mathrm{d}\mathbf{r} \tag{7.19}$$

where ρ_{el} is the electron density. ζ index runs over the QM nuclei, whose charges are named Z_ζ and whose positions are $\mathbf{R}_\zeta$. If the electronic density $\rho_{el}(\mathbf{r})$ is expanded in an atomic basis set $\{\chi_\mu\}$, the second terms in Eqs.7.18 and 7.19 become:

$$V_i^e(\mathbf{P}) = -\sum_{\nu\mu} P_{\mu\nu} \int_{\mathbb{R}^3} \mathrm{d}\mathbf{r} \frac{\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|} = \sum_{\mu\nu} P_{\mu\nu} V_{\mu\nu,i} \qquad \textbf{7.20}$$

$$\mathbf{E}_i^e(\mathbf{P}) = -\sum_{\mu\nu} P_{\mu\nu} \int_{\mathbb{R}^3} \frac{\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})(\mathbf{r}_i - \mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|^3} = \sum_{\mu\nu} P_{\mu\nu} E_{\mu\nu,i} \qquad \textbf{7.21}$$

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:

$$\begin{aligned} F(\mathbf{P}, \mathbf{q}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \mathrm{tr}\mathbf{h}\mathbf{P} + \frac{\mathbf{1}}{\mathbf{2}}\mathrm{tr}\mathbf{P}\mathbf{G}(\mathbf{P}) + \frac{\mathbf{1}}{\mathbf{2}}\mathbf{q}^\dagger\mathbf{T}^{\mathbf{qq}}\mathbf{q} + \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{\mathbf{q}\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q} + \\ + \mathbf{q}^\dagger\mathbf{V}(\mathbf{P}) - \boldsymbol{\mu}^\dagger\mathbf{E}(\mathbf{P}) \end{aligned} \qquad \textbf{7.22}$$

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:

$$\tilde{F}_{\mu\nu} = \frac{\partial\mathcal{E}}{\partial P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{V}_{\mu\nu}^\dagger\mathbf{q} - \mathbf{E}_{\mu\nu}^\dagger\boldsymbol{\mu} \qquad \textbf{7.23}$$

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.

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_\lambda & \mathbf{T}^{q\mu} \\ \mathbf{1}_\lambda^\dagger & 0 & \mathbf{0} \\ -\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \lambda \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ Q_{\text{tot}} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} -\mathbf{V}(\mathbf{P}) \\ 0 \\ \mathbf{E}(\mathbf{P}) \end{pmatrix} \Rightarrow \mathbf{DL}_\lambda = -\mathbf{C}_Q - \mathbf{R}(\mathbf{P}) \tag{7.24}$$

Notice that, with respect to Eq. 7.14, a new source term $\mathbf{R}(\mathbf{P})$ 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. This is a non linear term, because it contains the solution of the SCF equation, and can be computationally evaluated in an iterative way, i.e.:

1. Construct an initial density matrix $\mathbf{P}^{(i)}$
2. Calculate the electrostatic source term $\mathbf{S}(\mathbf{R}^{(i)})$ and solve for both charges and dipoles:
$$\mathbf{L}^{(i)} = -\mathbf{D}^{-1}\mathbf{R}(\mathbf{P}^{(i)}) \tag{7.25}$$
The $\mathbf{D}$ matrix only depends on the positions of the MM atoms, however it is independent of the eletronic density, so that matrix inversion can be performed only at the first step of the calculation.
3. Assemble the effective Fock matrix:
$$\tilde{F}_{\mu\nu}^{(i)} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}^{(i)}) + \mathbf{V}_{\mu\nu}^\dagger \mathbf{q}^{(i)} - \mathbf{E}_{\mu\nu}^\dagger \boldsymbol{\mu}^{(i)}$$
4. Solve Roothan equations and build-up a new density matrix:
$$\tilde{\mathbf{F}}^{(i)}\mathbf{C}^{(i+1)} = \mathbf{S}\mathbf{C}^{(i+1)}\tilde{\boldsymbol{\epsilon}}^{(i+1)}$$
$$\mathbf{P}^{(i+1)} = \mathbf{C}^{(i+1)}\mathbf{C}^{(i+1)\dagger}$$
5. Go back to step 2, and iterate until convergence is reached.

QM/FQFμ introduces two polarization sources: fluctuating charges and fluctuating dipoles. From Eq. 7.24 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 Eq. 7.23). 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.,[545] similarly to QM/DIM[226] and QM/CMM.[547] 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 Eqs. 7.5 and 7.10). 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/DIM and 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.[547] 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.[75]

7.2.3 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.[543] PQEq combines the Charge Equilibration model (QEq)[287] with the Drude Oscillator approach.[173] 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 7.6.2, given as SI). QM/PQEq Fock operator reads:

$$\tilde{F}_{\mu\nu} = \frac{\partial \mathcal{E}}{\partial P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{V}^{\dagger}_{\mu\nu,c}\mathbf{q} + \mathbf{V}^{\dagger}_{\mu\nu,c}\mathbf{Z} - \mathbf{V}^{\dagger}_{\mu\nu,s}\mathbf{Z} \qquad \textbf{7.26}$$

where, $\mathbf{q}$ and $\mathbf{Z}$ are the vectors containing fluctuating and fixed charges, respectively, whereas 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 Eq. 7.37 given as SI) introduces a non-linearity in the problem, which can be solved only by exploiting iterative techniques.[59,175–177] 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[54,55,58,149,150,162,165,169] by using the standard techniques of quantum chemistry.[550] 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.[178]

7.3 Computational Details

QM/FQFμ was implemented in a locally modified version of Gaussian16.[301] 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.,[156] ref.[162] and ref.[203] Non-electrostatic contributions, i.e. repulsion and dispersion, were modeled as reported in Ref.[197] All the classical Molecular Dynamics (MD) simulations were performed with the Gromacs package,[348,471–474] by keeping the same settings as previously reported by some of the present authors.[146,152,153] Details on MD simulations are given as SI. The Kitaura-Morokuma Energy Decomposition Analysis (KM-EDA)[475,476] was performed by using the GAMESS package.[477,478] Symmetry Adapted Perturbation Theory (SAPT)[440,441] calculations were performed by using Psi4 1.1.[479]

7.4 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 SAPT0/6-31+G* values.

7.4.1 Model Parametrization

QM/FQFμ is general enough to model any kind of external environment, pending an appropriate parametrization of the quantities entering Eqs. 7.22 and 7.23. 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 Eq. 7.22, selected water clusters taken from Kratz et al.[491] (see Figure 7.1), were studied. Reference full QM electrostatic energy

values of such clusters were calculated by performing a KM-EDA [475,476] calculation on each structure in Figure 7.1 at the HF/6-31+G* level, according to what has already been proposed in the literature. [197,482,483]

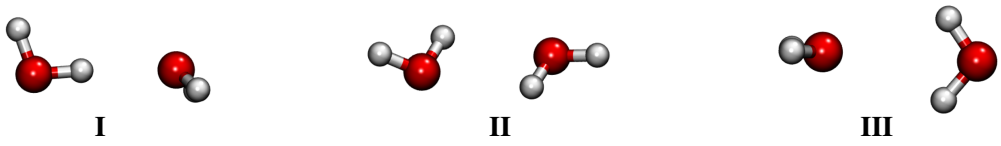


Figure 7.1. Structures of water dimers exploited in the parametrization of QM/FQFμ.

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, [197] 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 Eq. 7.24). Thus, QM/FQFμ electrostatic energies are compared with the sum of electrostatic and polarization KM-EDA energy contributions.
For the studied dimers, Eq. 7.22 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 7.3 in the SI. Electrostatic energies of the selected water dimers are reported in Table 7.4 in SI. 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.

7.4.2 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 7.2 (optimized at the MP2/aug-cc-pVQZ level) was exploited, and the distance d between the oxygen atoms was taken as reference.
In Figure 7.3, 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).

Figure 7.2. Structure of the water dimer used to study the dependence of electrostatic and interaction energies as a function of O-O intermolecular distance.

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 ($70 \sim 72$ %).

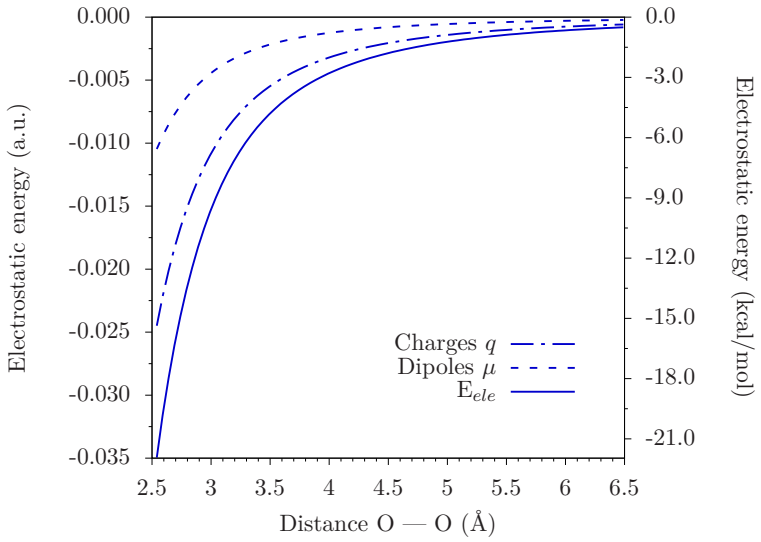


Figure 7.3. Plot of the QM/FQFμ electrostatic energy as a function of the O-O intermolecular distance for the water dimer depicted in Figure 7.2. Charge and dipole contributions to E$_e$*le* are also plotted.

In Figure 7.4, 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 Fig.7.1).
The inset in Figure 7.4 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).

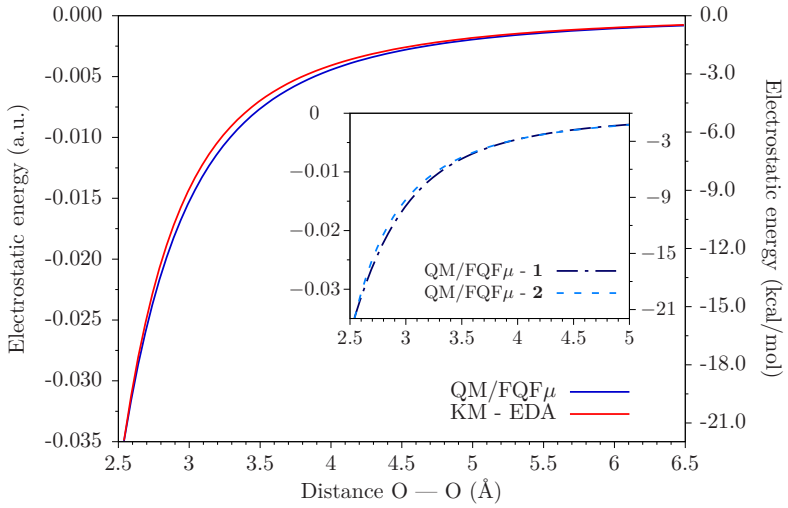


Figure 7.4. Plot of the electrostatic energy as a function of the O-O intermolecular distance for the water dimer depicted in Figure 7.2. QM/FQFμ values (HF/6-31+G* for the QM moiety) are compared to KM-EDA (HF/6-31+G*) calculations. In KM-EDA calculations, electrostatic and polarization contributions 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 7.2.

To end this discussion, the total B3LYP/aug-cc-pVTZ QM/FQFμ interaction energy as a function of d is plotted in Figure 7.5 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,[197] which formulates repulsion in terms of an auxiliary density on the MM portion, whereas QM/MM dispersion is obtained by extending the Tkatchenko-Scheffler approach to DFT.[198–202]
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.

7.4.3 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 7.2 with d = 2.94 Å is exploited. Thirteen different methods were used by following the recent literature,[147,493] ranging from HF to pure DFT functionals (LDA, PBE,[551] B97D,[494,495] R-TPSS[552]), to different classes of hybrid functionals (BLYP,[553] M06,[498] PBE0,[499] B3LYP,[496] M062X,[498] SOGGA11-X,[500] mPW1PW91[554]), also including long-range (CAM-B3LYP[347]). Each functional was coupled to several Pople-type basis sets (see Figure 7.6), in order to separate the contributions arising from polarization

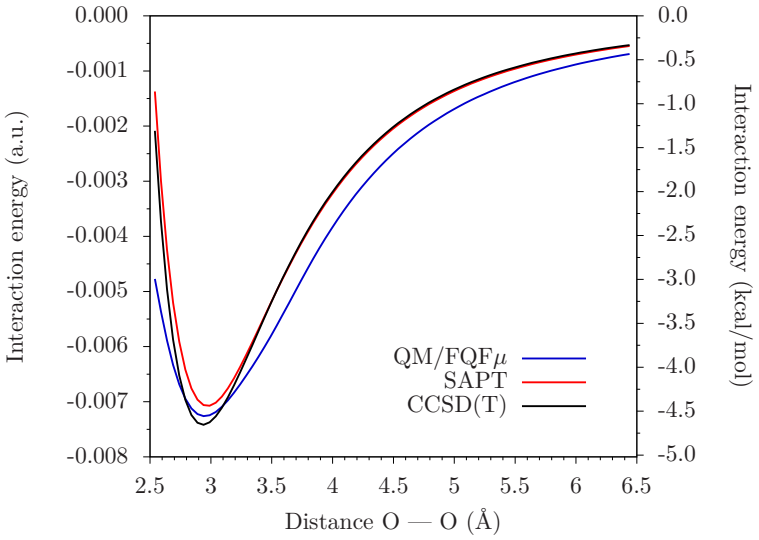


Figure 7.5. QM/FQFμ (B3LYP/aug-cc-pVTZ for the QM moiety), SAPT2+3(CCD) and CCSD(T)/aug-cc-pVTZ total interaction energies for the water dimer in Figure 7.2 as a function of the O-O distance.

and diffuse functions. In addition, correlation-consistent and augmented correlation-consistent basis sets were employed, up to aug-cc-pVQZ.[555]

Figure 7.6 schematically reports the observed trends.

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. 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

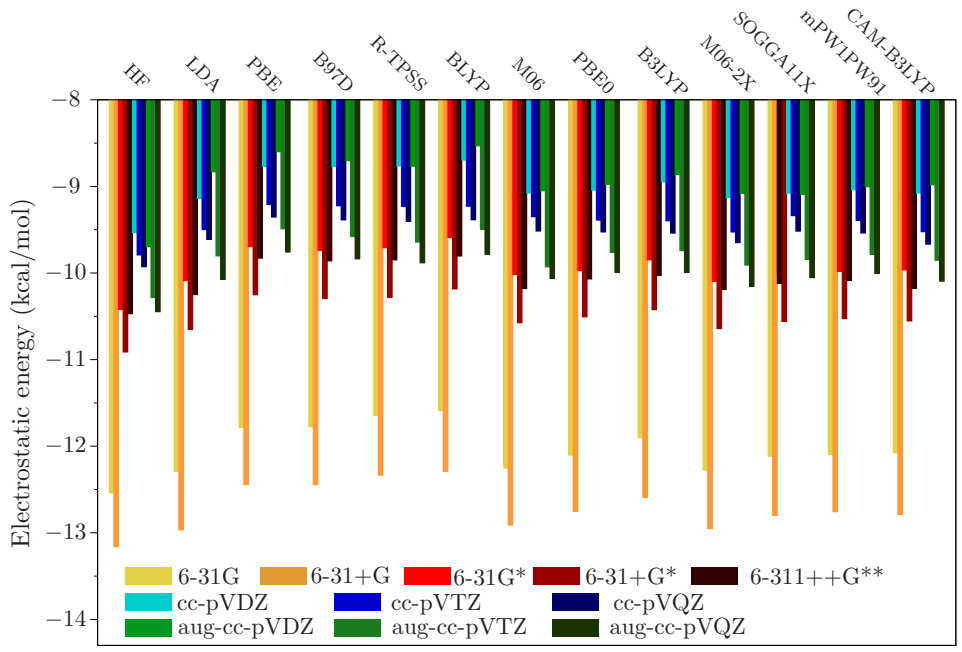


Figure 7.6. Dependence of E_{ele} on the choice of basis set and QM method for the water dimer depicted in Figure 7.2 with $d = 2.94$ Å.

and computational cost.

7.4.4 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 occurr, 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 7.6.4 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 7.7, where solute-solvent HBs are sketched. All extracted structures are reported in Figures 7.12-7.15 in SI. The choice of the spheres' radius is justified by the analysis of the Radial Distribution Functions $g(r)$ (see Figure 7.11 in SI), 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 7.1.

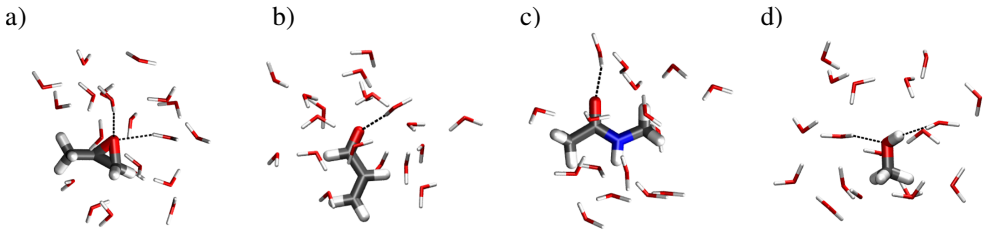


Figure 7.7. Sample structures obtained by cutting a sphere of 5.0 Å around **a)** (R)-methyloxirane; **b)** acrolein; **c)** N-methyl acetamide; **d)** methanol.

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 7.1. 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,[156] QM/FQ^b [162] and QM/FQ^c [203] were considered. QM/FQ and QM/FQFμ were compared with full-QM electrostatic energies calculated by exploiting SAPT0/6-31+G* (see Figure 7.8, the corresponding raw data are given in Tables 7.5-7.6 in SI). In both QM/FQ and QM/FQFμ calculations, the QM portion was described at the HF/6-31+G* level, and the charge constraint in Eq. 7.22 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 Table 7.5 in SI.

The comparison between polarizable QM/MM and SAPT0 E_{ele} are graphically depicted in Figure 7.8. RMSD, Maximum Absolute Error (MAE) and Relative Error (RE) on the ten selected structures are reported in Table 7.2. 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 7.12 in SI). 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.,[156] i.e. QM/FQ^a, whereas the best agreement is given by our recent parametrization,[203] i.e. QM/FQ^c (see also Table 7.2). 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.[156]) or QM atomic charges (FQ^b, ref.[162]). As depicted in Figure 7.8, QM/FQFμ overperforms QM/FQ. This is also confirmed by the data reported in Table 7.2, where a RMSD of only 1.67 kcal/mol and an RE of 6.44 % 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 7.6 in SI. 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 7.6.4 in SI.

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 7.8 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. Notice also that, among all tested FQ parametrizations, FQ^c does not give the best results for MeOH, as the MAE is 38 kcal/mol. 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 7.2 and Figure 7.8, 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 7.2, 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.[184]

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 7.12 and 7.15, given as SI) are analyzed. In Figures 7.9 and 7.10, 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).[556,557] $MOXY_1$ 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

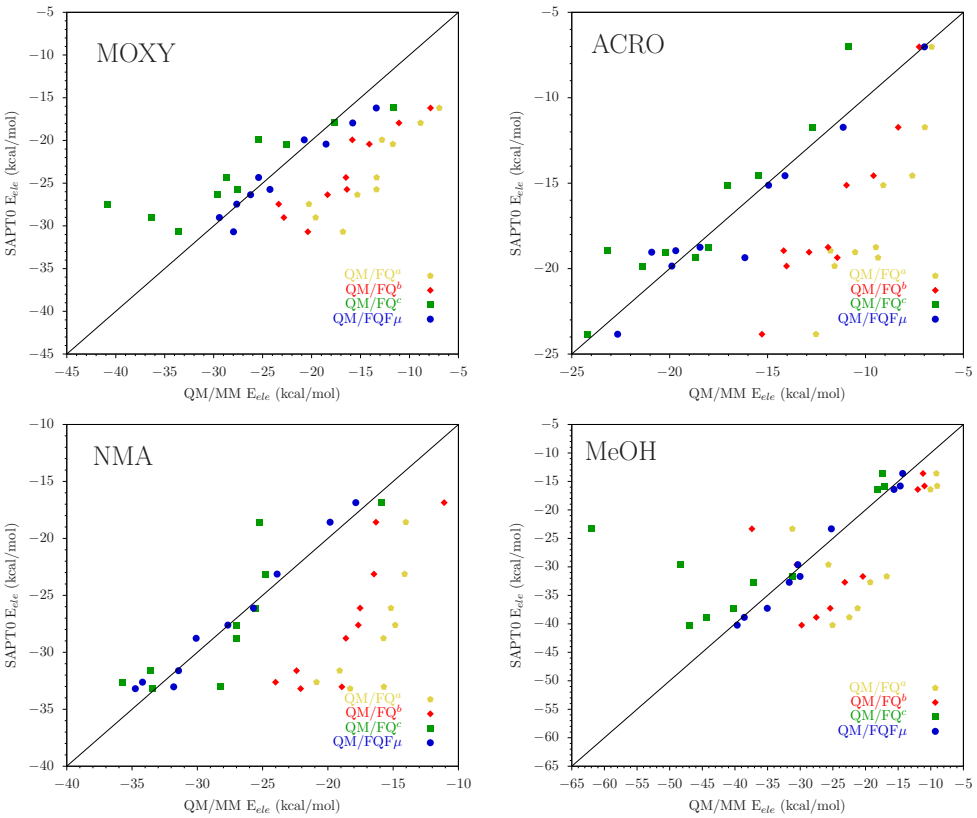


Figure 7.8. Comparison between calculated QM/FQ^a, QM/FQ^b, QM/FQ^c, QM/FQFμ E_{ele} (HF/6-31+G* level for the QM portion) and SAPT0/6-31+G* data. In case of SAPT0 calculations electrostatic and induction energy contributions are summed up. Raw data are given in Table 7.5 in SI. All data are reported in kcal/mol.
^{a} FQ parametrization taken from Ref. [156]
^{b} FQ parametrization taken from Ref. [162]
^{c} FQ parametrization taken from Ref. [203]

		MOXY	ACRO	NMA	MeOH	TOT
QM/FQ[a]	RMSD	10.12	7.83	11.93	11.59	10.49
	MAE	13.89	11.32	17.32	16.40	17.32
	RE	42.51%	40.66%	41.28%	37.28%	40.43%
QM/FQ[b]	RMSD	7.40	5.75	9.16	9.19	8.00
	MAE	10.31	8.55	14.10	14.09	14.10
	RE	31.00%	29.41%	31.10%	29.04%	30.14%
QM/FQ[c]	RMSD	5.75	2.08	2.96	14.04	7.80
	MAE	13.41	4.22	6.63	38.65	38.65
	RE	18.90%	12.75%	8.99%	33.03%	18.42%
QM/FQFμ	RMSD	1.67	1.29	1.07	1.25	1.34
	MAE	2.81	3.20	1.58	2.23	3.20
	RE	6.44%	4.67%	3.61%	4.41%	4.78%

Table 7.2. 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. SAPT0/6-31+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. [156]
[b] FQ parametrization taken from Ref. [162]
[c] FQ parametrization taken from Ref. [203]

acts as H-donor and a second one as H-acceptor. Figures 7.9 and 7.10 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.

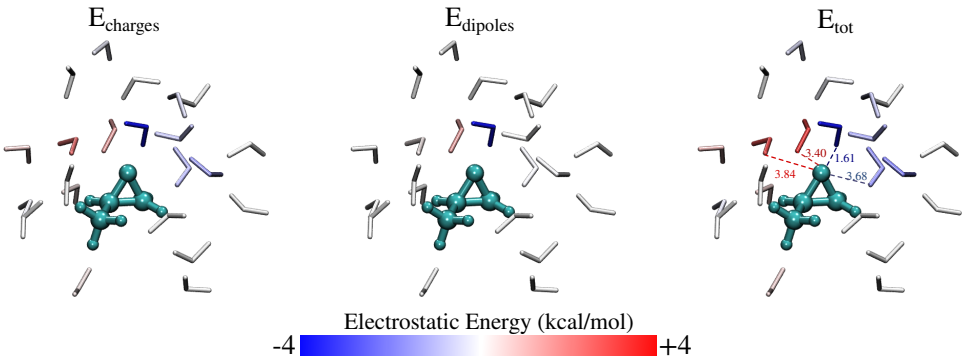


Figure 7.9. QM/FQFμ electrostatic energy contributions (kcal/mol) for $MOXY_1$. $E_{charges}$ and $E_{dipoles}$ indicate charge and dipole contributions to the total E_{ele}. All atoms in each water molecule are colored according to their contribution. The color maps saturate at ± 4 kcal/mol.

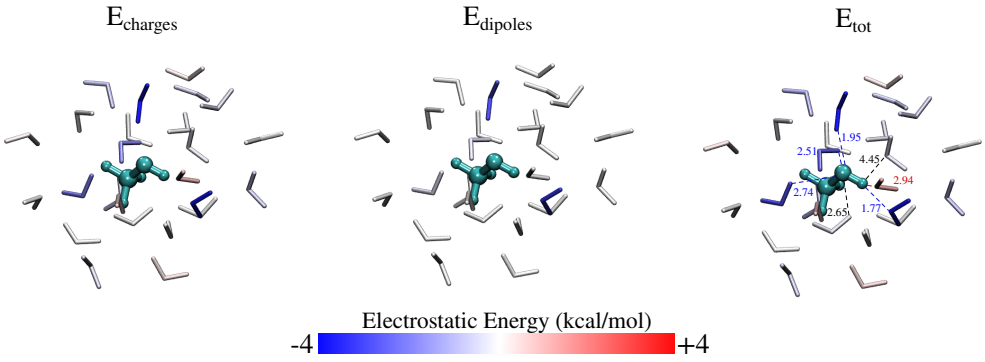


Figure 7.10. QM/FQFμ electrostatic energy contributions (kcal/mol) for MeOH-2. $E_{charges}$ and $E_{dipoles}$ indicate charge and dipole contributions to the total E_{ele}. All atoms in each water molecule are colored according to their contribution. The color maps saturate at ± 4 kcal/mol.

7.5 Summary and Conclusions

In this paper, a new polarizable force field, FQFμ, has been proposed and coupled to a QM SCF Hamiltonian. The peculiarity of QM/FQFμ stands in the fact the polarization of the MM portion is modelled 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 approach is an extension of the QM/FQ which we have developed in recent years,[54–56,149–151,161] in which only fluctuating charges are used to describe the polarization of the environment. Differences and analogies with previously developed methods and a comparison between Drude Oscillators and Fluctuating dipoles has been discussed, pointing out the novelty and the computational features of our approach.

QM/FQFμ has been parametrized in order to reproduce electrostatic energies of aqueous solutions. Then, such an approach has been tested against the reproduction of electrostatic energy of a water dimer as a function of the O-O distance, as well as its total interaction energy. QM/FQFμ has also been coupled with a model that we have recently proposed to account for non-electrostatic energy terms , and it has been shown to appropriately reproduce CCSD(T) equilibrium geometry and the corresponding interaction energy for the same water dimer.

Finally, QM/FQFμ has been applied to the calculation of electrostatic energies of four molecules in aqueous solution. Such molecules were chosen by considering the specific interactions that they can form with the surrounding water molecules (i.e. H-acceptor or H-donor). QM/FQFμ overcomes the limits of QM/FQ, giving a better agreement with reference full QM SAPT0 data. However, all the tested methods are in better agreement with full QM data than what has been shown for QM/AMOEBA on different aqueous systems.[184] The large errors reported for QM/AMOEBA E_{ele} values,[184] have been ascribed to the permanent electrostatic contribution (fixed charges and quadrupoles). Our results seem to show that charge polarization is indeed crucial to lower the errors with respect to full QM values, whereas the inclusion of dipole contributions refines the quality of the results.

To end this discussion we point out that, thanks to its variational formulation, QM/FQFμ can be extended to molecular properties/spectroscopies by following the same strategy which has been proposed by some of us for QM/FQ.[54,55,149,150,161] Such an extension, as well as model parametrization for non-aquous environments, will be the topic of future studies.

7.6 Supporting Information

Equations for FQFμ force field without charge transfer between MM molecules. QM/PQEq model. Details on the parametrization of aqueous solution. Raw data for the dependence on the level of theory of the water dimer. Details on MD runs of MOXY, ACRO, NMA and MeOH in aqueous solution. Structures of solute-solvent clusters. Raw data of electrostatic interaction energies for solute-solvent clusters. Analysis of MOXY-water cluster obtained by exploiting a cutting radius of 7 Å.

7.6.1 FQFμ model without charge transfer between MM molecules

If each molecule is constrained to assume a fixed, total charge Q_α, the energy functional F (eq. 14) can be written by exploiting some Lagrangian multipliers (λ_α), whose number is equal to the total number of molecules in the MM portion.

$$\begin{aligned} F(\mathbf{q},\boldsymbol{\mu},\lambda_\alpha) &= E(\mathbf{r},\mathbf{q},\boldsymbol{\mu}) + \sum_\alpha \left[\lambda_\alpha \sum_i (q_{\alpha i}) - Q_\alpha\right] = \\ &= \frac{1}{2}\sum_{i\alpha}\sum_{j\beta} q_{i\alpha}\mathrm{T}^{qq}_{i\alpha,j\beta}q_{j\beta} + \frac{1}{2}\sum_i\sum_j \boldsymbol{\mu}^\dagger_{i\alpha}\mathbf{T}^{\mu\mu}_{i\alpha,j\beta}\boldsymbol{\mu}_{j\beta} + \sum_i\sum_j q_{i\alpha}\mathbf{T}^{q\mu}_{i\alpha,j\beta}\boldsymbol{\mu}^\dagger_{j\beta} + \\ &+ \sum_{i\alpha} q_{i\alpha}\chi_{i\alpha} + \sum_\alpha \lambda_\alpha\left[\sum_i q_{\alpha i} - Q_\alpha\right] = \\ &= \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q} \end{aligned} \qquad \textbf{7.27}$$

where α and β runs over the molecules and the constraints λ_α are meant to preserve the total charge Q_α of every molecule. Therefore, the conditions for the constrined minimum are found by imposing the derivatives of F with respect all the variables to be zero, resulting in the following linear problem:

$$\begin{cases} \sum_{j\beta}\mathbf{T}^{qq}_{i\alpha,j\beta}q_{j\beta} + \lambda_\alpha + \sum_{j\beta}\mathbf{T}^{q\mu}_{i\alpha,j\beta}\boldsymbol{\mu}_{j\beta} = -\chi_{i\alpha} \\ \sum_{j\beta}\mathbf{T}^{\mu\mu}_{i\alpha,j\beta}\boldsymbol{\mu}_{j\beta} + \sum_{j\beta}\mathbf{T}^{q\mu}_{i\alpha,j\beta}q_{j\beta} = 0 \\ \sum_{i\alpha} q_{\alpha i} = Q_\alpha \end{cases} \qquad \textbf{7.28}$$

The whole system can be recast in a more compact form as:[542,547]

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_{\boldsymbol{\lambda}} & \mathbf{T}^{q\mu} \\ \mathbf{1}^\dagger_{\boldsymbol{\lambda}} & \mathbf{0} & \mathbf{0} \\ -\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ \mathbf{Q} \\ \mathbf{0} \end{pmatrix} \quad \Rightarrow \quad \mathbf{D}\mathbf{Q}_\lambda = -\mathbf{C}_Q \qquad \textbf{7.29}$$

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 $\mathbf{Q}_\lambda$ is a vector containing charges, dipoles and Lagrange muiltipliers.

7.6.2 QM/PQEq approach

PQEq force field consists of a combination of the Charge Equilibration model (QEq)[287] and the Drude Oscillator one.[173] Each atom is considered to be composed of core and shell, on which Gaussian charge distributions are placed. In particular, the core is constituited by a fluctuating charge (q) and a fixed one ($+Z$), which is connected trough an isotropic harmonic spring to the shell fixed charge ($-Z$), thus providing the possibility of having variable displacements of the electron charge distribution. The formal equations defining such a force field can be found in Ref.[543] In this section, we will strictly follow their notation.

The coupling of the PQEq force field with a QM density can be expressed through Coulomb interaction:

$$E_{\mathrm{QM/PQEq}} = \sum_i \phi[\rho_{\mathrm{QM}}](\mathbf{r}_{ic}) q_i + \sum_i \phi[\rho_{\mathrm{QM}}](\mathbf{r}_{ic}) Z_i - \sum_i \phi[\rho_{\mathrm{QM}}](\mathbf{r}_{is}) Z_i \qquad \textbf{7.30}$$

where $\phi[\rho_{\mathrm{QM}}](\mathbf{r}_i)$ is the electrostatic potential due to the QM charge density, calculated in the positions of the PQEq charges. $\mathbf{r}_{ic}$ and $\mathbf{r}_{is}$ indicate core and shell positions. The QM potential can be devided into the nuclear contribution and the electronic one as expressed in Eq. 18.
Thus, Eq. 7.30 can be rewritten as:

$$E_{\mathrm{QM/PQEq}} = \sum_i q_i V(\mathbf{P})_{ic} + \sum_i Z_i \left(V(\mathbf{P})_{ic} - V(\mathbf{P})_{is} \right)$$

where $V(\mathbf{P})_{ik}$ is the potential calculated in $\mathbf{r}_{ik}$ (where $k = c$ or $k = s$).

The energy of the whole QM/PQEq system is the sum of three terms, $E = E_{\mathrm{QM}} + E_{\mathrm{PQEq}} + E_{\mathrm{QM/PQEq}}$. Thus, the complete energy functional $\mathcal{F}$ reads:

$$\begin{aligned} \mathcal{F}(\mathbf{P}, \mathbf{q}, \lambda, \mathbf{r}_{is}) = & tr\mathbf{hP} + \frac{1}{2} tr\mathbf{PG}(\mathbf{P}) + \mathbf{q}^\dagger \boldsymbol{\chi} + \frac{1}{2} \sum_i \left[\eta_{ii} q_i^2 + K_s \left| \mathbf{r}_{ic} - \mathbf{r}_{is} \right|^2 \right] + \\ & + \lambda \mathbf{q} + \sum_{i>j} \left[C(\mathbf{r}_{ic,jc}) q_{ic} q_{jc} - C(\mathbf{r}_{ic,js}) q_{ic} Z_j - C(\mathbf{r}_{is,jc}) q_{jc} Z_i + C(\mathbf{r}_{is,js}) Z_i Z_j \right] + \\ & + \sum_i q_i V(\mathbf{P})_{ic} + \sum_i Z_i \left(V(\mathbf{P})_{ic} - V(\mathbf{P})_{is} \right) \end{aligned} \qquad \textbf{7.31}$$

Where the expression for the MM portion has been taken from ref.[543] $C(\mathbf{r}_{ik,jk}), k = c, s$ represents the interaction between core and shell charges. The lagrangian multiplier λ is imposed so that the charge of the whole MM portion is fixed to Q_{tot}. The contributions to the Fock matrix are obtained by differentiating Eq. 7.31 with respect to the density matrix elements:

$$\tilde{F}_{\mu\nu} = \frac{\partial \mathcal{E}}{\partial P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{V}^\dagger_{\mu\nu,c} \mathbf{q} + \mathbf{V}^\dagger_{\mu\nu,c} \mathbf{Z} - \mathbf{V}^\dagger_{\mu\nu,s} \mathbf{Z} \qquad \textbf{7.32}$$

where, $\mathbf{q}$ and $\mathbf{Z}$ are the vectors containing fluctuating and fixed charges, respectively, wheareas c and s pedix indicates core and shell positions where the QM potential is calculated.
The derivative with respect to the fluctuating charges q_i results in linear equation, similar to what has been reported for QM/FQ approach:[54]

$$\begin{pmatrix} \mathbf{H} & \mathbf{1} \\ \mathbf{1}^\dagger & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \lambda \end{pmatrix} = \begin{pmatrix} -\mathbf{A} \\ Q \end{pmatrix} + \begin{pmatrix} -\mathbf{V} \\ 0 \end{pmatrix} = \begin{pmatrix} -\mathbf{B} \\ Q \end{pmatrix} \qquad \textbf{7.33}$$

that is:

$$\mathbf{D}\mathbf{q}_\lambda = \mathbf{K} \qquad \textbf{7.34}$$

where, **H** and **A** are defined as:

$$H_{ij} = \eta_{ii}\delta_{ij} + (1 - \delta_{ij})C(\mathbf{r}_{ic,jc}) \quad \textbf{7.35}$$

$$A_i = \chi_i + \sum_{i>j} \left[C(\mathbf{r}_{ic,jc}) - C(\mathbf{r}_{ic,js})\right] Z_j \quad \textbf{7.36}$$

Notice that the equations are equal to the ones proposed in QM/FQ approach, with the only difference in **A**. In QM/FQ approach, in fact,only electronegativity χ enter in such a definition, whereas in QM/PQEq approach an additional polarization term arises from the other MM charges. By inverting **D**, both the fluctuating charges and the Lagrangian multiplier are calculated.
The derivative with respect to the shell positions is a further condition which needs to be satisfied to minimise the energy functional $\mathcal{F}$. This results in an equation in which the forces acting on each shell mobile charge must be equal to zero. The forces acting on each shell charges can be devided in three main contributions: $\mathbf{F}_{\text{intra}}$, the spring force, $\mathbf{F}_{\text{inter}}$, the electrostatic forces due to all the other MM charges, and $\mathbf{F}_{\text{QM}}$, the QM force, i.e. the QM electric field.

$$\mathbf{F}_{\text{intra}}(\mathbf{r}_{is}) + \mathbf{F}_{\text{inter}}(\mathbf{r}_{is}) + \mathbf{F}_{\text{QM}}(\mathbf{r}_{is}) = 0 \quad \textbf{7.37}$$

where

$$\mathbf{F}_{\text{QM}} = -\boldsymbol{\nabla}_{r_{is}} \sum_i V(\mathbf{P})_{is}$$

To obtain the electric field produced by the QM density acting on the PQEq charges, the derivative of the QM potential energy can be performed:

$$V_i^{(\xi)}(\mathbf{P}) = \sum_{\zeta=1}^{N_n} \frac{Z_\zeta(\mathbf{r}_i - \mathbf{R}_\zeta)}{|\mathbf{r}_i - \mathbf{R}_\zeta|^3} + \sum_{\mu\nu} P_{\mu\nu} \langle \chi_\mu \chi_\nu | \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} |\delta(\mathbf{r}' - \mathbf{r}_i)\rangle = \mathbf{E}_i(\mathbf{P}) \quad \textbf{7.38}$$

7.6.3 Parametrization

Atom	χ	η	α
O	0.321226	0.718474	1.645295
H	0.166401	0.921014	0.424619

Table 7.3. QM/FQFμ electronegativities (χ), chemical hardnesses (η) and polarizabilities (α) for Oxygen and Hydrogen atoms in water molecule.

Structure	$E^{ele}_{QM/FQF\mu}\|_q$	$E^{ele}_{QM/FQF\mu}\|_\mu$	$E^{ele}_{QM/FQF\mu}$	E_{KM-EDA}	Error %
Dimer 1_{HDon}	-0.011826 (70%)	-0.004980 (30%)	-0.016806	-0.017678	4.93
Dimer 1_{HAcc}	-0.013376 (72%)	-0.005139 (28%)	-0.018515	-0.017678	4.73
Dimer 2_{HDon}	-0.009051 (72%)	-0.003443 (28%)	-0.012494	-0.012329	1.34
Dimer 2_{HAcc}	-0.009052 (72%)	-0.003445 (28%)	-0.012496	-0.012329	1.36
Dimer 3_{HDon}	-0.007275 (71%)	-0.002986 (29%)	-0.010260	-0.010862	5.54
Dimer 3_{HAcc}	-0.007626 (70%)	-0.003218 (30%)	-0.010844	-0.010862	0.17

Table 7.4. Comparison of calculated QM/FQFμ and reference KM-EDA/6-31+G* electrostatic+polarization energies in a.u. HF/6-31+G* was used in QM/FQFμ calculations. $E^{ele}_{QM/FQF\mu}|_q$ and $E^{ele}_{QM/FQF\mu}|_\mu$ indicate charge and dipole contributions to E_{ele}, respectively.

7.6.4 Molecules in aqueous solution

Classical MD simulations

The geometry of (R)-methyloxirane (MOXY), Acrolein (ACRO), N-Methylacetamide (NMA) and Methanol (MeOH) was optimized at the B3LYP/ aug-cc-pVDZ level and the C-PCM[70] to represent the aqueous environment ($\varepsilon = 78.3553$). Following what has already been reported by some of the present authors,[146] in order to obtain a representative conformational sampling of solvated solutes, 10 ns MD simulations were performed in a pre-equilibrated box of 2865 SPC (Single Point Charge) water molecules in the NPT (isothermal-isobaric) ensemble using GROMACS.[348,471–474] All bonds were kept rigid using the Settle algorithm[558] for water; the geometry of the solute was kept rigid during the simulation. Electrostatic interactions were considered through the the Particle Mesh Ewald summation method.[559] The pressure was stabilized at 1 bar using the weak-coupling scheme with a coupling constant of 10 ps and an isotherm compressibility of $5 \cdot 10^{-5}\text{bar}^{-1}$. Each component of the system (both methyloxyrane and water) was coupled separately to a temperature bath at 300 K using the Berendsen thermostat[560] with a coupling constant of 0.5 ps. The all-atoms OPLS-AA (Optimized Potentials for Liquid Simulations - All Atoms) force field[297] was used for the solute

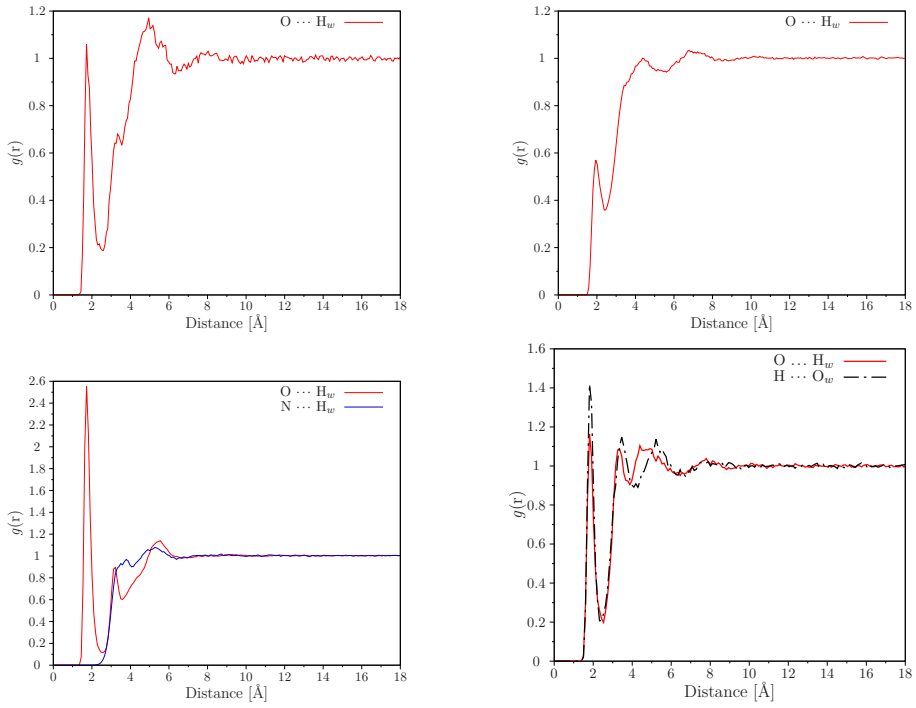


Figure 7.11. MOXY, ACRO, NMA, MeOH radial distribution functions between solulte nitrogen/oxygen atoms and water hydrogen.

Structures of solute-waters clusters

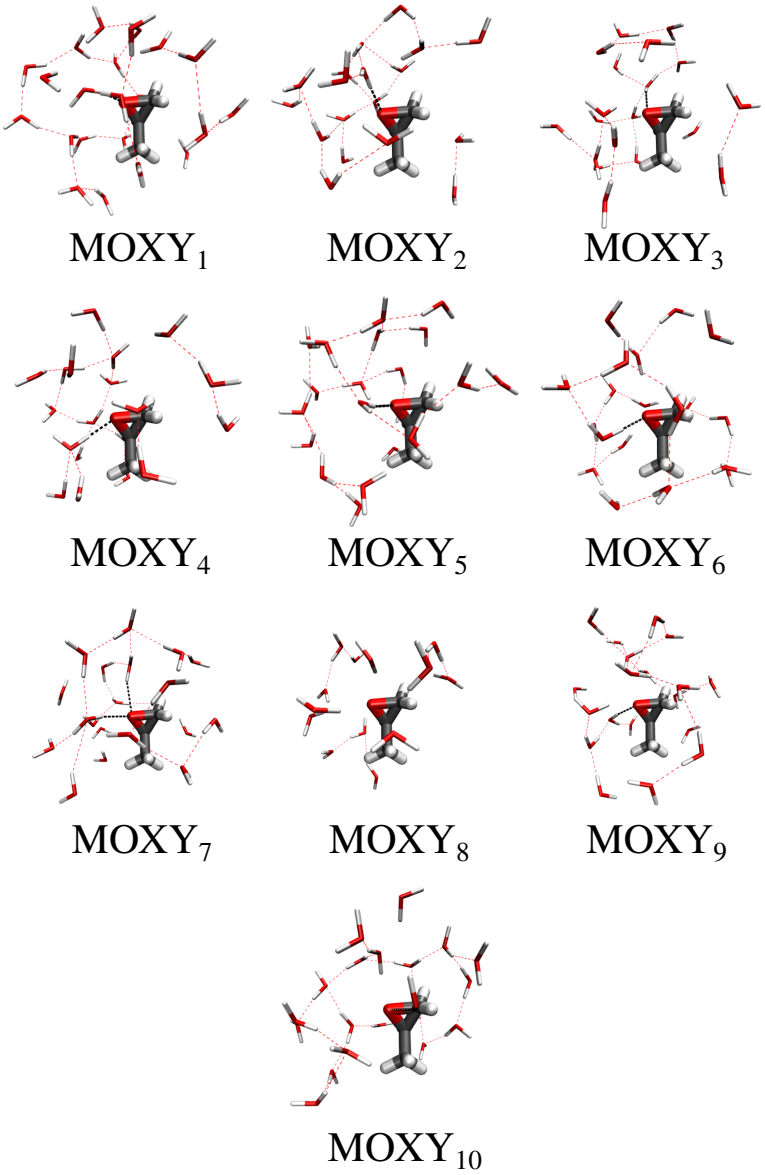


Figure 7.12. MOXY-water structures

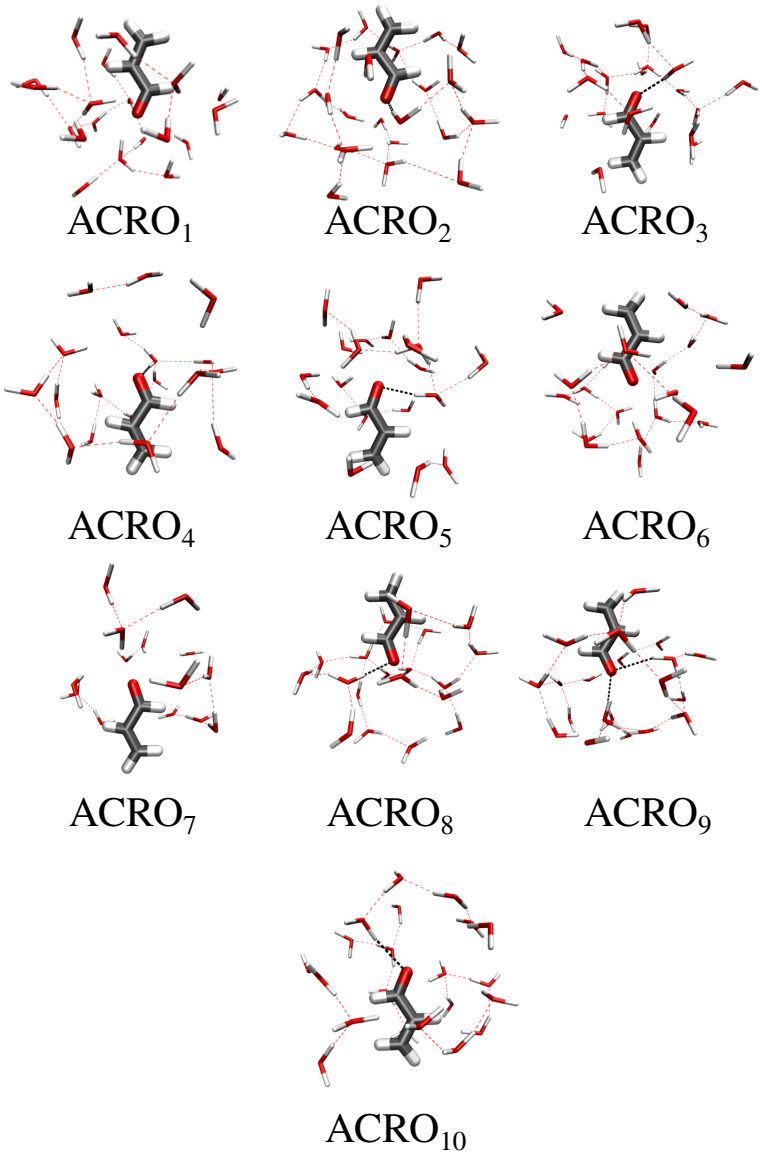


Figure 7.13. ACRO-water structures

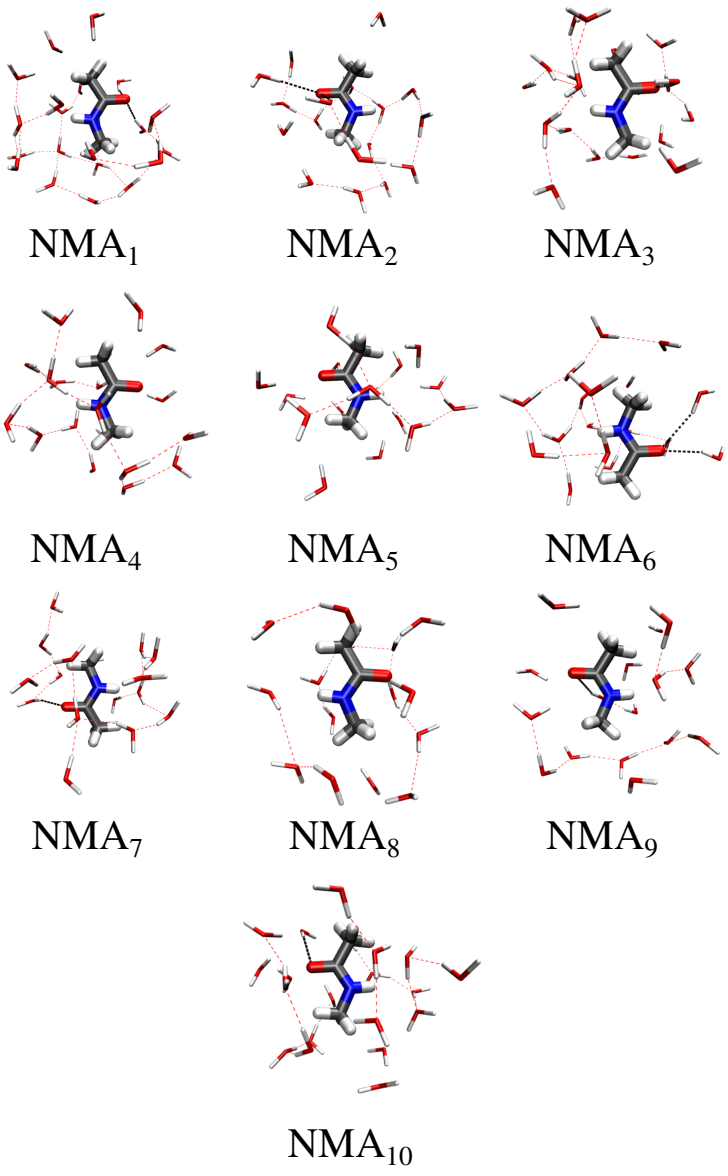


Figure 7.14. NMA-water structures

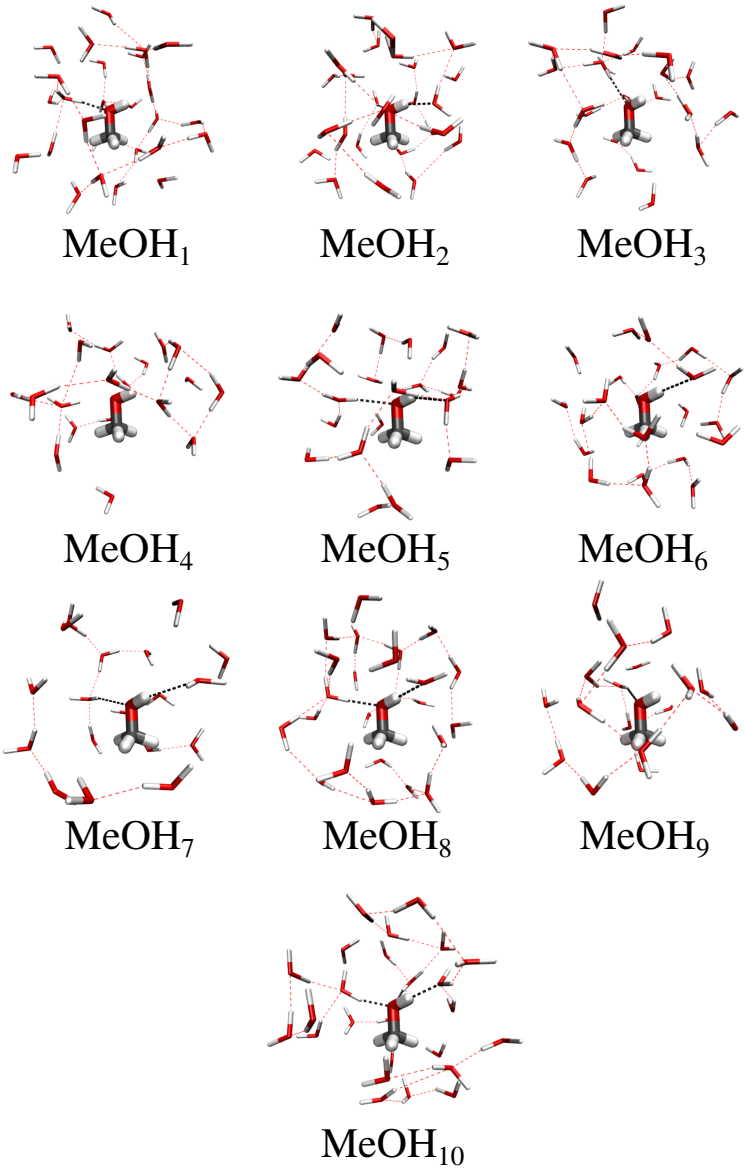


Figure 7.15. MeOH-water structures

Electrostatic energy with CT between solvent molecules

Structure	QM/FQF$\mu\|_q$	QM/FQF$\mu\|_\mu$	QM/FQF$\mu\|_{\text{tot}}$	SAPT0
$MOXY_1$	-21.98 (80%)	-5.66 (20%)	-27.64	-27.45
$MOXY_2$	-19.04 (75%)	-6.37 (25%)	-25.41	-24.33
$MOXY_3$	-21.55 (77%)	-6.43 (23%)	-27.97	-30.69
$MOXY_4$	-17.98 (74%)	-6.27 (26%)	-24.25	-25.73
$MOXY_5$	-22.76 (77%)	-6.65 (23%)	-29.41	-29.02
$MOXY_6$	-19.82 (76%)	-6.41 (24%)	-26.23	-26.35
$MOXY_7$	-15.75 (76%)	-5.00 (24%)	-20.75	-19.93
$MOXY_8$	-11.91 (75%)	-3.89 (25%)	-15.80	-17.95
$MOXY_9$	-14.38 (78%)	-4.13 (22%)	-18.52	-20.43
$MOXY_{10}$	-9.70 (72%)	3.69 (28%)	-13.39	-16.20
$ACRO_1$	-16.82 (74%)	-5.84 (26%)	-22.66	-23.84
$ACRO_2$	-15.01 (76%)	-4.67 (24%)	-19.69	-18.95
$ACRO_3$	-5.64 (81%)	1.34 (19%)	-6.99	-7.02
$ACRO_4$	-12.13 (75%)	-4.03 (25%)	-16.16	-19.36
$ACRO_5$	-15.12 (72%)	-5.80 (28%)	-20.92	-19.04
$ACRO_6$	-14.84 (75%)	-5.04 (25%)	-19.89	-19.85
$ACRO_7$	-13.60 (74%)	-4.85 (26%)	-18.45	-18.75
$ACRO_8$	-11.18 (75%)	-3.77 (25%)	-14.95	-15.12
$ACRO_9$	-10.24 (73%)	-3.87 (27%)	-14.11	-14.56
$ACRO_{10}$	-8.29 (74%)	2.86 (26%)	-11.14	-11.73
$MeOH_1$	-26.61 (76%)	-8.44 (24%)	-35.05	-37.27
$MeOH_2$	-28.82 (75%)	-9.74 (25%)	-38.57	-38.86
$MeOH_3$	-22.91 (75%)	-7.45 (25%)	-30.36	-29.62
$MeOH_4$	-10.89 (76%)	-3.42 (24%)	-14.31	-13.60
$MeOH_5$	-11.92 (76%)	-3.71 (24%)	-15.63	-16.40
$MeOH_6$	-23.89 (75%)	-7.79 (25%)	-31.68	-32.70
$MeOH_7$	-11.05 (75%)	-3.64 (25%)	-14.68	-15.80
$MeOH_8$	-30.17 (76%)	-9.50 (24%)	-39.66	-40.23
$MeOH_9$	-22.45 (75%)	-7.58 (25%)	-30.03	-31.68
$MeOH_{10}$	-19.59 (78%)	-5.64 (22%)	-25.23	-23.32
NMA_1	-20.76 (75%)	-6.89 (25%)	-27.66	-27.61
NMA_2	-22.61 (75%)	-7.48 (25%)	-30.09	-28.77
NMA_3	-26.03 (76%)	-8.17 (24%)	-34.20	-32.62
NMA_4	-23.77 (75%)	-8.04 (25%)	-31.81	-33.03
NMA_5	-13.50 (76%)	-4.36 (24%)	-17.86	-16.86
NMA_6	-26.43 (76%)	-8.32 (24%)	-34.75	-33.19
NMA_7	-19.79 (77%)	-5.90 (23%)	-25.70	-26.12
NMA_8	-15.43 (78%)	-4.40 (22%)	-19.83	-18.58
NMA_9	-18.40 (77%)	-5.49 (23%)	-23.89	-23.13
NMA_{10}	-24.30 (77%)	-7.14 (23%)	-31.44	-31.61

Table 7.5. Calculated electrostatic energies by exploiting QM/6-31+G*/FQFμ and SAPT0/6-31+G*. QM/GQ$\mu|_q$ represents the electrostatic contribution arising from the charges, whereas QM/GQ$\mu|_\mu$ that arising from the dipoles.

	QM/FQ[a]	QM/FQ[b]	QM/FQ[c]	QM/FQFμ	SAPT0
$MOXY_1$	-20.27	-23.36	-40.86	-27.64	-27.45
$MOXY_2$	-13.35	-16.49	-28.70	-25.41	-24.33
$MOXY_3$	-16.80	-20.38	-33.60	-27.97	-30.69
$MOXY_4$	-13.37	-16.39	-27.56	-24.25	-25.73
$MOXY_5$	-19.58	-22.83	-36.36	-29.41	-29.02
$MOXY_6$	-15.32	-18.37	-29.65	-26.23	-26.35
$MOXY_7$	-12.81	-15.84	-25.39	-20.75	-19.93
$MOXY_8$	-8.87	-11.07	-17.66	-15.80	-17.95
$MOXY_9$	-11.70	-14.10	-22.60	-18.52	-20.43
$MOXY_{10}$	-6.95	-7.85	-11.66	-13.39	-16.20
$ACRO_1$	-12.53	-15.29	-24.18	-22.66	-23.84
$ACRO_2$	-11.79	-14.18	-23.16	-19.69	-18.95
$ACRO_3$	-6.62	-7.26	-10.89	-6.99	-7.02
$ACRO_4$	-9.36	-11.44	-18.66	-16.16	-19.36
$ACRO_5$	-10.53	-12.88	-20.21	-20.92	-19.04
$ACRO_6$	-11.58	-14.03	-21.38	-19.89	-19.85
$ACRO_7$	-9.47	-11.91	-18.02	-18.45	-18.75
$ACRO_8$	-9.09	-10.97	-17.04	-14.95	-15.12
$ACRO_9$	-7.61	-9.59	-15.44	-14.11	-14.56
$ACRO_{10}$	-6.97	-8.33	-12.72	-11.14	-11.73
$MeOH_1$	-21.21	-25.40	-40.26	-35.05	-37.27
$MeOH_2$	-22.47	-27.54	-44.33	-38.57	-38.86
$MeOH_3$	-25.68	-30.52	-48.37	-30.36	-29.62
$MeOH_4$	-9.14	-11.19	-17.42	-14.31	-13.60
$MeOH_5$	-10.05	-11.99	-18.20	-15.63	-16.40
$MeOH_6$	-19.28	-23.17	-37.17	-31.68	-32.70
$MeOH_7$	-9.02	-10.97	-17.09	-14.68	-15.80
$MeOH_8$	-25.05	-29.78	-47.03	-39.66	-40.23
$MeOH_9$	-16.75	-20.40	-31.24	-30.03	-31.68
$MeOH_{10}$	-31.21	-37.40	-61.96	-25.23	-23.32
NMA_1	-14.84	-17.67	-26.98	-27.66	-27.61
NMA_2	-15.74	-18.62	-27.03	-30.09	-28.77
NMA_3	-20.87	-24.01	-35.71	-34.20	-32.62
NMA_4	-15.71	-18.93	-28.25	-31.81	-33.03
NMA_5	-9.61	-11.09	-15.89	-17.86	-16.86
NMA_6	-18.29	-22.09	-33.42	-34.75	-33.19
NMA_7	-15.17	-17.53	-25.55	-25.70	-26.12
NMA_8	-14.03	-16.32	-25.20	-19.83	-18.58
NMA_9	-14.12	-16.47	-24.76	-23.89	-23.13
NMA_{10}	-19.10	-22.40	-33.57	-31.44	-31.61

Table 7.6. Comparison between calculated QM/FQ[a], QM/FQ[b], QM/FQ[c], QM/FQFμ E_{ele} (HF/6-31+G* level for the QM portion) and SAPT0/6-31+G* data. In case of SAPT0 calculations electrostatic and induction energy contribution are summed up. All data are given in kcal/mol.
[a] FQ parametrization taken from Ref. [156]
[b] FQ parametrization taken from Ref. [162]
[c] FQ parametrization taken from Ref. [203]

Electrostatic energies without CT between solvent molecules

	QM/FQ[a]	QM/FQ[b]	QM/FQ[c]	QM/FQFμ	SAPT0
$MOXY_1$	-10.73	-13.10	-23.74	-18.46	-27.45
$MOXY_2$	-12.27	-16.62	-32.73	-22.86	-24.33
$MOXY_3$	-13.65	-17.58	-32.29	-24.28	-30.69
$MOXY_4$	-12.53	-16.72	-30.78	-22.73	-25.73
$MOXY_5$	-13.45	-16.61	-28.71	-22.93	-29.02
$MOXY_6$	-13.01	-16.90	-30.69	-22.53	-26.35
$MOXY_7$	-9.33	-12.06	-20.40	-16.78	-19.93
$MOXY_8$	-8.11	-10.97	-19.31	-14.32	-17.95
$MOXY_9$	-6.85	-8.40	-13.66	-13.13	-20.43
$MOXY_{10}$	-6.43	-8.53	-15.98	-12.33	-16.20
$ACRO_1$	-11.64	-15.54	-28.13	-21.13	-23.84
$ACRO_2$	-9.29	-12.16	-22.94	-16.05	-18.95
$ACRO_3$	-0.79	-0.55	-0.99	-2.15	-7.02
$ACRO_4$	-9.30	-12.44	-22.52	-15.65	-19.36
$ACRO_5$	-9.30	-12.33	-21.86	-18.06	-19.04
$ACRO_6$	-8.52	-11.26	-19.41	-15.76	-19.85
$ACRO_7$	-10.69	-14.46	-24.49	-18.98	-18.75
$ACRO_8$	-6.24	-7.72	-12.48	-11.64	-15.12
$ACRO_9$	-7.56	-10.28	-18.37	-13.61	-14.56
$ACRO_{10}$	-4.94	-6.51	-11.85	-8.79	-11.73
$MeOH_1$	-17.44	-22.26	-38.12	-30.56	-37.27
$MeOH_2$	-18.79	-24.48	-41.83	-34.18	-38.86
$MeOH_3$	-19.97	-25.04	-41.79	-35.97	-29.62
$MeOH_4$	-6.81	-8.82	-14.35	-11.86	-13.60
$MeOH_5$	-7.25	-9.27	-14.98	-12.90	-16.40
$MeOH_6$	-16.94	-21.67	-36.80	-28.89	-32.70
$MeOH_7$	-8.04	-10.45	-17.87	-13.56	-15.80
$MeOH_8$	-19.10	-24.09	-40.46	-33.69	-40.23
$MeOH_9$	-17.85	-23.61	-40.98	-30.20	-31.68
$MeOH_{10}$	-25.66	-32.80	-57.66	-44.25	-23.32
NMA_1	-10.83	-14.59	-25.86	-22.18	-27.61
NMA_2	-13.21	-17.70	-30.34	-25.02	-28.77
NMA_3	-12.14	-15.73	-26.72	-23.71	-32.62
NMA_4	-13.82	-19.27	-33.98	-27.63	-33.03
NMA_5	-7.77	-10.64	-18.94	-15.08	-16.86
NMA_6	-16.74	-22.49	-39.25	-30.98	-33.19
NMA_7	-12.22	-15.98	-27.98	-21.38	-26.12
NMA_8	-5.56	-7.35	-13.01	-11.89	-18.58
NMA_9	-9.17	-12.25	-21.52	-17.75	-23.13
NMA_{10}	-12.38	-15.89	-26.58	-22.66	-31.61

Table 7.7. Comparison between calculated QM/FQ[a], QM/FQ[b], QM/FQ[c], QM/FQFμ E_{ele} (HF/6-31+G* level for the QM portion) and SAPT0/6-31+G* data. In case of SAPT0 calculations electrostatic and induction energy contribution are summed up. All data are given in kcal/mol.
[a] FQ parametrization taken from Ref.[156]
[b] FQ parametrization taken from Ref.[162]
[c] FQ parametrization taken from Ref.[203]

MOXY$_1$ with a cutting radius of 7 Å

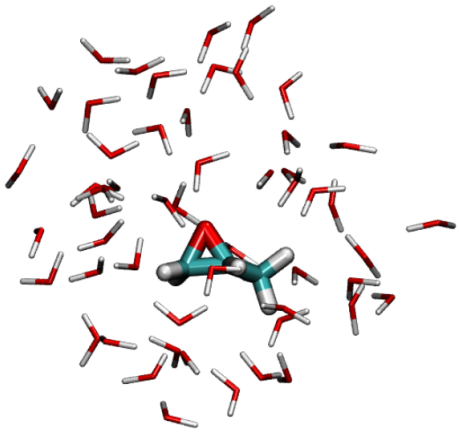


Figure 7.16. MOXY$_1$-cluster structure obtained by using a cutting radius of 7.0 Å. The number of water molecules included in the cluster is 53.

Structure	QM/FQF$\mu|_q$	QM/FQF$\mu|_\mu$	QM/FQF$\mu|_{\text{tot}}$	SAPT0
MOXY$_1$	-22.56	-5.74	-28.31	-27.74

Table 7.8. Calculated electrostatic energies for MOXY$_1$ obtained with a cutting radius of 7 Ådepicted in Figure 7.16 by exploiting QM/6-31G*/FQFμ and SAPT0/6-31G*. QM/GQ$\mu|_q$ represents the electrostatic contribution arising from the charges, whereas QM/GQ$\mu|_\mu$ that arising from the dipoles.

Acknowledgment

We are thankful for the computer resources provided by the high performance computer facilities of the SMART Laboratory (http://smart.sns.it/). CC gratefully acknowledges the support of H2020-MSCA-ITN-2017 European Training Network "Computational Spectroscopy In Natural sciences and Engineering" (COSINE), grant number 765739. T.G. acknowledges Dr. Franco Egidi (SNS) for helpful discussion.

Appendix

In this appendix, QM/FQFμ energy derivatives, together with linear response and electric/magnetic perturbations are formulated.
The resorting to the physical framework of the so-called "focused models" implies that external perturbations (i.e., electric/magnetic fields and/or a nuclear displacement) only act on the QM portion of the system, whereas the environment is only indirectly affected through the perturbation on the QM density. In geometrical derivatives, the

focusing on the QM portion of the system means that the geometric displacements of the MM molecules are not taken into account; this framework is well-defined within the Partial Hessian Vibrational Approach (PHVA).[89,292,293]

Linear Response

Here, we follow a linear response theory for SCF methods, without a complete derivation of QM/MM LR equations, which can be found elsewhere.[54] In this section, electronic transition energies ω_K and densities X_k,Y_k are obtained by solving the modified Casida's equations

$$\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^* & \tilde{\mathbf{A}}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{7.39}$$

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

$$\tilde{A}_{ai,bj} = (\epsilon_a - \epsilon_i)\delta_{ab}\delta_{ij} + \langle aj \| ib \rangle - \sum_{kl}^{N_q} \mathbf{R}^{\dagger}_{ia} \mathbf{D}^{-1} \mathbf{R}_{jb} \tag{7.40}$$

$$\tilde{B}_{ai,bj} = \langle ab \| ij \rangle - \sum_{kl}^{N_q} \mathbf{R}_{ia} \mathbf{D}^{-1} \mathbf{R}_{bj} \tag{7.41}$$

where, $\mathbf{D}$ is defined in Eq. 7.14, whereas $\mathbf{R}$ in Eq. 7.24. In particular, the modified response equations will include two contributions arising from the perturbed FQFμ charges ($\mathbf{V}^{\dagger}_{ai}\mathbf{q}^{(1)}$) and FQF$\mu$ dipoles ($\mathbf{E}^{\dagger}_{ai}\boldsymbol{\mu}^{(1)}$), where, again, the interaction potential/-field will be calculated at the cores positions. For the perturbed FQFμ charges/dipoles, a perturbed FQFμ equation is obtained:

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_{\lambda} & \mathbf{T}^{q\mu} \\ \mathbf{1}^{\dagger}_{\lambda} & 0 & \mathbf{0} \\ -\mathbf{T}^{q\mu\,\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q}^{(1)} \\ \lambda \\ \boldsymbol{\mu}^{(1)} \end{pmatrix} = \begin{pmatrix} -\mathbf{V}^{(1)}(\mathbf{P}) \\ 0 \\ \mathbf{E}^{(1)}(\mathbf{P}) \end{pmatrix} \Rightarrow \mathbf{D}\mathbf{L}^{(1)}_{\lambda} = -\mathbf{R}^{(1)}(\mathbf{P}) \tag{7.42}$$

where the perturbed potential/electric field are defined as in Refs.[54] and.[10]
Notice finally that, if we take the zero frequency limit of the response equations, the static coupled-perturbed Hartree–Fock (CPHF) equations are obtained.

Energy first derivatives

The energy first derivative of Eq. 7.22 can be expressed by means of the chain rule:[561]

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

The last two 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, is:

$$\frac{\partial \mathcal{E}}{\partial x} = \text{tr } \mathbf{h}^x \mathbf{P} + \frac{1}{2} \text{tr } \mathbf{G}^{(x)}(\mathbf{P})\mathbf{P} + \mathbf{q}^{\dagger}\mathbf{V}^{(x)}(\mathbf{P}) - \boldsymbol{\mu}^{\dagger}\mathbf{E}^{(x)}(\mathbf{P}) \tag{7.43}$$

where

$$V_i^{(x)}(\mathbf{P}) = \sum_{\mu\nu} P_{\mu\nu} V^x_{\mu\nu,i} + \text{nuclear contribution} \quad 7.44$$

$$= \frac{Z_\zeta}{|\mathbf{R}_\zeta - \mathbf{r}_i|^2} - \sum_{\mu\nu} \langle \frac{\partial(\chi_\mu \chi_\nu)}{\partial \mathbf{R}_\zeta} | \frac{1}{|\mathbf{r} - \mathbf{r}'|} | \delta(\mathbf{r}' - \mathbf{r}_i) \rangle P_{\mu\nu} \quad 7.45$$

$$E_i^{(x)}(\mathbf{P}) = \nabla_{r_i} V_i^{(x)}(\mathbf{P}) \quad 7.46$$

The term involving the derivatives of the density matrix can be computed starting from the idempotency condition, which gives rise to the usual energy-weighted density matrix contribution:

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

where the subscript *oo* denotes the occupied–occupied block of the matrix in the MO basis. Finally:

$$\mathcal{E}^x(\mathbf{P}, \mathbf{q}, \boldsymbol{\lambda}) = \text{tr } \mathbf{h}^x\mathbf{P} + \frac{1}{2}\text{tr } \mathbf{G}^{(x)}(\mathbf{P})\mathbf{P} + \mathbf{q}^\dagger \mathbf{V}^{(x)}(\mathbf{P}) - \boldsymbol{\mu}^\dagger \mathbf{E}^{(x)}(\mathbf{P}) - \text{tr}\mathbf{W}\mathbf{S}^x_{oo} \quad \mathbf{7.47}$$

Energy second derivatives

The energy second derivative with respect to a perturbation acting on the QM portion of the system can be obtained by differentiating eq. 7.47. By exploiting once again the chain rule and writing explicitly all contributions:

$$\mathcal{E}^{xy} = \sum_{\mu\nu} \left[h^{xy}_{\mu\nu} + \frac{1}{2} G^{(xy)}_{\mu\nu}(\mathbf{P}) + \mathbf{q}^\dagger \mathbf{V}^{xy}_{\mu\nu} - \boldsymbol{\mu}^\dagger \mathbf{E}^{xy}_{\mu\nu} \right] P_{\mu\nu} - \text{tr } \mathbf{W}\mathbf{S}^{xy} - \text{tr } \mathbf{W}^y \mathbf{S}^x$$
$$+ \sum_{\mu\nu} \left[h^x_{\mu\nu} + G^{(x)}_{\mu\nu}(\mathbf{P}) + \mathbf{q}^\dagger \mathbf{V}^x_{\mu\nu} - \boldsymbol{\mu}^\dagger \mathbf{E}^x_{\mu\nu} \right] P^y_{\mu\nu} + \sum_{\mu\nu} \mathbf{L}^{y\dagger} \mathbf{R}^x_{\mu\nu} P_{\mu\nu} \quad 7.48$$

It is hence necessary to compute the derivatives of the off-diagonal blocks of the density matrix and charges/dipoles, which can be done by means of the so-called Coupled Perturbed Hartree–Fock (CPHF) procedure. The FQFμ charge and dipole derivatives can be obtained by differentiating eq. 7.24

$$\mathbf{D}\mathbf{L}^x = -\mathbf{R}^{(x)}(\mathbf{P}) - \mathbf{R}(\mathbf{P}^x) \quad \mathbf{7.49}$$

The Fock matrix derivative is defined as:

$$\widetilde{F}^x_{\mu\nu} = \widetilde{F}^{(x)}_{\mu\nu} + G_{\mu\nu}(\mathbf{P}^x) - \mathbf{R}^\dagger_{\mu\nu} \mathbf{D}^{-1} \mathbf{R}(\mathbf{P}^x) \quad \mathbf{7.50}$$

where, coherently with the usual notation, the term:

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

collects all the explicit derivatives of the Fock matrix.
By rearrangement of the terms, the CPHF equations are obtained: In MO basis, CPHF equations reads:

$$\epsilon_i P^x_{ia} - \epsilon_a P^x_{ia} = -\tilde{Q}_{ia} + \sum_{jb} \left[\langle aj||ib\rangle - \mathbf{R}^\dagger_{ia} \mathbf{D}^{-1} \mathbf{R}_{jb} \right] P^x_{jb}$$
$$+ \sum_{jb} \left[\langle ab||ij\rangle - \mathbf{R}^\dagger_{ia} \mathbf{D}^{-1} \mathbf{R}_{bj} \right] P^x_{bj} \quad 7.51$$

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

$$\tilde{A}_{ia,jb} = (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab} + \langle aj||ib\rangle - \mathbf{R}^{\dagger}_{ia}\mathbf{D}^{-1}\mathbf{R}_{jb} \quad \textbf{7.52}$$

$$\tilde{B}_{ia,jb} = \langle ab||ij\rangle - \mathbf{R}^{\dagger}_{ia}\mathbf{D}^{-1}\mathbf{R}_{bj} \quad \textbf{7.53}$$

one obtains

$$\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^* & \tilde{\mathbf{A}}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{Q} \\ \mathbf{Q}^* \end{pmatrix} \quad \textbf{7.54}$$

where

$$\tilde{A}_{ia,jb} = (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab} + \langle aj||ib\rangle - \mathbf{R}^{\dagger}_{ia}\mathbf{D}^{-1}\mathbf{R}_{jb} \quad \textbf{7.55}$$

$$\tilde{B}_{ia,jb} = \langle ab||ij\rangle - \mathbf{R}^{\dagger}_{ia}\mathbf{D}^{-1}\mathbf{R}_{bj} \quad \textbf{7.56}$$

$$\tilde{Q}_{ia} = \tilde{F}^{(x)}_{ia} - G_{ia}(S^x_{oo}) - \tilde{\mathbf{F}}S^x_{ia} + \mathbf{R}^{\dagger}_{ia}\mathbf{D}^{-1}\mathbf{R}(\mathbf{S}^x_{oo}) \quad \textbf{7.57}$$

The solution of Eqs. 7.54 and 7.49 yields the derivatives of both the density matrix and the FQFμ charges/dipoles with respect to QM nuclear positions, thus allowing the calculation of energy second derivatives. This result is, of course, coherent to the zero-frequency limit of what was obtained above in case of linear response.
To summarize, the FQFμ contributions to analytical second derivatives can be grouped into three categories:

1. explicit contributions:

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

2. contributions to the explicit Fock matrix derivatives:

$$\mathbf{L}^{\dagger}\mathbf{R}^{x}_{\mu\nu} + \mathbf{L}^{(x)\dagger}\mathbf{R}_{\mu\nu}$$

3. contribution to the CPHF matrix:

$$-\mathbf{R}^{\dagger}_{ia}\mathbf{D}^{-1}\mathbf{R}_{jb}$$

The equations simplify when only electric perturbations are considered: only the CPHF contributions are to be added. However, if an oscillating electric field is considered, the static response picture is not formally justified.

Electric and Magnetic Perturbations

Electric Perturbations

In the presence of an external electric field $\mathbf{E}$ and assuming that the FQFμ charges and dipoles are affected by the field only through the response of the QM molecule, a perturbation term must be added to the energy functional:

$$V^{ele} = -\boldsymbol{\mu} \cdot \mathbf{E} = -(\boldsymbol{\mu}^n - \sum_{\mu\nu} P_{\mu\nu}\mathbf{M}_{\mu\nu}) \cdot \mathbf{E}$$

where $\mathbf{M}_{\mu\nu} = \langle\chi_\mu|\mathbf{r}|\chi_\nu\rangle$ are dipole integrals. An electric perturbation gives hence rise to a contribution to the mono-electronic part of the Fock operator of the system:

$$h^{ele}_{\mu\nu} = \mathbf{M}_{\mu\nu} \cdot \mathbf{E}$$

The second derivatives of the energy with respect to the electric field, that correspond to the static polarizability, reduce to:

$$\mathcal{E}^{xy} = \alpha_{xy} = \sum_{\mu\nu} M_{x,\mu\nu} P^y_{\mu\nu} \qquad \textbf{7.58}$$

From eq. 7.58 it is immediately clear that no explicit FQFμ contributions to the polarizability are involved. In the general case of an oscillating electric field, the derivative of the density matrix is obtained by exploiting the Frequency Dependent CPHF (FD-CPHF):

$$\begin{pmatrix} \tilde{\mathbf{Q}}_X \\ \tilde{\mathbf{Q}}_Y \end{pmatrix} + \begin{pmatrix} \tilde{\mathbf{A}} - \omega\mathbb{I} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}} & \tilde{\mathbf{A}} + \omega\mathbb{I} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = 0 \qquad \textbf{7.59}$$

Magnetic perturbations

In the presence of a static magnetic field, which we will assume be given by the sum of a homogeneous magnetic field $\mathbf{B}$ and of a field produced by the magnetic moment $\mathbf{m}_X$ of the nucleus X at position $\mathbf{R}_X$, the magnetic CPHF equations need to be modified. In case of working with London Orbitals, the CPHF Right Hand Side becomes (see Ref.,[161] for further details):

$$\tilde{Q}^{ele}_{ia} = h^{(x)}_{ia} + G^{(x)}_{ia}(\mathbf{P}) + \mathbf{L}^\dagger\mathbf{R}^x_{ia} - G_{ia}(S^x_{oo}) - \mathbf{F}S^x_{ia} \qquad \textbf{7.60}$$

where

$$R^x_{j,ia} = (V^x_{j,ia}, 0, E^x_{j,ia}) = (\langle\phi^x_i|\hat{V}_j|\phi_a\rangle + \langle\phi_i|\hat{V}_j|\phi^x_a\rangle, 0, \langle\phi^x_i|\hat{E}_j|\phi_a\rangle + \langle\phi_i|\hat{V}_j|\phi^x_a\rangle$$

FQFμ contributes only indirectly to the magnetic response.
The magnetic susceptibility is the second derivative of the energy with respect to the magnetic field:

$$\begin{aligned}\chi_{xy} = \frac{\partial^2\mathcal{E}}{\partial B_x \partial B_y} = &\sum_{\mu\nu}\left[h^{xy}_{\mu\nu} + \frac{1}{2}G^{(xy)}_{\mu\nu}(\mathbf{P}) + \mathbf{q}^\dagger\mathbf{V}^{xy}_{\mu\nu} - \boldsymbol{\mu}^\dagger\mathbf{E}^{xy}_{\mu\nu}\right]P_{\mu\nu} + \\ &- \text{tr }\mathbf{W}\mathbf{S}^{xy} - \text{tr }\mathbf{W}^y\mathbf{S}^x + \\ &+ \sum_{\mu\nu}\left[h^x_{\mu\nu} + G^{(x)}_{\mu\nu}(\mathbf{P}) + \mathbf{q}^\dagger\mathbf{V}^x_{\mu\nu} - \boldsymbol{\mu}^\dagger\mathbf{E}^x_{\mu\nu}\right]P^y_{\mu\nu}\end{aligned} \qquad \textbf{7.61}$$

There is an explicit FQFμ contribution, namely:

$$\mathbf{q}^\dagger\mathbf{V}^{xy}(\mathbf{P}) - \boldsymbol{\mu}^\dagger\mathbf{E}^{xy}(\mathbf{P})$$

It is also possible to calculate the NMR shielding tensor, which is defined as the second derivative of the energy with respect to the magnetic field and the nuclear magnetic moment:

$$\sigma_{xy}^{X} = \frac{\partial \mathcal{E}}{\partial B_x \partial m_y} = \mathrm{tr}(\mathbf{P}\mathbf{h}^{B_x m_y} + \mathbf{P}^{B_x}\mathbf{h}^{m_y}) \tag{7.62}$$

Here, there is no explicit FQFμ contribution as no FQFμ-related quantities depend on the nuclear magnetic moment.

Energy third derivatives

By exploiting the Placzek approach within the double harmonic approximation, Raman and ROA intensities are obtained in terms of the geometric derivatives of the electric dipole-electric dipole polarizability α^x, electric dipole-electric quadrupole polarizability A^x and electric dipole-magnetic dipole polarizability G'^x. In particular, Raman intensities depend only on α^x, whereas ROA intensities depend on α^x, A^x and G'^x. In the following equations the QM/FQFμ contributions to these quantities are reported. We refer the interested reader to Ref.[150] for further details on QM and QM/FQ analytical third derivatives.

$$\alpha^x_{QM/FQF\mu} = \sum_{\mu\nu}\left[\mathbf{L}^\dagger(\mathbf{P}^{e'}(\omega'))\mathbf{R}^x_{\mu\nu}P^e_{\mu\nu}(\omega)\right] + \sum_{\mu\nu}\left[\mathbf{L}^\dagger(\mathbf{P}^{e}(\omega))\mathbf{R}^x_{\mu\nu}P^{e'}_{\mu\nu}(\omega')\right] \tag{7.63}$$

$$A^x(\omega)_{QM/FQF\mu} = \sum_{\mu\nu}\left[\mathbf{L}^\dagger(\mathbf{P}^{e}(\omega))\mathbf{R}^x_{\mu\nu}P^q_{\mu\nu}\right] + \sum_{\mu\nu}\left[\mathbf{L}^\dagger(\mathbf{P}^{q})\mathbf{R}^x_{\mu\nu}P^e_{\mu\nu}(\omega)\right] \tag{7.64}$$

$$\begin{aligned} G'^x_{QM/FQF\mu} = &\sum_{\mu\nu}\left[\mathbf{L}^\dagger(\mathbf{P}^{m})\mathbf{R}^x_{\mu\nu}P^e_{\mu\nu}(\omega)\right] + \sum_{\mu\nu}\left[\mathbf{L}^\dagger(\mathbf{P}^{e}(\omega))\mathbf{R}^x_{\mu\nu}P^m_{\mu\nu}\right] + \\ &+ \sum_{\mu\nu}\left[\mathbf{L}^\dagger(\mathbf{R}^m(\mathbf{P}^e(\omega)))P^x_{\mu\nu}\mathbf{V}_{\mu\nu}\right] + \\ &+ \sum_{\mu\nu}\left[\mathbf{L}^\dagger\mathbf{R}^{m,x}_{\mu\nu}P^e_{\mu\nu}(\omega) + \mathbf{L}^{x\dagger}\mathbf{R}^m_{\mu\nu}P^e_{\mu\nu}(\omega)\right] \end{aligned} \tag{7.65}$$

Chapter 8

A Classical Picture of Subnanometer Junctions: an Atomistic Drude Approach to Nanoplasmonics

Tommaso Giovannini[†], Marta Rosa[‡], Stefano Corni[*,‡,♮], Chiara Cappelli[*,†]

[†] *Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy.*
[‡] *Department of Chemical Sciences, University of Padova, via Marzolo 1, Padova, Italy*
[♮] *CNR Institute of Nanoscience, via Campi 213/A, Modena, Italy*

E-mail: stefano.corni@unipd.it; chiara.cappelli@sns.it

Abstract The description of optical properties of subnanometer junctions is particularly challenging. Purely classical approaches fail, because the quantum nature of electrons needs to be considered. Here we report on a novel classical fully atomistic approach, ωFQ, based on Drude model for conduction in metals, classical electrostatics and quantum tunneling. We show that ωFQ is able to reproduce the plasmonic behavior of complex metal subnanometer junctions with quantitative fidelity to full *ab-initio* calculations. Besides the practical potentialities of our approach for large scale nanoplasmonic simulations, we show that a classical approach, in which the atomistic discretization of matter is properly accounted for, can accurately describe nanoplasmonics phenomena dominated by quantum effects.

A cornestone of nanoscience is that systems at the nanoscale have properties neither of the molecular nor of the macroscopic length scales.[217,562] Nanoplasmonics is a beautiful example of this: localized surface plasmons supported by metal nanostructures disappear in clusters with few atoms, and acquire different properties (surface plasmons polaritons) in extended surfaces.[217,562,563] The enormous progress of nanoscience has permitted a targeted control of the morphology of nanostructures at the nanometer and even subnanometer scales, thus allowing several applications in plasmonics and nanooptics.[564–568] Most properties of plasmonic nanostructures follow from the tunability of their optical response as a function of the their shape and dimensions; in case interparticle gaps are formed, the so-called "hot-spot" regions occur, in which localized surface plasmons can interact with molecules placed in the junctions, up to allow single molecule detection.[212,569–573]

The optical properties of nanostructures are generally treated, independent of the system's size/shape, by resorting to classical approaches.[209,226–230,546,574–580] However, when the size of the particles or junctions is only few nanometers or smaller, the quantum nature of electrons emerges, up to activate quantum tunneling effects across subnanometer interparticle gaps.[209,225,235,241,242,577,581,582,582–590] Tunnelling effects are not considered in classical models, so that quantum corrected approaches need to be applied.[241,588]

The theoretical study of the atomic-scale features in nanojunctions is still an almost unexplored field, because most phenomenological classical models do not address quantum effects. In fact, as reported by Urbieta et al.,[591] a proper description of atomic-scale effects would require a full quantum framework, accounting for the atomistic structure of the nanoparticles and the wave nature of electrons building up the plasmonic excitations.

By starting from the above considerations, in this paper we report on a fully atomistic classical model based on three very basic ingredients, i.e. Drude model for conduction in metal, classical electrostatics and quantum tunneling, which is able to reproduce with quantitative fidelity the optical properties of subnanometer junctions. In our approach, which we will call ωFQ (frequency dependent Fluctuating Charges), each atom of the nanostructure is endowed with an electric charge, which in not fixed but can vary as a response of the externally applied oscillating electric field.

Remarkably, here we go a step further with respect to other classical approaches. In fact, we are not using any experimental frequency dependent dielectric constant (possibly corrected for non-locality and electron scattering at the surface), but we let the dielectric response of the nanosystem to arise from atom-atom conductivity. Quantum tunnelling effects originate from a geometrical damping imposed on the atom-atom conductivity regime. The model is challenged to accurately reproduce complex *ab-initio* simulations on a stretched Na nanorod[234] and two approaching and retracting Na nanoparticles,[233] in which a single atom junction occurs, so that an atomistic description appears to be mandatory. ωFQ is the first ever classical approach succeeding at correctly modeling the optical properties of such systems, which up to now have been successfully treated only at the full DFT level. Remarkably, the results that we will show have relevant practical consequences, because we indeed provide a computationally viable model to investigate subnanometer junctions and complex nanostructures

of size well beyond what can be currently treated by *ab-initio* approaches.

8.1 Methods

The model we are introducing here, ωFQ, has its fundamentals on the Fluctuating Charges (FQ) force field, which is usually adopted for describing molecular systems.[56,146,150–152] FQ places on each atom of a molecular system a charge, which is not fixed but allowed to vary as a result of differences in atomic electronegativities. Charges are regulated by the atomic chemical hardnesses, that play the role of an atomic capacitance. From the mathematical point of view, FQ charges are obtained by minimizing the functional defining the energy of the system. ωFQ extends the basic formulation of the FQ model to take into account the interaction of the system with an external oscillating electric field $\mathbf{E}(\omega)$. In particular, each atom is assigned a charge, which is allowed to vary as a response of the polarization sources, which also include the external field $\mathbf{E}(\omega)$. Thus, being the electric field a complex quantity, calculated ωFQ charges become complex, being their imaginary value in quadrature with the field (if the field is real) and related to the absorption phenomenon. To build up the ωFQ approach, the time response of charges has to be related to external polarization sources. To this end, two alternative response regimes are set: (i) a conductive regime, in which the exchange of electrons between contiguous atoms is governed by the dynamics of the delocalized conduction electrons, giving rise to a damping; (ii) an alternative conductive regime, in which the exchange of electrons is also mediated by quantum tunneling effects. In this section we briefly discuss the main physical aspects of ωFQ: more details on the derivation of the equations and their implementation are given as SI.

The first regime is described by reformulating the Drude model of conductance[592] to treat charge redistribution between atoms. The key equation representing the Drude model reads:[592]

$$\frac{d\mathbf{p}}{dt} = \mathbf{E}(t) - \frac{\mathbf{p}}{\tau} \qquad \textbf{8.1}$$

where $\mathbf{p}$ is the momentum of the electron and τ a friction-like constant due to scattering events. The total charge derivative on the atom i can be written as:

$$\frac{dq_i}{dt} = \sum_j A_{ij}(n_j < \mathbf{p} > \cdot\hat{l}_{ji} - n_i < \mathbf{p} > \cdot\hat{l}_{ij}) \qquad \textbf{8.2}$$

where A_{ij} is an effective area dividing atom i by atom j, n_i is the electron density on atom i, $< \mathbf{p} >$ is the momentum of an electron averaged over the trajectories connecting i and j and $\hat{l}_{ji} = -\hat{l}_{ij}$ is the unit vector of the line connecting j to i. By assuming the total charge on each atom to be only marginally changed by the external perturbation, we can assume $n_i = n_j = n_0$. Therefore:

$$\frac{dq_i}{dt} = 2n_0 \sum_j A_{ij} < \mathbf{p} > \cdot\hat{l}_{ji} \qquad 8.3$$

$< \mathbf{p} > \cdot \hat{l}_{ji}$ needs to be estimated. To this end, it is convenient to consider a monochromatic applied electric field, so that eq.(8.3) translates to:

$$-i\omega q_i = 2n_0 \sum_j A_{ij} \frac{< \mathbf{E}(\omega) > \cdot \hat{l}_{ji}}{1/\tau - i\omega} \tag{8.4}$$

To proceed further, $< \mathbf{E}(\omega) > \cdot \hat{l}_{ji}$ (the total electric field averaged over the line connecting j to i) needs to be connected to atomic properties. This can be done by assuming $< \mathbf{E}(\omega) > \cdot \hat{l}_{ji} \approx (\mu_j^{el} - \mu_i^{el})/l_{ij}$, where μ_i^{el} is the electrochemical potential of atom i and l_{ij} the distance between atoms i and j. Therefore, Eq. 8.4 becomes:

$$\begin{aligned} -i\omega q_i &= \frac{2n_0}{1/\tau - i\omega} \sum_j \frac{A_{ij}}{l_{ij}} (\mu_j^{el} - \mu_i^{el}) \\ &= \sum_j K_{ij}^{\mathrm{dru}} (\mu_j^{el} - \mu_i^{el}) \end{aligned} \tag{8.5}$$

Eq. 8.5 can be rewritten collecting K_{ij}^{dru} in a $\mathbb{K}^{\mathrm{dru}}$ matrix, of which the definition is clearly evident from Eq. 8.5 itself. In order to make the model physically consistent, i.e. not to allow electron transfer between atoms that are too far apart, the pairs of atoms considered in Eq.8.5 have to be selected by exploiting a geometrical criterion, based on l_{ij}, i.e. to limit the interactions to nearest neighbors only.
To avoid any issue related to the specific definition of nearest neighbor atoms, a Fermi-like $f(l_{ij})$ damping function is introduced to weight the Drude conductive mechanism:

$$\begin{aligned} -i\omega q_i &= \sum_j (1 - f(l_{ij})) \cdot K_{ij}^{\mathrm{dru}} (\mu_j^{el} - \mu_i^{el}) \\ &= \sum_j K_{ij}^{\mathrm{tot}} (\mu_j^{el} - \mu_i^{el}) \end{aligned} \tag{8.6}$$

where:

$$f(l_{ij}) = \frac{1}{1 + \exp\left[-d\left(\frac{l_{ij}}{s \cdot l_{ij}^0} - 1\right)\right]} \tag{8.7}$$

In Eq.8.7 l_{ij}^0 is the equilibrium distance between two nearest neighbors in the bulk, whereas d and s are parameters determining the position of the inflection point and the steepness of the curve.
Eq.8.6 finally defines the ωFQ model. Whenever $f(l_{ij}) = 0$, the purely Drude conductive regime is recovered. For $f(l_{ij}) > 0$, Drude mechanisms exponentially turn off as l_{ij} increases, making electron transfer to enter in a second alternative regime. In this regime, the electric current exponentially decreases upon increasing the inter-atomic distance. Therefore, the typical functional form of tunneling exchange is recovered. [241]

Once ωFQ frequency-dependent charges are obtained by solving Eq. 8.6, the complex polarizability $\overline{\alpha}$ is easily calculated. From such a quantity, the absorption cross section is recovered:

$$\sigma_{\text{abs}} = \frac{4\pi}{3c}\omega \ \text{tr}\big(\alpha^*\big) \qquad \textbf{8.8}$$

where c is the speed of light, ω is the external frequency and α^* is the imaginary part of the complex polarizability $\overline{\alpha}$.
The ωFQ approach has been implemented in a stand alone Fortran 77 package. Eq. 8.6 is solved for a set of frequencies given as input. All computed spectra reported in the manuscript were obtained by explicitly solving linear response equations for steps of 0.01 eV. For all the studied Na nanosystems, the parameters given in Eqs. 8.5-8.6 were extracted from physical quantities recovered from the literature or numerically tested on single Na nanoparticles (see SI for more details). The parameters finally exploited are the following: $\tau = 3.2 \cdot 10^{-14}$ s,[593] $\sigma_0 = 2.4 \cdot 10^7$ S/m,[594] $A_{ij} = 3.38$ Å^2, $l^0_{ij} = 3.66$ Å,[594] $d = 12.00$, $s = 1.10$.

8.2 Results and Discussion

In order to test our newly developed ωFQ method, based on Drude model for conductivity in metals, classical electrodynamics and quantum tunneling, we shall compute the optical response of Na aggregates which are characterized by sub-nanometer gaps. We shall compare the results obtained by exploiting our model with those calculated at *ab-initio* level.
In particular, the optical absorption spectra of a metal nanorod pulled beyond the breaking point[234] and those of two small metal nanoparticles brought into contact[233] are studied because they are paradigmatic of a class of nanoplasmonic problems where *ab-initio* simulations seem mandatory. ωFQ is described with details in the Methods section and in the Supplementary Information (SI).

8.2.1 Stretched Sodium Nanorod

In this section the ωFQ approach (see Methods) is applied to a challenging system, i.e. a mechanically stretched sodium nanorod, which has been recently studied at the *ab-initio* level.[234] For such a system absorption cross sections as a function of the elongation distances at full *ab-initio* level have been reported,[234] and such data are taken in this paper as reference values to evaluate the quality of our fully atomistic, but classical ωFQ approach. It is worth noticing that as increasing the elongation distance, a sub-nanometer junction region occurs, in which quantum tunneling effects play a crucial role at determining the spectral features.[582,588,590] Therefore, the application of our model to such a challenging system will highlight its potentialities and limitations at describing such effects.
The nanorod structures, eight of which are depicted in Fig.8.1, were kindly provided to us by the authors of Ref.[234] They were obtained from an initially perfect Na_{261}

nanorod, which was adiabatically stretched, by allowing atomic positions of the central region to relax (see Ref.[234] for more details).

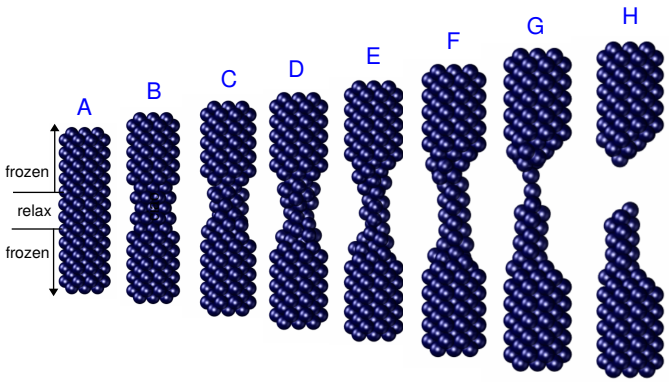


Figure 8.1. Selected structures obtained by Rossi et al.[234] by stretching a Na_{261} nanorod. The elongation distance d of the depicted structures are (from A to H): 0, 6, 10, 14, 20, 22, 26, and 28 Å.

For all 60 structures, absorption cross sections were calculated by exploiting the ωFQ model; Figure 8.2 reports the absorption spectra of selected 30 structures as a function of the elongation distance.

As depicted in Figure 8.1, the sodium nanorod is elongated and the atoms in the nanojunction region are relaxed until the structure breaks for distances longer than 26 Å, where the limit of mono atomic junction is reached.

These structural features are reflected by the calculated spectra (see Fig. 8.2 (a)); in fact, a clear discontinuity is evident at $d = 26$ Å (structure G). Let us focus on elongation distances $d < 26$ Å. The pristine nanorod structure (structure A) presents one intense excitation at 1.5 eV (dubbed Local Plasmon LP) and a less intense peak at 2.8 eV (LP2). Our atomistic model allows to identify the nature of such LPs, for instance by graphically plotting the imaginary part of atomic charges for each transition. Maps of the molecular electrostatic potential (MEP) obtained from such charges are reported in panel (a) of Figure 8.3 for structures A-H. The comparison of data in panels (a) of Figs. 8.2 and 8.3 clearly shows a first charge-transfer excitation and a second transition with a dipolar character. Therefore, by exploiting the same nomenclature used for metal dimers, LP will be renamed as Charge Transfer Plasmon (CTP), whereas LP2 as Boundary Dipolar Plasmon (BDP).[595–597] As the elongation distance increases, both CTP and BDP significantly redshift, and this feature is particularly evident for CTP. In addition, they behave in a complete different way: CTP intensity slowly decreases, whereas BDP becomes more and more predominant. Such a behavior is commonly identified in most nanoplasmonic dimers.[233,234,591,598,599]

When the elongation distance reaches 26 Å (structure G) a monoatomic junction is

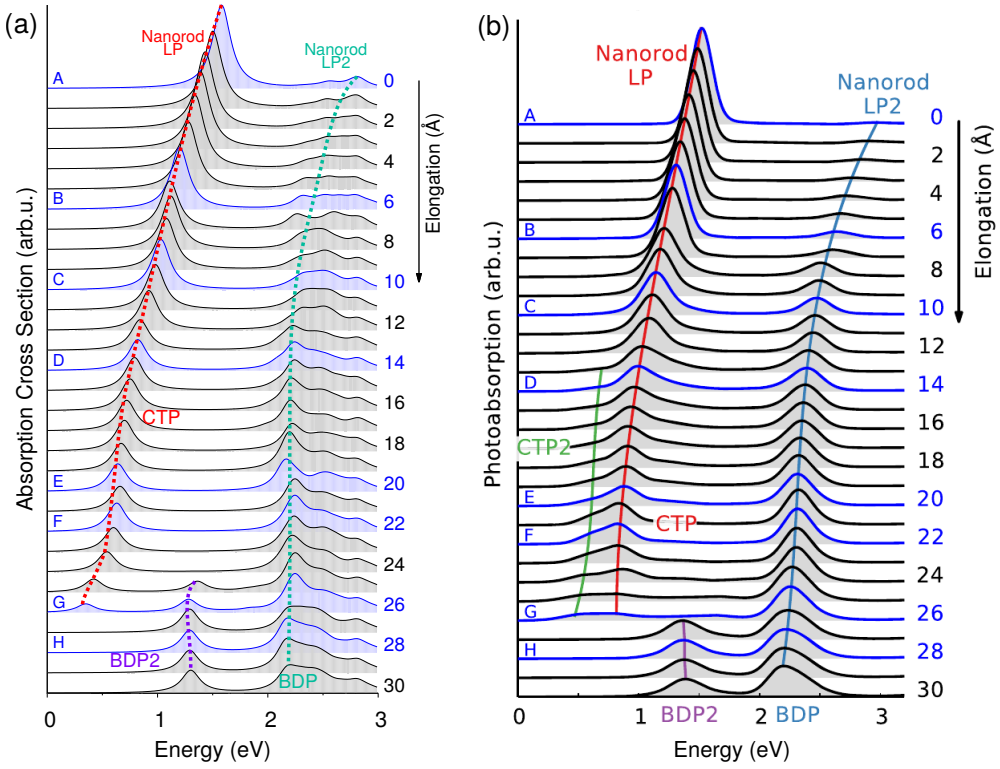


Figure 8.2. Evolution of the plasmonic response of a Na_{261} nanorod under stretching. Structures A-H in Fig. 8.1 are highlighted.**(a)** ωFQ absorption Cross Section as a function of the energy. **(b)** Reproduced from Ref. [234]

obtained, which is the limiting structure occurring just before the structure breaks. Such features are reflected by the absorption cross section. CTP occurs at about 0.5 eV and shows a very low intensity, because electrons can only transfer through a single atom. The BDP excitation becomes the most intense and shifts at 2.2 eV. The inspection of panel (b), structure G in Figure 8.3 shows that such an excitation has now a quadrupolar character. At such an elongation distance, a third excitation, which is actually already visible at 25 Å, arises at about 1.4 eV. The analysis of the MEP map suggests this transition to be due to an additional dipolar plasmon, BDP2.
We move now to comment spectra for $d > 26$ Å. CTP disappears, as expected, because the gap between the two nano-moieties is too large to allow electron tunneling between them. Thus, only BDP and BDP2 excitations are present. In particular, BDP2, which has a clear dipolar character as evidenced by the pictures reported in panel (b) of Fig. 8.3 for structure H ($d = 28$ Å), increases in intensity, although BDP still dominates the spectrum.

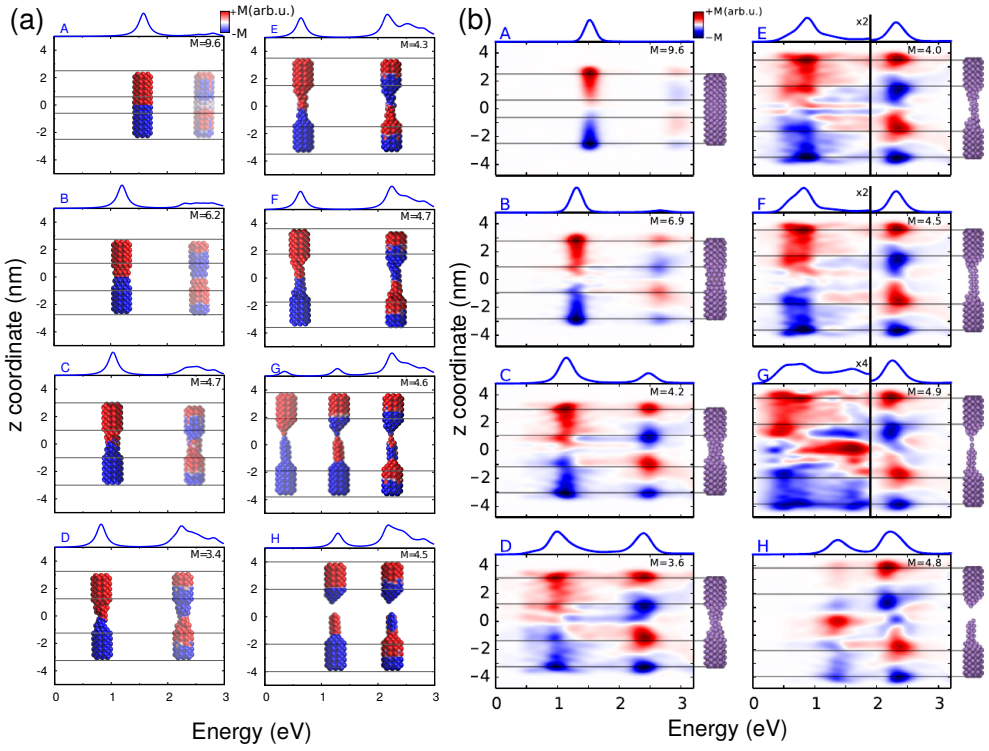


Figure 8.3. **(a)** ωFQ pictorial representation of the local plasmonic response for the 8 A-H selected structures. Blue color indicates a negative charge, whereas red color indicates a positive charge. **(b)** DFT results reproduced from Ref.[234]

To better show the capabilities of our atomistic classical model, our results can be directly compared with theoretical *ab-initio* calculations reported in Ref.[234] The resulting spectra are reproduced in panel (b) of Fig. 8.2. Our calculated absorption cross sections (panel (a) in Fig. 8.2) compare extremely well with *ab-initio* data, and only

minor discrepancies are present. *Ab-initio* spectra for structures with 13< d <26 Å show a really low peak at about 0.5 eV, associated to a CTP2. Such a peak, which however can hardly be identified in the *ab-initio* density maps (see panel (b) of Fig. 8.3), is not reproduced by ωFQ. In addition, the *ab-initio* BDP transition results in a narrower band. Such a difference can be justified by the atomistic nature of our model, which results in some kind of inhomogeneous broadening due to transitions with different nodal structure at the atomic scale, but corresponding to plasmons of similar nature. Despite such minor discrepancies, the agreement between ωFQ and DFT spectra is impressive. In fact not only excitation energies but also relative intensities are correctly reproduced in all the elongation range, but also red shifts for both CTP and BDP. Furthermore, ωFQ reproduces the *ab-initio* calculated redshift in the spectrum of structure with d =7 Å. Such a behavior can be due to the structural rearrangement of the nanorod as a result of the *ab-initio* geometry relaxation. The peculiar atomistic nature of ωFQ makes it capable to also catch such effects, resulting from tiny deformations of the nanostructure.

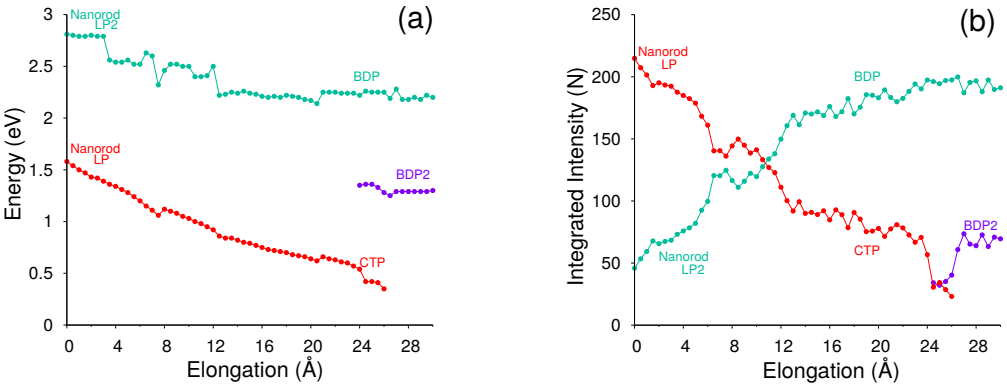


Figure 8.4. Detailed analysis of the plasmon modes during the nanorod stretching. (a) Peak energy and (b) integrated intensity (the area under the peak) of the plasmon modes as a function of the nanorod elongation d. The intensities are normalized so that the full spectrum integrates to the number of valence electrons (261).

To further analyze our results, in Figure 8.4 excitation energies and integrated intensities calculated by exploiting our model are shown for the three plasmons. Many discontinuity points are noticed for both CTP and BDP as a results of the stretching of the nanostructure. In particular, at about 7.5 Å the energies of CTP and BDP decrease of 0.1 eV. Integrated intensities also present a discontinuity point at such a distance. Such shifts and discontinuities, which as stated before are due to structural rearrangements, are also reported by DFT calculations.[234]

Our results, which also in this case are quantitatively comparable with DFT, show once again that our classical atomistic approach gives a correct description of the underlying physical phenomena.

8.2.2 Sodium NP dimer: approaching and retracting processes

As a second test to analyze the performances of ωFQ, the latter is challenged with the description of the optical properties of two Na_{380} icosahedral nanoparticles which are approached and retracted (see Fig. 8.5). This systems has been recently studied at full *ab-initio* level by Marchesin et al.,[233] who kindly provided us with the full set of model structures.

Two alternative processes will be considered: first, the two Na_{380} nanoparticles are placed at a distance of 16 Å (such a distance guarantees that they do not interact) and then they are drawn closer until they fuse (see Fig. 8.5, panel a). Then the two fused Na_{380} nanoparticles are retracted until the structure separates (see Fig.8.5, panel b), giving rise to a process which is similar to the case presented in the previous section. Two alternative situations of approaching and retracting were both considered, because as it has been already reported in Ref.[233] the two processes are physically different.

Let us start the discussion by considering the approaching process (Fig. 8.5, panel a). The imaginary part of the longitudinal polarizability, i.e. the component parallel to the dimer axis, has been computed as a function of the inter-nanoparticle distance; its values are reported in panel (a) of Figure 8.6.

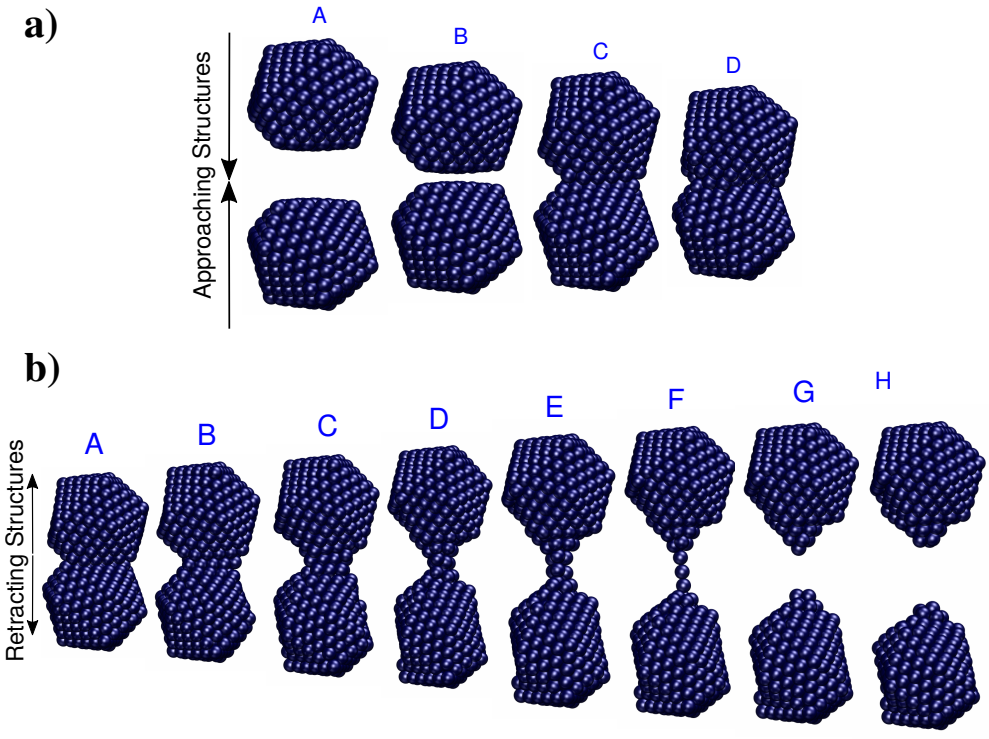


Figure 8.5. **a)** Selected structures resulting from the approaching of two Na_{380} nanoparticles. A: $d = 16$ Å; B: $d = 6.2$ Å; C: $d = 6.1$ Å; A: $d = 0$ Å. **b)** Selected structures resulting from the retracting of two Na_{380} nanoparticles. The nominal gap distances d are (from A to H): 3.7, 9.7, 14.7, 22.7, 25.9, 32.1, 32.3, 34.1 Å. Data taken from Ref.[233]

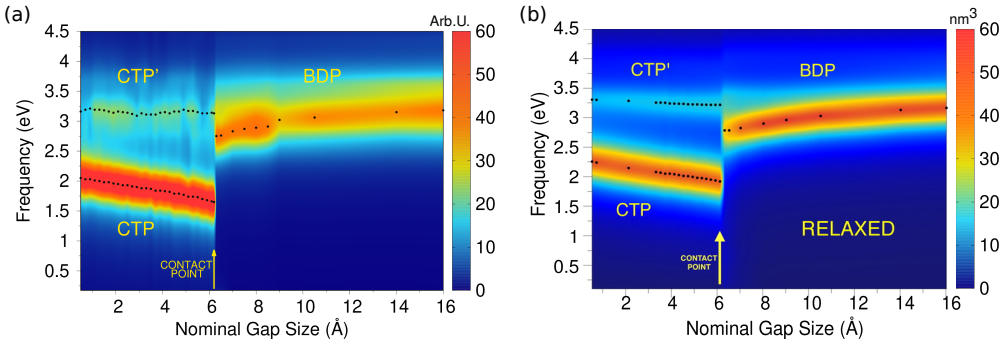


Figure 8.6. (a) ωFQ calculated 2D plots of the longitudinal imaginary polarizability (arbitrary units) as a function of the excitation energy and nominal gap size for the two approaching Na_{380} nanoparticles. (b) *ab-initio* values reproduced from Ref. [233]

The calculated 2D plots present a clear discontinuity between nominal gap sizes of 6.1 Å and 6.2 Å, i.e. between structures B and C in Figure 8.5 panel a, which correspond to a jump-to-contact instability. At higher inter-nanoparticles distances, plots are dominated by a single peak, which is placed at 3.19 eV at $d = 16$ Å , i.e. when the two nanoparticles are far apart. This band can be attributed to BDP. Induced charges and the corresponding MEP maps are depicted for the four selected significant structures in Fig.8.7. We clearly see that for structure A BDP is a dipolar plasmon. As expected, as the distance between the two nanoparticles decreases, the BDP redshifts due to the increasing of electrostatic interactions. When the two nanoparticles fuse (structure C) a clear discontinuity appears and for $d \leq 6.1$ Å , the 2D plot is characterized by two main peaks, namely CTP (1.66 eV) and CTP' (3.15 eV), corresponding to the plasmon excitations represented for structures C and D in Fig. 8.7. As the distance further reduces, CTP blueshifts whereas CTP' remains almost unchanged.

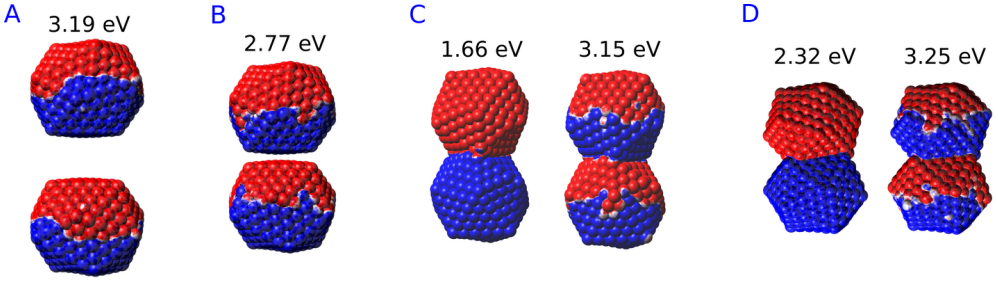


Figure 8.7. ωFQ MEP maps for plasmon excitations (eV) of selected structures. A: $d = 16$ Å; B: $d = 6.2$ Å; C: $d = 6.1$ Å; A: $d = 0$ Å. Blue color indicates a negative charge, whereas red color indicates a positive charge.

The inspection of the MEP maps in Fig. 8.7 shows that the higher order CTP' shows a dipolar character, similarly to BDP, which occurs for structures with $d > 6.1$ Å. The jump-to-contact structural instability is confirmed by the appearance of CTP, which is characterized by a net flux of charge between the two (fused) nanoparticles. Such

a flux gives rise to a conductive regime, resulting in an electric current. Notice that, as previously reported by Marchesin et al.,[233] the sudden occurrence of the junction bypasses the distance regime where quantum tunneling effects are relevant.
Moving back to Fig. 8.6, we remark the very good agreement between the results obtained by exploiting ωFQ approach and the *ab-initio* counterparts. Qualitatively, DFT results are perfectly reproduced, in fact all the three CTP, CTP' and BDP bands are described, their behavior as a function of the distance is correctly reproduced and bands relative intensities are qualitatively well described. Some minor discrepancies are present from the quantitative point of view. In fact, the behavior of BDP as a function of the distance is not perfectly described, e.g. ωFQ intensities remain almost constant along the approaching process. Also, CTP intensities are overestimated and the CTP' band seems broader. Such findings are in line with what has been found in the previous section and can be due to the atomistic nature of our approach, which does not smooth out inhomogeneities on the atomic scale.
We pass now to study the two fused Na_{380} nanoparticles which are retracted until the structure breaks (see Fig.8.5 panel b for representative structures). As it is evident the breaking process is gradually occurring. In fact, a monoatomic junction arises (structure F), which breaks as the distance increases further. Therefore, tunneling effects are expected to be relevant, thus resulting in a different behavior of the calculated spectrum with respect to what we have reported in the previous paragraphs. for the approaching process, and also found at the *ab-initio* level.[233]
Indeed, this is confirmed by ωFQ calculated values of the imaginary part of the longitudinal polarizability, i.e. the component parallel to the dimer axis; such data are reported in panel (a) of Figure 8.8 as a function of the elongation distance.
By starting from the fused A structure, we notice that, as expected, the spectrum consists of two bands, which can be related to CTP and CTP' excitations. Their nature can be understood by referring to Fig.8.9; CTP occurs at 1.84 eV and corresponds to a charge flux between the two nano-moieties. CTP' (3.18 eV) shows instead the anticipated dipolar character.
As the elongation distance increases, both CTP and CTP' redshift, and this is particularly evident especially for CTP. In addition the CTP band shrinks and its intensity decreases, whereas CTP' shows an opposite behavior, i.e. its intensity increases and the band broadens. Small discontinuities, characterized by sudden red- or blue-shift of the excitations, are visible. This behavior is similar to what we have found in the previous section for the stretched Na_{261} nanorod (see Fig.8.4, panel (a)), and can be reasonably due to the structural relaxation and the resulting thinning of the conductive channels as the structure stretches. As the limiting structure F is reached ($d = 32.1$ Å), a monoatomic junction arises (see Fig.8.9), resulting in the CTP band to occur at 0.25 eV and the CTP' at 2.89 eV. The inspection of the corresponding MEP maps shows that the nature of the associated plasmons is unchanged with respect to the initial A structure. Suddenly, the structure breaks (structure G, $d = 32.3$ Å), thus resulting in the disappearance of CTP and the convergence of CTP' towards BDP. The MEP associated, depicted in Fig.8.9, shows a dipolar character.
Moving back to Fig. 8.8, also for the elongation process a very good agreement between the results obtained by exploiting our classical atomistic ωFQ approach and the

reference *ab-initio* data[233] is noted. Qualitatively, DFT results are perfectly reproduced, in fact all the three CTP, CTP' and BDP bands are described, their behavior as a function of the distance is correctly reproduced and bands relative intensities are qualitatively well described. ωFQ intensities for the CTP band are slightly overestimated, and remain higher also as the nominal gap size increases. Furthermore, ωFQ well reproduces the discontinuities in the spectra, and specifically those marked as α, β and γ in Fig.8.8 panel (b). As already pointed out in the previous section, such a behavior can be due to the structural rearrangement of the nanostructure as a result of the *ab-initio* geometry relaxation. The classical but atomistic nature of our approach makes it able to correctly describe such effects.

The ωFQ imaginary charges for structures before and after the spectral jumps α, β and γ are depicted in Fig. 8.14, given as SI. The structural change associate to each spectral jump is reflected by differences in the corresponding plasmons, i.e. by changes on the charges of the junction atoms. Remarkably, our data are in agreement with DFT density distributions around the junction reported in Ref.,[233] thus showing once again the reliability of our classical atomistic model.

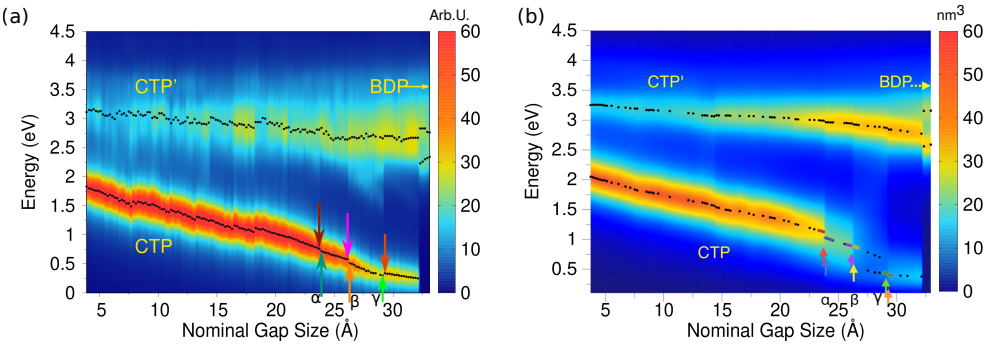


Figure 8.8. (a) ωFQ calculated 2D plots of the longitudinal imaginary polarizability (arbitrary units) as a function of the excitation energy and nominal gap size for the two approaching Na_{380} nanoparticles. (b) *ab-initio* values reproduced from Ref.[233]

To end the discussion and to further analyze the performance of the model, we report in Fig.8.10 the calculated ωFQ absolute values of the electric current through the plasmonic nanojunctions as a function of the elongation distance. Both the approaching and retracting processes are considered. The reported values were obtained at the excitation energies of each plasmon.

As expected, for the approaching process when the two nanoparticles do not interact, i.e. when spectra are dominated by BDP, no current flux is evidenced. As the jump-to-contact instability is reached, a discontinuity in the current arises, i.e. a net current flux is established. The current further increases as the inter-nanoparticle distance decreases.

For the retracting process, the CTP plasmon clearly dominates the charge flux. As the system is stretched, the current intensity slowly decreases, until it vanishes when the system breaks (structures F and G). Several discontinuities in the CTP current

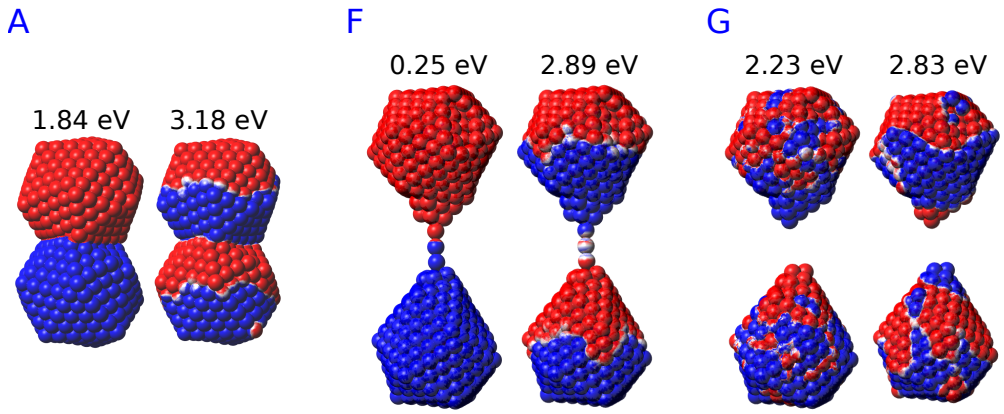


Figure 8.9. ωFQ MEP maps for plasmon excitations (eV) of selected structures. A: $d = 3.7$ Å; F: $d = 32.1$ Å; G: $d = 32.3$ Å. Blue color indicates a negative charge, whereas red color indicates a positive charge.

are present, similarly to what was already commented for the stretched nanorod in the previous section. Remarkably, the α β and γ spectral jumps can easily be identified in the current plot.

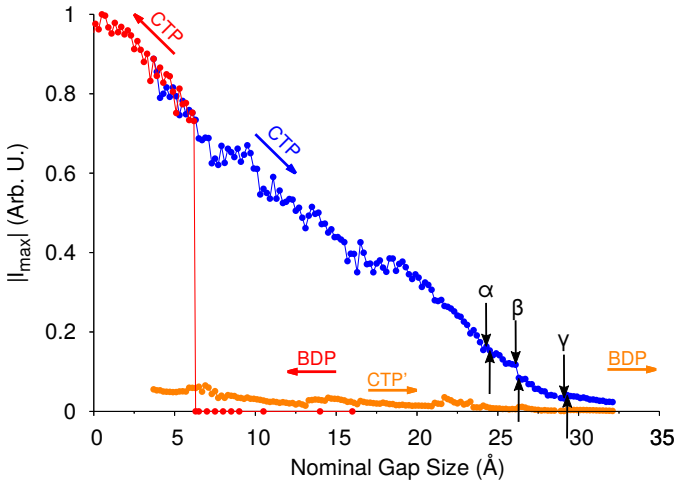


Figure 8.10. ωFQ absolute value of the electric current through the plasmonic nanojunction as a function of the elongation distance. Colored arrows indicate the direction of the process (approaching, in red, and retracting, in blue and orange). The current is computed by following Ref.[233] (see also SI). Black arrows indicate the position of the α β and γ spectral jumps. All values are given in arbitrary units.

8.3 Summary and Conclusions

In the present work, a novel atomistic model, ωFQ, based on textbook concepts (Drude theory, electrostatics, quantum tunneling) has been proposed. In such a model, the atoms of complex nanostructures are endowed only with an electric charge, which can vary according to the external electric field. The electric conductivity between the nearest atoms is modeled by adopting the simplest possible assumption, i.e. the Drude model which has been reformulated in terms of electric charges. Thus, only few physical parameters define our equations. Furthermore, the dielectric response of the system arises naturally from atom-atom conductivity. Remarkably, such a feature permits to avoid the use of any experimental frequency-dependent dielectric constant, which is adopted in the quantum corrected models.[241] Moreover, ωFQ takes also into consideration quantum tunneling effects by switching off exponentially conductivity between neighbor atoms.
ωFQ model was challenged to reproduced the optical response of complex Na nanoclusters which have been investigated previously at *ab-initio* level[233,234] and for which a QM description has been considered mandatory. The capability of our approach to reproduce the results of complex simulations has a relevant pratical consequence; in fact, due to its classical formulation, ωFQ can be applied to model nanoplasmonic systems of size well beyond what can be currently treated at the *ab-initio* level. Moreover, the good agreement between the *ab-initio* simulations and ωFQ results shows that the physics it encompasses (Drude model, electrostatics and a quantum tunneling correction) properly ported at the atomistic level, is dominating nanoplasmonic phenomena also in this small scale regime.
In this work, only Na clusters have been considered. However, ωFQ, properly extended to account for the atomic core polarizability that characterizes d-metals, has the potential to treat a great variety of plasmonic materials. Also, the formulation of the model in terms of electric charges and its manifest reliability shows that ωFQ has the potentialities to be coupled to fully QM molecular simulations within a QM/MM framework so to allow the modeling of spectral enhancement of molecules adsorbed on plasmonic nanostructures. This aspects will be treated in future communications.

Supplementary Information

Detailed derivation of the ωFQ model. Details on the calculation of the electric current. Model parametrization on single Na nanoparticles. Dependence of ωFQ absorption cross sections on model parameters. ωFQ MEP maps for plasmon excitations of selected structures. DFT values of the electric current as a function of the elongation distance.

Linear Response Equations

Static Response

In order to calculate static response properties, the FQ basic equations (see for instance Eq.2.1 has to be modified. In particular, in the case of one-metal nanoparticles, all the

atoms are of the same type, and thus no polarization occurs due to differences in atomic electronegativies, which in the pristine FQ model define the polarization "source" in the Maxwell's meaning of the term. However, the system polarizes under the effect of a static external electric field. Static response equations are derived by simply adding to the whole energy a term accounting for the interaction with the external electric field. It is worth noticing that this can be in principle achieved either by defining the external perturbation in terms of an electric field, or an associated electric potential. Since FQ is defined in terms of charges, the use of the potential V^{ext} is more convenient.
Thus, Eq. 2.1 becomes:

$$F(\mathbf{q},\lambda) = \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} + \mathbf{q}^{\dagger}\mathbf{V}^{ext} \qquad \textbf{8.9}$$

By minimizing $F(\mathbf{q},\lambda)$, the static response equations are obtained:

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C}_Q - \mathbf{V}^{ext} \qquad \textbf{8.10}$$

Once FQ charges are calculated by solving Eq. 8.10, the static polarizability can be calculated by evaluating the induced dipole moment.
Notice also that standard FQ equations can be reformulated by introducing the electrochemical potential μ_i^{el}:

$$F = \sum_i (q_i\chi_i + \frac{1}{2}q_i\eta_i q_i + V_i q_i) \qquad 8.11$$

$$V_i = V_i^{ext} + \sum_{k\neq i} J_{ik}q_k \qquad 8.12$$

$$\mu_i^{el} = \frac{\partial F}{\partial q_i} = \chi_i + q_i\eta_i + V_i \qquad 8.13$$

where J_{ik} is the proper electrostatic non-diagonal matrix element.

Frequency dependent Response

When nanoparticles of finite dimension are irradiated by an external oscillating electric field, local plasmons arise.
Because in the FQ approach a charge is placed on each atom, the frequency-dependent response of the system is defined once the variation of the charges in time is obtained. In our newly developed ωFQ model, we assume electron transfer to occur under two alternative regimes:

- Conductive regime: the exchange of electrons between contiguous atoms is governed by the dynamics of the delocalized conduction electrons, giving rise to a damping.
- Tunneling regime: the conductive exchange of electrons between the atoms is mediated by a quantum tunneling mechanism.

Conductive Regime

The dynamics of the electron exchange between contiguous atoms can be typically described by the Drude model.[592] This model is generally formulated in terms of induced dipole moments, therefore we need to adapt it to the atomistic description given by the FQ approach.
The key equation representing the Drude model reads:[592]

$$\frac{d\mathbf{p}}{dt} = \mathbf{E}(t) - \frac{\mathbf{p}}{\tau} \tag{8.14}$$

where $\mathbf{p}$ is the momentum of the electron and τ a friction-like constant due to scattering events. The total charge derivative on atom i can be written as:

$$\frac{dq_i}{dt} = \sum_j A_{ij}(n_j < \mathbf{p} > \cdot\hat{l}_{ji} - n_i < \mathbf{p} > \cdot\hat{l}_{ij}) \tag{8.15}$$

where A_{ij} is an effective area dividing atom i by atom j, n_i is the electron density on atom i, $< \mathbf{p} >$ is the momentum of an electron averaged over the trajectories connecting i and j and $\hat{l}_{ji} = -\hat{l}_{ij}$ is the unit vector of the line connecting j to i. By assuming the total charge on each atom to be only marginally changed by the external perturbation, we can assume $n_i = n_j = n_0$. Therefore:

$$\frac{dq_i}{dt} = 2n_0 \sum_j A_{ij} < \mathbf{p} > \cdot\hat{l}_{ji} \tag{8.16}$$

$< \mathbf{p} > \cdot\hat{l}_{ji}$ needs to be estimated. To this end, it is convenient to consider a monochromatic applied electric field, so that eq.(8.16) translates to:

$$-i\omega q_i = 2n_0 \sum_j A_{ij} \frac{< \mathbf{E}(\omega) > \cdot\hat{l}_{ji}}{1/\tau - i\omega} \tag{8.17}$$

To proceed further, $< \mathbf{E}(\omega) > \cdot\hat{l}_{ji}$ (the total electric field averaged over the line connecting j to i) needs to be connected to atomic properties. This can be done by assuming $< \mathbf{E}(\omega) > \cdot\hat{l}_{ji} \approx (\mu_j^{el} - \mu_i^{el})/l_{ij}$, where μ_i^{el} is the electrochemical potential of atom i and l_{ij} the distance between atoms i and j. Therefore, Eq. 8.17 becomes:

$$\begin{aligned} -i\omega q_i &= \frac{2n_0}{1/\tau - i\omega} \sum_j \frac{A_{ij}}{l_{ij}} (\mu_j^{el} - \mu_i^{el}) \\ &= \frac{2\sigma_0/\tau}{1/\tau - i\omega} \sum_j \frac{A_{ij}}{l_{ij}} (\mu_j^{el} - \mu_i^{el}) \\ &= \sum_j K_{ij}^{\mathrm{dru}} (\mu_j^{el} - \mu_i^{el}) \end{aligned} \tag{8.18}$$

where $n_0 = \sigma_0/\tau$ is follows from the relationship between the electron density n_0 and the static conductance σ_0. In Eq. 8.18 a Drude matrix K^{dru} is defined.

Conductive vs. Tunneling Regime

Eq. 8.18 can be rewritten by introducing a K^{dru} matrix, of which the definition is clearly evident from Eq. 8.18 itself. In order to make the model physically consistent, i.e. not to overestimate electron transfer effects, the number of atom pairs considered in Eq.8.18 has to be limited. The simplest way to achieve that is to exploit a purely geometrical criterion, based on l_{ij}, i.e. to limit the interactions to first neighbors only. To avoid any issue related to the specific definition of first neighbor atoms, a Fermi-like $f(l_{ij})$ damping function can be introduced to mediate the Drude conductive mechanism:

$$\begin{aligned} -i\omega q_i &= \sum_j (1 - f(l_{ij})) \cdot K_{ij}^{\mathrm{dru}} (\mu_j^{el} - \mu_i^{el}) \\ &= \sum_j K_{ij}^{\mathrm{tot}} (\mu_j^{el} - \mu_i^{el}) \end{aligned} \qquad \textbf{8.19}$$

where:

$$f(l_{ij}) = \frac{1}{1 + \exp\left[-d\left(\frac{l_{ij}}{s \cdot l_{ij}^0} - 1\right)\right]} \qquad \textbf{8.20}$$

In Eq.8.20 l_{ij}^0 is the equilibrium distance between two first neighbors, whereas d and s are parameters determining the position of the inflection point and the thickness of the curve.
Eq.8.19 finally defines the ωFQ model. Whenever $f(l_{ij}) = 0$, the purely Drude conductive regime is recovered. For $f(l_{ij}) > 0$, Drude mechanisms exponentially turn off as l_{ij} increases, making electron transfer to enter in a second alternative regime. The use of the $f(l_{ij})$ damping function guarantees electron exchange to only occur by adopting a conductive scheme, of which the intensity exponentially decreases as increasing the inter-atomic distance. Therefore, the typical functional form of tunneling exchange is recovered at the atomistic level.
Notice that from a computational point of view, Eq. 8.19 can be rewritten as:

$$\sum_j \left(-\sum_k K_{ik}^{\mathrm{tot}} D_{ij} + \sum_k K_{ik}^{\mathrm{tot}} D_{kj} + i\omega\delta_{ij} \right) q_j = \sum_j (V_i^{ext} - V_j^{ext}) K_{ij}^{\mathrm{tot}} \qquad \textbf{8.21}$$

Once ωFQ frequency-dependent charges are obtained by solving Eq. 8.19, the complex polarizability $\overline{\alpha}$ is easily calculated. In particular, starting from the charges, the complex electric dipole $\overline{\boldsymbol{\mu}}$ is calculated:

$$\overline{\boldsymbol{\mu}} = \sum_i q_i \cdot \mathbf{r}_i \qquad \textbf{8.22}$$

where $\mathbf{r}_i$ is the distance between the atom i-th and the origin. From the complex dipole moment, the complex polarizability $\overline{\alpha}$ is calculated by solving:

$$\overline{\alpha} = \frac{\partial \overline{\boldsymbol{\mu}}}{\partial \mathbf{E}} \qquad \textbf{8.23}$$

From such a quantity, the absorption cross section is recovered:

$$\sigma_{\text{abs}} = \frac{4\pi}{3c}\omega \ \text{tr}(\alpha^*) \qquad \textbf{8.24}$$

where c is the speed of light, ω is the external frequency and α^* is the imaginary part of the complex polarizability $\overline{\alpha}$.

Calculation of the electric current

Since we are dealing with a finite object we can use the continuity equation to define the current that flows across a plane perpendicular to the dimer axis and passing through the center of the junction (see Fig.8.11). ωFQ calculations yield complex response charges to an external monochromatic field in the frequency domain.

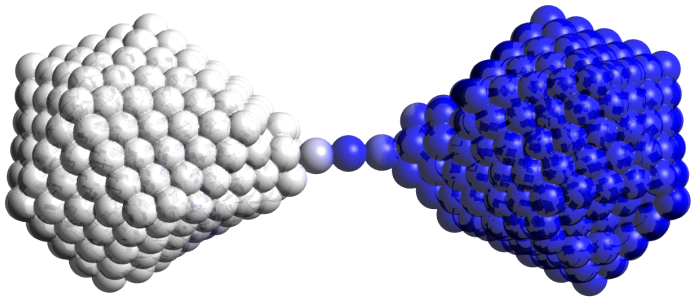


Figure 8.11. Pictorial view of a dimer junction. Charges on the blue atoms are those considered in the definition of the electric current (Eq.8.25).

The continuity equation gives a relation between the total induced charge and the current flowing across the junction $I(t)$ at time t. By referring for instance to the dimer junction in Fig.8.11, the most intuitive way of calculating the current is to only consider charges belonging to one of the two regions of the dimer, for instance those assigned to the blue atoms. In the frequency domain, the current can be obtained as:

$$I(\omega) = -i\omega \sum_i q_i(\omega) \qquad \textbf{8.25}$$

Therefore, the modulus of the current (maximum current) flowing across the junction in response to a given external electric field $\mathbf{E} = \mathbf{E_0}\cos(\omega t)$ reads:

$$|I_{max}(\omega)| = \omega \sqrt{\sum_i \Big(Re(q_i^2(\omega)) + Im(q_i^2(\omega)) \Big)} \qquad \textbf{8.26}$$

In this paper, the current is calculated for any of the plasmon resonances for each geometry of the two systems under investigation.

Computational Details

ωFQ has been implemented in a stand-alone program, named nanoFQ written in Fortran77. The final equation 8.21 is directly solved by an LU decomposition, however the

computational time can be further reduced by iterative methods for solving complex linear equations involving a sparse matrix. This will be considered in future works.

Model Parametrization

ωFQ was tested against sodium nanoparticles (see below), which have been chosen because of their simple electronic structure (only one valence electron is present on each atom). Eq. 8.19 depends on several parameters, which were recovered from the literature whenever possible or were fitted to reproduce reference *ab-initio* data. Such parameters are summarized in Table 8.1.

Parameter	Eq.	Value	Ref
η	8.10	0.292	t.w.
τ	8.18	1323	[593]
σ_0	8.18	5.21	[594]
A_{ij}	8.18	12.08	t.w.
d	8.20	12.0	t.w.
s	8.20	1.1	t.w.
l^0_{ij}	8.20	6.92	[594]

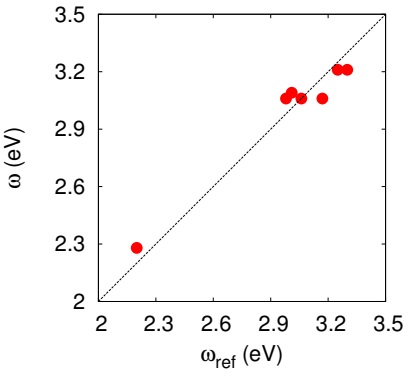
Table 8.1. ωFQ parameters for sodium nanoparticles used in this paper. t.w.: this work. All data are given in atomic units.

Single Sodium Nanoparticles

ωFQ was first tested against selected sodium clusters (see the following Table) in order to assign the parameters entering ωFQ equations. Such parameters were defined as the best set to reproduce reference *ab-initio* plasmon resonances (ω_{ref}). The differences between ωFQ and reference values can be inferred by data reported in the following table and also depicted in the following Figure:

Na_N	ω_{ref} (eV)	ω (eV)
59[a]	2.2	2.28
169[602]	3.01	3.09
331[599]	2.98	3.06
331[602]	3.06	3.06
331[599]	3.17	3.06
832[603]	3.25	3.21
1000[598]	3.3	3.21

[a] plan-wave DFT calculations performed with Quantum Espresso, LDA xc functional, ultrasoft pseudopotential, plane wave cut-off of 35 Ry (140 Ry on the density), TDDFT calculations with the Lanczos Liouville approach TurboTDDFT[600,601]



As it comes out evident, our choice of the model parameters makes ωFQ able to reproduce the plasmon resonance almost perfectly for most of the selected nanoparticles.

Almost perfect linearity with respect to the reference *ab-initio* data is also reported (the slope is of the dotted line is 0.99). It is also worth stressing that such results show that the free parameters of ωFQ, i.e. those which cannot recovered from the literature, can easily be determined once the optical properties of small nanoparticles are known. This means that the extension of the model to nanomaterials based on other kind of metal atoms, would simply required the knowledge of the associated plasmon resonances.

Dependence of ωFQ absorption cross sections on the choice of the model parameters

Among the several parameters defining our ωFQ model, n_0 is the one determining the absorption resonance (see Eq. 8.18). Thus, because n_0 is defined as the ratio between the static conductance σ_0 and the damping τ, such a ratio needs to remain constant in order to guarantee the plasmon resonance to stay the same. The results obtained by varying both σ_0 and τ, but keeping their ratio fixed are shown in the left panel of Figure 8.12. The different curves were obtained by multiplying σ_0 and τ by the factor f shown in the key. As expected, the curves become thinner and the limit of the stick spectrum is recovered as the ratio doubles or triplicates. On the other hand, if the ratio decreases by a factor 3 or 10, the curve broadens, thus recovering the artificial broadening usually employed to plot *ab-initio* calculations.

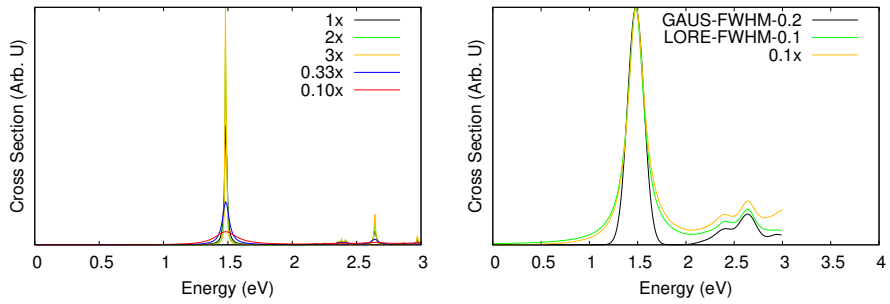


Figure 8.12. (left): calculated ωFQ absorption cross sections as a function of n_0; (right): Gaussian or Lorentzian convolution of ωFQ stick compared with ωFQ spectrum obtained with f=0.1

To further investigate the behavior of our model, ωFQ stick values were convoluted with a Lorentzian or a Gaussian-type function, with Full Width at Half Maximum (FWHM) of 0.1 and 0.2 eV, respectively. It is evident from the inspection of the right panel of Fig.8.12 that the Lorentzian function best fits ωFQ values obtained with f=0.1. This is not surprising, if Eq. 8.18 is inspected. On the basis of the data shown above, all the spectra reported in the paper were obtained by keeping σ_0 and τ to the literature values and to convolute each stick spectrum with a Lorentzian function. In this way, ωFQ spectra are coherent with the corresponding *ab-initio* values.

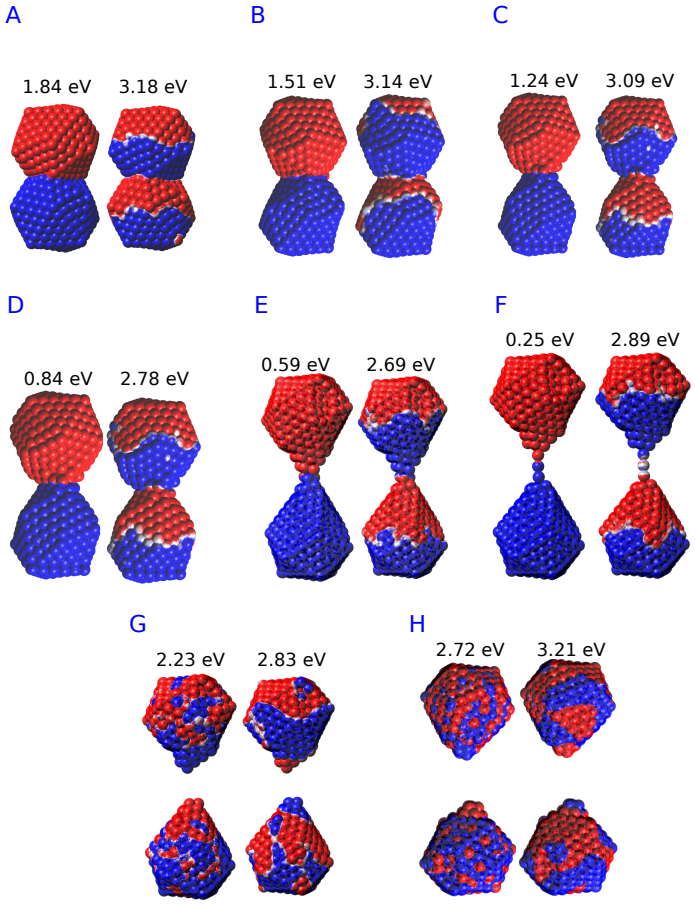


Figure 8.13. ωFQ MEP maps for plasmon excitations (eV) of selected structures A-H. The nominal gap distances d are (from A to H): 3.7, 9.7, 14.7, 22.7, 25.9, 32.1, 32.3, 34.1 Å. Blue color indicates a negative charge, whereas red color indicates a positive charge.

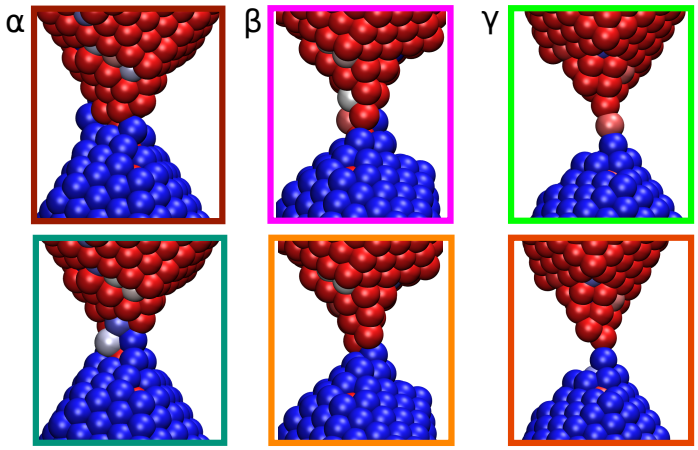


Figure 8.14. ωFQ imaginary charges on the junction atoms before (top) and after (bottom) the α, β and γ spectral jumps. Blue color indicates a negative charge, whereas red color indicates a positive charge.

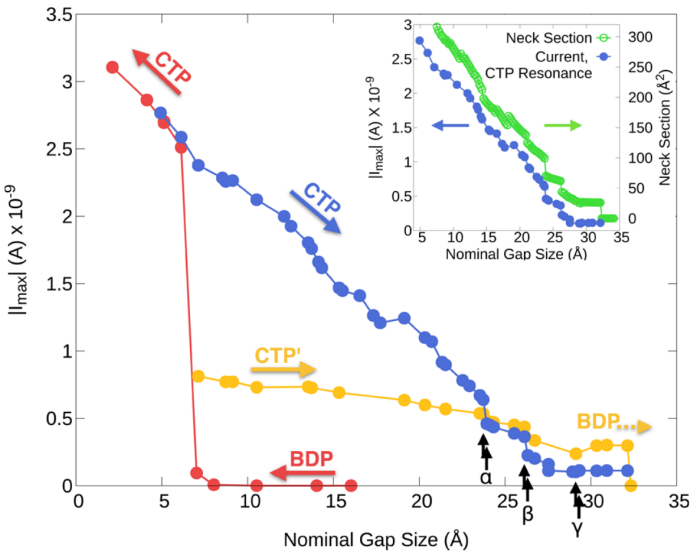


Figure 8.15. DFT absolute value of the electric current through the plasmonic nanojunction as a function of the elongation distance. Colored arrows indicate the direction of the process (approaching, in red, and retracting, in blue and orange). Black arrows indicate the position of the α β and γ spectral jumps. Data reproduced from Ref.[233]

Acknowledgments

We are greatly thankful to Prof. Risto Nieminen (Ålto University, Finland) and Prof. Daniel Sanchez-Portal (Centro de Fisica de Materiales CSIC-UPV/EHU Donostia-San Sebastian) for providing us the molecular structures of their studies. Financial

support from SNS "Progetti Interni Coordinati 2016" is acknowledged. C. C. gratefully acknowledges the support of H2020-MSCA-ITN-2017 European Training Network "Computational Spectroscopy In Natural sciences and Engineering" (COSINE), grant number 765739. S. C. acknowledges funding from ERC under the grant ERC-CoG-681285 TAME-Plasmons. We are thankful for the computer resources provided by the high performance computer facilities of the SMART Laboratory (http://smart.sns.it/).

Chapter 9

Summary, Conclusions and Future Perspectives

Whenever a theory appears to you as the only possible one,
take this as a sign that you have neither understood the theory
nor the problem which it was intended to solve
(Karl Popper)

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 work has been divided into three main steps: first, QM/Fluctuating Charge(FQ) developed in our group has been applied and extended to the calculation of different properties/spectroscopies of molecular systems in aqueous solution. Then, the purely electrostatic description given by commonly used QM/MM approaches has been extended so to include repulsion and dispersion energy terms into the QM Hamiltonian, thus allowing a future extension to molecular properties/spectroscopies. QM/FQ has also been extended so that an additional source of polarization, described in terms of fluctuating dipoles, is included in the MM portion. Finally, a novel FQ-based approach able to accurately describe the optical properties of a metal nanomaterial under the effect of an external electric oscillating field has been proposed. All the theoretical methods proposed in this thesis are expressed in a common framework in terms of the QM density. Due to this particular feature, all the proposed approaches can be further extended to molecular properties/spectroscopies by following what has already been done for QM/FQ.[56] In particular, once energy is formulated, molecular properties/spectroscopies can be obtained by appropriate (quasi)energy derivatives/response functions.
In Chapter 1, QM/FQ has been applied to the calculation of IR, Vibrational Circular Dichroism (VCD), Raman and Raman Optical Activity (ROA) spectra of (L)-Methyl Lactate and (S)-Glycidol in aqueous solution. Although the two molecules are rather small (15 and 12 atoms, respectively) they are characterized by several conformers in aqueous solution. Therefore, in order to obtain computed spectra which are in good agreement with their experimental counterparts, a good level of interplay between con-

formational and solvent effects is mandatory. QM/FQ shows an excellent agreement between computed and experimental spectra, and of higher quality if compared to standard continuum solvation approaches. Such discrepancies are not only due to the inaccurate description of HB interactions in the continuum approach, but also to a different sampling of the PES resulting from the static Polarizable Continuum Model (PCM) or dynamic QM/FQ+Molecular Dynamics (MD) approach. In Chapter 2, we have presented a computational study examining the merits and shortcomings of five different solvation models (QM/PCM, QM/QM$_w$/PCM, non polarizable QM/MM, QM/FQ, QM/QM$_w$/PCM) in modeling UV-Vis absorption spectra of organic chromophores in aqueous solution. The picture that has emerged shows that the performance of each model is highly dependent on the specific properties of each system, in particular on the extent of charge transfer character of the different excitations. However, a general conclusion that can be drawn is that the inclusion of solute-solvent polarization effects, whether using continuum or discrete models, can often be crucial and lead to a significant improvement in the results. In Chapter 3, the extension of QM/FQ approach to the calculation of SHG of selected organic acids in aqueous solution is reported. In the selected systems, the interaction with the surrounding environment is dominated by HB interactions. QM/FQ hows a good agreement in both the modeling of solvent effects and in the reproduction of experimental SHG data extracted from Hyper Rayleigh Scattering (HRS) experiments. In particular, also in case of SHG, polarization effects as reproduced by QM/FQ, are mandatory to get good agreement with experimental data.

On the other hand, QM/FQ is based on the assumption that electrostatic energy terms dominate QM/MM interactions. Although this can be a reasonable assumption in case of electrostatic-dominated systems, as aqueous solutions, this cannot be considered valid in case of other environments. In addition, in FQ only charges are placed in the MM portion. This poses some conceptual issues because only monopoles, i.e. zeroth order of the electrostatic Taylor expansion, are taken into consideration. These two conceptual limitations of QM/FQ have been overcome in Chapters 4-7, in which a model to include repulsion/dispersion contributions and a model which considers both fluctuating charges and fluctuating dipoles (FQFμ) have been proposed. Both novel approaches are based on the strong points of QM/FQ. In fact, they are both formulated in terms of the QM density, and the variational formalism of QM/FQ is kept.

The inclusion of repulsion/dispersion contributions into the molecular Hamiltonian, and the further extension of the electrostatic description in QM/FQFμ pave the way to the definition of a general QM/MM approach with *ab-initio* accuracy. As reported for EPR hyperfine coupling constants (see Chapter 6), an adequate inclusion of all the different contributions (i.e. electrostatics, repulsion and dispersion) can lead to results directly comparable with experimental findings. Furthermore, a reliable inclusion of all the different contributions to describe solute/solvent interaction can also permit an analysis of the influence of the different terms on molecular properties and spectroscopies. This is worth being investigated: in fact, several approach to decompose interaction energies have been proposed in the past, but to the best of our knowledge, a similar analysis focused on properties and spectroscopies has never been performed. In this thesis, both the theoretical model to include Pauli Repulsion and dispersion and

the novel QM/FQFμ approach have only been applied to aqueous solutions. However, their formulation is general enough to be applied to any kind of environment, pending a suitable parametrization. Such a parametrization can be effectively performed by following the strategy outlined in Sections 5.5.1 and 7.4.1. In this sense, molecule/environment couples dominated by non electrostatics can be treated. This is for instance the case of benzene solutions: our model, suitably extended to treat such an environment, can be used to verify and deeply understand the necessity of resorting to cluster approaches for an accurate descriptioin of molecular properties, such as the Optical Rotation of (R)-Methyloxirane in benzene solution.[604]

The last chapter of the thesis is dedicated to the novel ωFQ approach to describe the optical properties of nanoparticles and nanoaggregates, in particular the so-called sub-nanometer nanojunctions. ωFQ model is based on classical electrodynamics, effectively corrected to consider purely quantum effects such as electron tunneling. This is achieved trough a damping Fermi-like function which exponentially decreases the interaction between the charges.

The main future perspective of ωFQ is to describe optical properties and spectroscopies of molecules adsorbed on plasmonic nanomaterials. When such a system is irradiated by an external oscillating electric field, the properties and the spectroscopic signals of the molecule are enhanced, giving rise to signals which can be detected for instance by resorting to Surface Enhanced Raman Scattering (SERS).[210] The main reasons of such an enhancement are attributed to the enhancement of the electric field which acts on the molecule as a response of the optical absorption of the nanomaterial. Therefore, in order to appropriately reproduce SERS, an effective model to describe the nanomaterial and its spectroscopic absorption is a fundamental prerequisite. ωFQ is a good candidate to describe such features and the its application to SERS would be worth being investigated. In addition, ωFQ has been also applied to the challenging problem of nanojunctions, i.e. sub-nanometer gaps between two interacting nanoparticles, resulting in a nice agreement with *ab-initio* data. Nanojunctions are usually exploited in Surface Enhanced spectroscopy to obtain single molecule detection, due to the gargantuan enhancement of the electric field in the gap. To conclude this discussion, ωFQ has also the potentiality of being applied to 2D materials, such as graphene sheets. This can be done by including the features of such substrates in the theoretical formulation of the model, such as Fermi Level Energy, graphene 2D-density and its effective mass. In such a way, also newly experimental techniques, such as Graphene Enhanced Raman Scattering (GERS),[605] which have never been theoretically studied, could be investigated for the first time.

Bibliography

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

[2] Warshel, A.; Karplus, M. *J. Am. Chem. Soc.* **1972**, *94*, 5612–5625.

[3] Miertuš, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117–129.

[4] Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027–2094.

[5] Orozco, M.; Luque, F. J. *Chem. Rev.* **2000**, *100*, 4187–4226.

[6] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3094.

[7] Warshel, A. *Ann. Rev. Bioph. Biom.* **2003**, *32*, 425–443.

[8] Gao, J.; Xia, X. *Science* **1992**, *258*, 631–635.

[9] Vreven, T.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B.; Frisch, M. J. *J. ComputĊhem* **2003**, *24*, 760–769.

[10] Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. *J. Phys. Chem. B* **2011**, *115*, 3027–3037.

[11] Senn, H. M.; Thiel, W. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182–187.

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

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

[14] Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A. J.; Morokuma, K.; Frisch, M. J. *J.Chem. Theory and Comput.* **2006**, *2*, 815–826.

[15] Dapprich, S.; Komaromi, I.; Byun, K.; Morokuma, K.; Frisch, M. *J. Mol. Struct. - Theochem* **1999**, *461*, 1–21.

[16] Vreven, T.; Mennucci, B.; da Silva, C.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62–72.

[17] Rega, N.; Cossi, M.; Barone, V. *J. Am. Chem. Soc.* **1998**, *120*, 5723–5732.

[18] Barone, V.; Bloino, J.; Monti, S.; Pedone, A.; Prampolini, G. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10550–10561.

[19] Barone, V.; Biczysko, M.; Brancato, G. In *Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods*; Sabin, J. R., Canuto, S., Eds.; Advances in Quantum Chemistry; Academic Press, 2010; Vol. 59; pp 17–57.

[20] Rega, N.; Brancato, G.; Barone, V. *Chem. Phys. Lett.* **2006**, *422*, 367–371.

[21] Rega, N.; Brancato, G.; Petrone, A.; Caruso, P.; Barone, V. *J. Chem. Phys.* **2011**, *134*, 074504.

[22] Thompson, M. A. *J. Phys. Chem.* **1996**, *100*, 14492–14507.

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

[24] Gao, J.; Freindorf, M. *J. Phys. Chem. A* **1997**, *101*, 3182–3188.

[25] Martín, M. E.; Sánchez, M. L.; del Valle, F. J. O.; Aguilar, M. A. *J. Chem. Phys.* **2000**, *113*, 6308–6315.

[26] Cui, Q.; Karplus, M. *J. Chem. Phys.* **2000**, *112*, 1133–1149.

[27] Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2000**, *104*, 3721–3743.

[28] Cappelli, C.; Mennucci, B.; Monti, S. *J. Phys. Chem. A* **2005**, *109*, 1933–1943.

[29] Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

[30] Schwabe, T.; Olsen, J. M. H.; Sneskov, K.; Kongsted, J.; Christiansen, O. *J. Chem. Theory Comput.* **2011**, *7*, 2209–2217.

[31] Nielsen, C. B.; Christiansen, O.; Mikkelsen, K. V.; Kongsted, J. *J. Chem. Phys.* **2007**, *126*, 154112.

[32] Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. *J. Chem. Phys.* **2003**, *118*, 1620–1633.

[33] Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551–18560.

[34] Steindal, A. H.; Olsen, J. M. H.; Ruud, K.; Frediani, L.; Kongsted, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 5440–5451.

[35] Marini, A.; Muñoz-Losa, A.; Biancardi, A.; Mennucci, B. *J. Phys. Chem. B* **2010**, *114*, 17128–17135.

[36] Jacobson, L. D.; Herbert, J. M. *J. Chem. Phys.* **2010**, *133*, 154506.

[37] Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. *J. Chem. Theory Comput.* **2009**, *5*, 1838–1848.

[38] Scholes, G. D.; Curutchet, C.; Mennucci, B.; Cammi, R.; Tomasi, J. *J. Phys. Chem. B* **2007**, *111*, 6978–6982.

[39] Curutchet, C.; Kongsted, J.; Muñoz-Losa, A.; Hossein-Nejad, H.; Scholes, G. D.; Mennucci, B. *J. Am. Chem. Soc.* **2011**, *133*, 3078–3084.

[40] Bryce, R. A.; Vincent, M. A.; Malcolm, N. O. J.; Hillier, I. H.; Burton, N. A. *J. Chem. Phys.* **1998**, *109*, 3077–3085.

[41] Bryce, R. A.; Vincent, M. A.; Hillier, I. H. *J. Phys. Chem. A* **1999**, *103*, 4094–4100.

[42] Hillier, I. H. *J. Mol. Struct. - Theochem* **1999**, *463*, 45 – 52.

[43] Arora, P.; Slipchenko, L. V.; Webb, S. P.; DeFusco, A.; Gordon, M. S. *J. Phys. Chem. A* **2010**, *114*, 6742–6750.

[44] Li, H.; Gordon, M. S. *J. Chem. Phys* **2007**, *126*, 124112.

[45] Li, H. *J. Chem. Phys* **2009**, *131*, 184103.

[46] Gao, J. *J. Phys. Chem. B* **1997**, *101*, 657–663.

[47] Gao, J. *J. Chem. Phys.* **1998**, *109*, 2346–2354.

[48] Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Phys. Chem. B* **2008**, *112*, 14124–14131.

[49] Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2008**, *128*, 234108.

[50] Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 459–467.

[51] Song, L.; Han, J.; Lin, Y.-l.; Xie, W.; Gao, J. *J. Phys. Chem. A* **2009**, *113*, 11656–11664.

[52] Cembran, A.; Payaka, A.; Lin, Y.-l.; Xie, W.; Mo, Y.; Song, L.; Gao, J. *J. Chem. Theory Comput.* **2010**, *6*, 2242–2251.

[53] Gao, J.; Cembran, A.; Mo, Y. *J. Chem. Theory Comput.* **2010**, *6*, 2402–2410.

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

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

[56] Cappelli, C. *Int. J. Quantum Chem.* **2016**, *116*, 1532–1542.

[57] Boulanger, E.; Harvey, J. N. *Curr. Opin. Struc. Biol.* **2018**, *49*, 72–76.

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

[59] Boulanger, E.; Thiel, W. *J. Chem. Theory Comput.* **2012**, *8*, 4527–4538.

[60] Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. *J. Phys. Chem. A* **2015**, *119*, 5344–5355.

[61] Nåbo, L. J.; Olsen, J. M. H.; Holmgaard List, N.; Solanko, L. M.; Wüstner, D.; Kongsted, J. *J. Chem. Phys.* **2016**, *145*, 104102.

[62] Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. *J. Phys. Chem. Lett.* **2017**, *8*, 5949–5958.

[63] Nåbo, L. J.; Olsen, J. M. H.; Martínez, T. J.; Kongsted, J. *J. Chem. Theory Comput.* **2017**, *13*, 6230–6236.

[64] Mennucci, B., Cammi, R., Eds. *Continuum Solvation Models in Chemical Physics*; Wiley, New York, 2007.

[65] Cances, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

[66] Mennucci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101*, 10506–10517.

[67] Mennucci, B. *J. Phys. Chem. Lett.* **2010**, *1*, 1666–1674.

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

[69] Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210–3221.

[70] Barone, V.; Cossi, M. *J. Phys. Chem A* **1998**, *102*, 1995–2001.

[71] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem* **2003**, *24*, 669–681.

[72] Cossi, M.; Barone, V. *J. Chem. Phys.* **1998**, *109*, 6246–6254.

[73] Cammi, R.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1996**, *104*, 4611–4620.

[74] Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43–54.

[75] Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2010**, *132*, 114110.

[76] Lipparini, F.; Scalmani, G.; Mennucci, B.; Cancès, E.; Caricato, M.; Frisch, M. J. *J. Chem. Phys.* **2010**, *133*, 014106.

[77] Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett* **1995**, *246*, 122 – 129.

[78] Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824–19839.

[79] Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 2011–2033.

[80] Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2008**, *112*, 3024–3039.

[81] Liu, J.; Kelly, C. P.; Goren, A. C.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; Zhan, C.-G. *J. Chem. Theory Comput.* **2010**, *6*, 1109–1117.

[82] Mennucci, B. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6583–6594.

[83] Cammi, R. *Molecular Response Functions for the Polarizable Continuum Model: Physical basis and quantum mechanical formalism*; Springer Science & Business Media, 2013.

[84] Floris, F.; Tomasi, J.; Ahuir, J. P. *J. Comput. Chem.* **1991**, *12*, 784–791.

[85] Floris, F.; Tomasi, J. *J. Comput. Chem.* **1989**, *10*, 616–627.

[86] Amovilli, C.; Mennucci, B. *J. Phys. Chem. B* **1997**, *101*, 1051–1057.

[87] Mennucci, B.; Amovilli, C.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 221–225.

[88] Cupellini, L.; Amovilli, C.; Mennucci, B. *J. Phys. Chem. B* **2014**, *119*, 8984–8991.

[89] Biancardi, A.; Cammi, R.; Cappelli, C.; Mennucci, B.; Tomasi, J. *Theor. Chem. Acc.* **2012**, *131*, 1–10.

[90] Cammi, R.; Cossi, M.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1996**, *105*, 10556–10564.

[91] Cappelli, C.; Corni, S.; Mennucci, B.; Cammi, R.; Tomasi, J. *J. Phys. Chem A* **2002**, *106*, 12331–12339.

[92] Cappelli, C.; Corni, S.; Cammi, R.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2000**, *113*, 11270–11279.

[93] Mennucci, B. *J. Am. Chem. Soc.* **2002**, *124*, 1506–1515.

[94] Mennucci, B.; Cappelli, C.; Cammi, R.; Tomasi, J. *Chirality* **2011**, *23*, 717–729.

[95] Cammi, R.; Verdolino, V.; Mennucci, B.; Tomasi, J. *Chem. Phys.* **2008**, *344*, 135–141.

[96] Cammi, R.; Cappelli, C.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2012**, *137*, 154112.

[97] Tomasi, J.; Cammi, R.; Mennucci, B.; Cappelli, C.; Corni, S. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5697–5712.

[98] Cappelli, C.; Mennucci, B.; Cammi, R.; Rizzo, A. *J. Phys. Chem. B* **2005**, *109*, 18706–18714.

[99] Cappelli, C.; Rizzo, A.; Mennucci, B.; Tomasi, J.; Cammi, R.; Rikken, G. L.; Mathevet, R.; Rizzo, C. *J. Chem. Phys.* **2003**, *118*, 10712–10724.

[100] Cappelli, C.; Mennucci, B. *J. Phys. Chem. B* **2008**, *112*, 3441–3450.

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

[102] Jacquemin, D.; Planchat, A.; Adamo, C.; Mennucci, B. *J. Chem. Theory and Comput.* **2012**, *8*, 2359–2372.

[103] Jacquemin, D.; Mennucci, B.; Adamo, C. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16987–16998.

[104] Mennucci, B.; Cappelli, C.; Guido, C. A.; Cammi, R.; Tomasi, J. *The Journal of Physical Chemistry A* **2009**, *113*, 3009–3020.

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

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

[107] Corni, S.; Tomasi, J. *Surface-Enhanced Raman Scattering* **2006**, 105–123.

[108] Vukovic, S.; Corni, S.; Mennucci, B. *J. Phys. Chem. C* **2008**, *113*, 121–133.

[109] Angioni, A.; Corni, S.; Mennucci, B. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3294–3303.

[110] Cammi, R.; Cappelli, C.; Corni, S.; Tomasi, J. *J. Phys. Chem A* **2000**, *104*, 9874–9879.

[111] Corni, S.; Cappelli, C.; Cammi, R.; Tomasi, J. *J. Phys. Chem. A* **2001**, *105*, 8310–8316.

[112] Cappelli, C.; Corni, S.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 5531–5535.

[113] Corni, S.; Tomasi, J. *J. Chem. Phys.* **2002**, *116*, 1156–1164.

[114] Ferrighi, L.; Frediani, L.; Cappelli, C.; Sałek, P.; Ågren, H.; Helgaker, T.; Ruud, K. *Chem. Phys. Lett.* **2006**, *425*, 267–272.

[115] Pecul, M.; Lamparska, E.; Cappelli, C.; Frediani, L.; Ruud, K. *J. Phys. Chem.. A* **2006**, *110*, 2807–2815.

[116] Frediani, L.; Rinkevicius, Z.; Ågren, H. *J. Chem. Phys.* **2005**, *122*, 244104.

[117] Ferrighi, L.; Frediani, L.; Fossgaard, E.; Ruud, K. *J. Chem. Phys.* **2007**, *127*, 244103.

[118] Solheim, H.; Frediani, L.; Ruud, K.; Coriani, S. *Theor. Chem. Acc.* **2008**, *119*, 231–244.

[119] Ruud, K.; Frediani, L.; Cammi, R.; Mennucci, B. *Int. J. Mol. Sci.* **2003**, *4*, 119–134.

[120] Cappelli, C.; Lipparini, F.; Bloino, J.; Barone, V. *J. Chem. Phys.* **2011**, *135*, 104505.

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

[122] Cappelli, C.; Bloino, J.; Lipparini, F.; Barone, V. *J. Phys. Chem.. Lett.* **2012**, *3*, 1766–1773.

[123] Egidi, F.; Giovannini, T.; Piccardo, M.; Bloino, J.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput˙* **2014**, *10*, 2456–2464.

[124] Vidal, L. N.; Giovannini, T.; Cappelli, C. *J. Phys. Chem. Lett.* **2016**, *7*, 3585–3590.

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

[126] Santoro, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2011**, *7*, 1824–1839.

[127] Egidi, F.; Bloino, J.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2013**, *10*, 346–363.

[128] Avila, F. F.; Barone, V.; Cappelli, C.; Santoro, F. *J. Chem. Theory Comput.* **2013**, *9*, 3597–3611.

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

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

[131] Peng, B.; Lestrange, P. J.; Goings, J. J.; Caricato, M.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 4146–4153.

[132] Di Remigio, R.; Beerepoot, M. T.; Cornaton, Y.; Ringholm, M.; Steindal, A. H.; Ruud, K.; Frediani, L. *Phys. Chem. Chem. Phys.* **2017**, *19*, 366–379.

[133] Cammi, R.; Frediani, L.; Mennucci, B.; Tomasi, J.; Ruud, K.; Mikkelsen, K. V. *J. Chem. Phys.* **2002**, *117*, 13–26.

[134] Caricato, M.; Scalmani, G.; Trucks, G. W.; Frisch, M. J. *J. Phys. Chem. Lett.* **2010**, *1*, 2369–2373.

[135] Cammi, R. *Int. J. Quantum Chem.* **2010**, *110*, 3040–3052.

[136] Cammi, R. *J. Chem. Phys.* **2009**, *131*, 164104.

[137] Caricato, M.; Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2011**, *134*, 244113.

[138] Caricato, M. *Int. J. Quantum Chem.* **2018**, e25674.

[139] Remigio, R. D.; Repisky, M.; Komorovsky, S.; Hrobarik, P.; Frediani, L.; Ruud, K. *Mol. Phys.* **2017**, *115*, 214–227.

[140] Bugeanu, M.; Di Remigio, R.; Mozgawa, K.; Reine, S. S.; Harbrecht, H.; Frediani, L. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31566–31581.

[141] Remigio, R. D.; Bast, R.; Frediani, L.; Saue, T. *J. Phys. Chem. A* **2014**, *119*, 5061–5077.

[142] Toniolo, A.; Granucci, G.; Martínez, T. J. *J. Phys. Chem. A* **2003**, *107*, 3822–3830.

[143] Isborn, C. M.; Gotz, A. W.; Clark, M. A.; Walker, R. C.; Martínez, T. J. *J. Chem. Theory Comput.* **2012**, *8*, 5092–5106.

[144] Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martinez, T. J. *J. Phys. Chem. B* **2016**, *120*, 11381–11394.

[145] Punwong, C.; Owens, J.; Martínez, T. J. *J. Phys. Chem. B* **2014**, *119*, 704–714.

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

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

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

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

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

[151] Giovannini, T.; Del Frate, G.; Lafiosca, P.; Cappelli, C. *Phys. Chem. Chem. Phys.* **2018**, *20*, 9181–9197.

[152] Giovannini, T.; Ambrosetti, M.; Cappelli, C. *Theor. Chem. Acc.* **2018**, *137*, 74.

[153] Giovannini, T.; Macchiagodena, M.; Ambrosetti, M.; Puglisi, A.; Lafiosca, P.; Lo Gerfo, G.; Egidi, F.; Cappelli, C. *Int. J. Quantum Chem.* **2018**, e25684.

[154] Loco, D.; Cupellini, L. *Int. J. Quantum Chem.* **2018**, DOI: 10.1002/qua.25726.

[155] Jurinovich, S.; Curutchet, C.; Mennucci, B. *ChemPhysChem* **2014**, *15*, 3194–3204.

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

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

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

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

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

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

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

[163] Egidi, F.; Lo Gerfo, G.; Macchiagodena, M.; Cappelli, C. *Theor. Chem. Acc.* **2018**, *137*, 82.

[164] Thole, B. T. *Chem. Phys.* **1981**, *59*, 341–350.

[165] Caprasecca, S.; Jurinovich, S.; Viani, L.; Curutchet, C.; Mennucci, B. *J. Chem. Theory Comput.* **2014**, *10*, 1588–1598.

[166] Donati, G.; Wildman, A.; Caprasecca, S.; Lingerfelt, D. B.; Lipparini, F.; Mennucci, B.; Li, X. *J. Phys. Chem. Lett.* **2017**, *8*, 5283–5289.

[167] Jurinovich, S.; Pescitelli, G.; Di Bari, L.; Mennucci, B. *Phys. Chem. Chem. Phys.* **2014**, *16*, 16407–16418.

[168] Loco, D.; Jurinovich, S.; Cupellini, L.; Menger, M. F.; Mennucci, B. *Photochem. Photobiol. Sci.* **2018**, *17*, 552–560.

[169] Caprasecca, S.; Cupellini, L.; Jurinovich, S.; Loco, D.; Lipparini, F.; Mennucci, B. *Theor. Chem. Acc.* **2018**, *137*, 84.

[170] Lamoureux, G.; MacKerell Jr, A. D.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185–5197.

[171] Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025–3039.

[172] Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. *J. Chem. Theory Comput.* **2005**, *1*, 153–168.

[173] Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell Jr, A. D. *Chem. Rev.* **2016**, *116*, 4983–5013.

[174] Geerke, D.; van Gunsteren, W. F. *Mol. Phys.* **2007**, *105*, 1861–1881.

[175] Geerke, D. P.; Thiel, S.; Thiel, W.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2007**, *3*, 1499–1509.

[176] Boulanger, E.; Thiel, W. *J. Chem. Theory Comput.* **2014**, *10*, 1795–1809.

[177] Lu, Z.; Zhang, Y. *J. Chem. Theory Comput.* **2008**, *4*, 1237–1248.

[178] Huang, J.; Simmonett, A. C.; Pickard IV, F. C.; MacKerell Jr, A. D.; Brooks, B. R. *J. Chem. Phys.* **2017**, *147*, 161702.

[179] Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

[180] Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.

[181] Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. *J. Chem. Theory Comput.* **2013**, *9*, 4046–4063.

[182] Dziedzic, J.; Mao, Y.; Shao, Y.; Ponder, J.; Head-Gordon, T.; Head-Gordon, M.; Skylaris, C.-K. *J. Chem. Phys.* **2016**, *145*, 124106.

[183] Loco, D.; Buda, F.; Lugtenburg, J.; Mennucci, B. *J. Phys. Chem. Lett.* **2018**, *9*, 2404–2410.

[184] Mao, Y.; Shao, Y.; Dziedzic, J.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. *J. Chem. Theory Comput.* **2017**, *13*, 1963–1979.

[185] Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. *J. Chem. Phys.* **1996**, *105*, 1968–1986.

[186] Kairys, V.; Jensen, J. H. *J. Phys. Chem. A* **2000**, *104*, 6656–6665.

[187] Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177–193.

[188] Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293–307.

[189] Adamovic, I.; Gordon, M. S. *Mol. Phys.* **2005**, *103*, 379–387.

[190] Slipchenko, L. V. *Many-Body Effects and Electrostatics in Biomolecules*; Pan Stanford, 2016; pp 147–187.

[191] DeFusco, A.; Minezawa, N.; Slipchenko, L. V.; Zahariev, F.; Gordon, M. S. *J. Phys. Chem. Lett.* **2011**, *2*, 2184–2192.

[192] McWeeny, R. *Methods of molecular quantum mechanics*; Academic press: London, 1992.

[193] Stone, A. *The theory of intermolecular forces*; OUP Oxford, 2013.

[194] Grimme, S. *WIRES: Comput. Mol. Sci.* **2011**, *1*, 211–228.

[195] Lennard-Jones, J. E. *Proc. Phys. Soc.* **1931**, *43*, 461.

[196] Boulanger, E.; Huang, L.; Rupakheti, C.; MacKerell Jr, A. D.; Roux, B. *J. Chem. Theory Comput.* **2018**, *14*, 3121–3131.

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

[198] Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 073005.

[199] Tkatchenko, A.; Romaner, L.; Hofmann, O. T.; Zojer, E.; Ambrosch-Draxl, C.; Scheffler, M. *MRS bulletin* **2010**, *35*, 435–442.

[200] Tkatchenko, A.; DiStasio Jr, R. A.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012**, *108*, 236402.

[201] Hermann, J.; DiStasio, R. A.; Tkatchenko, A. *Chem. Rev.* **2017**, *117*, 4714–4758.

[202] Curutchet, C.; Cupellini, L.; Kongsted, J.; Corni, S.; Frediani, L.; Steindal, A. H.; Guido, C. A.; Scalmani, G.; Mennucci, B. *J. Chem. Theory Comput.* **2018**, *14*, 1671–1681.

[203] Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. *submitted*

[204] Xu, H.; Aizpurua, J.; Käll, M.; Apell, P. *Phys. Rev. E* **2000**, *62*, 4318.

[205] Halas, N. J. *Nano Lett.* **2010**, *10*, 3816–3822.

[206] Hao, E.; Schatz, G. C. *J. Chem. Phys.* **2004**, *120*, 357–366.

[207] Crozier, K.; Sundaramurthy, A.; Kino, G.; Quate, C. *J. Appl. Phys.* **2003**, *94*, 4632–4642.

[208] Polman, A.; Atwater, H. A. *Nat. Mater.* **2012**, *11*, 174–177.

[209] Halas, N. J.; Lal, S.; Chang, W.-S.; Link, S.; Nordlander, P. *Chem. Rev.* **2011**, *111*, 3913–3961.

[210] Stiles, P. L.; Dieringer, J. A.; Shah, N. C.; Van Duyne, R. P. *Annu. Rev. Anal. Chem.* **2008**, *1*, 601–626.

[211] Le Ru, E. C.; Etchegoin, P. G. *Annu. Rev. Phys. Chem.* **2012**, *63*, 65–87.

[212] Lim, D.-K.; Jeon, K.-S.; Kim, H. M.; Nam, J.-M.; Suh, Y. D. *Nat. Mater.* **2010**, *9*, 60–67.

[213] Eustis, S.; El-Sayed, M. A. *Chem. Soc. Rev.* **2006**, *35*, 209–217.

[214] Blaber, M.; Arnold, M.; Harris, N.; Ford, M.; Cortie, M. *Physica B* **2007**, *394*, 184–187.

[215] Aubry, A.; Lei, D. Y.; Fernández-Domínguez, A. I.; Sonnefraud, Y.; Maier, S. A.; Pendry, J. B. *Nano Lett.* **2010**, *10*, 2574–2579.

[216] Becker, J.; Trügler, A.; Jakab, A.; Hohenester, U.; Sönnichsen, C. *Plasmonics* **2010**, *5*, 161–167.

[217] Maier, S. A. *Plasmonics: fundamentals and applications*; Springer Science & Business Media, 2007.

[218] Roldán, A.; Viñes, F.; Illas, F.; Ricart, J. M.; Neyman, K. M. *Theor. Chem. Acc.* **2008**, *120*, 565–573.

[219] Kuisma, M.; Sakko, A.; Rossi, T. P.; Larsen, A. H.; Enkovaara, J.; Lehtovaara, L.; Rantala, T. T. *Phys. Rev. B* **2015**, *91*, 115431.

[220] Townsend, E.; Bryant, G. W. *Nano Lett.* **2011**, *12*, 429–434.

[221] Barbry, M.; Koval, P.; Marchesin, F.; Esteban, R.; Borisov, A.; Aizpurua, J.; Sánchez-Portal, D. *Nano Lett.* **2015**, *15*, 3410–3419.

[222] Garzón, I.; Michaelian, K.; Beltrán, M.; Posada-Amarillas, A.; Ordejón, P.; Artacho, E.; Sánchez-Portal, D.; Soler, J. *The European Physical Journal D-Atomic, Molecular, Optical and Plasma Physics* **1999**, *9*, 211–215.

[223] Sánchez-de Armas, R.; San-Miguel, M.; Oviedo, J.; Marquez, A.; Sanz, J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 1506–1514.

[224] Aizpurua, J.; Hanarp, P.; Sutherland, D.; Käll, M.; Bryant, G. W.; De Abajo, F. G. *Phys. Rev. Lett.* **2003**, *90*, 057401.

[225] Savage, K. J.; Hawkeye, M. M.; Esteban, R.; Borisov, A. G.; Aizpurua, J.; Baumberg, J. J. *Nature* **2012**, *491*, 574–577.

[226] Morton, S. M.; Jensen, L. *J. Chem. Phys.* **2010**, *133*, 074103.

[227] Morton, S. M.; Jensen, L. *J. Chem. Phys.* **2011**, *135*, 134103.

[228] Payton, J. L.; Morton, S. M.; Moore, J. E.; Jensen, L. *J. Chem. Phys.* **2012**, *136*, 214103.

[229] Jensen, L. L.; Jensen, L. *J. Phys. Chem. C* **2008**, *112*, 15697–15703.

[230] Chen, X.; Moore, J. E.; Zekarias, M.; Jensen, L. *Nat. Commun.* **2015**, *6*, 8921.

[231] Andreussi, O.; Corni, S.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2004**, *121*, 10190–10202.

[232] Trügler, A. *Optical properties of metallic nanoparticles*; Springer, 2011.

[233] Marchesin, F.; Koval, P.; Barbry, M.; Aizpurua, J.; Sánchez-Portal, D. *ACS Photonics* **2016**, *3*, 269–277.

[234] Rossi, T. P.; Zugarramurdi, A.; Puska, M. J.; Nieminen, R. M. *Phys. Rev. Lett.* **2015**, *115*, 236804.

[235] Romero, I.; Aizpurua, J.; Bryant, G. W.; De Abajo, F. J. G. *Opt. Expr.* **2006**, *14*, 9988–9999.

[236] Aubry, A.; Lei, D. Y.; Maier, S. A.; Pendry, J. *Phys. Rev. Lett.* **2010**, *105*, 233901.

[237] Prodan, E.; Radloff, C.; Halas, N. J.; Nordlander, P. *Science* **2003**, *302*, 419–422.

[238] Atay, T.; Song, J.-H.; Nurmikko, A. V. *Nano Lett.* **2004**, *4*, 1627–1631.

[239] Jain, P. K.; Huang, W.; El-Sayed, M. A. *Nano Lett.* **2007**, *7*, 2080–2088.

[240] Lassiter, J. B.; Aizpurua, J.; Hernandez, L. I.; Brandl, D. W.; Romero, I.; Lal, S.; Hafner, J. H.; Nordlander, P.; Halas, N. J. *Nano Lett.* **2008**, *8*, 1212–1218.

[241] Esteban, R.; Borisov, A. G.; Nordlander, P.; Aizpurua, J. *Nat. Commun.* **2012**, *3*, 825.

[242] Tan, S. F.; Wu, L.; Yang, J. K.; Bai, P.; Bosman, M.; Nijhuis, C. A. *Science* **2014**, *343*, 1496–1499.

[243] Chen, J.; Martínez, T. J. *Chem. Phys. Lett.* **2007**, *438*, 315–320.

[244] Chen, J.; Hundertmark, D.; Martínez, T. J. *J. Chem. Phys.* **2008**, *129*, 214113.

[245] Chen, J.; Martínez, T. J. *Advances in the Theory of Atomic and Molecular Systems* **2009**, 397–415.

[246] Barron, L. D. *Molecular light scattering and optical activity*; Cambridge University Press, 2004.

[247] Nafie, L. A. *Vibrational optical activity: principles and applications*; John Wiley & Sons: Chichester, 2011.

[248] Stephens, P. J.; Devlin, F. J.; Cheeseman, J. R. *VCD spectroscopy for organic chemists*; CRC Press, 2012.

[249] Losada, M.; Xu, Y. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3127–3135.

[250] Losada, M.; Tran, H.; Xu, Y. *J. Chem. Phys.* **2008**, *128*, 014508.

[251] Losada, M.; Nguyen, P.; Xu, Y. *J. Phys. Chem A* **2008**, *112*, 5621–5627.

[252] Yang, G.; Xu, Y. *J. Chem. Phys.* **2009**, *130*, 164506–164506.

[253] Poopari, M. R.; Zhu, P.; Dezhahang, Z.; Xu, Y. *J. Chem. Phys.* **2012**, *137*, 194308.

[254] Longhi, G.; Castiglioni, E.; Abbate, S.; Lebon, F.; Lightner, D. A. *Chirality* **2013**, *25*, 589–599.

[255] Berova, N.; Di Bari, L.; Pescitelli, G. *Chem. Soc. Rev.* **2007**, *36*, 914–931.

[256] Polavarapu, P. *J. Phys. Chem.* **1990**, *94*, 8106–8112.

[257] Jose, K. J.; Beckett, D.; Raghavachari, K. *J. Chem. Theory Comput.* **2015**, *11*, 4238–4247.

[258] Jose, K. J.; Raghavachari, K. *J. Chem. Theory Comput.* **2016**, *12*, 585–594.

[259] Barron, L. D. *Curr. Opin. Struct. Biol.* **2006**, *16*, 638–643.

[260] Merten, C.; Barron, L. D.; Hecht, L.; Johannessen, C. *Angew. Chem. Int. Ed.* **2011**, *50*, 9973–9976.

[261] Kurouski, D.; Lombardi, R. A.; Dukor, R. K.; Lednev, I. K.; Nafie, L. A. *Chem. Comm.* **2010**, *46*, 7154–7156.

[262] Oulevey, P.; Luber, S.; Varnholt, B.; Bürgi, T. *Angew. Chem. Int. Ed.* **2016**, *55*, 11787–11790.

[263] Ostovar pour, S.; Rocks, L.; Faulds, K.; Graham, D.; Parchaňskỳ, V.; Bouř, P.; Blanch, E. W. *Nature Chem.* **2015**, *7*, 591–596.

[264] Stephens, P.; Devlin, F.; Chabalowski, C.; Frisch, M. J. *J. Phys. Chem.*. **1994**, *98*, 11623–11627.

[265] Stephens, P. J. *J. Phys. Chem.* **1985**, *89*, 748–752.

[266] Stephens, P. *J. Phys. Chem.* **1987**, *91*, 1712–1715.

[267] Stephens, P.; Lowe, M. *Ann. Rev. Phys. Chem.* **1985**, *36*, 213–241.

[268] Amos, R.; Handy, N.; Jalkanen, K.; Stephens, P. *Chem. Phys. Lett.* **1987**, *133*, 21–26.

[269] Stephens, P. J.; Devlin, F. J.; Pan, J.-J. *Chirality* **2008**, *20*, 643–663.

[270] Stephens, P.; Devlin, F. *Chirality* **2000**, *12*, 172–179.

[271] Cheeseman, J.; Frisch, M.; Devlin, F.; Stephens, P. *Chem. Phys. Lett.* **1996**, *252*, 211–220.

[272] Autschbach, J. *Chirality* **2009**, *21*, E116–E152.

[273] Cheeseman, J. R.; Shaik, M. S.; Popelier, P. L.; Blanch, E. W. *J. Am. Chem. Soc.* **2011**, *133*, 4991–4997.

[274] Polavarapu, P. L. *Vibrational Spectra: Principles and Applications with Emphasis on Optical Activity: Principles and Applications with Emphasis on Optical Activity*; Elsevier, 1998; Vol. 85.

[275] Neugebauer, J.; Reiher, M.; Kind, C.; Hess, B. A. *J. Comput. Chem.* **2002**, *23*, 895–910.

[276] Reiher, M.; Brehm, G.; Schneider, S. *J. Phys. Chem. A* **2004**, *108*, 734–742.

[277] Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. *J. Chem. Phys.* **2014**, *141*, 034107.

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

[279] Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161–2200.

[280] Halgren, T. A.; Damm, W. *Curr. Opinion Struct. Biol.* **2001**, *11*, 236–242.

[281] Mao, Y.; Demerdash, O.; Head-Gordon, M.; Head-Gordon, T. *J. Chem. Theory Comput.* **2016**, *12*, 5422–5437.

[282] Stuart, S. J.; Berne, B. *J. Phys. Chem.* **1996**, *100*, 11934–11943.

[283] Gigante, D. M.; Long, F.; Bodack, L. A.; Evans, J. M.; Kallmerten, J.; Nafie, L. A.; Freedman, T. B. *J. Phys. Chem. A* **1999**, *103*, 1523–1537.

[284] Cassanas, G.; Morssli, M.; Fabregue, E.; Bardet, L. *J. Raman Spectrosc.* **1991**, *22*, 409–413.

[285] Badawi, H. M.; Ali, S. A. *Spectrochim. Acta A* **2009**, *74*, 558–562.

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

[287] Rappe, A. K.; Goddard III, W. A. *J. Phys. Chem* **1991**, *95*, 3358–3363.

[288] York, D. M.; Yang, W. *J. Chem. Phys.* **1996**, *104*, 159–172.

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

[290] Chelli, R.; Schettino, V.; Procacci, P. *J. Chem. Phys.* **2005**, *122*, 234107.

[291] Ohno, K. *Theor. Chim. Acta* **1964**, *2*, 219–227.

[292] Jin, S.; Head, J. D. *Surf. Science* **1994**, *318*, 204–216.

[293] Calvin, M. D.; Head, J. D.; Jin, S. *Surf. science* **1996**, *345*, 161–172.

[294] Cacelli, I.; Prampolini, G. *J. Chem. Theory and Comput.* **2007**, *3*, 1803–1817, PMID: 26627623.

[295] Grimme, S. *J. Chem. Theory Comput.* **2014**, *10*, 4497–4514, PMID: 26588146.

[296] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669–681.

[297] Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

[298] Dodda, L. S.; Vilseck, J. Z.; Cutrona, K. J.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2015**, *11*, 4273–4282.

[299] Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304–5314.

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

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

[302] Spencer, K. M.; Edmonds, R. B.; Rauh, R. D. *Appl. Spectrosc.* **1996**, *50*, 681–685.

[303] Borba, A.; Gómez-Zavaglia, A.; Lapinski, L.; Fausto, R. *Vib. Spectrosc.* **2004**, *36*, 79–88.

[304] Freedman, T. B.; Lee, E.; Nafie, L. A. *J. Phys. Chem. A* **2000**, *104*, 3944–3951.

[305] Aparicio, S. *J. Phys. Chem. A* **2007**, *111*, 4671–4683.

[306] Borho, N.; Xu, Y. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1324–1328.

[307] Sun, W.; Wu, J.; Zheng, B.; Zhu, Y.; Liu, C. *J. Mol. Struct.: THEOCHEM* **2007**, *809*, 161–169.

[308] Conrad, A.; Teumelsan, N.; Wang, P.; Tubergen, M. *J. Phys. Chem. A* **2009**, *114*, 336–342.

[309] Aparicio, S. *J. Phys. Chem. A* **2007**, *111*, 4671–4683, PMID: 17489567.

[310] Qiu, S.; Li, G.; Wang, P.; Jia, G.; Feng, Z.; Li, C. *J. Raman Spectrosc.* **2012**, *43*, 503–513.

[311] Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821–7839.

[312] Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

[313] Barone, V.; Cimino, P.; Pedone, A. *Magn. Reson. Chem.* **2010**, *48*, S11–S22.

[314] Stendardo, E.; Pedone, A.; Cimino, P.; Menziani, M. C.; Crescenzi, O.; Barone, V. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11697–11709.

[315] Scherrer, A.; Vuilleumier, R.; Sebastiani, D. *J. Chem. Theory Comput.* **2013**, *9*, 5305–5312.

[316] Barone, V.; Baiardi, A.; Biczysko, M.; Bloino, J.; Cappelli, C.; Lipparini, F. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12404–12422.

[317] Hättig, C.; Weigend, F. *J. Chem. Phys.* **2000**, *113*, 5154–5161.

[318] Egidi, F.; Segado, M.; Koch, H.; Cappelli, C.; Barone, V. *J. Chem. Phys.* **2014**, *141*, 224114.

[319] Improta, R.; Barone, V. *J. Am. Chem. Soc.* **2004**, *126*, 14320–14321.

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

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

[322] Jacquemin, D.; Perpete, E. A.; Scuseria, G. E.; Ciofini, I.; Adamo, C. *J. Chem. Theory Comput.* **2008**, *4*, 123–135.

[323] Laurent, A. D.; Jacquemin, D. *Int. J. Quantum Chem.* **2013**, *113*, 2019–2039.

[324] Goerigk, L.; Grimme, S. *J. Chem. Phys.* **2010**, *132*, 184103.

[325] Li, J.; Cramer, C. J.; Truhlar, D. G. *Int. J. Quantum Chem.* **2000**, *77*, 264–280.

[326] Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, 104103.

[327] Caricato, M.; Trucks, G. W.; Frisch, M. J.; Wiberg, K. B. *J. Chem. Theory Comput.* **2010**, *6*, 370–383.

[328] Isegawa, M.; Peverati, R.; Truhlar, D. G. *J. Chem. Phys.* **2012**, *137*, 244104.

[329] Isegawa, M.; Truhlar, D. G. *J. Chem. Phys.* **2013**, *138*, 134111.

[330] Alipour, M. *Theor. Chem. Acc.* **2016**, *135*, 67.

[331] Maier, T. M.; Bahmann, H.; Arbuznikov, A. V.; Kaupp, M. *J. Chem. Phys.* **2016**, *144*, 074106.

[332] Reichardt, C. *Chem. Soc. Rev.* **1992**, *21*, 147–153.

[333] Buncel, E.; Rajagopal, S. *Acc. Chem. Res.* **1990**, *23*, 226–231.

[334] Reichardt, C. *Chem. Rev.* **1994**, *94*, 2319–2358.

[335] Cannelli, O.; Giovannini, T.; Baiardi, A.; Carlotti, B.; Elisei, F.; Cappelli, C. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32544–32555.

[336] Prampolini, G.; Bellina, F.; Biczysko, M.; Cappelli, C.; Carta, L.; Lessi, M.; Pucci, A.; Ruggeri, G.; Barone, V. *Chemistry Eur. J.* **2013**, *19*, 1996–2004.

[337] Tosi, I.; Segado Centellas, M.; Campioli, E.; Iagatti, A.; Lapini, A.; Sissa, C.; Baldini, L.; Cappelli, C.; Di Donato, M.; Sansone, F.; Santoro, F.; Terenziani, F. *ChemPhysChem* **2016**, *17*, 1686–1706.

[338] Carlotti, B.; Cesaretti, A.; Cannelli, O.; Giovannini, T.; Cappelli, C.; Bonaccorso, C.; Fortuna, C. G.; Elisei, F.; Spalletti, A. *J. Phys. Chem. C* **2018**, *122*, 2285–2296.

[339] Lapini, A.; Fabbrizzi, P.; Piccardo, M.; di Donato, M.; Lascialfari, L.; Foggi, P.; Cicchi, S.; Biczysko, M.; Carnimeo, I.; Santoro, F.; Cappelli, C.; Righini, R. *Phys. Chem. Chem. Phys.* **2014**, *16*, 10059–10074.

[340] Budzák, S.; Laurent, A. D.; Laurence, C.; Medved', M.; Jacquemin, D. *J. Chem. Theory Comput.* **2016**, *12*, 1919–1929.

[341] Labat, F.; Le Bahers, T.; Ciofini, I.; Adamo, C. *Acc. Chem. Res.* **2012**, *45*, 1268–1277.

[342] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2013**, *9*, 3649–3659.

[343] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2014**, *119*, 958–967.

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

[345] Martínez-Fernández, L.; Pepino, A. J.; Segarra-Martí, J.; Banyasz, A.; Garavelli, M.; Improta, R. *J. Chem. Theory Comput.* **2016**, *12*, 4430–4439.

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

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

[348] Abrahama, M. J.; Murtola, T.; Schulz, R.; Pálla, S.; Smith, J. C.; Hess, B.; Lindahl, E. *SoftwareX* **2015**, *1-2*, 19–25.

[349] Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

[350] Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

[351] Sorin, E. J.; Pande, V. S. *Biophys. J.* **2005**, *88*, 2472 – 2493.

[352] Wang, L.-P.; Martinez, T. J.; Pande, V. S. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891.

[353] Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*.

[354] Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

[355] Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

[356] Antoon, L. H. *Annals Physics* **1881**, *248*, 127–136.

[357] Marcellin, B. *Comptes rendus hebdomadaires des séances de l'Académie des sciences* **1898**, *126*, 1703–1855.

[358] Mark, P.; Nilsson, L. *J. Phys. Chem. A* **2001**, *105*, 9954–9960.

[359] Brehm, M.; Kirchner, B. *J. Chem. Inf. Model.* **2011**, *51*, 2007–2023, PMID: 21761915.

[360] Egidi, F.; Trani, F.; Ballone, P. A.; Barone, V.; Andreoni, W. *Theoretical Chemical Accounts* **2016**, *135*, 264.

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

[362] Momeni, M. R.; Brown, A. *J. Chem. Theory Comput.* **2015**, *11*, 2619–2632.

[363] Le Guennic, B.; Jacquemin, D. *Acc. Chem. Res.* **2015**, *48*, 530–537.

[364] Arroyo, I. J.; Hu, R.; Merino, G.; Tang, B. Z.; Pena-Cabrera, E. *J. Org. Chem.* **2009**, *74*, 5719–5722.

[365] Politzer, I. R.; Crago, K. T.; Kiel, D. L.; Hampton, T. *Anal. Lett.* **1989**, *22*, 1567–1580.

[366] Bucsiová, L.; Hrdlovic, P. *J. Macromol. Sci. A* **2007**, *44*, 1047–1053.

[367] Hillig, K. W.; Morris, M. D. *Biochem. Bioph. Res. Comm.* **1976**, *71*, 1228 – 1233.

[368] Wolff, J. J.; Wortmann, R. *Adv. Phys. Org. Chem.* **1999**, *32*, 121–217.

[369] Innocenzi, P.; Lebeau, B. *J. Mater. Chem.* **2005**, *15*, 3821–3831.

[370] Boyd, R. W. *Nonlinear optics*; Academic press, 2003.

[371] Shen, Y.-R. *New York, Wiley-Interscience, 1984, 575 p.* **1984**,

[372] Marder, S. R.; Sohn, J. E.; Stucky, G. D. *Materials for nonlinear optics chemical perspectives*; 1991.

[373] Nalwa, H. S. *Appl. Organomet. Chem.* **1991**, *5*, 349–377.

[374] Chemla, D. S. *Nonlinear optical properties of organic molecules and crystals*; Elsevier, 2012; Vol. 1.

[375] Zyss, J. *Molecular nonlinear optics: materials, physics, and devices*; Academic press, 2013.

[376] Rizzo, A.; Cappelli, C.; Jansík, B.; Jonsson, D.; Sałek, P.; Coriani, S.; Wilson, D. J.; Helgaker, T.; Ågren, H. *J. Chem. Phys.* **2005**, *122*, 234314.

[377] Bishop, D. M.; Kirtman, B. *J. Chem. Phys.* **1991**, *95*, 2646–2658.

[378] Champagne, B. *Chem. Phys. Lett.* **1996**, *261*, 57–65.

[379] Bishop, D. M.; Luis, J. M.; Kirtman, B. *J. Chem. Phys.* **1998**, *108*, 10013–10017.

[380] Kirtman, B.; Bishop, D. M. *Chem. Phys. Lett.* **1990**, *175*, 601–607.

[381] Bishop, D. M.; Norman, P. *J. Chem. Phys.* **1999**, *111*, 3042–3050.

[382] Bishop, D. M.; Kirtman, B.; Kurtz, H. A.; Rice, J. E. *J. Chem. Phys.* **1993**, *98*, 8024–8030.

[383] Quinet, O.; Kirtman, B.; Champagne, B. *J. Chem. Phys.* **2003**, *118*, 505–513.

[384] Champagne, B.; Kirtman, B. *Chem. Phys.* **1999**, *245*, 213–226.

[385] Kongsted, J.; Christiansen, O. *J. Chem. Phys.* **2006**, *125*, 124108.

[386] Bishop, D. M.; Champagne, B.; Kirtman, B. *J. Chem. Phys.* **1998**, *109*, 9987–9994.

[387] Castiglioni, C.; Del Zoppo, M.; Zerbi, G. *Phys. Rev. B* **1996**, *53*, 13319.

[388] Luis, J. M.; Champagne, B.; Kirtman, B. *Int. J. Quantum Chem.* **2000**, *80*, 471–479.

[389] Benassi, E.; Egidi, F.; Barone, V. *J. Phys. Chem. B* **2015**, *119*, 3155–3173.

[390] Silva, D. L.; Fonseca, R. D.; Vivas, M. G.; Ishow, E.; Canuto, S.; Mendonca, C. R.; De Boni, L. *J. Chem. Phys.* **2015**, *142*, 064312.

[391] Vivas, M. G.; Silva, D. L.; Rodriguez, R. D.; Canuto, S.; Malinge, J.; Ishow, E.; Mendonca, C. R.; De Boni, L. *J. Phys. Chem. C* **2015**, *119*, 12589–12597.

[392] Dutra, A. S.; Castro, M. A.; Fonseca, T. L.; Fileti, E. E.; Canuto, S. *J. Chem. Phys.* **2010**, *132*, 034307.

[393] Salek, P.; Helgaker, T.; Vahtras, O.; Ågren, H.; Jonsson, D.; Gauss, J. *Mol. Phys.* **2005**, *103*, 439–450.

[394] Pawłowski, F.; Jørgensen, P.; Hättig, C. *Chem. Phys. Lett.* **2004**, *391*, 27–32.

[395] Rizzo, A.; Coriani, S.; Fernández, B.; Christiansen, O. *Phys. Chem. Chem. Phys.* **2002**, *4*, 2884–2890.

[396] Jensen, L.; van Duijnen, P. T.; Snijders, J. G. *J. Chem. Phys.* **2003**, *119*, 12998–13006.

[397] Norman, P.; Bishop, D. M.; Jensen, H. J. A.; Oddershede, J. *The Journal of chemical physics* **2005**, *123*, 194103.

[398] Luo, Y.; Norman, P.; Macak, P.; Ågren, H. *J. Chem. Phys.* **1999**, *111*, 9853–9858.

[399] Cammi, R. *Int. J. Quantum Chem.* **2012**, *112*, 2547–2560.

[400] Del Frate, G.; Bellina, F.; Mancini, G.; Marianetti, G.; Minei, P.; Pucci, A.; Barone, V. *Phys. Chem. Chem. Phys.* **2016**, *18*, 9724–9733.

[401] Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. *J. Chem. Phys.* **2004**, *120*, 3787–3798.

[402] Jensen, L.; van Duijnen, P. T. *J. Chem. Phys.* **2005**, *123*, 074307.

[403] Hidalgo Cardenuto, M.; Champagne, B. *J. Chem. Phys.* **2014**, *141*, 234104.

[404] Cardenuto, M. H.; Champagne, B. *Phys. Chem. Chem. Phys.* **2015**, *17*, 23634–23642.

[405] Olsen, J. M. H.; Kongsted, J. *Adv. Quantum Chem.* **2011**, *61*, 107–143.

[406] Wortmann, R.; Krämer, P.; Glania, C.; Lebus, S.; Detzer, N. *Chem. Phys.* **1993**, *173*, 99–108.

[407] Rice, J. E.; Amos, R. D.; Colwell, S. M.; Handy, N. C.; Sanz, J. *J. Chem. Phys.* **1990**, *93*, 8828–8839.

[408] Rice, J. E.; Handy, N. C. *Int. J. Quantum Chem.* **1992**, *43*, 91–118.

[409] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

[410] Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1950–1958.

[411] Macchiagodena, M.; Mancini, G.; Pagliai, M.; Del Frate, G.; Barone, V. *Chem. Phys. Lett.* **2017**, *677*, 120–126.

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

[413] Macchiagodena, M.; Mancini, G.; Pagliai, M.; Cardini, G.; Barone, V. *Int. J. Quantum Chem.* **2017**, DOI:10.1002/qua.25554.

[414] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

[415] Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.

[416] Ray, P. C.; Das, P. K.; Ramasesha, S. *The Journal of chemical physics* **1996**, *105*, 9633–9639.

[417] Clays, K.; Persoons, A. *Phys. Rev. Lett.* **1991**, *66*, 2980.

[418] Clays, K.; Persoons, A. *Rev. Sci. Instrum.* **1992**, *63*, 3285–3289.

[419] Hendrickx, E.; Clays, K.; Persoons, A. *Acc. Chem. Res.* **1998**, *31*, 675–683.

[420] Botek, E.; Spassova, M.; Champagne, B.; Asselberghs, I.; Persoons, A.; Clays, K. *Chem. Phys. Lett.* **2005**, *412*, 274–279.

[421] Bersohn, R.; Pao, Y.-H.; Frisch, H. *J. Chem. Phys.* **1966**, *45*, 3184–3198.

[422] Plaquet, A.; Guillaume, M.; Champagne, B.; Castet, F.; Ducasse, L.; Pozzo, J.-L.; Rodriguez, V. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6223–6232.

[423] Limacher, P. A.; Mikkelsen, K. V.; Lüthi, H. P. *J. Chem. Phys.* **2009**, *130*, 194114.

[424] Pluta, T.; Sadlej, A. J. *Chem. Phys. Lett.* **1998**, *297*, 391–401.

[425] Luo, Y.; Ågren, H.; Vahtras, O.; Jørgensen, P.; Spirko, V.; Hettema, H. *The Journal of chemical physics* **1993**, *98*, 7159–7164.

[426] Giese, T. J.; York, D. M. *J. Chem. Phys.* **2005**, *123*, 164108.

[427] Giese, T. J.; York, D. M. *J. Chem. Phys.* **2007**, *127*, 194101.

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

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

[430] Woo, T.; Margl, P.; Deng, L.; Cavallo, L.; Ziegler, T. *Catal. Today* **1999**, *50*, 479–500.

[431] Sinnecker, S.; Neese, F. *J. Comput. Chem.* **2006**, *27*, 1463–1475.

[432] De Mitri, N.; Monti, S.; Prampolini, G.; Barone, V. *J. Chem. Theory Comput.* **2013**, *9*, 4507–4516.

[433] Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. *Chem. Rev.* **2009**, *110*, 949–1017.

[434] Cavalli, A.; Carloni, P.; Recanatini, M. *Chem. Rev.* **2006**, *106*, 3497–3519.

[435] Lodola, A.; De Vivo, M. *Adv. Protein Chem. Struct. Biol* **2012**, *87*, 337–362.

[436] Jacquemin, D.; Perpète, E. A.; Laurent, A. D.; Assfeld, X.; Adamo, C. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1258–1262.

[437] Monari, A.; Very, T.; Rivail, J.-L.; Assfeld, X. *Comput. Theor. Chem.* **2012**, *990*, 119–125.

[438] Sherrill, C. D. *Reviews in Computational Chemistry*; John Wiley & Sons, Inc., 2009; pp 1–38.

[439] Tschumper, G. S. *Reviews in Computational Chemistry*; John Wiley & Sons, Inc., 2009; pp 39–90.

[440] Szalewicz, K. *WIRES: Comput. Mol. Sci.* **2012**, *2*, 254–272.

[441] Hohenstein, E. G.; Sherrill, C. D. *WIRES: Comput. Mol. Sci.* **2012**, *2*, 304–326.

[442] Amovilli, C.; McWeeny, R. *Chem. Phys.* **1990**, *140*, 343–361.

[443] Burke, K.; Perdew, J. P.; Ernzerhof, M. *J. Chem. Phys.* **1998**, *109*, 3760–3771.

[444] Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

[445] Henderson, T. M.; Janesko, B. G.; Scuseria, G. E. *J. Chem. Phys.* **2008**, *128*, 194105.

[446] Janesko, B. G.; Scuseria, G. E. *J. Chem. Phys.* **2007**, *127*, 164117.

[447] Janesko, B. G. *Int. J. Quantum Chem.* **2013**, *113*, 83–88.

[448] Janesko, B. G. *J. Chem. Phys.* **2011**, *134*, 184105.

[449] Janesko, B. G. *J. Chem. Phys.* **2010**, *133*, 104103.

[450] Janesko, B. G.; Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2014**, *141*, 034103.

[451] Obara, S.; Saika, A. *J. Chem. Phys.* **1986**, *84*, 3963–3974.

[452] Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular electronic-structure theory*; John Wiley & Sons, 2014.

[453] Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3072–3077.

[454] Operstein, V. *J. Approx. Theory* **1996**, *85*, 233–235.

[455] Amovilli, C.; McWeeny, R. *Chem. Phys. Lett.* **1986**, *128*, 11–17.

[456] Klimeš, J.; Michaelides, A. *J. Chem. Phys.* **2012**, *137*, 120901.

[457] Johnson, E. R.; Mackie, I. D.; DiLabio, G. A. *J. Phys. Org. Chem.* **2009**, *22*, 1127–1135.

[458] Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.

[459] Andersson, Y.; Hult, E.; Rydberg, H.; Apell, P.; Lundqvist, B. I.; Langreth, D. C. *Electronic Density Functional Theory*; Springer, 1998; pp 243–260.

[460] Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.

[461] Thonhauser, T.; Cooper, V. R.; Li, S.; Puzder, A.; Hyldgaard, P.; Langreth, D. C. *Phys. Rev. B* **2007**, *76*, 125112.

[462] Román-Pérez, G.; Soler, J. M. *Phys. Rev. Lett.* **2009**, *103*, 096102.

[463] Vydrov, O. A.; Van Voorhis, T. *Phys. Rev. Lett.* **2009**, *103*, 063004.

[464] Vydrov, O. A.; Van Voorhis, T. *Phys. Rev. A* **2010**, *81*, 062708.

[465] Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[466] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

[467] Bader, R. F. W. *Atoms in Molecules A Quantum Theory*; OUP Oxford, 1990.

[468] Hirshfeld, F. L. *Theor. Chem. Acc.* **1977**, *44*, 129–138.

[469] Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2005**, *123*, 024101.

[470] Olasz, A.; Vanommeslaeghe, K.; Krishtal, A.; Veszprémi, T.; Van Alsenoy, C.; Geerlings, P. *J. Chem. Phys.* **2007**, *127*, 224105.

[471] Berendsen, H. J.; van der Spoel, D.; van Drunen, R. *Computer Physics Communications* **1995**, *91*, 43–56.

[472] Lindahl, E.; Hess, B.; Van Der Spoel, D. *Journal of Molecular Modeling* **2001**, *7*, 306–317.

[473] Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. *J. Comput. Chem* **2005**, *26*, 1701–1718.

[474] Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

[475] Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325–340.

[476] Morokuma, K.; Kitaura, K. *Chemical applications of atomic and molecular electrostatic potentials*; Springer, 1981; pp 215–242.

[477] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

[478] Gordon, M. S.; Schmidt, M. W. *Theory and Applications of Computational Chemistry: the first forty years*; Elsevier, 2005; pp 1167–1189.

[479] Parrish, R. M. et al. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197.

[480] Abascal, J. L.; Vega, C. *J. Chem. Phys.* **2005**, *123*, 234505.

[481] Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 296.

[482] Jensen, J. H.; Gordon, M. S. *Mol. Phys.* **1996**, *89*, 1313–1325.

[483] Jensen, J. H.; Gordon, M. S. *J. Chem. Phys.* **1998**, *108*, 4772–4782.

[484] Thierfelder, C.; Assadollahzadeh, B.; Schwerdtfeger, P.; Schäfer, S.; Schäfer, R. *Phys. Rev. A* **2008**, *78*, 052506.

[485] Chu, X.; Dalgarno, A. *J. Chem. Phys.* **2004**, *121*, 4083–4088.

[486] Chu, X.; Dalgarno, A.; Groenenboom, G. C. *Phys. Rev. A* **2007**, *75*, 032723.

[487] Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.

[488] Eisenberg, D.; Kauzmann, W. *The structure and properties of water*; Oxford University Press on Demand, 2005.

[489] Guidez, E. B.; Gordon, M. S. *J. Phys. Chem. A* **2017**,

[490] Su, P.; Li, H. *J. Chem.Phys.* **2009**, *131*, 014102.

[491] Kratz, E. G.; Walker, A. R.; Lagardère, L.; Lipparini, F.; Piquemal, J.-P.; Andrés Cisneros, G. *J. Comput. Chem.* **2016**, *37*, 1019–1029.

[492] Temelso, B.; Archer, K. A.; Shields, G. C. *J. Phys. Chem. A* **2011**, *115*, 12034–12046.

[493] Jacquemin, D.; Moore, B.; Planchat, A.; Adamo, C.; Autschbach, J. *J. Chem. Theory Comput.* **2014**, *10*, 1677–1685.

[494] Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554–8560.

[495] Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624–9631.

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

[497] Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. *Phys. Rev. B* **1992**, *46*, 6671.

[498] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[499] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

[500] Peverati, R.; Truhlar, D. G. *J. Chem. Phys.* **2011**, *135*, 191102.

[501] Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

[502] Hopmann, K. H.; Ruud, K.; Pecul, M.; Kudelski, A.; Dračínský, M.; Bour, P. *J. Phys. Chem. B* **2011**, *115*, 4128–4137.

[503] Lahiri, P.; Wiberg, K. B.; Vaccaro, P. H.; Caricato, M.; Crawford, T. D. *Angew. Chem.* **2014**, *126*, 1410–1413.

[504] Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. *Annu. Rev. Phys. Chem.* **2013**, *64*, 553–578.

[505] Gokcan, H.; Kratz, E. G.; Darden, T. A.; Piquemal, J.-P.; Cisneros, G. A. *J. Phys. Chem. Lett.* **2018**, *9*, 3062–3067.

[506] Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Chem. Phys.* **2005**, *123*, 044109.

[507] Piquemal, J.-P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. *J. Chem. Phys.* **2006**, *124*, 104101.

[508] Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 184101.

[509] Pavone, M.; Cimino, P.; De Angelis, F.; Barone, V. *J. Am. Chem. Soc.* **2006**, *128*, 4338–4347.

[510] Barone, V. *Chem. Phys. Lett.* **1996**, *262*, 201–206.

[511] Improta, R.; Barone, V. *Chem. Rev.* **2004**, *104*, 1231–1254.

[512] Pavone, M.; Cimino, P.; Crescenzi, O.; Sillanpaeae, A.; Barone, V. *J. Phys. Chem. B* **2007**, *111*, 8928–8939.

[513] Barone, V.; Cimino, P.; Crescenzi, O.; Pavone, M. *J. Mol. Struc. THEOCHEM* **2007**, *811*, 323–335.

[514] Pavone, M.; Sillanpää, A.; Cimino, P.; Crescenzi, O.; Barone, V. *J. Phys. Chem. B* **2006**, *110*, 16189–16192.

[515] Barone, V.; Brustolon, M.; Cimino, P.; Polimeno, A.; Zerbetto, M.; Zoleo, A. *J. Am. Chem. Soc.* **2006**, *128*, 15865–15873.

[516] Schöneboom, J. C.; Neese, F.; Thiel, W. *J. Am. Chem. Soc.* **2005**, *127*, 5840–5853.

[517] Moon, S.; Patchkovskii, S.; Salahub, D. R. *J. Mol. Struc. THEOCHEM* **2003**, *632*, 287–295.

[518] Barone, V.; Bencini, A.; Cossi, M.; Matteo, A. D.; Mattesini, M.; Totti, F. *J. Am. Chem. Soc.* **1998**, *120*, 7069–7078.

[519] Berliner, L. J. *Spin labeling: theory and applications*; Academic, New York, 1976.

[520] Kocherginsky, N.; Swartz, H. M. *Nitroxide spin labels: reactions in biology and chemistry*; CRC Press, New York, 1995.

[521] Buchaklian, A. H.; Klug, C. S. *Biochemistry* **2005**, *44*, 5503–5509.

[522] Engström, M.; Vahtras, O.; Ågren, H. *Chem. Phys. Lett.* **2000**, *328*, 483–491.

[523] Rinkevicius, Z.; Frecus, B.; Murugan, N. A.; Vahtras, O.; Kongsted, J.; Ågren, H. *J. Chem. Theory Comput.* **2011**, *8*, 257–263.

[524] de Almeida, K. J.; Rinkevicius, Z.; Hugosson, H. W.; Ferreira, A. C.; Ågren, H. *Chem. Phys.* **2007**, *332*, 176–187.

[525] Di Valentin, C.; Pacchioni, G.; Selloni, A.; Livraghi, S.; Giamello, E. *J. Phys. Chem. B* **2005**, *109*, 11414–11419.

[526] Anpo, M.; Shioya, Y.; Yamashita, H.; Giamello, E.; Morterra, C.; Che, M.; Patterson, H. H.; Webber, S.; Ouellette, S. *J. Phys. Chem.* **1994**, *98*, 5744–5750.

[527] Clément, J.-L.; Ferré, N.; Siri, D.; Karoui, H.; Rockenbauer, A.; Tordo, P. *J. Org. Chem.* **2005**, *70*, 1198–1203.

[528] Giner, E.; Tenti, L.; Angeli, C.; Ferré, N. *J. Chem. Theory Comput.* **2017**, *13*, 475–487.

[529] Malkina, O. L.; Vaara, J.; Schimmelpfennig, B.; Munzarová, M.; Malkin, V. G.; Kaupp, M. *J. Am. Chem. Soc.* **2000**, *122*, 9206–9218.

[530] Kaupp, M.; Remenyi, C.; Vaara, J.; Malkina, O. L.; Malkin, V. G. *J. Am. Chem. Soc.* **2002**, *124*, 2709–2722.

[531] Barone, V.; Cimino, P. *Chem. Phys. Lett.* **2008**, *454*, 139–143.

[532] Barone, V.; Cimino, P.; Stendardo, E. *J. Chem. Theory Comput.* **2008**, *4*, 751–764.

[533] Barone, V. *Recent Advances In Density Functional Methods: (Part I)*; World Scientific, 1995; pp 287–334.

[534] Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–725.

[535] Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

[536] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327–341.

[537] Keana, J. F.; Lee, T. D.; Bernard, E. M. *J. Am. Chem. Soc.* **1976**, *98*, 3052–3053.

[538] Rockenbauer, A.; Korecz, L.; Hideg, K. *J. Chem. Soc., Perkin Trans. 2* **1993**, 2149–2156.

[539] Puzzarini, C.; Barone, V. *Acc. Chem. Res.* **2018**, *51*, 548–556.

[540] Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. *J. Chem. Phys.* **2014**, *140*, 094106.

[541] Boys, S. F.; Bernardi, F. d. *Mol. Phys.* **1970**, *19*, 553–566.

[542] Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730–4737.

[543] Naserifar, S.; Brooks, D. J.; Goddard III, W. A.; Cvicek, V. *J. Chem. Phys.* **2017**, *146*, 124117.

[544] Oppenheim, J. J.; Naserifar, S.; Goddard III, W. A. *J. Phys. Chem. A* **2018**, *122*, 639–645.

[545] Mayer, A. *Phys. Rev. B* **2007**, *75*, 045407.

[546] Jensen, L. L.; Jensen, L. *J. Phys. Chem. C* **2009**, *113*, 15182–15190.

[547] Rinkevicius, Z.; Li, X.; Sandberg, J. A.; Mikkelsen, K. V.; Ågren, H. *J. Chem. Theory Comput.* **2014**, *10*, 989–1003.

[548] Jakobsen, S.; Jensen, F. *J. Chem. Theory Comput.* **2014**, *10*, 5493–5504.

[549] Jakobsen, S.; Jensen, F. *J. Chem. Theory Comput.* **2016**, *12*, 1824–1832.

[550] Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. *Chem. Rev.* **2012**, *112*, 543–631.

[551] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396–1396.

[552] Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J. *Phys. Rev. Lett.* **2009**, *103*, 026403.

[553] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[554] Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664–675.

[555] Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[556] Parrish, R. M.; Parker, T. M.; Sherrill, C. D. *J. Chem. Theory Comput.* **2014**, *10*, 4417–4431.

[557] Parrish, R. M.; Thompson, K. C.; Martínez, T. J. *J. Chem. Theory Comput.* **2018**, *14*, 1737–1753.

[558] Miyamoto, S.; Kollman, P. A. *J. Comput. Chem* **1992**, *13*, 952–962.

[559] Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

[560] Berendsen, H. J.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

[561] Frisch, M.; Head-Gordon, M.; Pople, J. *Chem. Phys.* **1990**, *141*, 189–196.

[562] Stockman, M. I. *Opt. Expr.* **2011**, *19*, 22029–22106.

[563] Morton, S. M.; Silverstein, D. W.; Jensen, L. *Chem. Rev.* **2011**, *111*, 3962–3994.

[564] Moskovits, M. *Rev. Mod. Phys.* **1985**, *57*, 783.

[565] Nie, S.; Emory, S. R. *Science* **1997**, *275*, 1102–1106.

[566] Atwater, H. A.; Polman, A. *Nat. Mater.* **2010**, *9*, 205.

[567] Santhosh, K.; Bitton, O.; Chuntonov, L.; Haran, G. *Nat. Commun.* **2016**, *7*, 11823.

[568] Anker, J. N.; Hall, W. P.; Lyandres, O.; Shah, N. C.; Zhao, J.; Van Duyne, R. P. *Nat. Mater.* **2008**, *7*, 442–453.

[569] Muehlschlegel, P.; Eisler, H.-J.; Martin, O. J.; Hecht, B.; Pohl, D. *Science* **2005**, *308*, 1607–1609.

[570] Giannini, V.; Fernández-Domínguez, A. I.; Heck, S. C.; Maier, S. A. *Chem. Rev.* **2011**, *111*, 3888–3912.

[571] Maier, S. A.; Kik, P. G.; Atwater, H. A.; Meltzer, S.; Harel, E.; Koel, B. E.; Requicha, A. A. *Nat. Mater.* **2003**, *2*, 229.

[572] Kneipp, K.; Wang, Y.; Kneipp, H.; Perelman, L. T.; Itzkan, I.; Dasari, R. R.; Feld, M. S. *Phys. Rev. Lett.* **1997**, *78*, 1667.

[573] Neuman, T.; Esteban, R.; Casanova, D.; García-Vidal, F. J.; Aizpurua, J. *Nano Lett.* **2018**, *18*, 2358–2364.

[574] Jin, R.; Cao, Y.; Mirkin, C. A.; Kelly, K.; Schatz, G. C.; Zheng, J. *Science* **2001**, *294*, 1901–1903.

[575] Myroshnychenko, V.; Rodríguez-Fernández, J.; Pastoriza-Santos, I.; Funston, A. M.; Novo, C.; Mulvaney, P.; Liz-Marzan, L. M.; de Abajo, F. J. G. *Chem. Soc. Rev.* **2008**, *37*, 1792–1805.

[576] Hao, F.; Nehl, C. L.; Hafner, J. H.; Nordlander, P. *Nano Lett.* **2007**, *7*, 729–732.

[577] Pérez-González, O.; Zabala, N.; Borisov, A.; Halas, N.; Nordlander, P.; Aizpurua, J. *Nano Lett.* **2010**, *10*, 3090–3095.

[578] Ciracì, C.; Hill, R.; Mock, J.; Urzhumov, Y.; Fernández-Domínguez, A.; Maier, S.; Pendry, J.; Chilkoti, A.; Smith, D. *Science* **2012**, *337*, 1072–1074.

[579] Payton, J. L.; Morton, S. M.; Moore, J. E.; Jensen, L. *Acc. Chem. Res.* **2013**, *47*, 88–99.

[580] Liu, P.; Chulhai, D. V.; Jensen, L. *ACS Nano* **2017**, *11*, 5094–5102.

[581] Zuloaga, J.; Prodan, E.; Nordlander, P. *Nano Lett.* **2009**, *9*, 887–891.

[582] Scholl, J. A.; García-Etxarri, A.; Koh, A. L.; Dionne, J. A. *Nano Lett.* **2013**, *13*, 564–569.

[583] Marinica, D.; Kazansky, A.; Nordlander, P.; Aizpurua, J.; Borisov, A. G. *Nano Lett.* **2012**, *12*, 1333–1339.

[584] Huck, C.; Neubrech, F.; Vogt, J.; Toma, A.; Gerbert, D.; Katzmann, J.; H artling, T.; Pucci, A. *ACS Nano* **2014**, *8*, 4908–4914.

[585] Taylor, R. W.; Lee, T.-C.; Scherman, O. A.; Esteban, R.; Aizpurua, J.; Huang, F. M.; Baumberg, J. J.; Mahajan, S. *ACS Nano* **2011**, *5*, 3878–3887.

[586] Panaro, S.; Ciracì, C. *ACS Photonics* **2016**, *3*, 2467–2474.

[587] Teperik, T. V.; Nordlander, P.; Aizpurua, J.; Borisov, A. G. *Opt. Expr.* **2013**, *21*, 27306–27325.

[588] Zhu, W.; Esteban, R.; Borisov, A. G.; Baumberg, J. J.; Nordlander, P.; Lezec, H. J.; Aizpurua, J.; Crozier, K. B. *Nat. Commun.* **2016**, *7*, 11495.

[589] Sinha-Roy, R.; Garcia-Gonzalez, P.; Weissker, H.-C.; Rabilloud, F.; Fernandez-Dominguez, A. I. *ACS Photonics* **2017**, *4*, 1484–1493.

[590] Esteban, R.; Zugarramurdi, A.; Zhang, P.; Nordlander, P.; García-Vidal, F. J.; Borisov, A. G.; Aizpurua, J. *Faraday Discuss.* **2015**, *178*, 151–183.

[591] Urbieta, M.; Barbry, M.; Zhang, Y.; Koval, P.; Sánchez-Portal, D.; Zabala, N.; Aizpurua, J. *ACS Nano* **2018**, *12*, 585–595.

[592] Bade, W. *J. Chem. Phys.* **1957**, *27*, 1280–1284.

[593] Palik, E. D. *Handbook of optical constants of solids*; Elsevier, 1997.

[594] Haynes, W. *CRC Handbook of Chemistry and Physics*; CRC Press: Boca Raton, 2014.

[595] Liu, L.; Wang, Y.; Fang, Z.; Zhao, K. *J. Chem. Phys.* **2013**, *139*, 064310.

[596] Pérez-González, O.; Zabala, N.; Aizpurua, J. *New J. Phys.* **2011**, *13*, 083013.

[597] Duan, H.; Fernández-Domínguez, A. I.; Bosman, M.; Maier, S. A.; Yang, J. K. *Nano Lett.* **2012**, *12*, 1683–1689.

[598] Varas, A.; García-González, P.; Feist, J.; García-Vidal, F.; Rubio, A. *Nanophotonics* **2016**, *5*, 409–426.

[599] Zhang, P.; Feist, J.; Rubio, A.; García-González, P.; García-Vidal, F. *Phys. Rev. B* **2014**, *90*, 161407.

[600] Rocca, D.; Gebauer, R.; Saad, Y.; Baroni, S. *J. Chem. Phys.* **2008**, *128*, 154105.

[601] Malcıoğlu, O. B.; Gebauer, R.; Rocca, D.; Baroni, S. *Comput. Phys. Commun.* **2011**, *182*, 1744–1754.

[602] Li, J.-H.; Hayashi, M.; Guo, G.-Y. *Phys. Rev. B* **2013**, *88*, 155437.

[603] Weick, G.; Ingold, G.-L.; Jalabert, R. A.; Weinmann, D. *Phys. Rev. B* **2006**, *74*, 165421.

[604] Mukhopadhyay, P.; Zuber, G.; Wipf, P.; Beratan, D. N. *Angewandte Chemie International Edition* **2007**, *46*, 6450–6452.

[605] Ling, X.; Xie, L.; Fang, Y.; Xu, H.; Zhang, H.; Kong, J.; Dresselhaus, M. S.; Zhang, J.; Liu, Z. *Nano Lett.* **2009**, *10*, 553–561.

Acknowledgments

There are many people I need to acknowledge for letting the three years of my PhD be an amazing experience. First let me thank my supervisor prof. Chiara Cappelli, for sharing my whole academic life, for introducing me to the fascinating world of quantum chemistry and for allowing me to follow and develop my own ideas but keeping me always in the right direction. Thanks to her brilliant supervision I had the opportunity to attend several international conferences, and to take part into stimulating, both from the human and scientific point of view, collaborations which brought me from Padova to above the Arctic Circle. Such collaborations would not be so positive if it wasn't for Luca, Stefano, their groups and Roberto, who always feel like home.

I also express my gratitude to all my co-authors and to the SMART group led by prof. Vincenzo Barone. In particular, I owe a debt of gratitude to Franco, who has provided continuous help and stimulating scientific (and not) discussions. I would not have arrived to the end of these years without all the past and present people of EmbedLab group of prof. Chiara Cappelli: Piero, Matteo, Alessandra, Luca, Rosario, Laura, Marta, who have always been a support and a stimulus during the three years. I am also much obliged to all the people of the student administrative staff.

A particular thank goes also to Marco, Oliviero and Emanuele, with whom I unfortunately shared only partially these years, but who are always present even if they are in different countries.

Finally, my deepest gratitude goes to my family, who always supports me with passion, criticisms and suggestions. In particular, my father and my brother for being two, different, models to look up, and my mother for her unconditional love. A special thanks goes to Irene, with whom I shared, I am sharing and I will share my life.