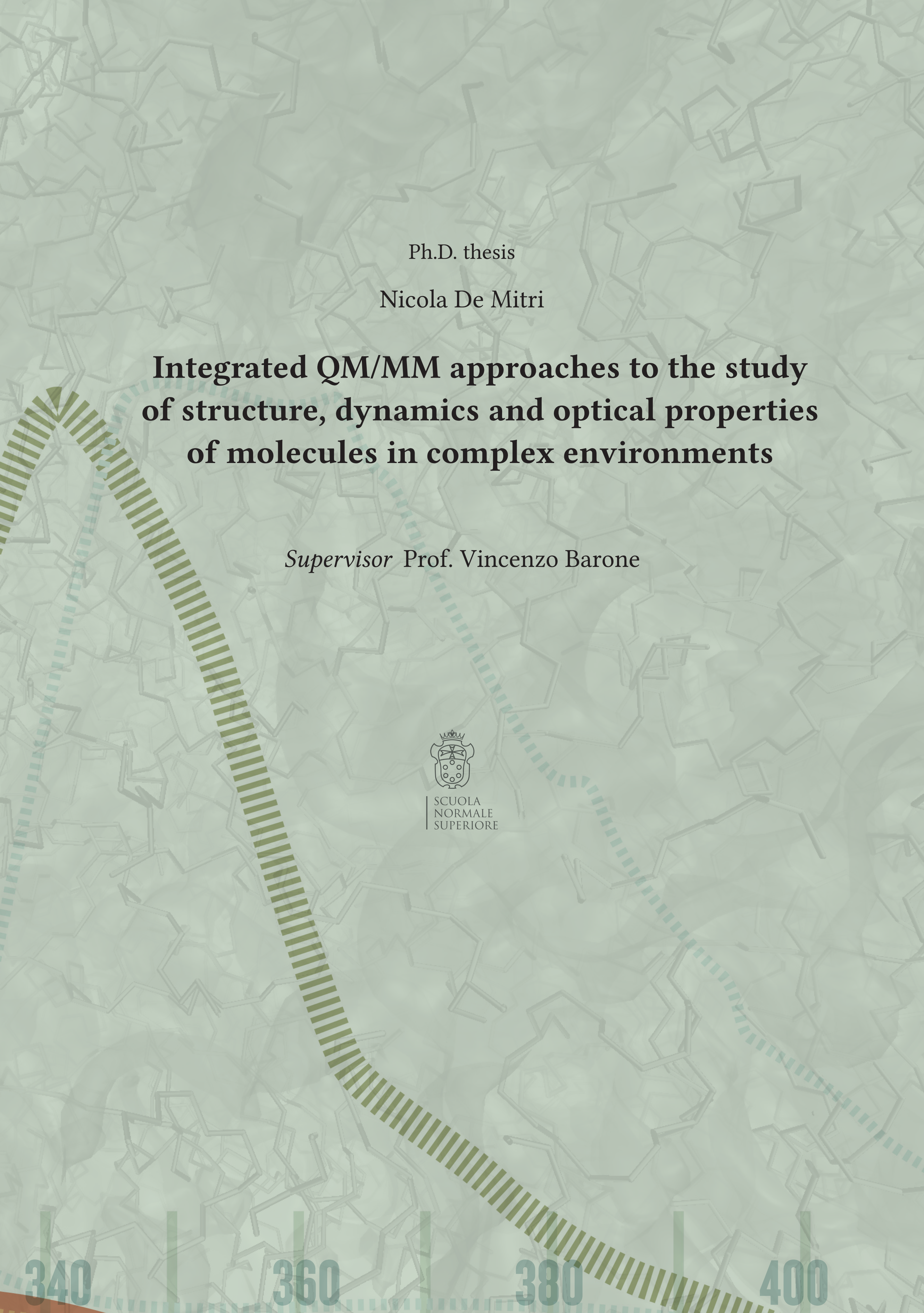
Ph.D. thesis

Nicola De Mitri

Integrated QM/MM approaches to the study of structure, dynamics and optical properties of molecules in complex environments

Supervisor Prof. Vincenzo Barone

SCUOLA
NORMALE
SUPERIORE



This thesis was accepted by the referees in a digital version, and
printed with minor adjustments.
Printed before the defense session (Scuola Normale Superiore,
Pisa, 21-05-2015) at Campano printing house in Pisa.

Set in Linux Libertine and Linux Biolinum (GPL and OFL free licenses).
Cover design by Claudia Flandoli (www.claudiaflandoli.com).
Molecule's render generated in UIUC's VMD.

Author contact: nicola.dmt@gmail.com

Integrated QM/MM approaches to the study of structure, dynamics and optical properties of molecules in complex environments

Nicola De Mitri
Ph.D. thesis

Supervisor
Prof. Vincenzo Barone

Scuola Normale Superiore
Pisa

Contents

Chapter 1

Introduction

This thesis reports on the most recent efforts in the determination of the properties of medium-sized organic molecules and of their interaction with the chemical environment; in particular, the integrated use of quantum mechanical (QM) methods, classical atomistic models (molecular mechanics, MM) and molecular dynamics simulation (MD), has been studied and implemented in a combined approach, validated through spectroscopic results.

Improvements made on top of the current state methodologies concern the accounting for: (*i.*) the molecular flexibility through accurate intramolecular force fields (FF); (*ii.*) the role of complex structured environments on structure and optical properties of molecular probes; (*iii.*) the mutual polarization of the target molecule and the surroundings.

* * *

In the last decades, computational chemistry and, in particular, computational spectroscopy have established themselves as common sets of tools for the interpretation of experimentally determined data in terms of chemical physical effects [1, 2]. From this point of view, the density functional theory (DFT) has had the greatest success among the quantum chemical methods, the reason being its efficiency and ease, coupled with adequate predictivity, for quite large molecules [3]. DFT can be used to simulate a wide variety of molecular properties regarding the electronic ground state, while its time-dependent formulation, TDDFT, can tackle excited-state spectroscopic properties like the UV-vis spectra [4] and even chiroptical properties, including circular dichroism (CD) [5, 6].

As a matter of fact, most spectroscopic measurements occur in solution, therefore QM calculations performed on the isolated molecule may fail to compare with the experiments. In fact, the range of environments in which a molecule can be observed is huge, ranging from the almost isolated case of gases of atmospheric interest [7] to well-structured biomolecular matrices [8], and from solutions of small molecules (including water) [9] to the insertion in synthetic polymers or nanoparticles [10]. The effect of these environments is manifold: on the one end, the solute undergoes a redistribution of its electron density due to the electric field exerted by the surroundings (polarization), on the other hand its geometrical conformation may depend on the interaction with the physical encumbrance of the solvent itself: these effects are regarded as electrostatic and mechanical embedding, respectively.

A proper account for the embedding within the DFT framework raises a number of theoretical and computational matters [9]. At first, while it's meaningful, to some extents [11], to rely on the conformation of minimal energy for the calculation of the isolated molecule's properties, no such thing as an equilibrium structure exists for a disordered system, like a solution or a polymer. Moreover, even if a good solvent representative structure existed, the insertion of the first solvation shells in the quantum mechanical calculation ("explicit treatment") would require a substantial increase of the computational cost, and would even bring up a further issue: boundary treatment.

For this reason, different yet approximated routes exist to deal with solvation. The so-called [12] "continuum method" aims to describe the solvent in the most implicit way, by only defining its spatial extent and bulk electrostatic properties. One of the most common implementation of this idea is the polarizable continuum model (PCM) [13]: a polarizable, continuum medium, characterised by its dielectric constant, is defined around the target molecule, making use of atom-centred spheres. This cavity is segmented in tiles, each endowed with a local charge: the surface charges and the electronic density, which is subject to the external polarization, are determined in a self consistent way.

Nevertheless, the continuum methods lack a description for specific local interactions that depend on the exact relative position of solute and solvent [14]. An alternate route overcomes this defect, by giving a microscopic description of the solvent through its actual coordinates. These methods [15, 16] require a procedure for the generation of the solvent coordinates, which is usually achieved by molecular dynamics, Monte Carlo and, less frequently, *ab initio* simulations. At the time of property calculation, the system is partitioned in a inner subsystem that is treated quantum mechanically and a outer MM one. The choice of the partitioning must take care of the nature of the interactions between the target and the environment: covalent bonds and strong non covalent interactions may require additional QM “link” atoms. Indeed, the QM/MM interaction is, in principle, limited to the electrostatics and one may need to describe dispersion or stacking effects within the high-level layer.

As stated before, this “explicit” treatment raise the matter of boundary description, which can be easily tackled by an outer continuum layer (QM/MM/PCM) [17, 18]; in fact, in order to achieve only the coordinates belonging to a target-centred cavity, the underlying simulations can be also performed in non periodic boundary conditions (NPBC) [19], exploiting the polarizable cavity already at this stage. Finally, the layered formulation described above retains a gap in the description of polarizability, being the inner subsystem and the outer continuum both polarizable, while the explicit MM subsystem is not. An effort has been put in bridging this gap [18]: in particular, within our group, the polarizable fluctuating charge approach (FQ [20]) has been coupled to DFT and PCM to provide a fully polarizable model [21, 22, 23], whose validation for the calculation of optical properties is accounted for in the following pages.

A final feature of the molecular system that may be considered relevant to its measurable properties is the actual geometry that it takes at room temperature. The existence of nuclear motion both influences the definition of the equilibrium structure [11] and brings up couplings between electronic excitations and vibrations. In this case, it is advisable to take nuclear motions into the proper account by means of vibrationally resolved spectra [24, 25, 26], which can be easily coupled to a continuum solvent description.

However, in many cases involving floppy molecules, whose conformations strongly affect the optical features, it is constraining to rely only on a few structures of minimal energy, as the vibronic methods would in principle require. In these cases one can again lean on MD simulations to produce statistically uncorrelated structures of the solute and the solvent, and to average the TDDFT results obtained for each of them.

It is undeniable that such a procedure is grounded on the reliability of the force field (FF) used for the simulations, as it has the role of predicting a consistent and accurate distribution of the most relevant geometric features.

1.1 An integrated "statistical approach"

In this thesis a general approach will be presented, that combines most of the methodologies listed above, exploiting polarizable TDDFT/MM/PCM calculations on the basis of simulations performed with *ab-initio* based FF.

The description of the computational method will be accompanied by validation studies that have been carried out for the whole protocol or portions of it.

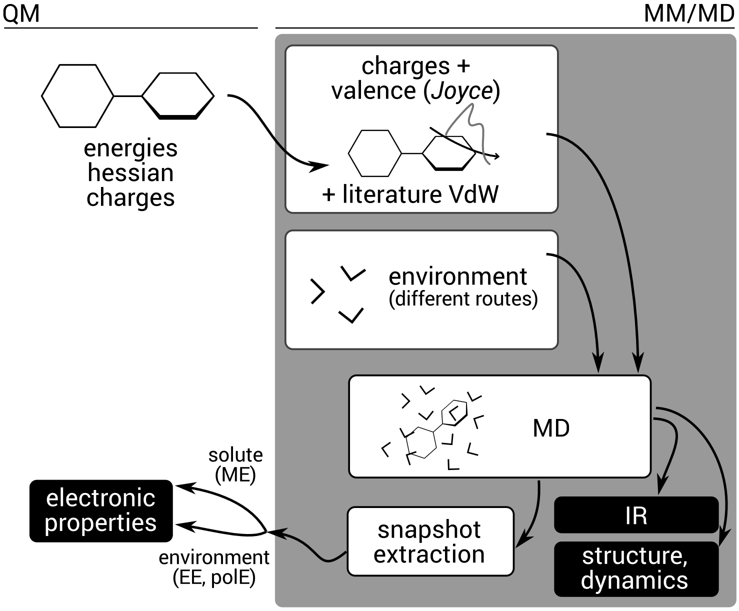


Figure 1.1: The general approach "from QM to MM/MD and back"

The approach is sketched in Fig. 1.1. It originates from a QM-level study of the spectroscopically relevant molecule of interest[1]: in particular, a DFT method is used to identify the equilibrium structure of the molecule, to derive its atomic charges and to compute the hessian matrix, *i.e.* the second derivatives of the internal energy with respect to nuclear Cartesian displacements. These data can be gained for any electronic

---

[1] For the sake of brevity, the target molecule will be often referred as as "solute" and the environment as "solvent" throughout this thesis, although the presented procedure is not limited to solvated systems.

state, and the inclusion of implicit solvation may be relevant, above all, for the polarization of the atomic charges, as will be discussed.

Moving to the classical (MM) realm, a FF is then built by combining standard Lennard-Jones parameters based on the atom-types, DFT-derived charges and *ad hoc* parametrised valence terms (stretching, bending and torsions), obtained through a devoted software called Joyce [27], making use of the outputs of the quantum mechanical calculations collected before. Extensive explanation of this optimisation algorithm will be provided in the devoted chapter, as improvements and validation of this code have been a major goal of the presented work.

Usually a more modest effort is devoted to the description of the environment during MD simulations: parameters are taken from literature when they are available with reasonable fidelity (it is seldom the case for polymer matrices). Once the whole topology of the composite system is built, MD simulation are carried out in the NPT ensemble exploiting periodic boundary conditions.

At this point, several insight can be already taught by the analysis of the trajectories [28], including semi-classically derived vibrational spectra [29]. However, it is needed to revert to the QM description to obtain electronic spectra like UV-vis absorption, emission, and ECD: these are eventually calculated by averaging and smoothing the TDDFT signals from a set of statistically independent snapshots extracted from the trajectories. The inclusion of the solvent can be done at different levels of approximation: in ascending order of complexity:

ME.	“Mechanical embedding”: the atoms from the environment are entirely neglected when calculating transitions. However they were present during the MD run, therefore their effect is still present in the solute conformations.
EE (QM/MM/PCM).	“Electrostatic embedding”: an explicit sphere of solvent of adequate size is represented as point charges, and the entire system is inserted in a PCM cavity.
POL.E (QM/FQ/PCM).	Like EE, but the solvent molecules are made polarizable using the fluctuating charges method.
QM∗/FQ/PCM.	Like above, but the most closely interacting solvent molecules are inserted in the QM layer.

1.2 Related papers

Methodology and results accounted for by this thesis gain in part from the following manuscripts:

i. Lipparini, F.; Cappelli, C.; Scalmani, G.; De Mitri, N.; Barone, V. "Analytical First and Second Derivatives for a Fully Polarizable QM/Classical Hamiltonian" *J. Chem. Theory Comput.* **8**, 2012, 4270–4278

ii. Barone, V.; Cacelli, I.; De Mitri, N.; Licari, D.; Monti, S.; Prampolini, G. "JOYCE and ULYSSES: integrated and user-friendly tools for the parameterization of intramolecular force fields from quantum mechanical data" *Phys. Chem. Chem. Phys.*, **15**, 2013, 3736–3751

iii. De Mitri, N.; Monti, S.; Prampolini, G.; Barone, V. "Absorption and Emission Spectra of a Flexible Dye in Solution: A Computational Time-Dependent Approach" *J. Chem. Theory Comput.* **9**, 2013, 4507–4516

iv. Prampolini, G.; Monti, S.; De Mitri, N.; Barone, V. "Evidences of long lived cages in functionalized polymers: Effects on chromophore dynamic and spectroscopic properties" *Chem. Phys. Lett.* **601**, 2014, 134–138

v. De Mitri, N.; Prampolini, G.; Monti, S.; Barone, V. "Structure, Dynamic and Photophysical Properties of a Fluorescent Dye Incorporated in an Amorphous Hydrophobic Polymer Bundle", *Phys. Chem. Chem. Phys.*, **16**, 2014, 16573–16587

The TOC graphic accompanying the last mentioned manuscript is reported in Fig. 1.2, and it provides a visual clue of the overall method and of the achievable results.

The work leading to this manuscript has also led to some remarks about data treatment and visualisation, that are included in the chapter mentioned below but are only peripherally touched upon in the following pages:

vi. Salvadori, A; Licari, D.; Mancini, G.; Brogni, A.; De Mitri, N.; Barone, V. "Graphical Interfaces and Virtual Reality for Molecular Sciences", in *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*, Elsevier 2014

Finally, most of the results presented in sections 5 and 6 are still unpublished.

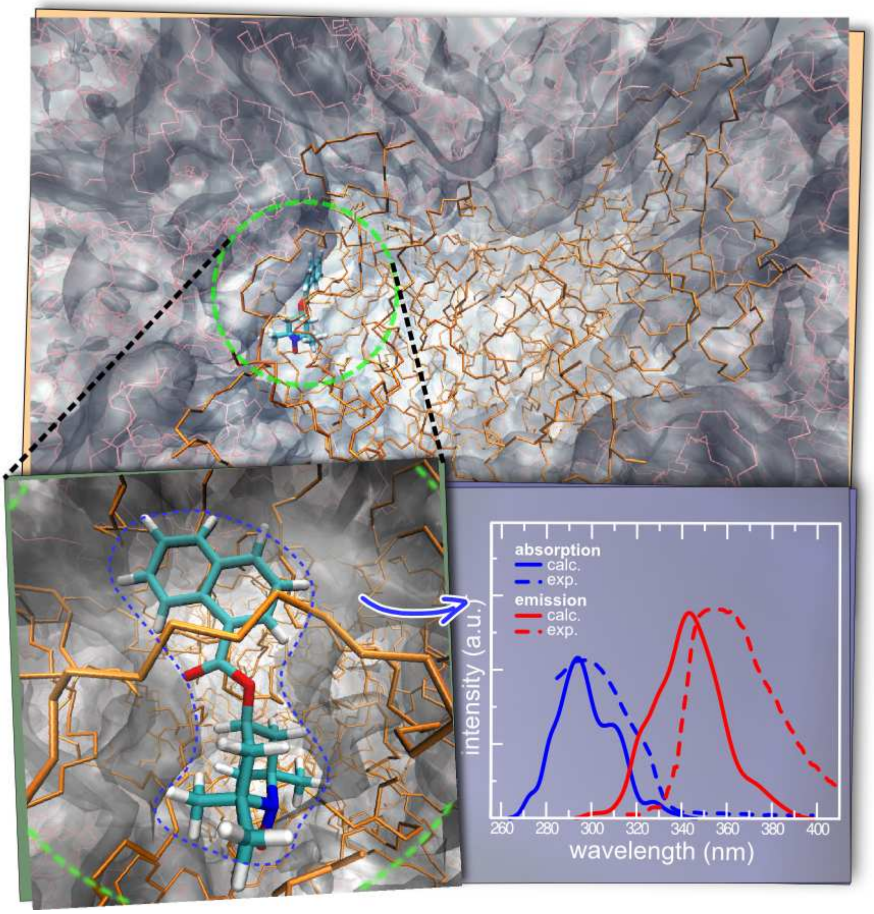


Figure 1.2: TOC graphic for the manuscript featured in Phys. Chem. Chem. Phys. **16**: it represents the “statistical” simulation of optical spectra for a flexible dye incorporated in a polymer bundle, by means of a layered QM/MM description. Details in chapter 7.

Chapter 2

Overview on theory and methods

2.1 Parametrisation of force fields

2.1.1 Joyce

A wide variety of force fields exist in literature [30], that differentiate as regards functional form, classes of atom types, parametrisation strategies and so on.

In broad terms, two different philosophies can be told apart: on one side the completely general force fields, among them UFF ("universal force field"), aiming to handle any compound using an averaged set of parameters for each element, regardless of its local chemical environment, *i.e.* there is roughly one "atom type" for each element of the periodic table. On the other side there are the popular force fields AMBER, CHARMM, OPLS, COMPASS, which come with a set of specific parameters, tailored on selected classes of molecules (usually, but not limited to, organic compounds of biochemical interest). In these FFs, distinctions are made among atoms of the same element, treating for example a methyl and an aromatic carbon as two separate atom types, thus assigning them different force functionals and constants. Such specific force fields are accurate in describing properties in condensed phase, both by molecular dynamics (MD) and Monte Carlo simulations, while the UFF family is mostly used for fast optimisation of structures in the gas phase only.

Pictorially, a double-ended line can be traced: one one end, the "general" force fields address the goals of generality and transferability; on the opposite end the "specific" force fields are focused on predictivity. The well-known force fields for biological macromolecules lay near this end, but not at the far extreme.

In the recent years, parametrisation protocols and relative software packages have been distributed (PARAFREQ [31], JOYCE [32] among others), to develop or refine the force parameters within a given class of FF functional terms, using *ab initio* data as benchmarks: these FFs shoot for the most accurate and predictive description of the molecule, at a disadvantage of transferability.

Among the parametrisation protocols to develop *ad hoc* force fields, one particular scheme has been developed and implemented in collaboration with our group in the 2007 Fortran code Joyce [32]. This protocol performs a parameter optimisation for the intermolecular potential, using DFT data as input.

Its workflow is briefly depicted in Fig. 2.1. At first, energies, gradients and hessian matrices, calculated at the DFT level of theory of user's choice, are provided to the software, as outputs from a Gaussian (09 or later) calculation [33]; the software therefore generates a suitable potential form, that can be modified by the user to his needs.

As of now, the intramolecular potential can be described as a sum of the following terms:

$$\begin{aligned} E_{\text{intra}} = & \sum_{\mu\in\text{bonds}} \underbrace{\frac{1}{2}k^s_\mu(b_\mu - b^0_\mu)^2}_{V_{\text{str}}(b_\mu)} + \\ & + \sum_{\mu\in\text{angles}} \underbrace{\frac{1}{2}k^b_\mu(\theta_\mu - \theta^0_\mu)^2}_{V_{\text{bend}}(\theta_\mu)} + \\ & + \sum_{\mu\in\text{h.dih.}} \underbrace{\frac{1}{2}k^t_\mu(\phi_\mu - \phi^0_\mu)^2}_{V_{\text{htors}}(\phi_\mu)} + \\ & + \sum_{\mu\in\text{p.dih.}} \underbrace{\sum_{j}^{N^\mu_{\cos}} k^d_{j\mu}\left[1+\cos\left(n^\mu_j\delta_\mu - \gamma^\mu_j\right)\right] +}_{V_{\text{tors}}(\delta_\mu)} \\ & + \sum_{i,j} \underbrace{\left(c_{ij}\frac{q_iq_j}{r_{ij}} + 4c_{ij}\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]\right)}_{V_{\text{nb}}(i,j)} \end{aligned} \tag{2.1}$$

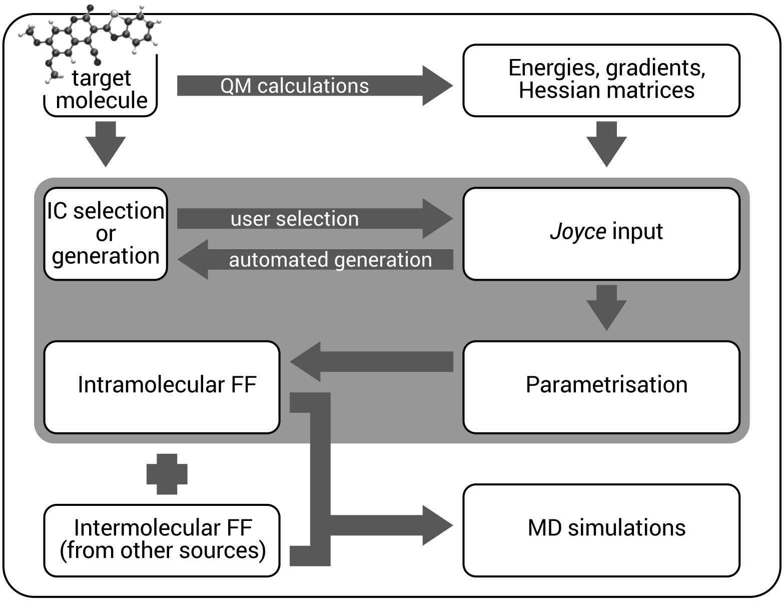


Figure 2.1: Joyce parametrisation scheme

where R_μ ($Q \in \{b, \theta, \phi, \delta\}$) are the instantaneous values of the so-called valences: bond lengths, angles, harmonic and periodic dihedrals, respectively, Q^0_μ their equilibrium values; $k^{s,b,t}_\mu$ are the force constants for the harmonic IC's, and $k^d_{j\mu}$ for the periodic dihedrals. In particular, $N^\mu_{\cos}$ cosines with increasing multiplicities n^μ_j and variable phases γ^μ_j are employed for the "proper" torsions. Finally, q_i are the point charges assigned to each atom, r_{ij} the distance between them, and ε and σ are the Lennard-Jones parameters, combined by means of the OPLS combination rules [34]. The shielding factor $0 \leq c_{ij} \leq 1$ is used to trace out the pairs of atoms that are less than three bonds away from each other, and to possibly modulate the inclusion of the other ("1-4+") pairs.

As shown, harmonic potential are employed for stetchings, bendings and the "rigid" torsions, while a periodic function, namely a sum of cosines, can be chosen for some selected "flexible" dihedrals. All the valence coordinates Q_μ that appear in equations 2.1 are together called the internal coordinates of the system (IC), and they are generally redundant, this meaning that the transformation between a set of ICs and Cartesian or normal coordinates is overdetermined, provided that all the occurring bonds, angles and dihedral are taken into account in the potential energy definition. Broadly speaking, also the non-bonded distances r_{ij} in the last equation are ICS for the system, but they can be always expressed in terms of the previous ones.

Once the functional form has been defined, on the basis of the molecular connectivity and user imposed definitions, Joyce stores the energy U, gradient $\mathbf{G}$ and Hessian matrix H given by the Gaussian 09 output file. These data, achieved from a single optimised conformation, are enough to parametrise the harmonic terms, while for the fitting of a flexible torsion, whose functional is expressed as a cosine sum, eventually the energies of several twisted conformations should be given as input as well.

Being E^{intra} the potential energy calculated at the Joyce FF level, g an index labelling the optimised ($g = 0$) and twisted ($g > 0$) conformations, and W's a set of purposely selected weights, a merit function can be defined as follows, provided that K and L run over the normal coordinates:

$$I^{intra} = \sum_{g=0}^{N_{geom}} \left[W_g \left[U_g - E_g^{intra} \right]^2 \right] + \sum_{K}^{3N-6} W' \left[G_K - \left(\frac{\partial E^{intra}}{\partial Q_K} \right) \right]_{g=0}^2 + \tag{2.2}$$

$$+ \sum_{K \leq L}^{3N-6} \frac{2W''_{KL}}{(3N-6)(3N-5)} \left[H_{KL} - \left(\frac{\partial^2 E^{intra}}{\partial Q_K \partial Q_L} \right) \right]_{g=0}^2$$

The minimisation of this error function, after performing the needed coordinate transformation, corresponds to a linear problem when the free parameters appears at first order in the energy expression, this is the case of $k^s, k^b, k^t, k_j^d, \epsilon_{ij}$ in Eq. 2.1. The first term of the sum in Eq. 2.2 must be only evaluated for the fitting of k_j^d, and only if "floppy" dihedrals are involved: in this case, the torsional energy fitting is performed within the Frozen Internal Rotation Approximation (FIRA). The FF-level energy is calculated, in correspondence to the twisted conformations (E_g^{intra}) by keeping all the moieties of the molecule rigid, except for the dihedral of interest. Since, on the contrary, the quantum mechanical counterparts U_g are calculated after optimising the twisted geometries, FIRA assumes that the molecule will experience no huge rearrangements along the torsion, except the torsion itself.

As a result of the parametrisation process, an intramolecular FF is produced and written in a format suitable to be used by a popular MM/MD package, Gromacs [35].

During the last three years, the Joyce package has been improved within our group, leading to the distribution of version 2.0 and to new validated results [36].

Version 2.0 comes with the graphical user interface Ulysses, a software that can handle all the stages of the development of a force field. Given the Gaussian required files, Ulysses provides tools for visualising the molecule, identifying the ICS, selecting their functional forms and putting them into the topology file; it extends the Joyce capabilities to a more straightforward way of treating non-bonded parameters, endowing a tool for the calculation of the atomic partial charges in the CM5 scheme, and a direct link to the OPLS set of LJ parameters. Moreover, it directly plot the comparison between QM and FF-derived normal modes and torsional energy profiles. Finally, Ulysses belongs to a family of tools developed in our group and devoted to put easy visualisation techniques at molecular science's service [37].

The other main improvement made to the original code, is the inclusion of the coupling terms ("off diagonal") to the potential energy, in a way suitable for MD simulations with Gromacs. The inclusion of these couplings makes the Joyce potential energy formulation a "Class II force field" [38].

These couplings are expressed in terms of already defined ICs. Let Q_μ and Q_ν be a couple of internal coordinates: the coupling term between them reads

$$E_{\mu\nu}^{coupl} = k_{\mu\nu}(Q_\mu - Q_\mu^0)(Q_\nu - Q_\nu^0) \qquad \text{if both are harmonic ICs} \tag{2.3}$$

$$E_{\mu\nu}^{coupl} = \sum_j k_{\mu\nu j}(Q_\mu - Q_\mu^0)\sin(n_j^\nu Q_\nu - \gamma_\nu) \qquad \text{if one IC is a flex. dih.}$$

$$E_{\mu\nu}^{coupl} = \sum_{ij} k_{\mu\nu ij}\sin(n_i^\mu Q_\mu - \gamma_\mu)\sin(n_j^\nu Q_\nu - \gamma_\nu) \qquad \text{if both are flex. dih.}$$

Since the parameters $k_{\mu\nu}$, $k_{\mu\nu j}$ $k_{\mu\nu ij}$ enter the potential energy function as multiplying scalars, the minimisation problem remains linear and the protocol must not be upgraded.

As regards the compatibility with other packages, it must be pointed out that Gromacs only reads stretching-stretching and stretching-bending couplings, provided that they involve exactly three adjacent atomic sites. The related topology specification is then entered within the "angles" directive in the *.top* format, and this has been made compatible within Joyce. However, a new "couplings" directive was added to a Joyce custom *.top* format that handles all the generalised coupling terms defined above but, as of now, they can not be used within MD simulations.

2.1.2 Parametrisation protocol

10 molecules were studied by means of their classical model. They are all summarised in Figure 4.1; details about the studies carried out on each compound will be given starting from chapter 4.

For each of them, the Joyce parametrisation route was followed:

i. DFT is employed to identify the conformational minima of the molecule, and to compute energy, gradient and hessian matrix at the stationary points. The choice of the exchange-correlation functional vary from case to case and will be detailed in the following sections. For two of the investigated molecules, the energy and its derivatives were calculated also for the first electronically excited state (EES), by means of the TDDFT theory.

ii. If flexible dihedrals are present, which is the case for six of the considered molecules, an energy scan is performed around each of them (unless they are identical and therefore the parameters transferable). Several conformations are produced by constraining the torsion to a value that is increased by fixed steps (15° or 30°) and optimising the rest of the molecule. The DFT energy is then evaluated for each conformation. The energy scan is sometimes omitted for -CH_3 terminations.

iii. Joyce and Ulysses are used to define the FF functional form. The torsional functions (harmonic or periodic), and the inclusion of improper dihedrals and non-bonded interacting pairs are chosen at this point.

iv. The collected QM data about the stationary points are used by Joyce to parametrise the harmonic part of the FFS (Eq. 2.1(1-3)). Possibly, off-diagonal terms (Eq. 2.3) are set between coupled ICS. Then, in a second run, all the torsional scans are evaluated to fit the $k^{d}_{j\mu}$ parameters in Eq. 2.1(4).

v. The parametrised FFS are compared to the QM benchmark. Three kinds of comparison are done. A first value, indicating the goodness of the optimised parameters, is the square root of the merit function $(I^{intra})^{\frac{1}{2}}$ (Eq. 2.2), which is minimised by Joyce. Secondarily, the normal mode wavenumbers are calculated at the MM level and compared to the QM ones: the frequencies are sorted in a descending order (although also a sorting based on the overlap between the projection of the normal modes on the ICS is made possible by Joyce/Ulysses [32, 36, 37]), then the standard deviation is computed as

$$\sigma_{freq} = \left[\frac{1}{N_{freq}} \sum_{i=1}^{N_{freq}} (\nu_i^{QM} - \nu_i^{MM})^2 \right]^{\frac{1}{2}} \tag{2.4}$$

where N_{freq} is the total number of considered frequencies while ν_i^{QM} and ν_i^{MM} are the i^{th} frequency computed at QM and MM level, respectively. Finally, when torsional scans are performed, a combined plot of the DFT and FF energies along each torsion makes it possible to assess the quality of the fit.

As point (i) is concerned, while the exploitation of reference calculations *in vacuo* might be of use of transferability, it has the disadvantage of neglecting substantial polarization that can occur in solution. For this reason, a change to the protocol was made on the way, including the PCM in the calculations performed at this stage. It is apparent that the resulting force field will be biased toward solvents with the same, or similar, dielectric constant used by the PCM algorithm: this change goes indeed in the direction of an higher accuracy, although going back a little from the transferability point of view. See the case of nicotine (chapter 6) for a discussion about that.

2.2 MD sampling and statistical approach

2.2.1 Molecular dynamics simulations

MD simulations *in vacuo* and in solution were performed for three of the investigated molecules: nicotine, the NfO-TEMPO dye and acrolein (*vide infra* for details), using a parallel version of the Gromacs software. The FFS used for the solute are the ones described above. For NfO-TEMPO-Me both the ground and excited state (GS, EES) FFS were used. For acrolein, in addition to the Joyce FF, also a totally frozen description was adopted.

The simulation timestep was set to 0.5 fs in all the simulations, while the total time vary. For the condensed phase simulations, carried out with periodic boundary conditions (PBC) in all directions, the solute molecule was inserted in a cavity in a pre-equilibrated box of solvent. TIP3P water [39] was used for acrolein, toluene for NfO-TEMPO, methanol and TIP3PFB water [40] for nicotine. Methanol and toluene are modelled using OPLS parameters. The NfO-TEMPO compound was also simulated in combination with a chain of polyethylene (PE), to which it was covalently bonded. The equilibration of PE is detailed in section 7.5. For all the simulations, the system was coupled to a thermostat at 300 K and to a pressure bath at 1 atm, with time constant of 0.1 and 1.0 ps respectively, by means of the Berendsen coupling scheme (only T coupling is used in gas phase). For the long-range interactions, the PME method was employed with a cutoff radius of 10 Å. Positions and velocities are printed at least every 0.5 ps into a trajectory file.

2.2.2 Snapshot extraction and UV-vis spectra calculation

In the case of NfO-TEMPO and acrolein, snapshots spaced by constant time intervals have been extracted from MD trajectories in order to calculate the UV-Vis absorption and, when the EES MD is sampled, emission spectra.

From each snapshot only a spherical portion of solvent was retained for further calculation, centred in the solute centre of mass and with radius R_{cut}: discussion on the choice of the radius is presented up-ahead. In the case of the PE chain, the cut was made at R_{cut} retaining, for all the carbon atoms, also the hydrogen atoms bound to it. The extracted coordinates were used for TDDFT calculations of the electronic transitions: $h\nu_{c,i}$ and $f_{c,i}$ are the energies and oscillator strength of the i-th transition computed on the c-th snapshot.

While the solute is always treated at a full-QM level, different strategies for the inclusion of the environment have been used, namely ME, EE, POL.E, QM*/FQ/PCM, as described before.

The transition energies computed on each set of snapshots were convoluted with Gaussian functions in the energy domain with a custom half width at half maximum (HWHM Δ_ν). Thus, the spectrum computed for the c-th frame is

$$\epsilon_c(\nu) \propto \sum_{i \in \text{states}_c} \frac{f_{c,i}}{\Delta_\nu} \exp\left[\left(\frac{\nu - \nu^0_{c,i}}{\sigma_\nu}\right)^2\right] \tag{2.5}$$

where $\sigma_\nu = [2\sqrt{2\ln(2)}]^{-1} \cdot \Delta_\nu$. The energy spectrum can be reported in the wavelength domain ($\epsilon_c(\lambda)$), for a better comparison with experimental data expressed in nm. The signals originating from each snapshot are then averaged according to the equation:

$$\bar{\epsilon}(\lambda) = \sum_{c \in \text{snap}} \frac{\epsilon_c(\lambda)}{N_{\text{snaps}}} \tag{2.6}$$

to achieve the final "statistical" UV-Vis spectrum.

2.3 Dynamic properties determination

From the molecular dynamics simulations of the NfO-TEMPO molecule, some dynamic properties were calculated that exploit ensemble averages of time-resolved quantities. This is the case of the mean square displacements (MSD) used to describe translational diffusion, and time auto-correlation functions, used for both diffusion parameters and vibrational mode analysis.

Since it will be used in the following, it is convenient to remember that the time auto-correlation function of a quantity A is defined as

$$c_{AA}(\tau) = \frac{\langle \delta A(t_0 + \tau) \cdot \delta A(t_0) \rangle}{\sigma(A)^2} \tag{2.7}$$

where δA is the deviation of A from its mean value, and σ its mean squared value [28]. The non normalised ACF is

$$C_{AA}(\tau) = c_{AA}(\tau) \cdot \sigma(A)^2 \tag{2.8}$$

The averaging $\langle \cdots \rangle$ is done over all the configurations in an ensemble: in the discretised picture of a MD simulation it means that during a "averaging loop" the time origin t_0 is updated at each step, while the correlation product inside the brackets is evaluated in a inner loop for every $t > t_0$. Thus, the denominator in the average is smaller for bigger τs, if the simulation time is finite, it being the reason why auto-correlation statistics are less significant for longer times, and usually it is computed for at most $\tau \simeq \frac{1}{2} t_{\mathrm{tot}}$.

Additional MD simulations devoted to these calculations were performed in the NVE ensemble, starting from equilibrated configurations extracted from the preliminary NPT or NVT (condensed and gas phase, respectively) runs. The output positions and velocities were stored in a binary file obtained with a Gromacs *longformat* (8 decimal places) print [35].

2.3.1 Translational diffusion

For the GS of NfO-TEMPO-Me in solution and NfO-TEMPO in PE, the MSD was calculated as

$$\mathrm{MSD}(\tau) = \langle [\mathbf{r}(t_0 + \tau) - \mathbf{r}(t_0)]^2 \rangle \tag{2.9}$$

where $\mathbf{r}(t)$ is the position of the centre of mass of the solute at time t. The averaging $\langle \cdots \rangle$ is performed like in the ACF case above.

If an asymptotic linear regime is reached for the MSD, the isotropic translational diffusion coefficient is computed as

$$D = \lim_{\tau \to \infty} D(\tau) = \lim_{\tau \to \infty} \frac{1}{6t} \mathrm{MSD}(\tau) \tag{2.10}$$

2.3.2 Rotational diffusion

Also the rotational behaviour of the molecule has been studied, namely the rotational diffusion coefficients, obtained through the angular velocities auto-correlation functions (AVAF), and the inertia axis reorientations.

First, three principal inertia axes (a, b, c) are calculated for each configuration, including only the dye into the inertia tensor definition. Then, the molecular angular velocity ω is calculated and projected onto the axes. The rotational diffusion coefficients are then defined as

$$D_k^R = \int_0^\infty < \omega_k(t_0 + \tau) \cdot \omega_k(t_0) > d\tau \;\; ; \;\; k = a, b, c \tag{2.11}$$

To quantify the reorientation of the inertia axes, the angle $\phi_k(\tau)$ between each axis k and itself at times $t_0 + \tau$ and t_0 are stored, then the first and second Legendre polynomial of $\cos(\phi_k(\tau))$ are calculated. Those polynomial read:

$$P_k^1(\tau) = \cos(\phi_k(\tau)) \qquad ; \qquad P_k^2(\tau) = \frac{1}{2}(3\cos^2(\phi_k(\tau)) - 1)$$

and the value of their integral are called the first and second reorientation time for axis k.

2.3.3 Vibrations

Given an atomistic MD, for which simulation timesteps as short as fractions of a femtoseconds must be used, one can have the availability of the positions and velocities of a simulated system or a relevant portion of it (*e.g.* the solute), with the printout time being shortest than the characteristic time of the fastest nuclear motion, which is usually in the order of 10 fs.

From those data, information about the vibrational behaviour of the system can be obtained, included simulated IR spectra. However the model characteristics, such as whether polarization was included or not, can play a crucial role in determining the goodness of these results.

The infrared absorption intensity per unit volume, at frequency ω and temperature $T \equiv \frac{1}{\beta}$ can be written as [29]

$$\alpha(\omega) = \left[\frac{4\pi^2\omega}{3V\hbar c n(\omega)}\right] \left(e^{-\beta\hbar\omega}\right) I(\omega)$$

being the "lineshape function" $I(\omega)$ the Fourier transformed auto-correlation function (ACF, see Eq. 2.8) of the dipole moment operator $\mathbf{M}$:

$$I(\omega) = \frac{1}{2\pi} \int \mathrm{d}t e^{i\omega t} C_{\mathbf{MM}}(t) = \frac{1}{2\pi} \int \mathrm{d}t e^{i\omega t} \langle \mathbf{M}(0)\mathbf{M}(t) \rangle$$

A rough approximation of this last function within the classical picture would be to substitute the operator $\mathbf{M}$ with its classical counterpart

$$\mathbf{M} \to \mathbf{M}_{cl} \equiv \sum_{i \in \mathrm{atoms}} q_i \mathbf{r}_i$$

where qs are the partial atomic charges and $\mathbf{r}$s the atomic positions (note that this quantity doesn't depend on the coordinates origin if $\sum q_i = 0$, which is always the case in this work). Nevertheless, when doing this substitution, the temperature dependence existent in the quantum formalism is lost. Since $I(-\omega) = e^{-\beta\hbar\omega} I(\omega)$ must hold [41], a corrective pre-factor Q_{QC} has to be inserted. Among many factors proposed, the so called "harmonic correction"

$$Q^h_{QC} = \frac{\beta\hbar\omega}{1 - e^{-\beta\hbar\omega}}$$

will be of our choice, following the suggestion of Ref. [41]. Being I_{cl} the lineshape function calculated using $\mathbf{M}_{cl}$, the final lineshape function shall be expressed as

$$I(\omega) \simeq Q^h_{QC} I_{cl}(\omega) \tag{2.12}$$

The goodness of an IR spectrum is limited, in the case studies presented, by the absence of polarization, that affects both the MD simulation itself and its analysis, since the partial atomic charges enter the definition of classical dipole moments. The latter issue, however, only influences the IR intensities: when it comes to the peak positions, they are still correctly determined by the Fourier transform.

That said, the task of extracting and comparing only the frequencies of the vibrational modes can be accomplished by calculating the Fourier transform of other auto-correlation function. Indeed also the mode analysis of the velocity auto-correlation (VAF) was used to obtain a "power spectrum".

2.4 Hydrogen bond analysis

In the case of simulations in polar solvents, *e.g.* in water, that can act as hydrogen bond donor for electronegative atoms like oxygen (in pyridine) or nitrogen (nicotine), the characterisation of the number and geometry of hydrogen bonded molecules is of interest. Sure enough, the very strong and specific character of this solute-solvent interaction may strongly affect the features of the molecular orbitals responsible for the electronic excitation; moreover, obtaining an hydrogen bond dynamics which is in agreement with the prior-existing chemical knowledge is crucial to validate the force field, with particular respect to the treatment of explicit lone pairs (LP).

The standard criterion to define an hydrogen bond is a discrete geometrical threshold: all the acceptor (A) - water pairs for which the following hold:

$$\begin{aligned} \mathrm{A\cdot\cdot H} &< r^{\mathrm{thr}} \\ \mathrm{A\cdot\cdot O} &< r^{\mathrm{thr}}_{\mathrm{A-D}} \\ \mathrm{A\cdot\cdot O-H} &< \theta^{\mathrm{thr}} \end{aligned} \tag{2.13}$$

are considered as hydrogen bonded pairs. Sometimes either the first or the second distance criterion is omitted, and in general the numerical parameters may vary, as the operational definition of hydrogen bonding is still not unanimous [42]. However, in the case of nitrogen and water (nitrogen is treated as an acceptor when it's not protonated), usually $r^{\mathrm{thr}} = 2.6$ Å, $r^{\mathrm{thr}}_{\mathrm{A-D}} = 3.5$ Å and $\theta^{\mathrm{thr}} = 30°$ are used [43].

However, a continuous formulation has been proposed [44] to take into account the graduality of the decay of typical radial and angular distribution functions, substituting the step-like thresholds with Gaussian slopes.

By doing so, every donor $\cdot\cdot$ H_2O pair can be considered an hydrogen bond only partly, with a "score" between 0 and 1 given by

$$f(r,\theta) = g(r; r_e, r_{\mathrm{hw}}) \times g(\theta; \theta_e, \theta_{\mathrm{hw}}) \tag{2.14}$$

where r is the acceptor-hydrogen distance, θ the acceptor-O-H angle, and

$$g(x; x_e, x_{\mathrm{hw}}) = \begin{cases} 1 & \text{if } x \le x_e \\ \exp\left(-\frac{(x-x_e)^2}{2x_{\mathrm{hw}}^2}\right) & \text{if } x > x_e. \end{cases} \tag{2.15}$$

Here x_e, x_{hw} are the maximum and half-width values relative to the distribution function of the considered variable, and evidently require a preliminary fit of the descending slopes of the radial and angular distribution functions, which depend on the system under examination.

This formulation has been used for the analysis reported in section 6.2.2 for the nicotine molecule.

2.5 Embedding schemes for the property calculation

As previously said, dealing with solvents or more complex environment, particularly when they are responsible for specific interactions, may require an explicit treatment of a portion of the solvent when performing the TDDFT property calculation. For either quantum mechanical, classical or polarizable description (QM∗, MM, FQ) that can be chosen for the embedding, an additional outer layer of PCM is beneficial for the boundary treatment.

2.5.1 Fluctuating charges

Polarizable classical force fields are used in computational chemistry to reproduce the redistribution of a molecule's electron density, during both MD simulations and single point calculations, while retaining an atomistic model based on the nuclear positions [45].

Many methods exist in literature that include polarization, and almost every force field has its polarizable companion. Anyway, three fully classical groups of methods are noticeable: the first one is based on oscillating point charges (Drude oscillators or COS), the second exploits explicit induced dipole terms (TID); a third scheme is the so called "Fluctuating Charges" (FQ), that allow the point charges on each atomic site to flow according to the Electronegativity Equalisation Principle (EEP [20]), and has been exploited, for example, in CHARMM.[46] In our group, the coupling of FQ with PCM has been presented in 2011 for both static and dynamic approaches. The insertion of a QM layer has been added and validated in two 2012 papers [22, 23].

In the FQ model, a point charge q_i is placed on each atomic site and its value is allowed to vary depending on the environment. The contribution to the electrostatic energy of the molecular system depends on the charges in a way that can be expressed, up to the 2$^{\text{nd}}$ order, as

$$E^{FQ} = \sum_i \left(\chi_i q_i + \eta_i q_i^2 + \sum_{j<i} J_{ij} q_i q_j \right) \tag{2.16}$$

where $\{\chi_i\}$ and $\{\eta_i\}$ are respectively the Mulliken electronegativity and the chemical hardness of the isolated i-th atom, and $\{J_{ij}\}$ describes the Coulomb interaction of two charges and depends on their distance. For convenience, the boundary value is imposed:

$J_{ii}(r_{ij}=0) \equiv \eta_i$; the distance-dependency of $J_{ij}(r_{ij})$ is then calculated by means of one among different ways that exist in literature. They include an overlapping integral of a pair of spherical charges with Slater distribution [47] and shielded negative power-laws [48], which retain a lower computational cost.

According to the EEP, the electronegativity at each atom site,

$$\tilde{\chi}_i = \frac{\partial E^{FQ}}{\partial q_i}$$

should be the same. This can be achieved requiring E^{FQ} to be at its minimum with a constraint on the total charge, that has to remain unchanged within each molecule. This constrained minimum can be obtained by a set of Lagrangian multipliers $\{\lambda_\alpha\}$, where Greek indexes run on molecules:

$$\mathcal{E}(\mathbf{q},\boldsymbol{\lambda}) = \sum_{\alpha,i} q_{\alpha i}\chi_{\alpha i} + \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j} q_{\alpha i}J_{\alpha i,\beta j}q_{\beta j} + \sum_{\alpha}\lambda_\alpha \sum_i (q_{\alpha i} - Q_\alpha)$$

In a matrix form:

$$\mathcal{E}(\mathbf{q},\boldsymbol{\lambda}) = \mathbf{q}^\dagger\boldsymbol{\chi} + \frac{1}{2}\mathbf{q}^\dagger\mathbf{J}\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q} \tag{2.17}$$

($\boldsymbol{q}, \boldsymbol{\chi}$ are column vectors of size N_{at}, $\boldsymbol{J}$ is a $N_{\text{at}} \times N_{\text{at}}$ matrix, $\boldsymbol{\lambda}$ is a N_{mol} sized vector). By taking the derivatives of F with respect to charges and Lagrangian multipliers and setting them to zero, one finds that

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C} \tag{2.18}$$

where the equations have been written in a compact form: the vector $\boldsymbol{q}_{\boldsymbol{\lambda}}$ contains both the q's and the λ's, the matrix $\boldsymbol{D}$ is equivalent to $\boldsymbol{J}$ for the atomic part, and accounts for Lagrange multipliers for the molecules part, and $\boldsymbol{C}$ is made out of the electronegativities and the molecular charge constraints.

This variational formulation easily couples to a SCF description of the QM layer, since also Kohn-Sham DFT and Hartree-Fock are expressed in a variational form; moreover, this is also true for a recent formulation of the PCM. That said, a total energy functional $\mathcal{E}$ (Eq. 2.17) for a QM/FQ/PCM layered model is developed, and a coupling term appear, as an additive contribution, in both the FQ equation (Eq. 2.18) and the Fock matrix[1].

---

[1] The total energy functional $\mathcal{E}$ for QM/FQ/PCM is composed by the FQ part in Eq. 2.17, plus the diagonal and cross terms relative to the other two layers:

- the SCF energy expressed in terms of a single-particle reduced density matrix $\boldsymbol{P}$: $\text{tr}\boldsymbol{hP} + \frac{1}{2}\text{tr}\boldsymbol{PG(P)}$, being $\boldsymbol{h}$ and $\boldsymbol{G}$ respectively the one and two-electron matrices;

With the aim of calculating the classical forces for an MD simulation and optimising the geometry, analytical first derivatives with respect to the nuclear coordinates are needed, while second derivatives can also allow for the calculation of the IR spectra and polarizability. The analytical derivatives have been presented, implemented and validated in [23]. Also a linear response theory for the calculation of excited states properties has been carried out in our group [22]. See the details in the cited manuscripts.

2.5.2 Polarizable models on top of the Joyce force field

The Joyce parametrised FFS can be integrated with polarizable models. Since the electrostatic term remains constant during the whole parametrisation procedure, the FF can be forced to inherit directly the polarizable electrostatic calculated with other software

---

- the variational expression for PCM: $\frac{1}{2f(\epsilon)}\boldsymbol{\sigma}^\dagger\boldsymbol{S}\boldsymbol{\sigma}$, where the surface charge distribution is replaced by the set of discrete charges $\boldsymbol{\sigma}$, interacting via the Coulomb interaction matrix $\boldsymbol{S}$;
- the QM/FQ interaction: $\boldsymbol{q}^\dagger\boldsymbol{V}(\boldsymbol{P})$ where $\boldsymbol{V}$ is a classical electrostatic potential at the FQ atomic site caused by the QM density of charges;
- the FQ/PCM interaction $\boldsymbol{\sigma}^\dagger\boldsymbol{\Omega}\boldsymbol{q}$, where $\boldsymbol{\Omega}$ is the Coulomb interaction between the two point charges $\boldsymbol{\sigma}$ and $\boldsymbol{q}$;
- the QM/PCM interaction $\boldsymbol{\sigma}^\dagger\boldsymbol{V}(\boldsymbol{P})$.

Using an atomic basis set $\{\chi_\mu\}$ for the electrons of the QM layer, the classical electrostatic potential $\boldsymbol{V}$ on the i-th FQ or PCM point charge can be expressed as

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

which defines the "uncontracted" potential $\mathbf{V}_{\mu\nu}$ and makes the $\boldsymbol{\sigma}^\dagger\boldsymbol{V}(\boldsymbol{P})$ and $\boldsymbol{q}^\dagger\boldsymbol{V}(\boldsymbol{P})$ terms variational in both their variables.

The PCM contributions only modify Eq. 2.18 to

$$\begin{pmatrix} \mathbf{D} & \boldsymbol{\Omega}^\dagger \\ \boldsymbol{\Omega} & \mathbf{S}/f(\varepsilon) \end{pmatrix} \begin{pmatrix} \mathbf{q}_{\boldsymbol{\lambda}} \\ \boldsymbol{\sigma} \end{pmatrix} = \begin{pmatrix} -\mathbf{C} \\ \mathbf{0} \end{pmatrix} \tag{2.19}$$

Taking the derivatives of the functional with respect to charges and Lagrange multipliers, the effect of the QM/FQ coupling on the FQ equation is additive:

$$\mathbf{D}\mathbf{q}_{\boldsymbol{\lambda}} = -\mathbf{C} - \boldsymbol{V}(\boldsymbol{P}) \tag{2.20}$$

On the other hand, the derivative with respect to the density matrix gives a formulation for the effective density matrix

$$\tilde{F}_{\mu\nu} = \frac{\partial\mathcal{E}}{\partial P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{q}^\dagger\mathbf{V}_{\mu\nu} \tag{2.21}$$

where the tilde means a correction to the Fock matrix due to the FQs.

(here, a development version of Gaussian09).

Being U^{pol}, G^{pol} and H^{pol} the polarization contributions to energy, gradient and Hessian matrix, computed at the same QM optimised geometries given to Joyce, the uncorrected values can be defined as

$$\Delta U = U - U^{pol} \ ; \ \Delta G = G - G^{pol} \ ; \ \Delta H = H - H^{pol}$$

For this modified parametrisation route, the functional I^{intra} changes to

$$\begin{aligned} I^{intra} &= \sum_{g=0}^{N_{geom}} \left[W_g \left[\Delta U_g - E_g^{intra} \right]^2 \right] + \sum_{K}^{3N-6} W'_g \left[\Delta G_K - \left(\frac{\partial E^{intra}}{\partial Q_K} \right) \right]_{g=0}^2 \\ &+ \sum_{K \leq L}^{3N-6} \frac{2W''_{KL}}{(3N-6)(3N-5)} \left[\Delta H_{KL} - \left(\frac{\partial^2 E^{intra}}{\partial Q_K \partial Q_L} \right) \right]_{g=0}^2 \end{aligned} \quad (2.22)$$

where all symbols have the same meaning as in equation (2.2) and, after that, the same minimisation strategy is exploited.

The FF resulting from the minimisation of I^{intra} can be recoupled to the analytical expression of the electrostatic term U^{pol} within a MM software that supports it, as it's being done within Gaussian 09 in our group.

Since Joyce reads the polarization contribution to the Hessian matrix, it can also reconstruct the effective Hessian of the "polarizable Joyce force field" (FFpol), defined over the normal coordinates as

$$H_{KL}^{FFpol} = \left(\frac{\partial^2 E^{intra}}{\partial Q_K \partial Q_L} \right) + H_{KL}^{pol} \quad (2.23)$$

The diagonalisation of H^{FFpol} leads to a set of vibrational frequencies that are directly comparable to the QM ones obtained from $H \equiv H^{pol} + \Delta H$.

Results are shown in chapter 4 for pyrimidine.

Chapter 3

Property determination by means of polarizable models

In the manuscript Ref. [49], coworkers have devoted effort to the development of the theory and the implementation of the first and second derivatives of the composite QM/FQ/PCM functional with respect to nuclear and electronic perturbations: the availability of a complete hessian matrix is needed for the computation of geometry optimised structures and of vibrational spectra.

On the basis of this enforcement, several tests have been run as proof of concepts: a brief smattering of the results for first and second derivatives is presented here. However, in the following chapters, mainly time dependent response properties have been exploited from the multi layered model, rather than vibrational ones, so long as they are needed for the UV-vis spectra calculations: the theory, implementation and proof of concept of the TD aspects is presented by coworkers in [50].

The results below are presented fundamentally to show how is FQ able to highlight differences from a fixed-charge MM model, in the right direction of the QM benchmark. Moreover, the problem of a cutoff radius for the explicit treatment of solvent environment is tackled for the case study of the frequency dependent isotropic polarizability of pyrimidine in water. This result will be beneficial for the choice of a cutoff in later applications.

3.1 Convergence with respect to explicit solvent radius

The convergence study has been carried out on a pyrimidine molecule in aqueous solution; the computed quantity is the isotropic polarizability at 589nm. A single snapshot was taken from a molecular dynamics simulation [51], so the computation has no statistical relevance as regards the final value, but is anyway good to assess the role of the solvent on each individual multilayer process.

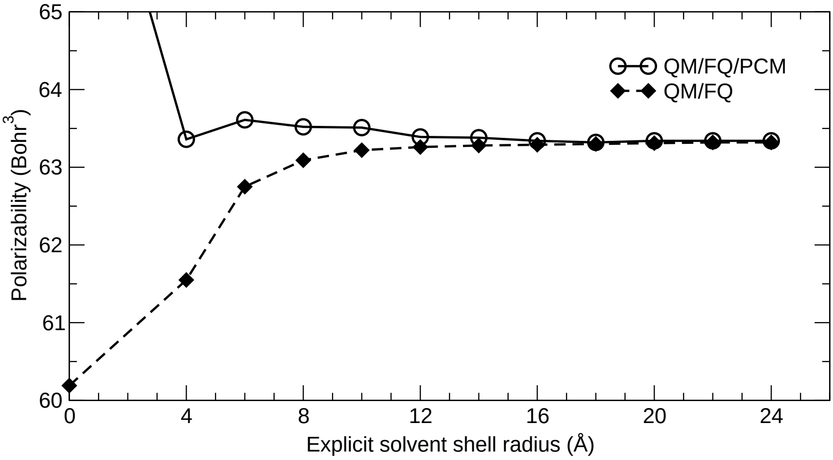


Figure 3.1: Frequency dependent isotropic polarizability at 589 nm for the pyrimidine molecule, as a function of the explicit solvent cutting radius. Solid and dashed curve: with and without the PCM embedding respectively. $R = 0$ represents QM/PCM only, and is off scale.

Concentric spheres ($R^{\text{cutoff}} = 0 \ldots 24$ Å) were cut around pyrimidine, retaining all the water molecules whose oxygen atom was inside the sphere. For the QM/FQ/PCM a single spherical PCM cavity was employed: its radius is 1.5 Å larger than R^{cutoff}. The chosen level of theory for the solute is B3LYP as exchange-correlation functional and aug-cc-pVTZ [52] as basis set.

At last, the FQ model needs a set of electronegativities and chemical hardnesses for the water oxygen and hydrogen. As a first approximation, the ones given in Ref. [47] were used, coupled to the description of electrostatic interactions ($J_{ij}(r_{ij})$, see Eq. 2.16) by means of the Ohno's kernel [48].

In Fig. 3.1 the results are shown for the two schemes QM/FQ and QM/FQ/PCM, $R = 0$

corresponding to the calculation without explicit solvent molecules: *in vacuo* or with a molecule shaped PCM cavity, respectively.

For the case of isotropic polarizability, the application of the purely continuum approach overestimates by a great quantity the value – computed with the explicit treatment – which is believed to be the most accurate one. It is thus evident that the short-range treatment goes in the opposite direction with respect to the averaged, long-range ones.

The number of explicit molecules needed to achieve property convergence is quite large - a cutting radius of 12-14 Å is necessary, however, as expected, the interplay between specific and long-range interactions is beneficial. In fact, in the full multilayered treatment, the calculated data are already near convergence when only the first solvation shell is explicitly considered.

This behaviour shows that a good balance between long and short range effects must be taken into account so long as the averaged treatment of PCM, which is known to be reliable for long-range treatment, gives scarce results here, while the inclusion of even only one explicit shell brings the values to an almost converged result.

It should be noted that the role of the explicit treatment is more evident in the present case than for the vertical excitation energies [50]. For this reason, cutoff radii of about $R = 14$ Å will be used with rather confidence in the applications of the following chapters that are focused on optical response.

3.2 Geometry optimisation of a molecule in a biological matrix

First derivatives with respect to the geometrical displacements were employed for the calculation of minimal energy geometries of the retinal chromophore in a rhodopsin mimic, starting from the known crystal structure of the molecule's pocket.

As before, different models have been used to account for the environment, while the retinal has constituted the QM inner subsystem, treated at the B3LYP/6-31G$*$ DFT level. Here the retinal is the only portion of the composite which is allowed to move, while the biological matrix is held fixed.

Three models are compared for rhodopsin: a continuum PCM with a dielectric constant of 15, fixed charges based on Amber, and FQs with parameters from CHARMM [46]. The differences between the various models – as it is demonstrated by the table in Fig. 3.2, where selected bond distances and angles are reported – are not negligible, despite that the embedding particles are in a fixed position.

	PCM	MM	FQ
Distances (Å)			
l_a: terminal C-O	1.231	1.226	1.235
l_b: chain C-C	1.365	1.365	1.363
l_c: chain C-H	1.086	1.084	1.084
l_d: chain C-CH_3	1.510	1.514	1.513
l_e: ring C-C	1.513	1.510	1.510
l_f: ring C-CH_3	1.518	1.518	1.514
l_g: chain length	12.140	12.061	12.228
Angle (°)			
θ_a: chain orientation	90	104	94
Dihedral (°)			
δ_a: chain orientation	-176	-174	-168

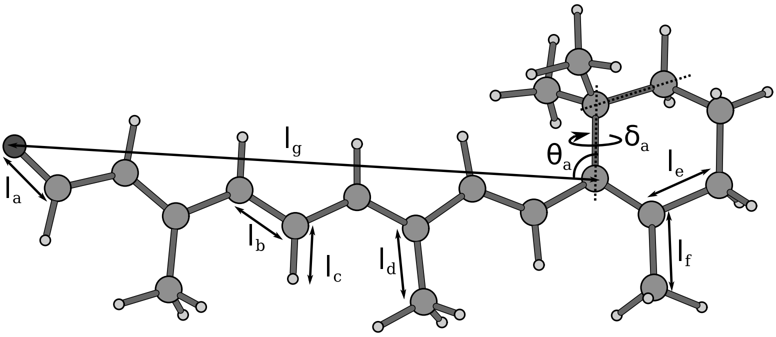


Figure 3.2: Top: calculated geometric features of the retinal in rhodopsin mimic as obtained at DFT level with different models for the environment. Bottom: Schematic view of retinal with the indication of selected bond lengths and angles.

3.3 IR harmonic spectra

Finally, for both the investigated systems, IR spectra have been calculated, making use of QM/FQ/PCM second derivatives. To be more exact, even if an hessian matrix as large as the sum of QM and MM atoms is built, only the QM-QM portion is considered when diagonalising and extracting the normal modes. This approximation, exploited in Ref.

[49], is called Partial Hessian Vibrational Approach (PHVA) and is suitable for a general "focused" methodology, where the interest stands in the properties of the solute and not of the system as a whole.

Fig. 3.3 (top) reports calculated IR spectra of pyrimidine in the gas phase and in aqueous solution as described by means of a hierarchy of approaches; starting from the snapshot used in section 3.1, a preliminary optimisation was run at the different levels by keeping the water molecules fixed Intensities were scaled by $1/n$, with n being the refractive index of water, for the spectra in solution.

The inclusion of environmental effects causes, as expected, a shift in the calculated frequency values, which results in either an increase or decrease of the values depending on the mode under investigation. The relative intensities of the bands also change, and in general they are more affected by the environment than the corresponding frequencies. All the solvation approaches give consistent results, not only for frequencies, but noticeably also for intensities. As expected, the PCM intensities are generally larger that the corresponding values in vacuo. The QM/FQ/PCM intensities are generally slightly larger than the corresponding non-polarizable QM/MM/PCM ones, which is consistent with the introduction of polarization effects.

The same procedure for the isolated molecule and the QM/FQ protocol has been applied to the said biological framework of a rhodopsin with the retinal inside, and the results are reported in Fig. 3.3 (bottom). The effect of the environment is noticeable and do not simply results in a scaling of the spectrum of the isolated molecule; in general relative intensities are more affected by the environment than frequency values. The bands more affected by the environment are that around 1600 cm^{-1}, due to chain motions and the one around 1700 cm^{-1}, due to the C=O stretching mode.

By comparing all the three embedding schemes (two of them are not shown in the plot), the C=O line is shifted by -39 cm^{-1} with PCM and -34 cm^{-1} with QM/FQ, whereas the QM/MM approach gives a shift of +6 cm^{-1}. It should be noted from Tab. 3.2 that the bond length is in effect subject to a strong embedding effect already upon geometry optimisation. These findings demonstrate the crucial role of polarization in the vibrational modes of the target molecule.

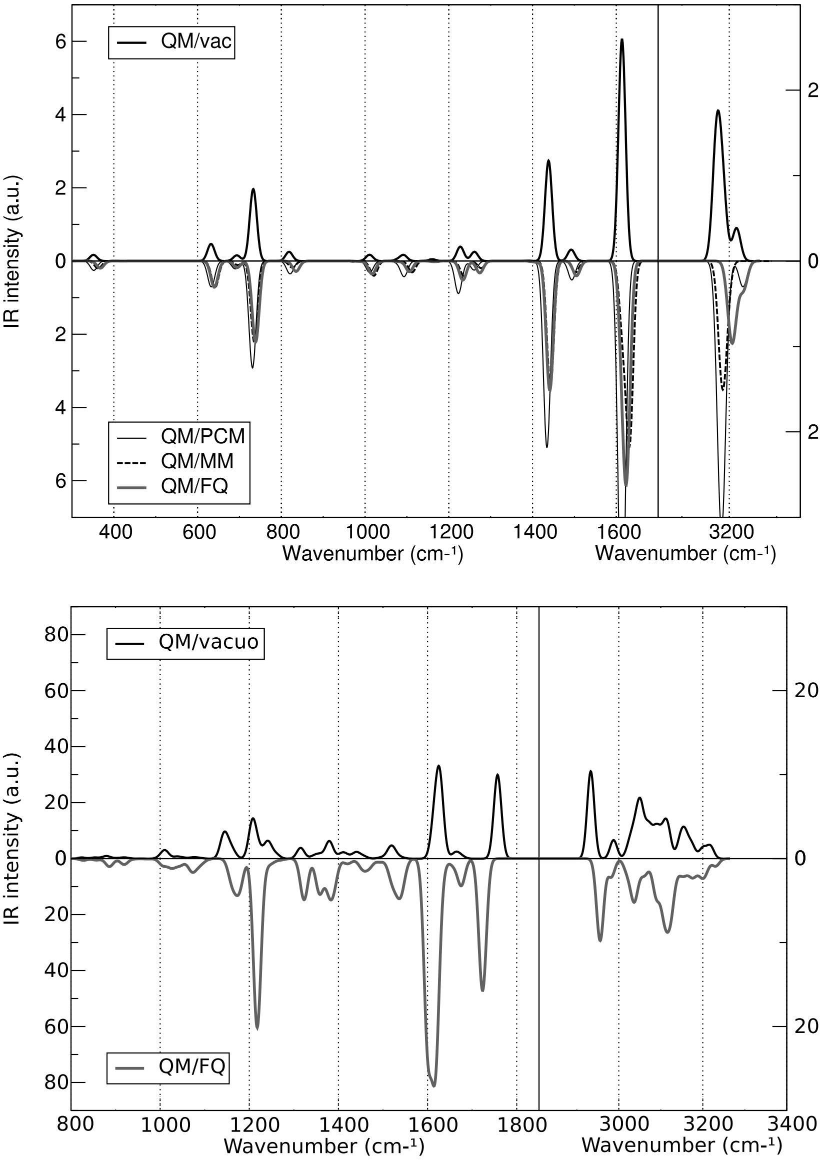


Figure 3.3: IR spectra of pyrimidine in water (top) and retinal in rhodopsin (bottom) calculated for the isolated "solute" (top halves) and for the embedded system. Different schemes are employed for the embedding.

Chapter 4

Validation set for the Joyce force field

In this chapter, a first set of developed force fields is itemised, allowing to introduce several methods and capabilities available within our modeling strategy. All the investigated molecules treated in this and in the following chapters are summarised in fig. 4.1, yet some of them are studied into an embedding consisting of a solvent or a polymer matrix, not reported in the figure.

This chapter will present the main features of the parametrisation protocol and of the physical chemical insights that can be drawn from the developed force fields alone, while the following ones will present three case studies, where MD simulations have been performed and eventually the statistical spectra have been computed.

All the force fields that were developed are reported in the Appendix.

4.1 FF-level computation of vibrational frequencies

A validation study of the Joyce2.0/Ulysses method has been presented in [36].

A first test was devised for the validation of Joyce capabilities to reproduce the vibrational frequencies and modes of the selected target molecule. The target set corresponds to fig. 4.1(a-e), and was chosen as benchmark for two reasons. First it contains small molecular models characterised by different degrees of flexibility, ranging from

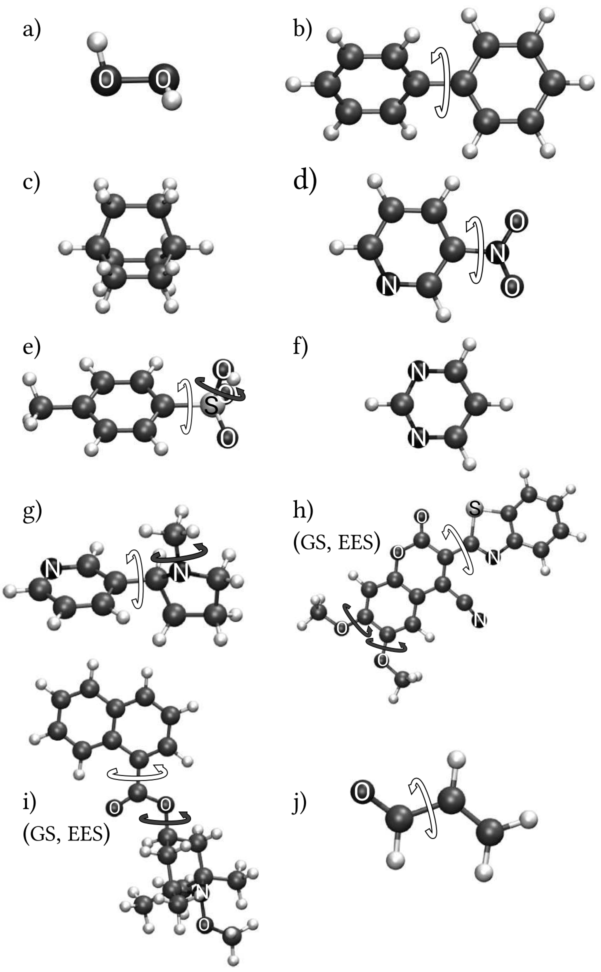


Figure 4.1: The molecules investigated throughout this document: a) hydrogen peroxide; b) biphenyl; c) bicyclo[2.2.2]octane; d) 3-nitro-pyridine; e) toluensulfonic acid; f) pyrimidine; g) nicotine (see also chapter 6); h) coumarin-based dye; i) NfO-TEMPO-Me (chapter 7); j) acrolein (chapter 5). The dihedrals studied by means of a torsional scan are evidenced with light (δ_1) and dark ($\delta_{n>1}$) arrows. For h) and i) an FF was parametrised also for the EES.

the very rigid bicyclo[2.2.2]octane to the more "soft" toluensulfonic acid. In second place this set was previously employed to test the performances of three different FFs [31], allowing for a direct comparison of the different parametrisation procedure. Furthermore, to show the flexibility of Joyce program, several parametrisation strategies are followed for different molecules, which will be discussed in some detail in the following. The general outline of section 2.1.2 is followed.

H-peroxide

The only QM calculations performed in this case are the evaluation of the minimum energy conformation and the vibrational frequencies computed on it. Considering H-peroxide's simple structure, it was chosen to adopt a purely diagonal FF, employing only terms in Eq. 2.1(1-3) and turning off both intramolecular non-bonded interactions and off-diagonal couplings terms (Eq. 2.1(5) and 2.3).

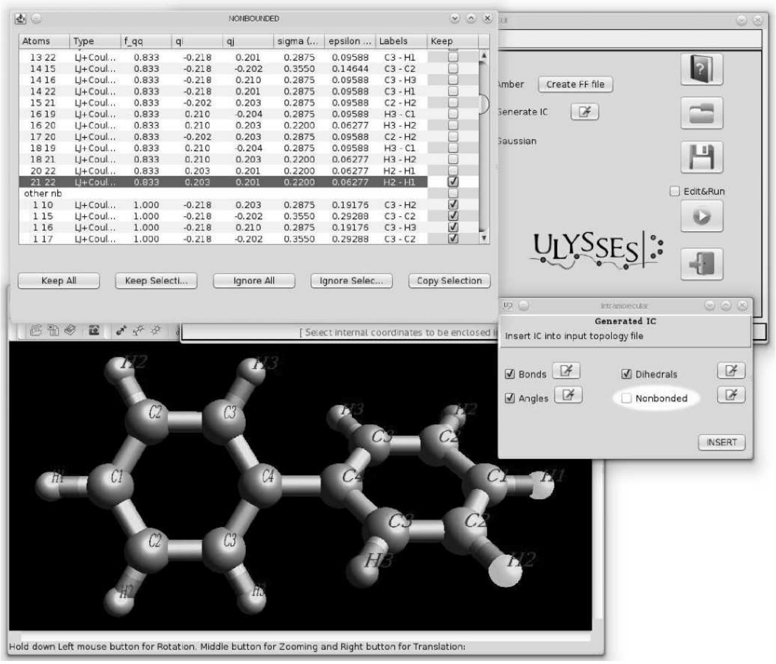
Joyce were used to find the best parameter set and the parametrised FF was employed to re-compute the frequencies at MM level. Although Joyce is capable of making this calculation, the same software, Gaussian 09, employed for the QM calculation was used for consistency, making use of the input file automatically produced by Joyce. The frequency related standard deviation (σ_{freq}) with respect to the QM values was computed: see Eq. 2.4. Eventually, the resulting σ_{freq} is compared (see Table 4.1) to those obtained with the ν_i^{MM} values reported in the "parafreq" article [31]. It may be worth mentioning that since Joyce parametrisation procedure is guided by the functional (2.2) and by the overall standard deviation $I^{\frac{1}{2}}$ (which in this case it is solely proportional to the difference between the QM and MM Hessian matrices) there is no guarantee that the resulting σ_{freq} is the best available for the selected set of ICs. Notwithstanding this observation, the resulting value of 49.8 cm^{-1} is found to be better than the one obtained with the *parafreq* approach (65.6 cm^{-1}).

Biphenyl

The biphenyl molecule is often employed as a prototype for flexible molecules, as it is composed by two fairly rigid moieties as the phenyl rings, connected through a partially conjugated bond, which allows the molecular skeleton to undergo large amplitude oscillations. Its non planar minimum geometry is the result of a delicate balance between the repulsion among the hydrogen atoms closer to the inter-ring bond and the energy gain due to ring-ring conjugation, maximised by the planar conformation. In this case, to mimic the steric repulsion between hydrogen atoms, the non-bonded intramolecular term of Eq. 2.1(5) was added to the diagonal FF. In particular, atomic point charges and LJ parameters were assigned with the Ulysses GUI, by taking Mulliken charges from the

QM optimisation and OPLS [53] parameters from the literature. The initial redundant set of IC generated by Joyce, in addition to bond lengths, angles and dihedrals, also includes all possible intramolecular distances between atoms more than three bonds apart. However, considering the rigidity of the planar rings, the 1-4 interactions were switched off, leaving those between atoms separated by four or more bonds (*i.e.* 1-5, 1-6, ...). Figure 4.2 is reported to show how the IC selection/deletion is made possible in a visual way through Ulysses [37].

The large amplitude libration, already active at room temperature, makes the harmonic approximation for the "soft" degree of freedom rather questionable. For this reason, besides the optimised geometries and relative Hessian matrix, a QM energy scan along the main dihedral, highlighted by a light arrow in fig. 4.1(b), was performed,



NONBOUNDED

Atoms	Type	f_qq	qi	qj	sigma (...	epsilon ...	Labels	Keep
13 22	LJ+Coul...	0.833	-0.218	0.201	0.2875	0.09588	C3 - H1	☐
14 15	LJ+Coul...	0.833	-0.218	-0.202	0.3550	0.14644	C3 - C2	☐
14 16	LJ+Coul...	0.833	-0.218	0.210	0.2875	0.09588	C3 - H3	☐
14 22	LJ+Coul...	0.833	-0.218	0.201	0.2875	0.09588	C3 - H1	☐
15 21	LJ+Coul...	0.833	-0.202	0.203	0.2875	0.09588	C2 - H2	☐
16 19	LJ+Coul...	0.833	0.210	-0.204	0.2875	0.09588	H3 - C1	☐
16 20	LJ+Coul...	0.833	0.210	0.203	0.2200	0.06277	H3 - H2	☐
17 20	LJ+Coul...	0.833	-0.202	0.203	0.2875	0.09588	C2 - H2	☐
18 19	LJ+Coul...	0.833	0.210	-0.204	0.2875	0.09588	H3 - C1	☐
18 21	LJ+Coul...	0.833	0.210	0.203	0.2200	0.06277	H3 - H2	☐
20 22	LJ+Coul...	0.833	0.203	0.201	0.2200	0.06277	H2 - H1	☐
21 22	LJ+Coul...	0.833	0.203	0.201	0.2200	0.06277	H2 - H1	☑
other nb								☐
1 10	LJ+Coul...	1.000	-0.218	0.203	0.2875	0.19176	C3 - H2	☑
1 15	LJ+Coul...	1.000	-0.218	-0.202	0.3550	0.29288	C3 - C2	☑
1 16	LJ+Coul...	1.000	-0.218	0.210	0.2875	0.19176	C3 - H3	☑
1 17	LJ+Coul...	1.000	-0.218	-0.202	0.3550	0.29288	C3 - C2	☑

Generated IC

Insert IC into input topology file

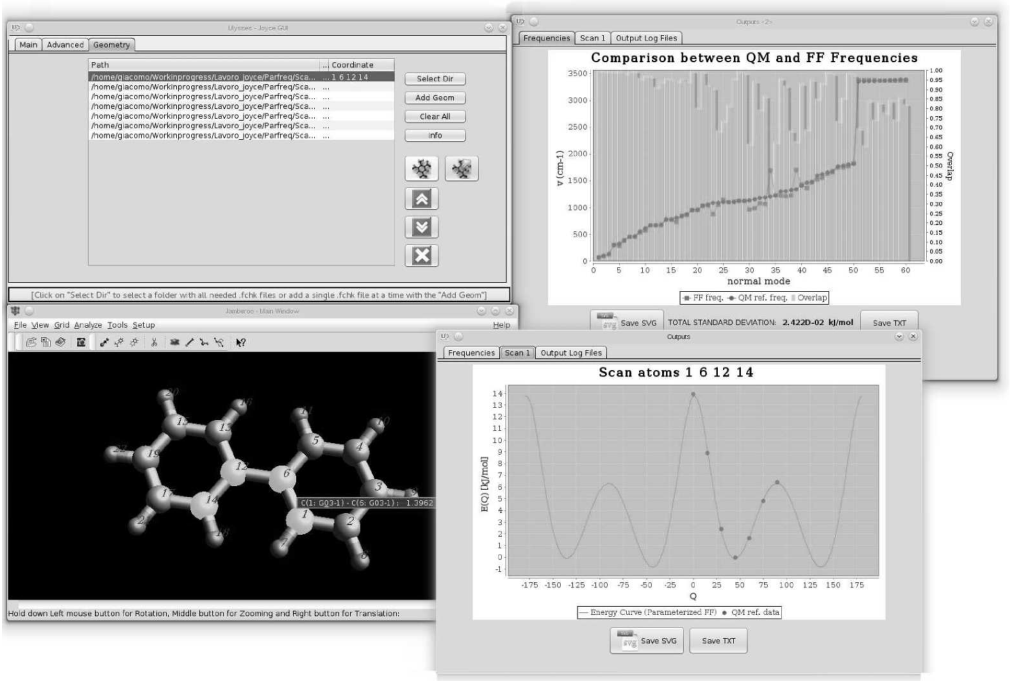
☑ Bonds ☑ Dihedrals

☑ Angles ☐ Nonbonded

Hold down Left mouse button for Rotation, Middle button for Zooming and Right button for Translation:

Figure 4.2: Snapshot of Ulysses during biphenyl parametrisation. The 1-4 intramolecular interactions are turned off by un-checking the small boxes on the right. The 3D visualisation helps the user to quickly localise the 1-4 pairs in the molecule.

and the resulting values were employed in the parametrisation. At variance with the intra-ring torsions, where a single cosine (n=2, γ=180°) for each dihedral was employed in equation 2.1(4), the energy connected to the inter-ring rotation was described using a sum of three cosines (n = 0, 2, 4; γ_0=γ_2=γ_4=0°). The whole parametrisation yielded an overall standard deviation σ of 0.003 kJ/mol. Parametrisation results can be



Path	... Coordinate
/home/giacomo/Workinprogress/Lavoro_joyce/Parfreq/Sca...	... 1 6 12 14
/home/giacomo/Workinprogress/Lavoro_joyce/Parfreq/Sca...	...
/home/giacomo/Workinprogress/Lavoro_joyce/Parfreq/Sca...	...
/home/giacomo/Workinprogress/Lavoro_joyce/Parfreq/Sca...	...
/home/giacomo/Workinprogress/Lavoro_joyce/Parfreq/Sca...	...
/home/giacomo/Workinprogress/Lavoro_joyce/Parfreq/Sca...	...
/home/giacomo/Workinprogress/Lavoro_joyce/Parfreq/Sca...	...

[Click on "Select Dir" to select a folder with all needed .fchk files or add a single .fchk file at a time with the "Add Geom"]

TOTAL STANDARD DEVIATION: 2.422D-02 kJ/mol

C(1: G03-1) - C(6: G03-1) : 1.3962

Hold down Left mouse button for Rotation, Middle button for Zooming and Right button for Translation:

Figure 4.3: Snapshot of Ulysses during biphenyl parametrisation. The parametrisation results are shown in terms of frequencies and torsional scans. In the right top corner the graphical comparison between QM (dots) and FF (squares) frequencies is shown, the underlying histograms representing the overlap between the relative normal modes. On the right bottom the comparison between the QM (dots) and FF (red) torsional profiles can also be visualised.

visualised through Ulysses too. As an example, in Figure 4.3, the comparison between QM and MM computed data is shown in separate windows. In the right top corner the good agreement between QM and MM vibrational frequencies appears from the graphical comparison, whereas the overlap between QM and MM normal modes is indicated by the underlying histograms. On the bottom-right panel, the Joyce capability of well reproducing the torsional energetic profile is apparent.

Finally, to further validate FF reliability, the same procedure employed for Hydrogen peroxide was followed, calculating the vibrational frequencies at the FF level and the resulting σ_{freq}, then comparing to literature values. Since the parametrisation performed in *parafreq* for biphenyl does not include the "soft" ICs, to make the comparison more strict, the σ_{freq} was also computed on a restricted set of frequencies, namely those corresponding to normal modes having null projection over dihedrals. Both deviations, together with literature values, are reported in Table 4.1. As in the previous case, Joyce performance (62.9 cm^{-1}) reveals a good agreement with the parent QM values; the agreement is even better when only the "stiff" subset (37.3 cm^{-1}) is considered.

Bicyclooctane

As regards bicyclo[2.2.2]octane, to describe its fairly rigid structure, at variance from biphenyl, a purely diagonal FF was employed, turning off intramolecular non-bonded interactions, which could cause too strong repulsions within the cyclic structure, ending up in unphysical distortion of the MM optimised structure. As in previous cases, σ_{freq} was computed and reported in table 4.1, confirming the reliability of the present approach.

Nitro-pyridine and toluensulfonic acid

The last two molecules of the test set, considered here for comparison with Ref. [31], are both (Fig. 4.1(d,e)) characterised by the presence of a "stiff" aromatic moiety linked to a substituent which can rotate around the bond connecting it to the ring. As in the case of biphenyl, the inclusion of intramolecular non-bonded terms can help to better reproduce the QM torsional energy profile, so the same parametrisation strategy could be followed. However, with the aim of showing some of the possible parametrisation routes that can be handled by Joyce, a slightly different protocol was adopted. In fact, since in both molecules the most important intramolecular interactions are expected to occur between the α Hydrogens and the oxygen atoms on the two substituents (*i.e.* –NO_2 and –SO_3H, respectively), all non-bonded intramolecular interactions were turned off except those between these sites. As for biphenyl, only one cosine term in Eq. 2.1(4) was assigned to ring dihedrals, whereas multiple cosines where employed for the scanned ones. Joyce parametrisation was performed, yielding an overall standard deviation of 0.2 and 0.08 kJ/mol for nitro-pyridine and toluensulfonic acid, respectively.

By looking at Table 4.1, it is evident that the parametrised FFS are able to reproduce rather accurately the vibrational behaviour of the tested molecules, giving frequency related standard deviations in agreement or even better than the best values reported in the literature.

However, as the highest σ_{freq} is found for the 3-nitro-pyridine molecule, to further test Joyce flexibility, the effect of the inclusion of coupling terms (2.3) was also evaluated. In fact Joyce automatically identifies, through the calculation of appropriate projection coefficients, the most important ICS involved in each normal mode and this can be exploited to select the most important coupling terms to be included in the FF. In the 3-nitro-pyridine case, when terms 2.3(1) between ring's harmonic IC are added to the standard diagonal FF, the overall standard deviation drops from 0.2 to 0.04 kJ/mol and σ_{freq} decreases from 63.2 to 50.9 cm^{-1}.

As the torsional profiles are concerned, parametrisation results are shown in Figure 4.4. It appears that Joyce is capable of reproducing with very good accuracy the QM data for both pyridine and toluensulfonic acid molecules, accounting for all minima and barriers. It should be stressed that both vibrational frequencies and torsional profiles are expected to be important when employing the parametrised FF in condensed phase simulations (as for solvated molecules or bulk phases) or when very accurate molecular geometries are requested, as when MD snapshots are to be employed in QM calculations of electronic properties. These kinds of calculations were done for different molecules, as shown in further sections.

Molecule	GAFF [54]	Seminario [55]	Parafreq [31]	Joyce	
				all IC	stiff IC$^{(a)}$
H-peroxide	-	-	65.6	49.8	-
biphenyl	130.0	80.7	62.2	62.9	37.3
bicyclo[2.2.2]octane	73.5	129.4	68.6	52.9	-
3-nitro-pyridine	156.5	128.7	78.7	63.2	61.1
toluensulfonic acid	158.8	108.4	46.7	41.6	21.1

Table 4.1: Frequency related standard deviations (σ_{freq}, cm^{-1}, see Eq. 2.4) with respect to QM reference values obtained from MM vibrational frequencies computed with different FFS. Values obtained with GAFF [54] parameters, those derived with the method proposed by Seminario [55] or with the parafreq code [31] were all taken from Ref. [31]. (a) In the last column, for comparison purposes, standard deviations are computed only for normal modes with null projection over "soft" ICS.

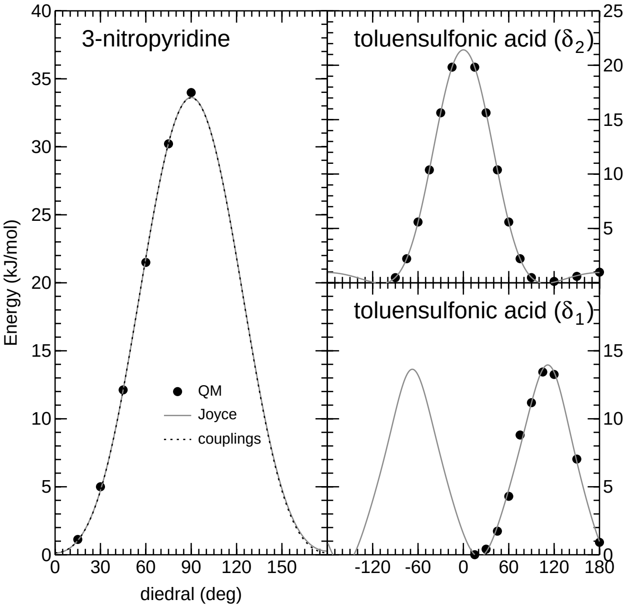


Figure 4.4: Torsional profiles comparison between QM and Joyce scans for 3-nitro-pyridine (left) and toluensulfonic acid (right). For the latter, dihedrals δ_1 and δ_2 are those indicated in Figure 4.1, light and dark arrows respectively. QM data are reported with black dots, whereas grey lines are employed for the Joyce parametrised FF. For the 3-nitro-pyridine molecule, the torsional profile obtained when coupling terms are included is also reported with a dotted line (almost indistinguishable).

4.2 Polarizable force field for pyrimidine

Pyrimidine (fig. 4.1(f)) is the sixth modelled molecule investigated, for which a polarizable FF (FFpol, see section 2.5.2) was developed, as detailed in the following.

Indeed, the spectroscopic behaviour of the molecule has been recently investigated [51] at different levels of theory, including the parametrisation of a classical FF by means of the Joyce software. It turned out that a time dependent approach, based on MD trajectories obtained with the parametrised FF, succeeded in reproducing pyrimidine electronic absorption in CCl_4 and water solvents. However, for the latter solvent, a better agreement with the experiment was achieved when some water molecules were handled at QM level during spectra calculations. It can be thus speculated that results of similar accuracy could be obtained if a polarizable description was used for the solvent. Indeed, as it has been shown in the previous chapter (above all Fig. 3.3 (top)), the polarizable description of the environment has at least a remarkable effect on the vibrational features of the molecule, stretching individual intensities and slightly shifting frequencies of the modes calculated with the non polarizable environment.

To make this treatment possible together with a Joyce force field, the parametrised FF has to be integrated with the polarizable FQ model, as detailed in section 2.5.2.

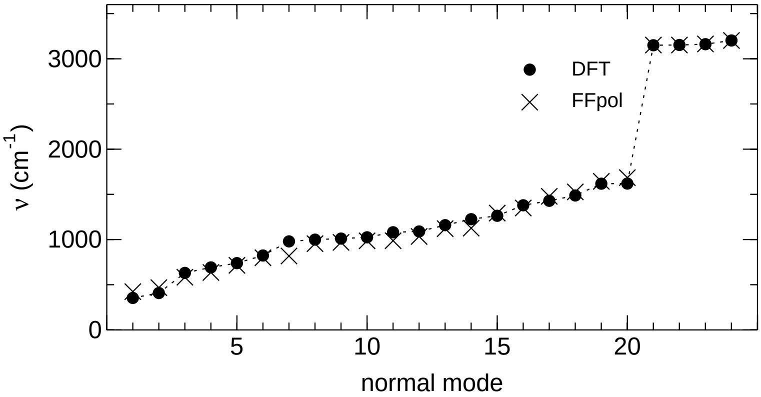


Figure 4.5: Comparison between QM (DFT, full circles) and FFpol (crosses) vibrational frequencies, computed for pyrimidine.

The QM optimised structure of pyrimidine was employed to compute energy, gradients and Hessian matrix by using the FQ model [23]. The electronegativity and hardness for each pyrimidine atom were taken from CHARMM [46] and are reported in the Appendix, whereas UFF [56] parameters where used for the valence FF terms (*e.g.* stretching or bending).

The resulting data (*i.e.* energy, gradients and Hessian matrix) were subtracted from original QM ones and parametrisation performed according to equation (2.22), yielding an overall standard deviation of 0.02 kJ/mol, *i.e.* comparable to those obtained for the already discussed molecules.

Moreover, when vibrational frequencies are recomputed by means of equation 2.23 a frequency related standard deviation σ_{freq} of 58 cm^{-1} is obtained. A comparison between QM and FFpol computed frequencies is reported in figure 4.5.

It can be interesting to compare FF valence terms obtained by Joyce with or without the polarization term. The largest changes are found in stretching constants of the C-N bonds, which increase of about 6% when the polarizable FQ term is subtracted from QM reference data. Conversely, C-C and C-H constants result almost unchanged. Bending and dihedral constants are also almost unaffected, with the exclusion of the N-C-N bending constant, which changes from 1600 kJ/mol rad^{-2} to 1394 kJ/mol rad^{-2} when polarization is accounted for. Despite a deeper interpretation calls for further investigation, it can be hypothesised that the part most sensitive to polarization effects is the portion of the aromatic ring containing nitrogen atoms, and care should be taken in the parametrisation of electronegativity and hardness for these sites.

4.3 Sampling of the nicotine conformers

The molecular structure of the nicotine molecule is shown in figure 4.1(g). It consists of two building blocks, namely a pyridine rigid ring and a aliphatic five-membered ring. Two main "soft" ICs can be identified: the dihedral around the bond connecting the two aforementioned moieties (δ_1, light arrow) and the one driving the rotation of the methyl group linked to the aliphatic Nitrogen (δ_2, dark arrow).

The resulting overall flexibility is expected to cause a number of different conformations to be populated at room temperature, and it would be desirable that an accurate FF could reliably represent most of the possible conformers. In fact, preliminary results on nicotine's dichroic behaviour in solution, obtained with several QM calculations performed in our laboratory, have shown that the conformational equilibrium's taking place in solution have to be accounted for rather accurately, to obtain accurate spectral shape. Indeed, more than two conformers should be taken into account, since the weighted average of only two spectra (computed for the absolute (**A**: $\delta_1 \sim 135^\circ$; $\delta_2 = \pm 60^\circ$) and first local (**B**: $\delta_1 \sim -45^\circ$; $\delta_2 = \pm 60^\circ$) minima) did not yield satisfactory results. It is in situations like this one, that the time dependent “statistical approach” – employing MD trajectories – may be beneficial for the sampling of the target molecule phase space.

Throughout the last two years, two slightly different routes were employed to deal with the simulation of nicotine in solvent. The first one is accounted for below, as it is part of the Joyce validation set [36]: in this case, QM benchmark calculations did not benefit from the PCM treatment. By doing so, the FF is capable of describing the molecule's internal flexibility and is rather transferable to several simulation schemes involving non polar solvents, but it gives a poor description of the interaction with polar ones when it comes to specific short-range behaviours.

A new force field for nicotine, based on DFT/PCM electronic densities, has been developed at a later stage, and is reported in chapter 6.

The FF parametrisation was performed making use of the QM computed energy, gradient and Hessian matrix of the absolute minimum conformation and all the energies and geometries obtained by scanning dihedrals δ_1 and δ_2.

As the selection of the IC set and the FF functional forms are concerned, with respect to those designed for the "parafreq" test cases, a yet different strategy was adopted for nicotine. First, considering the rigidity of the aromatic ring, the harmonic form (Eq. 2.1(3)) was used for all pyridine dihedrals. Conversely, to account for the interconversion among the different conformers of the 5-membered ring, equation 2.1(4) was employed, with a single cosine term (n=3). Finally, the "soft" δ_1 and δ_2 were described with five (n = 0-4) and one (n=3) cosine terms, respectively. Moreover, non-bonded intramolecular interactions described in equation 2.1(5) were all included, except all 1-4 interactions and those involving the hydrogen atoms of the methyl group. LJ intramolecular parameters were taken from the OPLS FF, whereas CM5 [57] charges were employed. The final overall standard deviation, after optimisation of all other parameters, was 0.14 kJ/mol.

By looking at top panels of Figure 4.6, where the QM and Joyce torsional energy profiles are reported for both "soft" dihedrals, it appears that the MM curves well agree with the reference QM profiles. In the top left panel it is shown that the two conformations corresponding to the global **A** and local **B** energy minima can interconvert one into another by a rotation around the δ_1 dihedral, from 135° to -45° , if a barrier of $\sim$ 25 kJ/mol is surmounted. Conversely, δ_2 presents two degenerate minima at ±60° , separated by a lower barrier ($\sim$ 15 kJ/mol).

The parametrised FF, together with an intermolecular part composed of CM5 charges for the Coulomb interactions and OPLS LJ parameter, was employed in a MD simulation of nicotine in methanol solvent. A system of one nicotine and 944 methanol molecules was simulated for 10 ns (after a 2 ns equilibration period) at 1 atm and 298 K, resulting in an average density of 0.77 g cm^{-3}. To investigate the conformational behaviour of the solute, population distribution of both δ_1 and δ_2 dihedrals were computed after 1, 5 and 10 ns of simulation and reported in the lower panels of Figure 4.6. At the end of the simulation run (grey dotted line), both dihedral distributions show two peaks, corresponding to the QM energy minima. The comparison of different time averages reveals that the δ_2 rotation is faster, as equilibrium populations are reached already after 1 ns. Conversely, δ_1 movement is much more hindered, and 10 ns are needed to

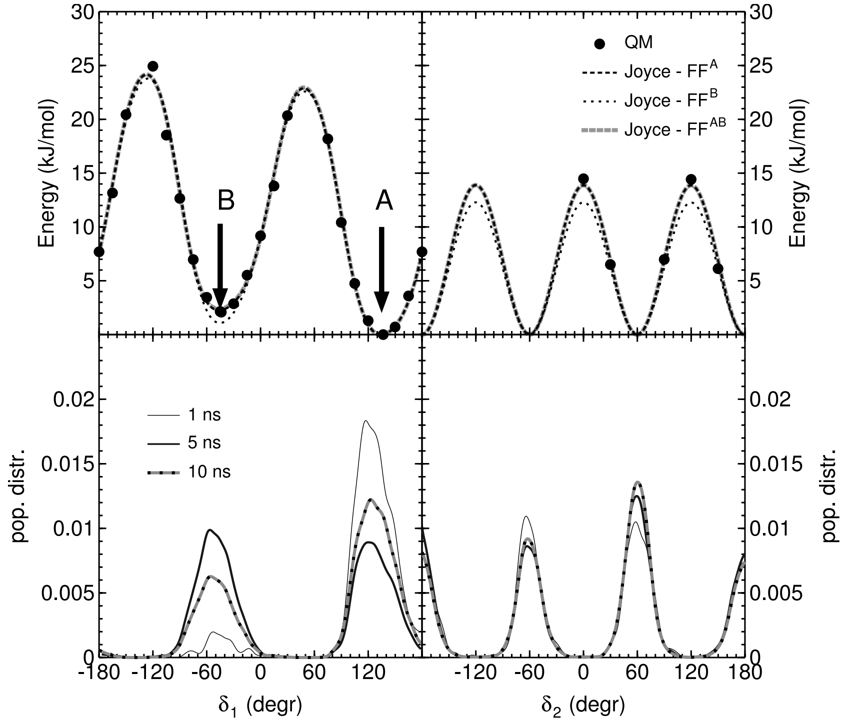


Figure 4.6: Top panels: Joyce fitting results for torsional profiles of the nicotine molecule performed according to the FF^A, FF^B and FF^{AB} schemes; δ_1 δ_2 dihedrals are reported on left and right, respectively. Bottom panels: dihedral distributions computed from MD trajectories of nicotine obtained with FF^A in methanol solvent at 298 K and 1 atm; distributions have been computed after 1 ns (black thin), 5 ns (black thick) and 10 ns (grey dotted line). Labeled arrows indicate the two nicotine's conformers, **A** and **B**.

reach the correct Boltzmann distribution. Indeed, the almost equal populations found at 5 ns (thick black curve) show that when an interconversion occurs, it takes some time to cross the barrier backwards. In other words, δ_1 oscillates around 120° for almost all the first ns, thereafter it jumps to -60° remaining there for about 3 ns and eventually comes back to its original value. More important, all four peaks are rather broad, indicating that both dihedrals undergo large amplitude motions ($\sim \pm 30°$ for δ_1 and $\sim \pm 15°$ for

δ_2).

Finally, it may be worth investigating if the current FF, based on the Hessian matrix of conformer **A** only (and thereafter labelled FF^{A}), is capable to accurately describe the vibrations and structure of conformer **B**, which is not negligibly populated at room temperature. With these aims a further QM calculation was performed, and the Hessian matrix and harmonic vibrational frequencies have been computed for the **B** conformer. These data were subsequently used for the parametrisation of two supplementary FFS: FF^{B} parameters were obtained by using the **B** conformer as reference geometry in functional (2.22), whereas FF^{AB} was parametrised through an extended Joyce merit function, including both the energy sets with equal weights. The overall standard deviation (σ) was computed for all three FF, and reported in the table below, while the resulting torsional FF are sketched in Figure 4.6.

FF	σ (kJ/mol)
FF^{A}	0.14
FF^{B}	0.25
FF^{AB}	0.10

It appears that all three FFS are efficient to reproduce with good accuracy nicotine's vibrational behaviour in both its conformers, accounting for an accurate description of the local motion around the two minima. Furthermore, as appears from Figure 4.6, also the interconversion between them is well described along the torsional reaction coordinate, all FF profiles being in good agreement with the QM computed ones.

In chapter 6 an upgrade of the present FF will be shown, characterised also in relation to the specific solute - solvent (water) interactions, and designed to be used in statistical spectra calculations.

4.4 Effect of flexibility on the absorption and emission spectra of a coumarin dye

The organic dye, whose molecular structure is shown in figure 4.1(h), is a coumarin based fluorophore, which has been recently synthesised and studied by means of experimental and computational techniques [58]. It has attracted much attention thanks to the sensitivity of its fluorescence behaviour to the polarity of the surrounding media, that is of potential interest for use in high-resolution cell imaging. Signore and coworkers [58] report that the computed maximum absorption wavelength (λ_{abs}) of the ground to excited state (GS→ EES) transition seems to depend on the dihedral δ_1 (green arrow), that drives the rotation of the coumarin moiety with respect of the benzo-thiazolic group. In fact λ_{abs} was computed for both a planar and a twisted (δ_1 = 20-30°) conformation, yielding 470 and 459 nm, respectively. Interestingly, it is the twisted geometry that best agrees with the experimental measure (455 nm) [58], although DFT calculations indicate that the minimum energy conformer is planar. It can be therefore speculated that the absorption (and emission) spectra arise from a statistical mixture of several conformers around δ_1=0° with remarkable contributions of non planar conformations. To face this problem one could resort to MD simulations and extract a number of geometries from the trajectory, to form a statistical sample.

A simulation has not been done for this dye in the presented work, although some insights are obtained from the developed FFs for the GS and EES.

The QM energy, gradient and Hessian matrix were computed for the absolute minimum conformation of both GS and EES. The torsional energy profiles were computed separately in both states for dihedrals δ_1 and δ_2 (the same torsional profile is transferred to both the "dark arrow" dihedrals of fig. 4.1(h)) The coupling between δ_2 and δ_3 was accounted through a non-bonded intramolecular interaction between the Carbon atoms of the methoxy group. This guarantees that unphysical overlaps of the two methyl groups do not occur in the simulation. No other intramolecular non-bonded interaction was included. All aromatic dihedrals were described through the harmonic form, "soft" δ_1, δ_2, and δ_3 dihedrals were described through periodic potentials with five (n = 0-4) terms. Finally one single cosine (n=3) was employed to describe the rotation of the methyl groups in the methoxy substituents. The final overall standard deviations were 0.06 kJ/mol and 0.16 kJ/mol for GS and EES, respectively.

The energy profile, excitation wavelengths and oscillator strengths obtained for several values of the two relevant dihedral angles are displayed in Figure 4.7. Both QM and MM calculations are reported for comparison. In all QM calculations the geometries are fully optimised except for δ_1 and δ_2, which were taken fixed, according to the Relaxed Energy Scan method applied separately to GS and EES. All these conformers with assigned δ_1 (δ_2), were re-optimised through the FF. The resulting MM optimised geometries are then used to compute the vertical excitation energy which corresponds to absorption and emission, depending on the FF used for geometry optimisation.

The minimum energy conformation is planar for both electronic states: δ_1 = 180° (*i.e.* with the two nitrogen atoms on the same side with respect to the coumarin - benzothiazol linkage) and δ_2 = = 0° (*i.e.* with the two methoxy groups pointing in opposite directions). The MM energy profile is almost coincident with the QM data for both electronic states. The behaviour of the GS and EES energies vs. δ_1 is in line with the already

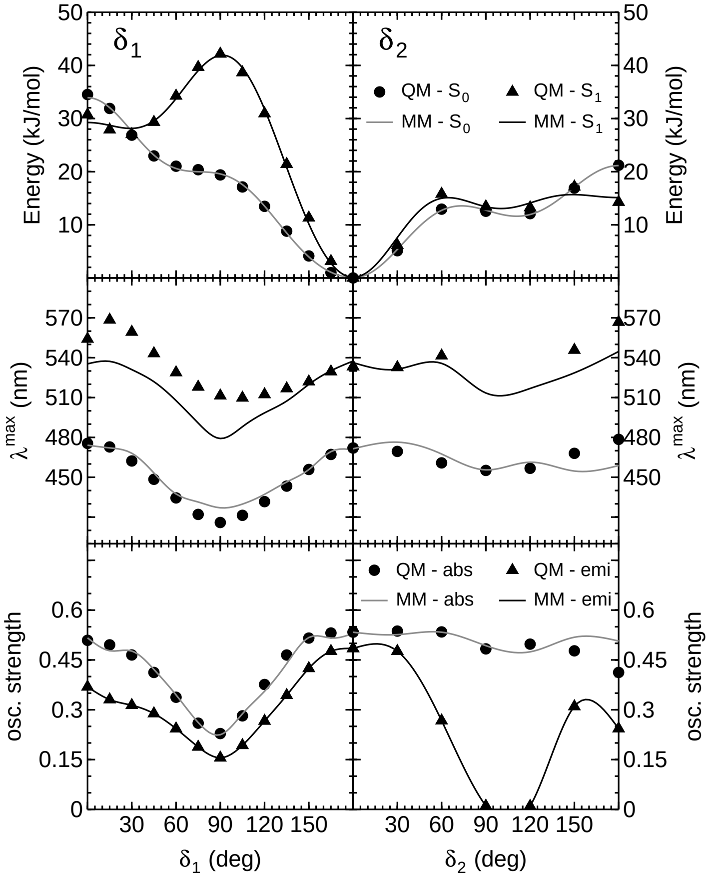


Figure 4.7: Quantities relevant for the absorption and emission spectra of the coumarin dye vs. torsional angles. In all panels QM data are reported with black symbols, whereas MM computed energy are reported with solid lines. Upper panels: torsional energy profiles of GS and EES. Middle panels: QM computed absorption (circles and grey lines) and emission (triangles and black lines) wavelengths at QM and MM optimised geometries. Lower panels: QM computed oscillator strengths for absorption (circles and grey lines) and emission (triangles and black lines) at QM and MM optimised geometries.

mentioned blue shift observed with the increase of δ_1 [58]. In fact the transition energy tends to increase for deviations from the equilibrium geometry and, consequently, the wavelength decreases (left middle panel).

This shift is estimated in 5 − 10 nm, in agreement with literature data, while the effect of δ_2 seems to be less relevant.

In the middle panels the high accuracy of MM λ around the equilibrium geometry of both states is apparent, with moderate discrepancies for distorted conformations which, according to the energy profiles, should be little populated at room temperatures. Finally, however, by looking at the bottom panels it appears that both absorption and emission oscillator strengths well agree with those computed for the QM optimised geometries. This means that the such geometrical discrepancies involve internal coordinated not relevant for the HOMO - LUMO transition which characterise the S_1 state.

In conclusion, the obtained FFS, that have been exploited for geometry optimisation of both the considered electronic states, possess the accuracy requested for absorption/emission spectra calculations, and some small discrepancies are uninfluential on the intensity and position of the bands.

Chapter 5

Statistical approach case study: acrolein

The acrolein molecule, fig. 4.1(j), is a small planar aldehyde. The Oxygen atom of the carbonyl termination exposes two lone pairs, and the resulting hydrogen bond structure, that occurs in water solution, is expected to tune the molecule's excitation bands [59]. Secondarily, the planarity of the molecule should be discussed as deviation from $\delta = 0$ geometries might affect the optical properties.

By means of accurate MD simulations, both the questions about lone pairs and molecular flexibility have been addressed, and the optical features on top of the resulting models have been evaluated.

5.1 Lone pairs treatment

At a first stage, a totally rigid description of the molecule was employed, assuming a high relative stability of the *trans* conformer.

Two MD simulations were performed for 4 ns in a box of TIP3P water molecules with length 4.3 nm. Configurations were saved every 0.25 ps. More details are in section 2.2.1. The Lennard-Jones parameters for acrolein are taken from OPLS, for the atomic charges, a QM reference calculation at the optimised geometry is used, employing the B3LYP/6-311++G** DFT level.

In the first simulation picture (see fig. 5.1), the oxygen atom is treated as an isotropic

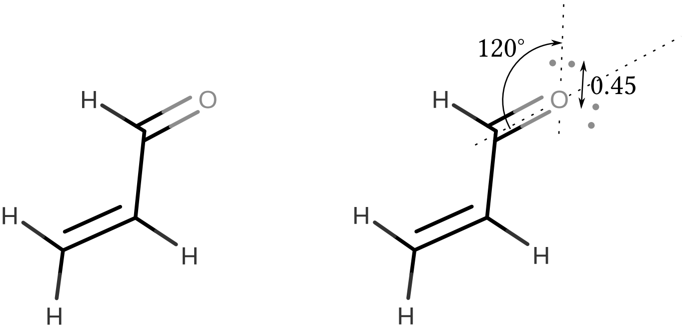


Figure 5.1: Two rigid MM models for the acrolein molecule

site, while in the second picture the two lone pairs are explicitly included in the simulation as two virtual sites without mass located in the molecular plane at a distance of 0.45 Å from the oxygen site. They formed an angle of ± 120° with the C = O bond direction. The total CM5 charge of the oxygen is split between the two virtual sites, while the nucleus retains the LJ potential.

The water distribution around the molecule, specifically the carbonyl group, was studied for the two simulation pictures. In Fig. 5.2 cross sections on the x-y plane, at z = 0, of the spatial distribution function (SDF) are shown for the water oxygen and hydrogen atoms around the acrolein molecule, with the two different descriptions for the carbonyl oxygen.

It can be noticed that water approaches acrolein at a shorter distance from the O site than from the other portions of the molecule, with a depletion area $\simeq$ 1.6 Å wide around the oxygen, and $\simeq$ 1.8 Å around the remaining part, for both systems. The substantial difference between the two models is essentially the shape of the first coordination shell: around the isotropic O site, both hydrogen and oxygen atoms form an hemispherical shell. On the other hand, when the lone pairs are present, hydrogen and oxygen atoms mark denser spots in their direction. In particular water molecules are distributed within two cones around the Oxygen - lone pair axes, being the hydrogen atoms constrained in a narrower volume than the oxygen atoms.

The 3D shape of those coordination shells can also be observed in Fig. 5.3, where selected isosurfaces are plotted around the acrolein molecule for the computed distribution functions. Two clearly separated sites are present around the oxygen - lone pairs directions in the first image, corresponding to the most intense areas in Fig. 5.2(c-d), while only a compact, hemispherical cap is placed over the oxygen in the case of

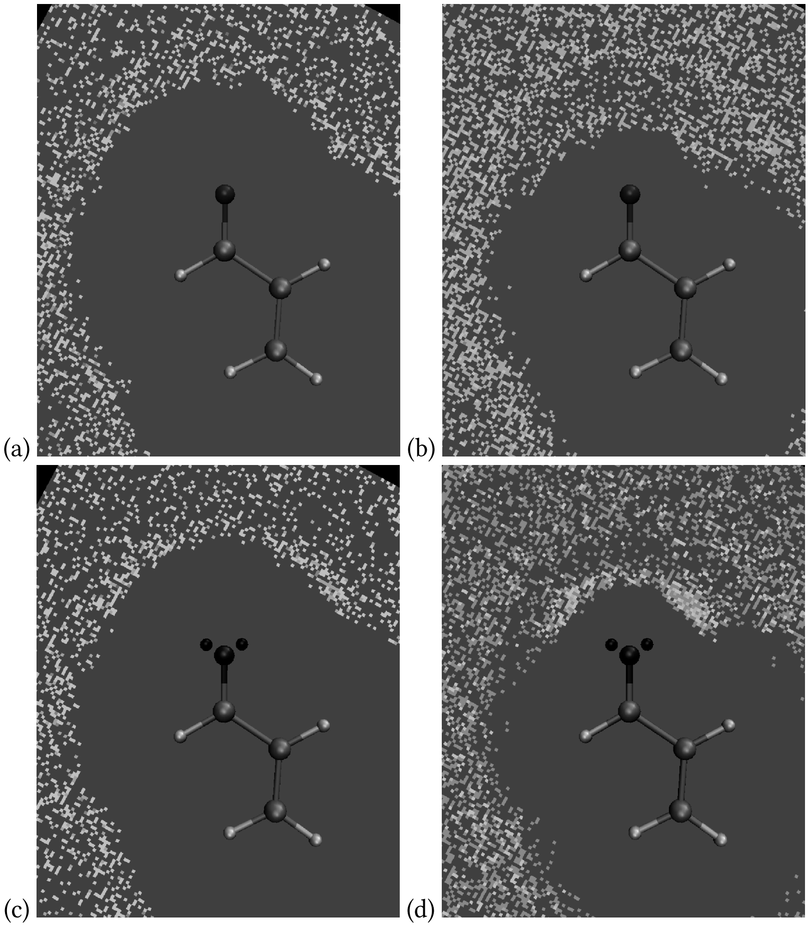


Figure 5.2: x-y cross sections of the normalised SDFs. (a) isotropic O, O··OW distribution; (b) isotropic O, O··HW distribution; (c) lone pairs on O, O··OW distribution; (d) lone pairs on O, O··HW distribution; brightness increases from dark grey (zero density) to white (maximum density).

isotropic treatment.

The differences between the two solvation structures can also be seen from the radial distribution function $g(r)$ of the water hydrogens around the acrolein oxygen, which

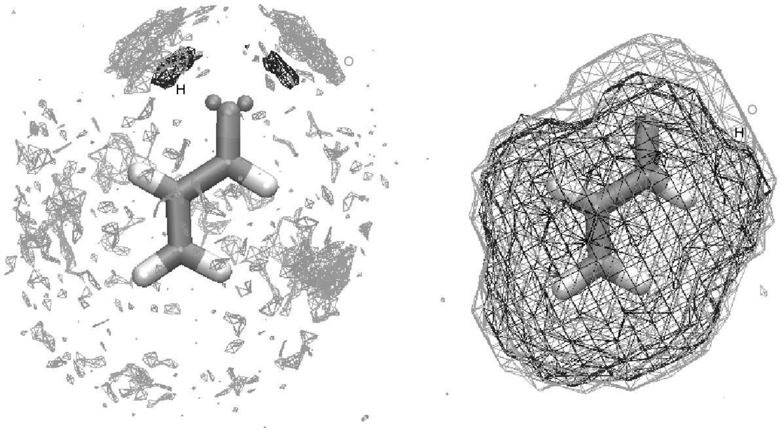


Figure 5.3: Selected isosurfaces of the hydrogen (black) and oxygen (grey) SDFs around the acrolein molecule.

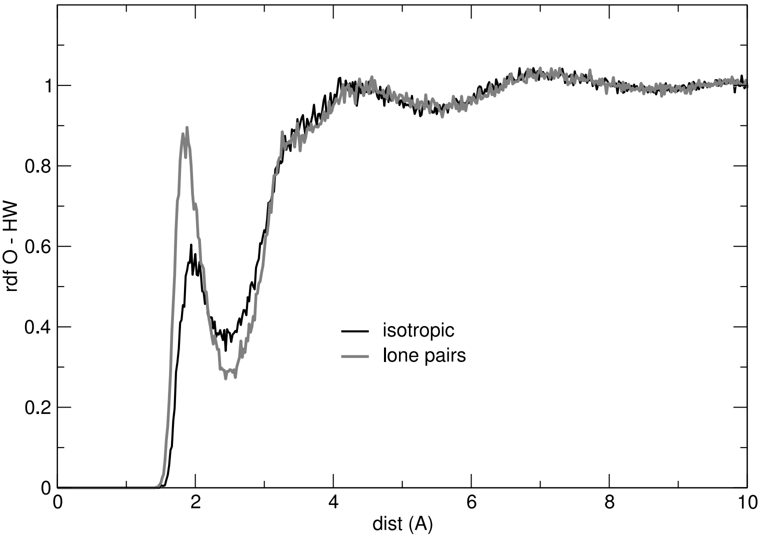


Figure 5.4: $g(r)$ for the O - HW pairs in the two simulation pictures.

are shown in Fig. 5.4: more solute atoms reach the first shell when explicit lone pairs are included. The first coordination numbers corresponding to the HW configurations are 0.35 for isotropic and 0.48 for lone pairs simulations.

5.1.1 Transition lineshapes

Following the route explained in section 2.2.2, 100 snapshot were extracted from the two MD simulations. On each snapshot a multi-layered TDDFT calculation (QM/FQ/PCM) was performed to obtain the main electronic transition energies and oscillator strengths. The average number of water molecules within the FQ sphere of radius 15 and its standard deviation are 336 ± 6 for both the isotropic and lone pair pictures.

The spectral lines arising from the TDDFT calculation were broadened with Gaussian functions in the energy domain, with HWHM of 0.02 eV, then they were averaged over all the snapshots to give the complete spectral shape.

Two transitions were considered: the former is the transition placed around 3.6 eV, identified as a HOMO - LUMO excitation with a $n \to \pi^*$ character; the latter is a $\pi \to \pi^*$ transition involving the HOMO-1 - LUMO molecular orbitals, with energy around 6.4 eV.

The broadened spectra, plotted in the wavelength domain, are shown in Fig. 5.5, and their peak wavelengths are compared to the ones obtained with a calculation *in vacuo*.

A proper description of the oxygen lone pairs seems to affect the UV-VIS spectral shape and position. As it is particularly clear in the $n \to \pi^*$ transition, the arrangement of water molecules due to the inclusion of lone pairs is shown to blueshift the main peaks with respect to the isotropic case, and to enhance the underlying features of the lineshape. Comparing this spectra with the one obtained by Brancato *et al.* in [59], the presence of several peaks in the $n \to \pi^*$ UV-VIS line is coherent with the results from the QM/MM dynamics reported therein.

5.2 Molecular flexibility

An intramolecular FF was developed for acrolein as well as for all the other molecules presented in the previous chapter, using B3LYP/6-311++G** calculations as QM benchmarks. At variance with the previous calculations, the QM benchmarks were here prepared both with *in vacuo* calculation and using a PCM solvation scheme. These choices

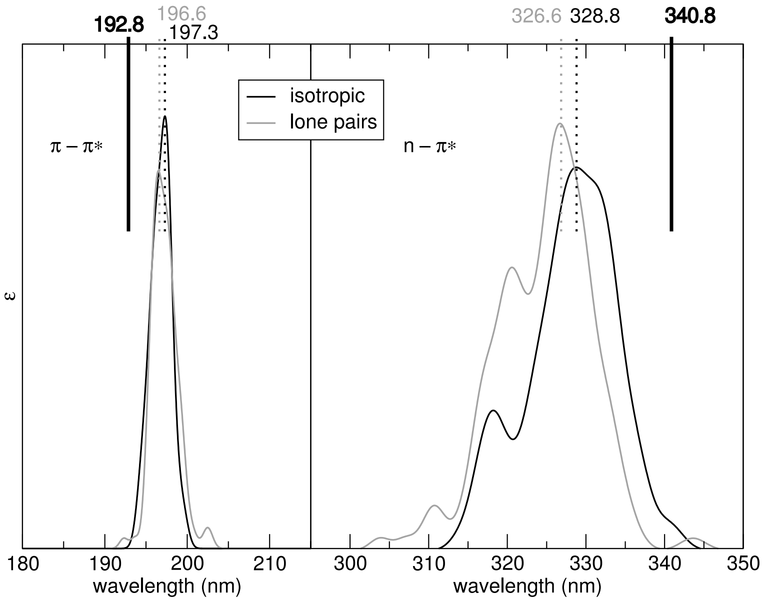


Figure 5.5: UV-VIS spectra for acrolein in water solution, considering the two different rigid MM descriptions. The position of *in vacuo* peak wavelengths calculated at the same level of theory are represented by the tick lines. The $n \rightarrow \pi^*$ transition is 5000 times magnified with respect to the $\pi \rightarrow \pi^*$.

lead to the determination of two different FFS: *vacff* and *pcmff*, respectively. The whole FFS are reported in the Appendix.

The harmonic terms of Eq. 2.1(1-3) were employed (harmonic torsions are used for improper dihedrals), while the main dihedral δ (fig. 4.1(j), light arrow) was described by means of a 4 cosine sum, for both the FFS. The minimisation process gave a frequency related standard deviation of 54.9 and 52.2 cm^{-1} for *vacff* and *pcmff*, respectively. The torsional energy profile of the parametrised force field is plotted in fig. 5.6: since the energy profiles are very similar in the *vacff* and *pcmff* cases, the same FF parameters are used for this IC.

It should however noted that the *cis* conformation is much less favoured than the *trans* one (the difference being as high as 10 kJ/mol), thus the torsional energy scan is more important to sample the shape of the main well, rather than quantifying the almost forbidden conformational change.

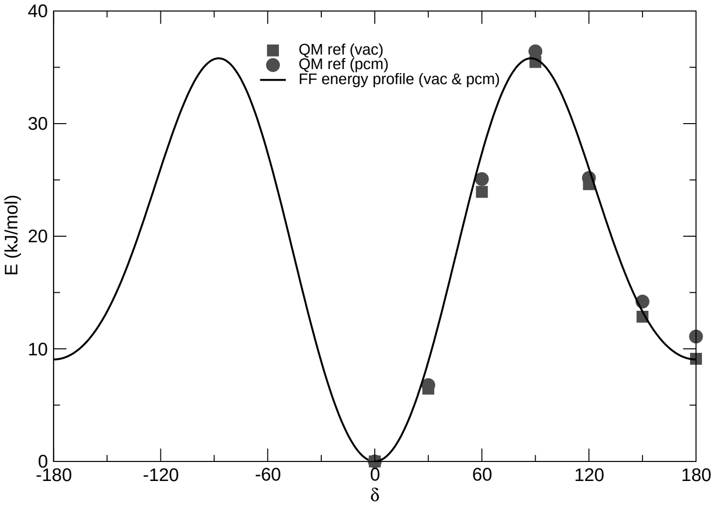


Figure 5.6: Torsional energy profile along δ: QM reference points *in vacuo* (boxes) and in PCM (circles), along with the scan obtained with the FF description.

Finally, the isotropic description for oxygen has been discarded, and lone pairs are employed as shown before.

The two FFs' equilibrium geometries are not negligibly different, as shown in Tab. 5.1 (see the Appendix for the whole FF, which reports the equilibrium values of all the ICs).

In particular, considerable differences concern two ICs in particular: the $C = O$ and the $C - C$ (single bond) length. The averaged PCM effect of the solvent increases the former length and concurrently decreases the latter. Though small, these differences may reveal a slight delocalisation of the π orbital on the carbonyl toward the central region of the molecule.

Three molecular dynamics simulations (*vacff* in gas and solvated phase, *pcmff* in

IC	Opt. *in vacuo*	Opt. in PCM
C = O (Å)	1.211	1.219
C − C (Å)	1.474	1.466
C = C (Å)	1.335	1.338
O = C − C (deg)	124	124
δ (deg)	0	0

Table 5.1: Geometric features of the optimised acrolein geometry

solvated phase) with periodic boundary conditions were performed in the NPT ensemble for 4 ns each.

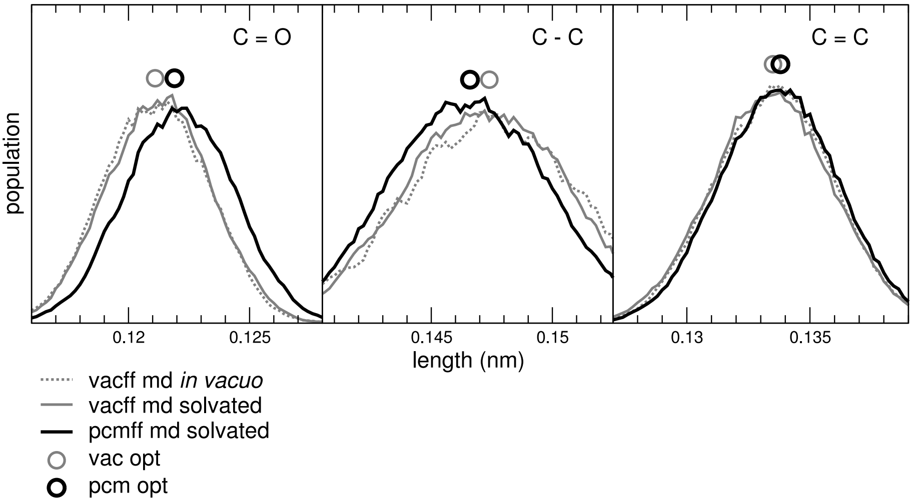


Figure 5.7: Distribution of the main bonded lengths along three MD simulations: vacff in gas and solvated phase, pcmff in solvated phase. The values at the optimised geometries are reported as circles.

It can be seen from figure 5.7 that *vacff* is not capable of describing the changes in the bond lengths discussed above, being the distributions arising from the simulation that employ *vacff* almost identical, regardless of the solvent, and centred in the same reference values. For this reason, only the *pcmff* simulation in solution will be used for snapshot extraction, as it includes a peculiar solvent effect even in the reference geometries.

Finally, although the δ dihedral has been described by means of periodic functions, the energy barrier between the 0° and 180° region is steep and the molecule never explores the twisted conformation, which is as high as 4 times the thermal energy. In fact it remains almost planar, with zero population for $|\delta| > 30^\circ$ (distributions are not shown here).

5.2.1 Statistical approach

Three different methods are compared for the "statistical" calculation of the electronic transitions, each one applied to the same 400 snapshots (spaced by 10 ps) from the *pcmff* simulation in water.

They come with an increasing level of detail, and are referred as (see section 1.1) ME, POL.E, QM*/FQ/PCM, the latter using the ten nearest water molecules within the QM layer. In any case, the CAM-B3LYP functional coupled to the cc-pvDz basis set was used: a discussion about this choice will be presented in section 7.5.1.

The relevant TDDFT transitions are identified by the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ contributions (always between 0.65 and 0.70) in the first two schemes. When the QM layer is extended to 10 water molecules, the MOs are more difficult to identify, so an empiric criterion on the energies and oscillator strengths was used.

Before performing convolution and averaging, the set of computed transitions can be analysed further.

In particular, we looked for a correlation between the value of some ICs and the transition energies and oscillator strengths.

While the C − C distance, the O = C − C angle and δ dihedral do not show remarkable correlation, the value of the C = O and C = C distance seem to be related to the excitation energies. Specifically, C = O correlates to the $n \rightarrow \pi^*$ and the $\pi \rightarrow \pi^*$ energies, C = C only to the latter.

Linear regression of the wavelengths against the ICs values are highlighted in figure 5.8 : it clearly appears that the absorption redshifts when each of the double bond lengths become longer.

Restricting one's attention to the most statistically relevant intervals (within the full-width-half-maximum FWHM of the distribution of the bond lengths) the ranges of energies spanned by the regression curves can be summarised as in Tab. 5.2.

By inspection of the table, it can be pointed out that the thermal effect on the molecular geometry is important in the determination of the transition lines, since within the

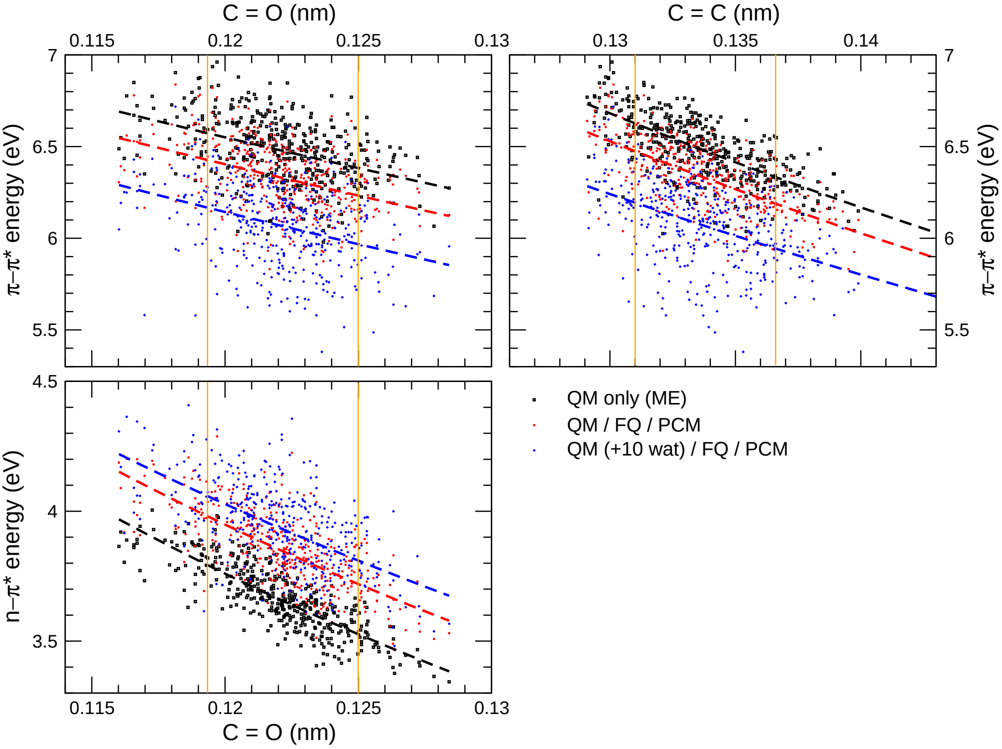


Figure 5.8: Plot of the main correlations found between TD results and IC values as resulted from the MD simulations. Black, red and blue symbols refer to the three different approaches investigated (see legend). Linear regression fit are performed in the wavelength domain, and superimposed to the scatter plots. Orange bars delimit the FWHM of the statistical distributions of each IC's value.

populated region it is responsible for shifts as large as of 0.29 eV and 0.27 eV for the two considered lines.

Moreover, it can be observed once more that the inclusion of the direct electrostatic

Transition	Method	Range (eV) varying C=O	Range (eV) varying C=C
$\pi \to \pi^*$	ME	6.39 to 6.59	6.33 to 6.63
	QM/FQ/PCM	6.23 to 6.42	6.20 to 6.49
	QM10/FQ/PCM	5.96 to 6.17	5.93 to 6.20
$n \to \pi^*$	ME	3.53 to 3.80	
	QM/FQ/PCM	3.72 to 3.97	
	QM10/FQ/PCM	3.81 to 4.07	

Table 5.2: Effect of the double-bond lengths on the transition energy for the two considered excitations

effect of the solvent, in addition to the ME treatment, blueshifts the $n \to \pi^*$ line while redshifting the $\pi \to \pi^*$. This "electrostatic embedding" do not converge when including only FQ and PCM water: on the contrary, it still remarkably increases when a layer of QM water is added, revealing the imperativeness of such inclusion.

The computed statistical lineshapes, based on the three mentioned methods, are reported in Fig. 5.9, top panel.

All the [energy, osc. strength] couples arising from each snapshot were broadened with Gaussian functions and averaged with all the others. The half-width-half-maximum is set to 0.05 eV, larger than 0.02 used for the previous spectra (Fig. 5.5), since the flexible geometry leads to more disperse lines.

5.2.2 Vibronical approach

The line shape and broadening obtained by means of a statistical sampling of the configuration space of acrolein in water were compared to the vibronic contribution to the same excitations.

All the calculations were performed *in vacuo*. The geometry of the molecule, in the three considered electronic states, was optimised with the CAM-B3LYP DFT functional and the CC-PVDZ basis set, while imposing a constraint on the δ dihedral to remain planar. The Hessian matrices were calculated at the same level of theory, though negative frequencies were found in the $\pi \to \pi^*$ state because of the geometry constraints.

The one-photon absorption spectra were evaluated with [60] the Adiabatic Shift (AS-FCHT) and Adiabatic Hessian (AHFCHT) methods (the latter limited to the $n \to \pi^*$ excitation), and then convoluted with Lorentzian functions with half-width-half-maximum 300 cm^{-1} ($\simeq 0.037$ eV).

5.2.3 Computed UV lineshapes

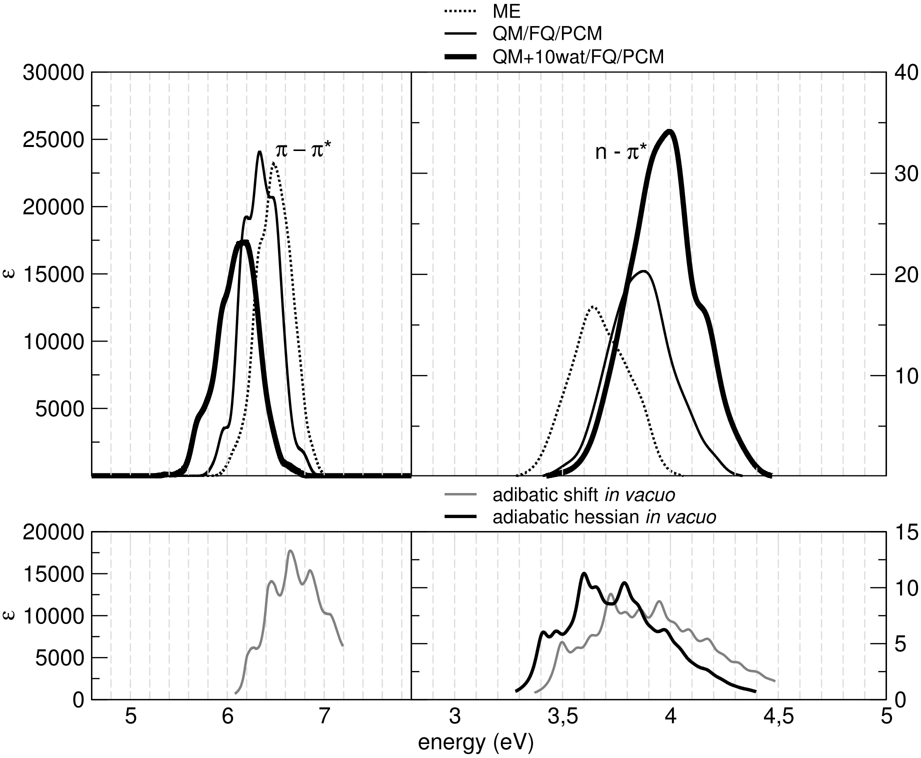


Figure 5.9: UV line shapes of the first two electronic transitions, labelled $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$. The arbitrary units used for the absorption intensity ϵ are the same for all graphs: the lines in the right panel are actually 1000 times less intense than in the left one. Curves of different colours result from different approaches (see legend and text)

The broadened and averaged line shapes are shown in figure 5.9. As regards the spectra based on MD simulations, the three approaches are compared in the top panel. The spectral lines obtained with a flexible description for the acrolein molecule are far more broadened than in the rigid-molecule picture, as a result of the sampling of the molecule's configurations.

From the comparison between the three TD approaches, a specific trend can be observed moving from a coarse ("mechanical" only) to a more precise representation of the embedding. The energy shifts are in the same direction as described in the section above.

In the bottom panel of figure 5.9, the vibrationally resolved electronic spectra are shown for comparison. The overall bandwidths obtained with the inclusion of the vibronic contributions is in line with the MD sampled ones. In particular, when comparing the best vibronic approach (adiabatic shift in the $\pi \rightarrow \pi^*$ case, adiabatic hessian in the $n \rightarrow \pi^*$) with the ME statistical method (in which the solvent effect is retained the less), a good agreement can be observed both in the peak energies and in the asymmetry of the excitation decay. Since the vibronic spectra were computed *in vacuo*, no conclusions can be drawn about solvent effect.

Approach	$\pi \rightarrow \pi^*$ peak (eV)	$n \rightarrow \pi^*$ peak (eV)
Fixed geom; isolated	6.431	3.638
Fixed geom; QM/FQ/PCM	6.306	3.796
Flex geom; ME	6.474	3.642
Flex geom; QM/FQ/PCM	6.332	3.877
Flex geom; QM+10wat/FQ/PCM	6.159	3.993
adiabatic shift *in vacuo*	6.649	3.724
adiabatic hessian *in vacuo*		3.601

Table 5.3: Summary of the transitions computed with different force fields and embedding schemes

To summarise how several models lead to different peak values, all the data were collected in Tab. 5.3. All the TDDFT calculations are performed at the same level of theory. From the table it appears, as regards the MD sampled calculations, that the two calculations with only the inner layer (fixed isolated and flexible ME) lead to peak energies that are quite similar: indeed, the flexibility mostly contribute in broadening the spectra, but not moving their peak positions. The inclusion of two or three water layers in the electronic calculations produce the already discussed shifts, more relevant if the closest solvent molecules are treated at the QM level.

Since the vibronic spectra are calculated *in vacuo*, they can only be compared to the ME scheme, where the solvent is only used to drive the exploration of the configuration space. Indeed, the peak energies are in good agreement, as can be seen from the table, with the maximal discrepancy being below 0.2 eV.

Chapter 6

Accurate molecular dynamics simulations: the case of nicotine

A preliminary study on the Nicotine molecule (Fig. 4.1(g)) was presented in section 4.3, together with the results obtained from a simulation in methanol.

The FF employed therein was based on DFT calculations *in vacuo* and, even though it was solid enough to reproduce a good distribution of the intramolecular features, like the δ_1 dihedral distribution, it was still perfectible as regards the intermolecular treatment.

The new FF presented below makes use of PCM to calculate the reference geometry and charges, and is specifically tailored to the simulation in water solvent. Although this formulation is surely less transferable, it is optimal for the purpose of a statistical calculation of electronic properties. Indeed, the model and the simulations were done to this end; the calculation of the spectra will anyway follow the publication of this manuscript.

Among many choices [61], a recently published model, called TIP3P-FB [40], was used for water. This is a three-site, fixed-geometry model, that has greatly improved the predictivity of TIP3P as regards density, calorimetry and diffusion; moreover, as local interactions are of particular interest in the present work, the experimental $g(r)$ are reproduced quite exactly by the employed water model.

6.1 Nicotine force field

A minimum energy geometry was identified by the DFT method coupled to the implicit PCM treatment of the solvent.

Similarly to what has been shown for calculations *in vacuo*, the most stable conformer (labelled "A") is found with the NCH_3 group of methylpyrrolidine pointing in the opposite direction with respect to the N in pyridine: the dihedral angle δ_1 is 133.1°.

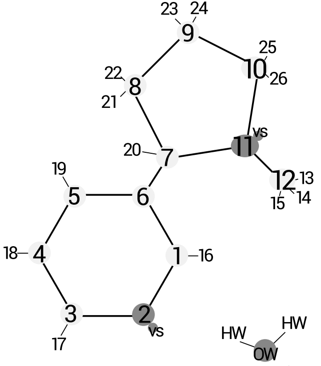
Most of the FF was tailored to this specific configuration: namely the atomic charges and all the harmonic terms of Eq. 2.1, plus the improper periodic dihedrals. These and all the other values constituting the force field are listed in Table 6.1. It should be pointed out that the employed set of charges already reflects the presence of water as a surrounding medium: the charge on N_{pyrid} (*i.e.* N6) and on N_{pyrro} (N5), carried by the virtual sites in the MM description, are both around -0.42, while the CM5 method algorithm on a calculation without PCM would give -0.18 and -0.11 respectively. This force field is thus applicable to simulations in solvents that are as polar as water. A set of 22 twisted geometries, with δ_1 spanning the [-180°, 180°] interval , and their relative energy reveals the presence of at least one other stable conformer near -45° ("B"), whose energy is 1.47 kJ/mol higher than the "A" geometry, which is separated by a 22.5 kJ/mol high energy barrier: see the circles in Fig. 6.1.

This energy scan, together with a similar one regarding the methyl termination, was taken into account in the development of the force field. After the parametrisation, the intramolecular energy (Eq. 2.1), on varying δ_1, is reported as a black solid curve in Fig. 6.1: it correctly describes the DFT/PCM benchmark. It should be pointed out that the contributing terms to the overall torsional profile are basically three: the sum of all the non bonded (Coulomb and LJ) interactions, the energy term associated with δ_1, and the energy term associated to the complementary quadruplet δ_1^c ($V_{\text{tors}}(\delta_1)$ and $V_{\text{tors}}(\delta_1^c)$ in Eq. 2.1). The last two terms were summed, after having imposed that $\delta_1^c \simeq \delta_1 - 60°$, and plotted as a dashed line in the plot. It can be observed that the force field terms inherent to the dihedrals, if taken alone, would stabilise the "B"-like conformation more than the "A" one, whilst the non bonded interactions shift and readjust the relative energies of the two wells, and recreate the reference picture.

6.1.1 Energy and stability of the hydrogen bonds

Having assigned the CM5 charges from DFT/PCM calculations to nicotine and the OPLS LJ parameters, a single variable is left that can be optimised in order to reproduce the

Atom		charge	ϵ (kJ mol^{-1})	σ (Å)	v.s. dist (Å)	v.s. charge
1	C1	0.040	0.276	3.55		
2	N6	0.0	0.711	3.25	0.346	-0.420
3	C1	0.039	0.276	3.55		
4	C2	-0.095	“	“		
5	C3	-0.073	“	“		
6	C2	-0.002	“	“		
7	C4	0.025	“	3.50		
8	C5	-0.155	“	“		
9	C5	-0.159	“	“		
10	C6	-0.058	“	“		
11	N5	0.0	0.711	3.25	0.450	-0.416
12	CT	-0.136	0.276	3.50		
13	HT	0.104	0	0		
14	HT	0.099	“	“		
15	HT	0.084	“	“		
16	H1	0.114	“	“		
17	H1	0.116	“	“		
18	H2	0.116	“	“		
19	H3	0.114	“	“		
20	H4	0.097	0.126	2.50		
21	H5	0.094	“	“		
22	H5	0.091	“	“		
23	H5	0.093	“	“		
24	H5	0.092	“	“		
25	H6	0.086	“	“		
26	H6	0.104	“	“		
	OW	-0.848	0.652	3.18		
	HW	0.424	0	0		



Bond	Å	kJ mol^{-1}Å^{-2}
C1 - N6	1.34	3020.8
C1 - C2	1.40	3011.7
C2 - C3	1.39	3198.1
C2 - C4	1.51	2246.7
C4 - C5	1.55	1832.6
C5 - C5	1.55	1954.7
C5 - C6	1.54	2096.0
C6 - N5	1.47	2343.3
C4 - N5	1.47	2133.4
N5 - CT	1.46	2335.1
CT - HT	1.09	3048.0
C1 - H1	1.09	3283.7
C2 - H2	1.09	3387.4
C3 - H3	1.09	3354.3
C4 - H4	1.11	2753.1
C5 - H5	1.09	3156.5
C6 - H6	1.10	2928.3

Angle	°	kJ mol^{-1}rad^{-2}
C1 - N6 - C1	117.44	603.2
C1 - C2 - C3	118.00	531.5
C1 - C2 - C4	120.79	485.1
C1 - C2 - H2	120.14	329.5
N6 - C1 - C2	124.00	877.2

Angle (cont.)	°	kJ mol^{-1}rad^{-2}
C2 - C3 - C2	119.32	666.7
C2 - C4 - C5	114.21	546.7
C2 - C4 - N5	113.28	356.7
C2 - C1 - H1	120.00	287.4
C2 - C3 - H3	120.22	323.6
C2 - C4 - H4	107.71	420.8
C3 - C2 - C4	122.12	614.0
C4 - C5 - C5	104.88	617.4
C4 - N5 - CT	114.39	774.8
C4 - N5 - C6	104.96	797.0
C4 - C5 - H5	110.4	409.8
C5 - C5 - C6	104.27	714.9
C5 - C4 - N5	103.26	540.1
C5 - C6 - N5	104.26	586.9
C5 - C4 - H4	108.42	375.0
C5 - C6 - H6	112.00	372.8
C5 - C5 - H5	111.8	392.8
C6 - N5 - CT	113.39	696.7
C6 - C5 - H5	111.00	402.8
N5 - CT - HT	110.3	448.1
N5 - C6 - H6	110.55	451.4
N5 - C4 - H4	109.82	458.1
N6 - C1 - H1	116.27	408.5
HT - CT - HT	108.12	326.3
H5 - C5 - H5	107.42	319.1
H6 - C6 - H6	107.53	329.7

Dihedral	Functional
Improper	Harmonic in pyrid.; cos(3θ) in pyrrol.
C4 - N5 - CT - HT	7.097 [kJ/mol] $(1 + \cos(3\theta[\text{rad}]))$
C1 - C2 - C4 - N5	$\sum_{n=1}^{4} k_n(1 + \cos(n\delta - \phi_n))$: see Fig. 6.1

Nb	Functional
1-4	off
1-5+	LJ and q-q

Water parameter	settle value
OW - HW (Å)	1.0118
HW - OW - HW (°)	108.15

Table 6.1: FF specifications, following the atom numbers and names in inset figure.

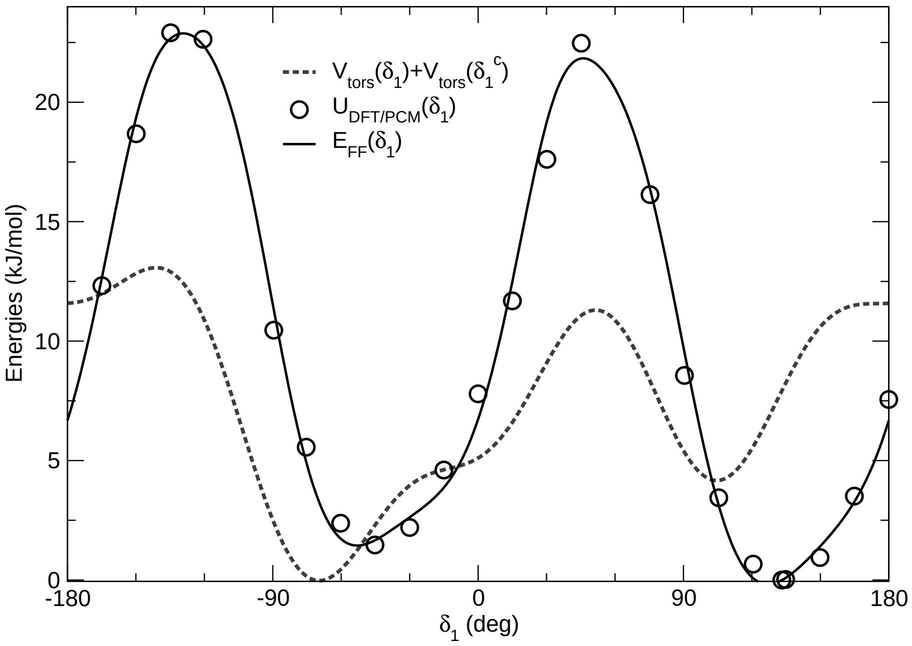


Figure 6.1: Dashed grey line: sum of the two torsional potential energy terms, concerning the δ_1 internal coordinate and its pair quadruplet δ_1^c, in the fitted FF; solid black line: total FF energy along the torsion. DFT/PCM level scan shown for comparison (circles)

hydrogen bond features on the N6 atom: *i.e.* d_{N-vs}, the distance between the nucleus and the virtual site (vs) representing its lone pairs.

Using *ab initio* calculations, the capability of making such non covalent interactions has been assessed, as summarised in Table 6.2. Geometry optimisation *in vacuo* and in PCM has identified stable Nicotine - water ("Nw") structures in a interacting configuration (hydrogen bond). The energy of formation of this compound is calculated by considering the reaction $\{w_2 + N \rightarrow w + Nw\}$ where w_2 represents the water dimer, and the hydrogen bonded configuration is found energetically favoured by more than 8 kJ/mol. On the other hand, the choice to include PCM is relevant to the establishment and stability of the "Nww" compound, with 2 water molecules hydrogen-bonded to the same nitrogen atom of nicotine. Indeed, the existence of such a structure is apparent only with PCM, whilst the calculation *in vacuo* makes the second water molecule mi-

System	DFT	DFT/PCM	MM ($d_{\mathrm{N-vs}}$ = 0.346Å)
Nicotine + 1 water			
structure	h.bond	h.bond	h.bond
$E(w, Nw) - E(w_2, N)$	-8.92 kJ/mol	-8.39 kJ/mol	-8.40 kJ/mol
Nicotine + 2 water			
structure	2$^{\mathrm{nd}}$ shell	2 h.bonds	2$^{\mathrm{nd}}$ shell
$E(w, Nww) - E(w_2, Nw)$	/	+8.99 kJ/mol	/

Table 6.2: Structure of the two composite systems resulted from three methods: "h.bond" refers to one or two non covalent bondings of the water hydrogen to the pyridine nitrogen; "2$^{\mathrm{nd}}$ shell" means that the second water molecule interacts with the first but not with nicotine. Energies of formation are reported for the h. bonded structures.

grate to the second shell. It is believed that the PCM method of adjusting the solvent cavity during an energy optimisation is biased toward more compact minima, and that would explain the prediction of 2 hydrogen bonds; as things stand, the absence of a second hydrogen bond, predicted by the DFT alone, seems more trustworthy. Still the formation of the second hydrogen bond is unfavoured even with the PCM treatment, as the energy of formation is around +9 kJ/mol.

The structure of the system composed of a nicotine molecule and one water molecule was also minimised at the MM level. The nicotine FF and the TIP3P-FB water were employed, while trying different values for $d_{N_{\mathrm{pyrid}}-vs}$.

By means of a binary search, a suitable value for d_{N-vs} was obtained, such that the system forms an hydrogen bond, and its formation energy is within 0.01 kJ/mol from the benchmark DFT/PCM value. Moreover, the force field exploiting this distance (0.346 Å, see Table 6.2), assures that a second hydrogen bond won't stabilise. In fact, the minimum energy configuration for a second water molecule is found in the second shell, in agreement with the DFT calculations.

6.2 Molecular dynamics simulations

6.2.1 Configurations of the main dihedral

Two simulations with different initial conformations were performed in the NPT ensemble. The simulations, having a total production time of 150 ns and 126 ns respectively, were preceded by several ns of equilibration. The value taken by the main dihedral during the simulations are reported in Fig. 6.2 (top). It can be observed that most of the time the value lays in either the [90° - 180°] or in the [-90° - 0°] range, with negligible transients. The two-well structure of Fig. 6.1 is thus reproduced in solution.

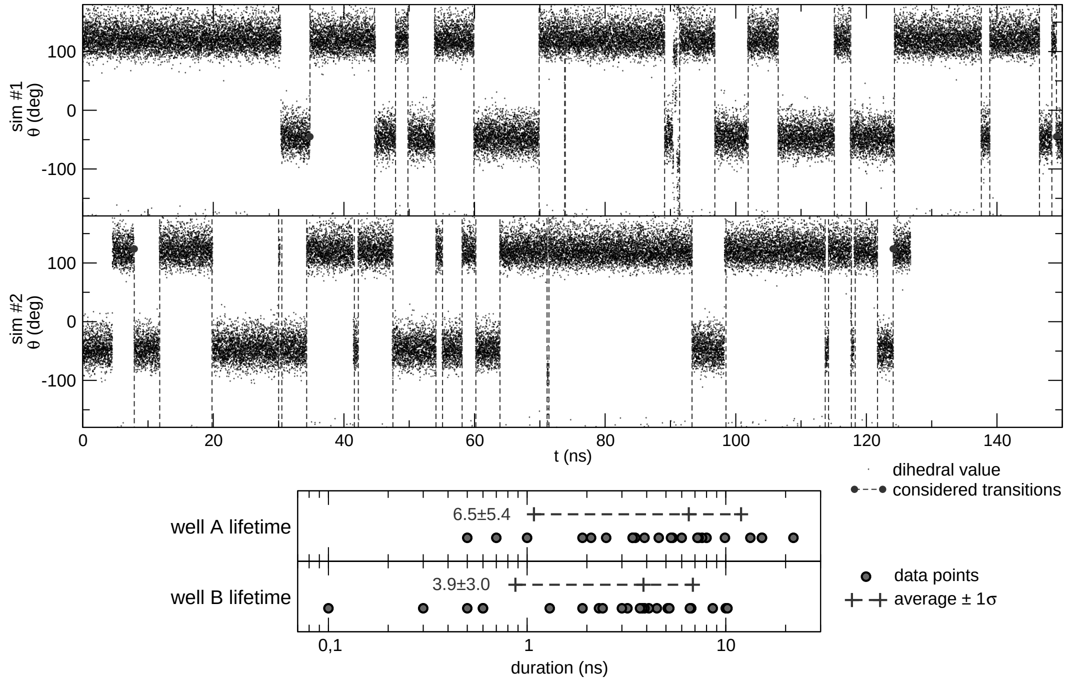


Figure 6.2: Top: evolution of the δ_1 dihedral throughout the two production simulations: a black dot is placed every 3.75 ps. Conformational changes are marked with vertical dashed lines. Bottom: distribution of the two wells lifetimes, their average and standard deviation.

During 276 ns of simulation time, 24 A→B and 24 B→A transitions can be observed. The distribution of the lifetimes of either the two configurations is shown on Fig. 6.2 (bottom); the average timelifes are indicated, plus or minus one standard deviation. It appears that the main well remains populated on average for 6.5 ns, while the secondary well is populated for 3.9 ns. The distributions are quite broad, so that several residence times for well A exceed 10 ns, and well B experiences population times as short as tenths of a ns.

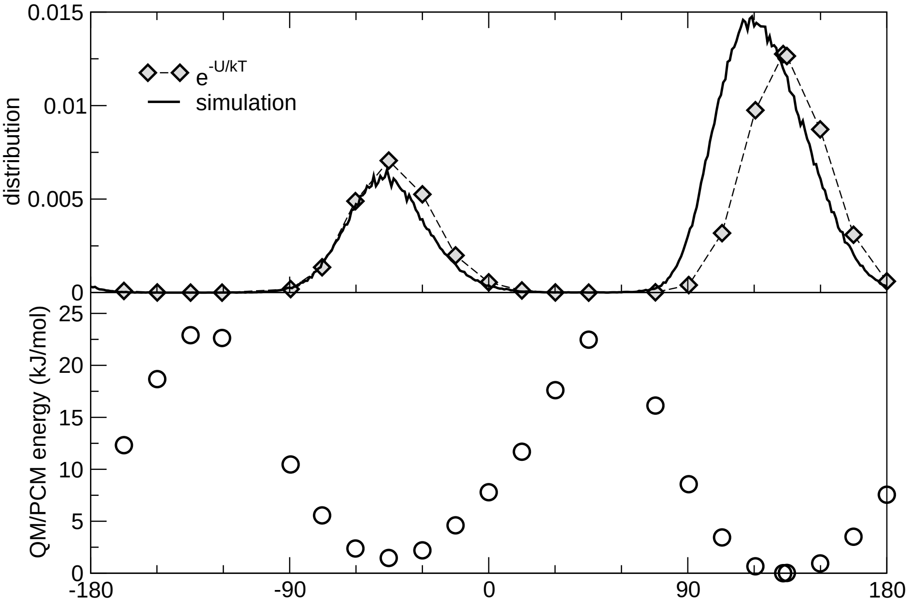


Figure 6.3: Top panel, solid line: distribution of the δ_1 dihedral from the simulation of nicotine in TIP3P-FB water. For comparison: the Boltzmann inverse $e^{-U/kT}$ of the DFT/PCM energy associated with the torsion (diamonds and dashed line). The DFT/PCM energy itself is reported in the bottom panel: it's the same as the circles in Fig. 6.1.

As regards the resulting distribution, it is displayed in Fig. 6.3, top panel, solid line. It shows two sharp peaks at 120° and -46°, the former being twice as high as the latter. The dataset reported with a dashed line alongside the distribution is obtained by evaluating $e^{-U(\delta_1)/kT}$, where U is the DFT/PCM energy profile (circles from the bottom panel) already used for the parametrisation. It can be observed that the peaks' relative heights and their width are almost the same in the DFT/PCM and in the solvated MD picture. The position of the main peak is shifted from the reference calculation by -12°, and is slightly narrower, which is justifiable as a consequence of an explicit treatment of the solvent, whose spatial configurations may result in a slightly different equilibrium value of the dihedral.

6.2.2 Hydrogen bonds

The number of hydrogen bonds that form between water and the pyridine nitrogen atom is analysed on a 500 ps long portion (2000 snapshots separated by 0.25 ps) of the whole MD simulations, by means of the definition given in Eq. 2.14. The time interval is abundantly larger than the hydrogen bonding dynamics, and it has been checked in advance that the conformational changes, that occur on longer timescales, are not correlated to the number of bonds (a visual clue of that is given by Fig. 6.7; see later).

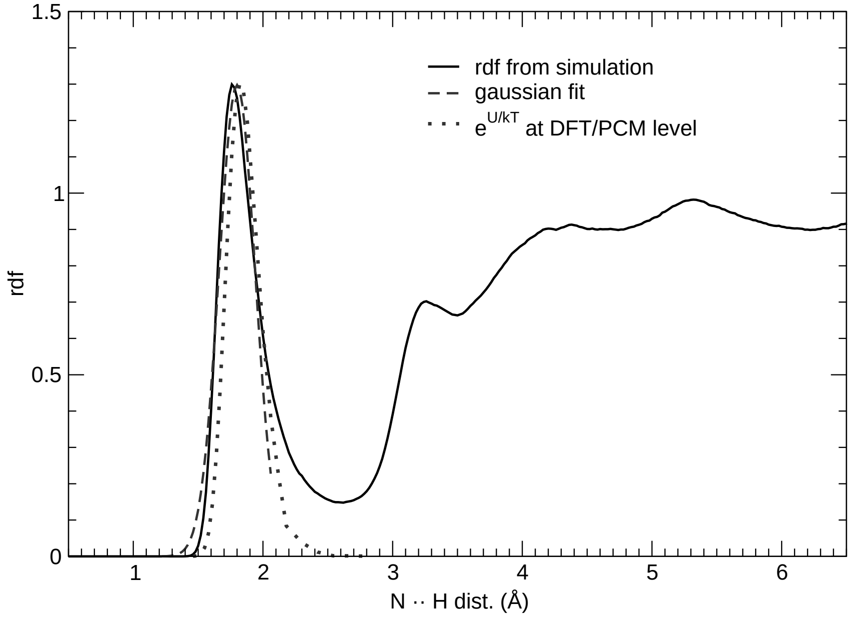


Figure 6.4: N · · H radial distribution function (black solid) and Gaussian fit of its first peak (dashed). For comparison, the Boltzmann inverse $e^{-U/kT}$ of the DFT/PCM energy U for the Nicotine+1water system at increasing distances is reported with a dotted line (arbitrary units).

As regards the nitrogen atom in the pyridine moiety, the N-H radial distribution function is shown in Fig. 6.4. Its first peak has been fitted with a Gaussian function (dashed line) to obtain the values of r_e, r_{hw} to be used in Eq. 2.14; these values are found to be 1.800 Å and 0.139 Å respectively.

In the same figure, it is reported (dotted line) the expected population profile of the hydrogen-bonded water molecules in first shell, as obtained from a scan of the DFT/PCM energy $U(r)$ along the $r = \mathrm{N} \cdot \cdot \mathrm{H}$ distance, and reported by means of a Boltzmann inversion $P(r) \propto e^{-\frac{U(r)}{kT}}$. The peak position and the width of the curve are found in very good agreement with the results from the simulation (black solid curve).

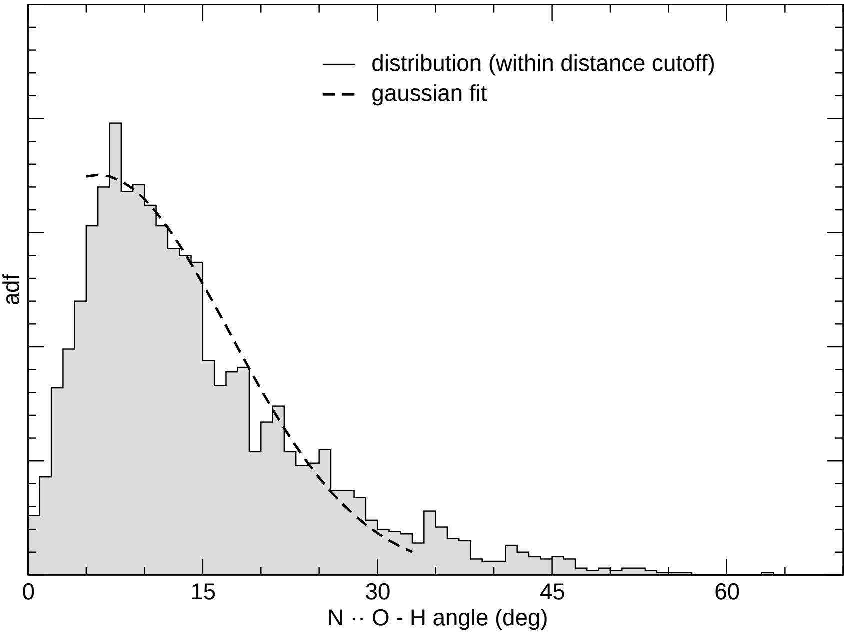


Figure 6.5: N ·· O — H angular distribution function (histogram) limited to the water - acceptor pairs below the geometric thresholds; Gaussian fit of the descending slope (dashed line).

In order to perform an hydrogen bond analysis with Eq. 2.14, the $\theta_e, \theta_{\mathrm{hw}}$ parameters are obtained by fitting a Gaussian slope to the angular distribution (Fig. 6.5) of the nitrogen - water pairs roughly belonging to the first shell, as defined by the classical discrete distance criteria. The maximum and half width are found to be 6° and 11°.

Having introduced all the parameters required, the hydrogen bonds have been counted for every frame: it should be noted that the number of hydrogen bonds for each snapshot can take non-integer values, as a fractional score is assigned to every nitrogen - water pair. An excerpt from the trajectory is shown in Fig. 6.6, left panel: in black, the number of pairs lying beneath the thresholds; in red, the total score assigned with the $f(r,\theta)$ functional. Its value oscillates much and is usually lower than the "discrete" one (apparently the usual parameters are too broad); anyway, when the classical method recognises a very stable single hydrogen bond, like in the interval between 63 and 69 ps, also the continuous definition agrees with a solid "1" scored most of the time. To explain why the "continuous" count is often lower than the traditional one, one can refer to the picture at the bottom of Fig. 6.6: here two water molecules fall below the thresholds, so that $n_{\text{hb}}^{\text{discr}}$=2; nevertheless, they'd be assigned a low score by the $f(r,\theta)$ functional, because one of them is highly misaligned ($\theta = 28°$), and the other one is at a longer distance than the N ·· H peak value. In this case, a fractional value $n_{\text{hb}}^{\text{cont}}$ between 0 and 2, rather than only 2, gives more information about the "quality" of the hydrogen bonds, that are likely to be in a unstable transient.

By looking at the histograms on the right panel of Fig. 6.6, thanks to the parcelization of the functional-based definition (maroon curve), a fine outline is revealed: the number of hydrogen bonds $n_{\text{hb}}^{\text{cont}}$ can take the integer values of 1, 0 or 2 with noticeable peaks of decreasing intensities; for the rest, transients exist during which a water molecule is either settling in or leaving the first shell, with ambiguous orientations that result in fractional scores. On the contrary, the $n_{\text{hb}}^{\text{discr}}$ definition overestimates the average number of hydrogen bonds, as explained before.

6.3 Snapshots

The number of snapshot extracted from the two trajectories, separated by 300 ps of simulation time, is 934. In Fig. 6.7 the configuration space made by the two investigated variables, n_{hb} and δ_1, is reported. It is apparent that two most abundant configurations are present, both having around 1 hydrogen bond, and falling in the middle of the A and B well defined by the intramolecular torsion. However, the inclusion of a large set of snapshots allows to sample also the nearby regions where either the variables can change significantly, assigning to them the correct statistical weight. As a matter of fact, n_{hb} and δ_1 show no relevant correlation as a result of the MD simulation, although they can play a cooperative role when it comes to the optical properties.

Since the water structure, used throughout this chapter, doesn't have a polarizable companion yet, effort is being put in our group into the parametrisation of a suitable one. TDDFT calculations for the simulation of UV-vis and ECD spectra will be performed as soon as a suitable polarizable model is ready for production.

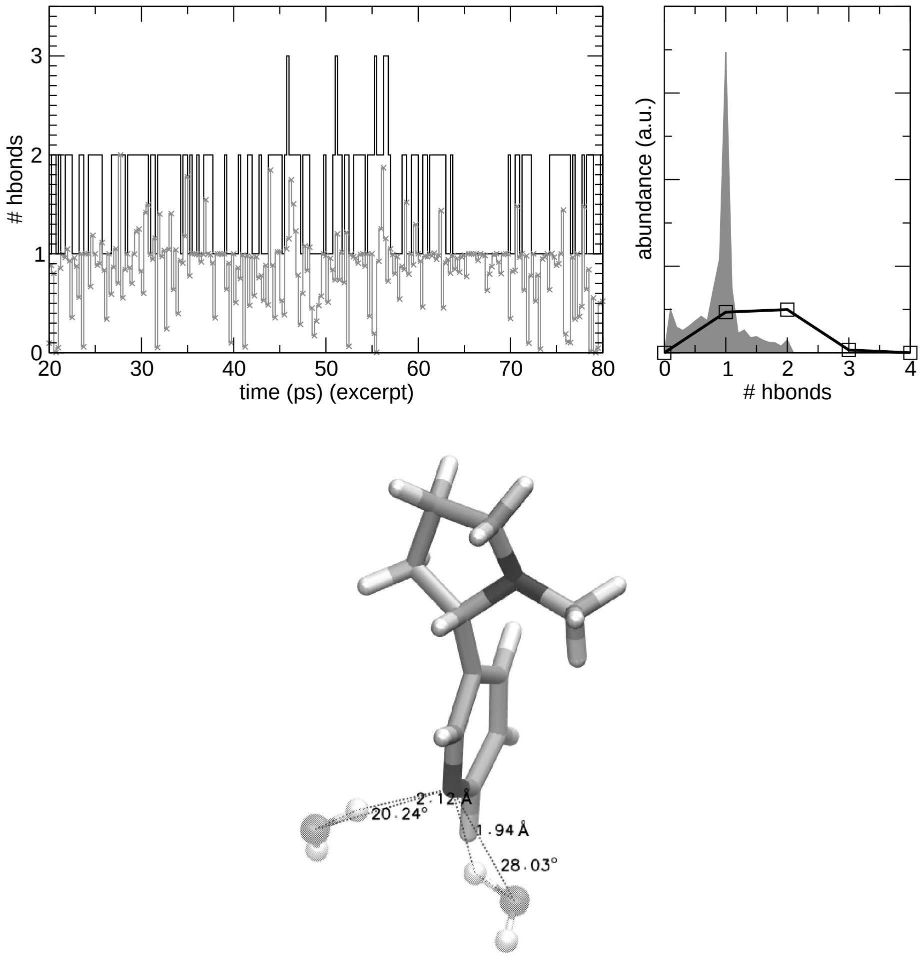


Figure 6.6: Top, left panel: number of hydrogen bonds over time, counted with the geometric threshold criterion (black line) and with the continuous functional of Eq. 2.14 (grey line with crosses). Right panel: normalised histograms counting the distribution of the number of hydrogen bonds: in the continuous case (grey, filled) it is possible to include fractional values. Bottom: an example scenario. r, θ are displayed for two possible hydrogen bonded pairs.

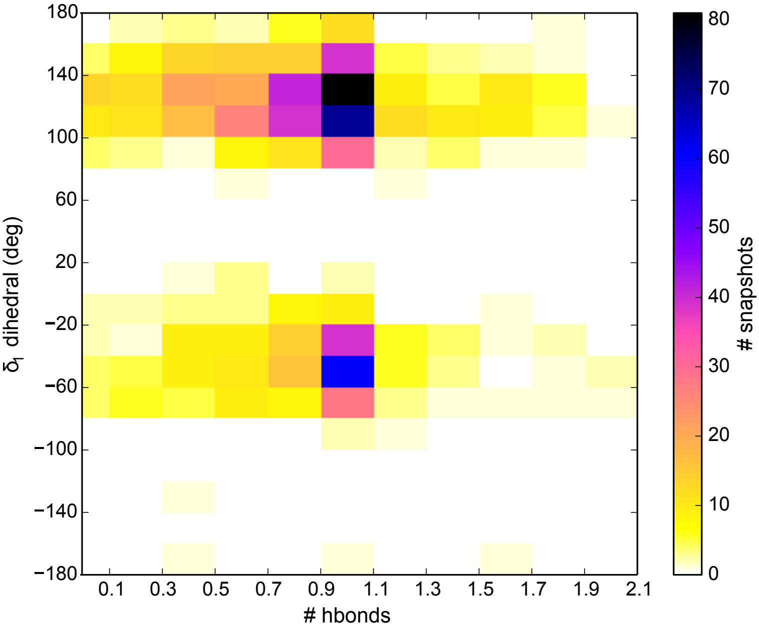


Figure 6.7: Distribution of the 934 configurations that will be used for the property calculation. The number of snapshots belonging to given ranges of n_{hb} (defined by Eq. 2.14) and δ_1 is colour coded.

Chapter 7

Statistical approach case study: NfO-TEMPO in solution and in polymer matrix

In the present chapter, the "statistical" presented in section 1.1 is applied, from beginning to end, to compute the electronic spectra of 4-naphthoyloxy-1-methoxy- 2,2,6,6- tetramethylpiperidine (NfO-TEMPO-Me) molecule, depicted in Fig. 4.1(i). This compound has been selected because of its potential application in the design of doped polymers with fluorescent attributes, as proven by the experimental characterisation obtained through spectroscopic techniques and reported in recent papers [62, 63, 64].

More specifically, the TEMPO unit, composed by the substituted piperidine moiety, has been recently used, in combination with functionalizing groups, for grafting the fluorescent probe to polyethylene derivatives via post-reactor modification [62, 63]. The NfO-TEMPO derivative, substituted with a methyl group at the grafting site, has been studied from both experimental and theoretical points of view in toluene solution [63, 64]. The theoretical analysis [64] depicted a complex conformational landscape, containing six stable conformers, characterised by different values of the three flexible dihedral angles describing the relative orientation of the naphthyl moiety and the TEMPO ring system.

Preliminary TDDFT studies have been carried out with a vertical excitation approach, *i.e.* a single point calculation, which succeeded in reproducing absorption and emission lines (whose experimental peak wavelengths are 299 and 361 nm, respectively) but lacked the detail of the spectral lineshape. This studies have identified the naphtoic moiety as the location of the HOMO-LUMO molecular orbitals involved in the fluorescence transition.

The results are shown below for the statistical approach applied to two systems: the dye in toluene solution and the dye covalently bonded to a polyethylene (PE) matrix. Since UV-vis absorption and fluorescence from the first EES are tackled, two FFS have been developed for the two involved electronic states, and have been exploited for simulations in both the considered environments.

7.1 Force field parametrisation

The atom types chosen to model the NfO-TEMPO molecule, for both GS and EES, are reported in Figure 7.1.

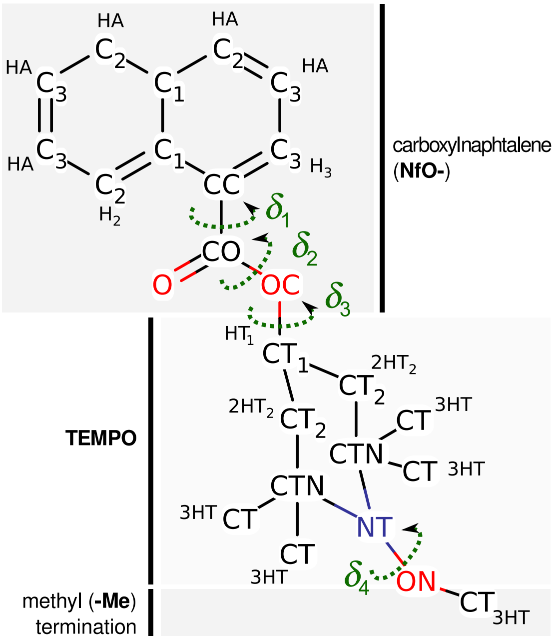


Figure 7.1: The NfO-TEMPO molecule. The atom types employed in the FF and the four main flexible dihedrals are shown.

The intramolecular FF parametrisation was performed by minimising the functional of Eq. 2.2, achieving standard deviations of 0.172 kJ/mol and 0.165 kJ/mol for the two considered states.

7.1.1 Differences between the electronic states

Inspection of Table 7.1 (where a selected set of the optimal parameters is reported) shows that the parametrised FF is able to capture even subtle differences between similar internal coordinates, that could not be distinguished by general purpose FFS. In the GS parametrisation, for instance, a non negligible difference is found between the C1-C2 and the C2-C3 aromatic bonds in the naphthalene moiety, which are described by slightly different equilibrium distances (1.42 and 1.38 Å) and remarkably different force constants (2954 and 3897 kJ/mol $Å^{-2}$, respectively). Structure rearrangements taking place upon electronic transition are also accounted for, as indicated by the changes shown by FF parameters between the GS and EES FFS. For instance the CC-CO equilibrium bond distance decreases from 1.49 Å to 1.45 Å and the relative force constant increases by $\sim$ 500 kJ/mol $Å^{-2}$, parallelling the partial π character that the bond assumes in the EES. It is worth noticing that in the EES, as a result of the increased delocalisation of the naphthalene π cloud toward the carboxyl region, the naphthalene structure reduces its conjugate character and consequently its rigidity, as can be seen from the decrease by $\sim$ 50 kJ/mol of the harmonic improper dihedral terms describing the out-of-plane deviations in the region.

Broadly speaking, the force constants relative to the internal coordinates of the carboxylnaphthalene group show rather substantial differences between GS and EES: changes as large as 22%, 47% and 31% are found for the stretching, bending and torsion harmonic terms, respectively. Instead, negligible modifications were found in the parameters relative to the TEMPO-Me region, with deviations on the stretching and bending harmonic terms lower than 2% and 8%, respectively. A tentative explanation lies in the fact that NfO-TEMPO-Me HOMO and LUMO orbitals are essentially localised on the carboxylnaphthalene portion of the molecule [64], that can thus be considered the most relevant part in the description of the optical behaviour of this compound.

7.1.2 Torsional energy scan

As far as the the two flexible dihedral angles, δ_1 and δ_3, are concerned, they are expected to affect the conformation of the NfO moiety, and therefore, for each dihedral, the

		GS		EES	
Moiety	**Bonds**	r^0	k^s	r^0	k^s
NfO	C2-C3	1.38	3897	1.42	3033
	C3-C3	1.41	3341	1.38	3801
	C1-CC	1.44	2554	1.44	2498
	C3-CC	1.38	3507	1.44	2763
	CC-CO	1.49	2034	1.45	2489
	CO-O	1.21	6703	1.23	5682
	CO-OC	1.35	2936	1.36	2848
	OC-CT1	1.44	2390	1.43	2441
TEMPO-Me	CT1-CT2	1.52	2276	1.52	2226
	CT2-CTN	1.54	1936	1.54	1924
	CTN-NT	1.49	2053	1.49	2036
	NT-ON	1.42	2444	1.42	2416
	ON-CT	1.41	2491	1.41	2503
Moiety	**Angles**	θ^0	k^b	θ^0	k^b
NfO	C2-C1-C1	118	581	118	383
	C2-C1-CC	124	805	123	709
	C2-C3-C3	121	648	119	581
	C1-CC-CO	122	420	122	234
	O-CO-OC	122	968	121	1013
TEMPO	CO-OC-CT1	117	727	117	661
Moiety	**Harmonic dihedrals**	ϕ^0	k^t	ϕ^0	k^t
NfO	C*-C*-C*-C* (*cis*)	0	53.6	0	47.6
	C*-C1-C1-C2 (*trans*)	180	53.6	180	47.6
	H*-C*-C*-C* (o.o.p.)	0	554.2	0	490.4
	O-CO-OC-CT1 (δ_2)	0	30.3	0	30.3
	CC-O-OC-CO (o.o.p.)	0	928.6	0	670.5
Moiety	**Periodic dihedrals**	n^j	k^d_j	n^j	k^d_j
TEMPO	OC-CT1-CT2-CTN	3	8.95	3	7.77

Table 7.1: Selected FFS parameters for NfO-TEMPO-Me as optimised by Joyce. The atom types are reported in fig. 7.1. The equilibrium coordinates r^0, θ^0, ϕ^0 (reported in Å and deg, respectively) and the force constants k^b, k^b, k^t, k^d_j (in kJ/mol Å^{-2}, kJ/mol rad^{-2} and kJ/mol) have been defined in equations 2.1.

reference QM torsional energy scan was performed separately for GS and EES. For all dihedrals, both FFs accurately match their reference energy profile. In the top panel of Figure 7.2 the QM energies calculated for the optimised geometries at a given torsional angle are compared to their MM counterparts, showing a very good agreement for both the states. As could be expected considering the localisation of the HOMO and LUMO orbitals, major differences between GS and EES profiles were found only for δ_1, whereas the δ_3 torsional curve does not undergo noticeable changes upon electronic transition. The different steepness around the δ_1 minimum is also in agreement with the partial π character of the CC-CO bond revealed by the force constants analysis.

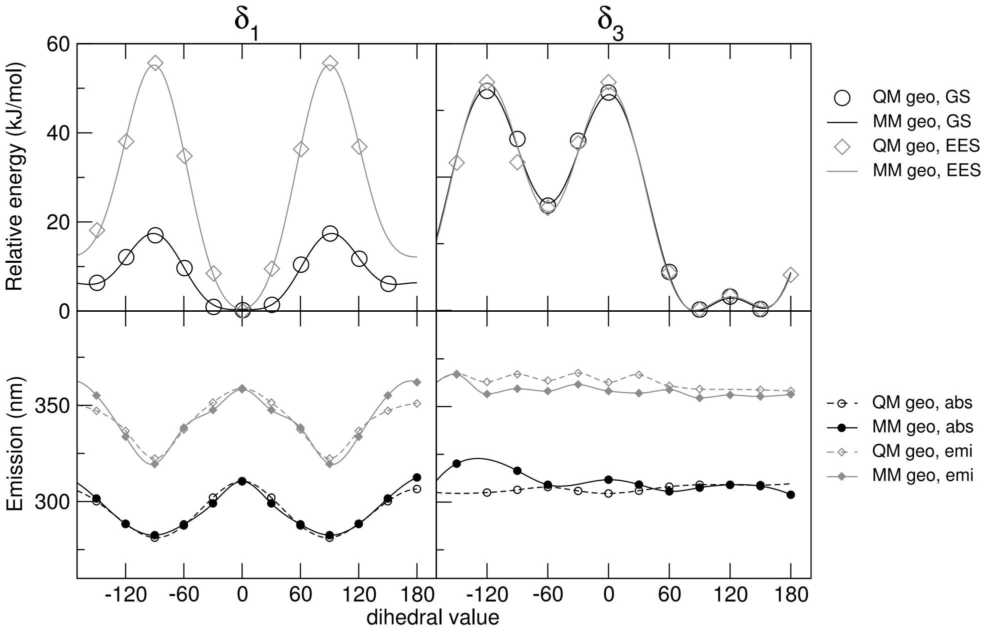


Figure 7.2: FF description of the dihedral angles. Top: energy profiles along the δ_1 and δ_3 internal dihedrals for GS and EES, at QM and FF level. Bottom: single point excitation wavelengths at partially optimised QM and MM geometries. Excited states properties are reported in grey.

It is expected that the dihedral description strongly affects the optical properties of the molecule, and for that reason the absorption and emission wavelength $\lambda_{\mathrm{VE}}^{\mathrm{abs/emi}}$ have been calculated, at the TDDFT level, for all the QM and MM geometries, obtained in the torsional energy scans. The results are shown in the bottom panel of Figure 7.2. As a matter of fact, both absorption and emission wavelengths show a strong dependence on the δ_1 angle, which is responsible for the relative orientation of the naphthyl and carboxyl moieties within the NfO portion of the molecule. The values computed for the MM geometries are in good agreement with those identified by the QM calculations, with slightly larger differences (+ 6.0 nm for absorption, and + 11.1 nm for emission) in the case of $\delta_1 = 180°$. Nonetheless, the latter conformation is expected to be scarcely populated at room temperature, considering its relative stability ($\simeq$ 12 kJ/mol) and high interconversion barrier ($\simeq$ 60 kJ/mol) with respect to the absolute energy minimum (at $\delta_1 = 0\,\mathrm{deg}$).

As long as δ_3 is concerned, its effect on $\lambda_{\mathrm{VE}}^{\mathrm{abs/emi}}$ is less pronounced, with some minor variations found in emission for $\delta_3 \sim$ 30° . This scarce sensitivity is not surprising, considering that the orbitals involved in the transition are localised in molecular regions not affected by the δ_3 torsion. As for δ_1, the FF mimics with rather good accuracy this behaviour, with slightly larger deviations for absorption, found in high energy conformations ($\delta_3 <$ -60°).

7.2 Simulations in gas and in solution

The GS and EES parametrised FFS were first employed in two sets of MD simulations, performed on the isolated dye at 300 K and on the solvated NfO-TEMPO at 1 atm and 300 K. All four systems (*i.e.* GS and EES *in vacuo* and solvated GS and EES) were equilibrated for 1 ns, and subsequently simulated for 1.2 and 3 ns (for the *in vacuo* and solvated phase, respectively), saving trajectories every 0.5 ps.

The dye internal structure, along the four different simulations has been first evaluated by looking at the distributions of the two main flexible dihedrals, reported in Figure 7.3. As regards the GS simulation in gas phase (black dotted line) δ_1 remains constrained to the main potential well, and its distribution is symmetrically broadened around 0°, the QM reference value being within this well, at about $-12°$.

The solvated system shows similar distributions within $\delta_1 \simeq 0°$ region, but the interaction with the solvent allows for exploring also the local minimum at about 180°. Conversely, due to the increased steepness around the δ_1 minimum region found in the torsional profile, the EES population distribution is much narrower for both isolated and

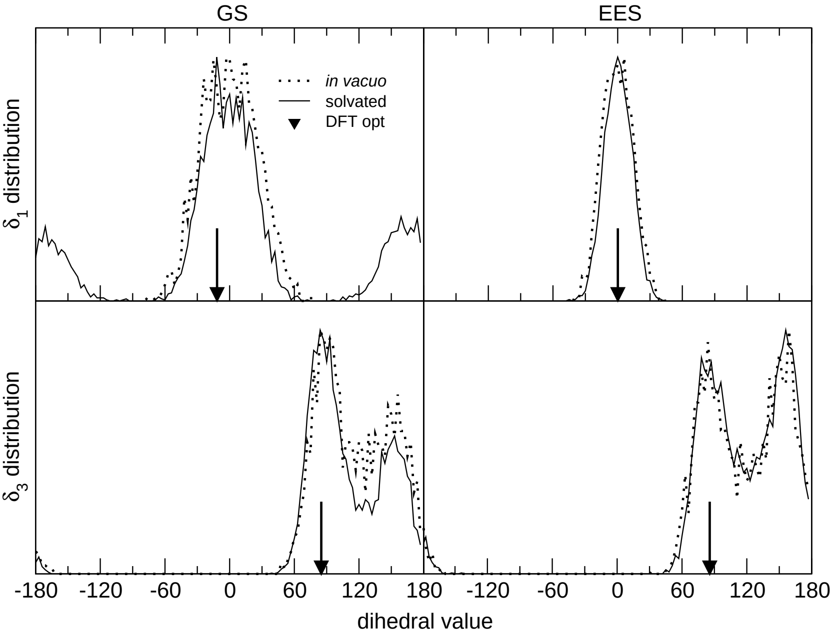


Figure 7.3: Distributions of the δ_1 (top) and δ_3 (bottom) dihedrals during the simulations of NfO-TEMPO in toluene, compared to the QM optimised geometry. The curves are normalised to their maximum value. For NfO-TEMPO-Me the ground state (left) and the electronically excited state (right) are considered.

solvated systems, and the 180° region is never explored. This can be considered, in facts, as an effect of the increased π delocalisation to the C-C bond described by δ_1 that in the EES retains the peculiar stiffness of a higher-order bond.

The δ_3 distribution shows, instead, a double peak for both GS and EES simulations, with the first one corresponding to the QM value of about 85° and the second one located at $\delta_3 \simeq 150°$ *i.e.* close to the second local minimum found in the torsional energy scan reported in Fig 7.2.

Furthermore, as the solvated systems are concerned, the local solvent structure around the dye has been monitored by computing the pair correlation function between the NfO-TEMPO and toluene centres of mass, for both GS and EES. Inspection of Figure 7.4 shows that some differences arise in the first solvation shell for GS end EES, the number of first neighbours being larger in the latter case. On the other hand, the differences become negligible for the second solvation shell and at distances larger than 13 Å both pair correlation functions converge to an homogeneous density. For this reason, when the EE extraction scheme, described in section 2.2.2, is to be used, the cut radius is set to $R_{cut} = 15$. After this region, indeed, the continuum description for the solvent describes reasonably well the bulk properties.

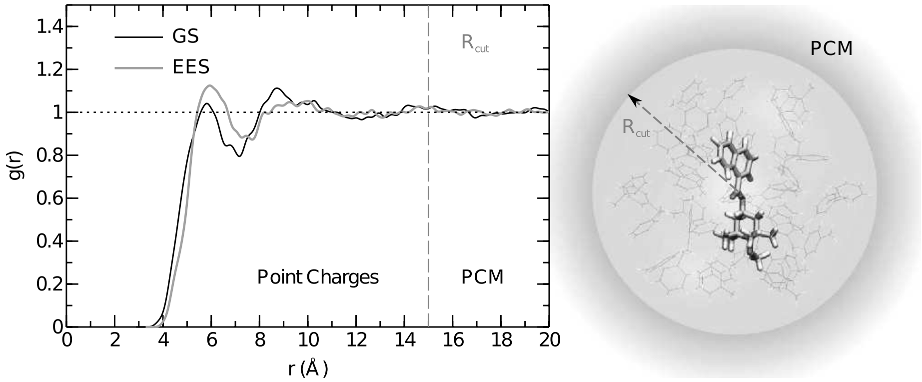


Figure 7.4: Left panel: solute – solvent pair correlation functions relative to the GS and EES simulations. Right panel: layered hybrid description of the dye in solution. The dye is described at QM level, a 15 Å large sphere of explicit toluene atomic charges is placed around it, and is embedded in a spherical PCM cavity.

The included solvent molecules amount on average to 30, with a standard deviation of 2, for both GS and EES simulations.

7.3 Modeling of the polyethylene and simulation of the grafted dye

The NfO-TEMPO dye has been used in experimental works to functionalyze, by covalent bonding, polyethylene derivatives [62, 63], in order to transfer to the polyolefins its spectral properties. Although an apolar polymer won't affect much the UV-Vis spectra by the electrostatic embedding, a dynamic influence is expected. To assess the role of the polymeric matrix in a dye+polymer compound, a polyethylene (namely, Linear Low Density PolyEthylene, LLDPE) model have been developed and coupled to the NfO-TEMPO molecular FF afore described.

7.3.1 The polymer model

The steps to construct the model polymer matrix are sketched in fig. 7.5.

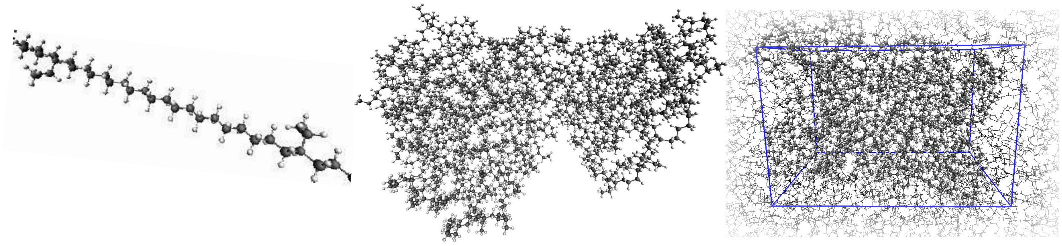


Figure 7.5: Snapshots from the polyethylene equilibration process. From left to right: a portion of the linear chain; entanglement obtained in a NVT run; NPT run with PBCS.

A model polymer matrix was constructed by randomly branching a single linear aliphatic hydrocarbon chain composed by more than 1400 Carbon atoms, described at full atomic level. Each branching was achieved by replacing a methylene H atom with an ethylene moiety every 9, 10 or 11 CH_2 monomer units of the principal chains, thus obtaining a model polymeric branched chain described by 4655 atoms. At this level, the polymer is spread in a linear chain with all the dihedrals in their *trans* configuration (first panel in fig. 7.5)

The FF for the model polymer was first described by transferring literature parameters [34, 54] for aliphatic hydrocarbons (see the Appendix for the force field). However, for a first run, the torsional barriers of the chain dihedrals were artificially lowered to increase the *trans* to *gauche* conversion rate. An effective entangling of the chain was achieved with short equilibration runs in the NVT ensemble, with the PBC box walls enough far from the polymer to avoid self-interactions at this stage.

When the model polymeric matrix was judged sufficiently entangled, a pressure coupling was introduced and the PBC box was tightened onto the polymer chain. The torsional parameters were restored to their values from OPLS FF.

After the first equilibration NPT step, the final density was found lower than 850 kg/m^3. To achieve a mass density closer to the average experimental value reported for LLDPE (910-940 kg/m^3) and ULDPE (Ultra-Low Density PE, 860-900 kg/m^3) all carbon's and hydrogen LJ parameters σs were reduced by 5%, achieving a final value of 890 kg/m^3.

7.3.2 Complete force field of the grafted system

The NfO-TEMPO dye was attached to the equilibrated model of LLDPE, by linking TEMPO Oxygen atom to a branched site of the polyethylene chain, after removal of the corresponding H atom.

As the FF are concerned, three regions were defined into the molecular assembly, as reported in the top panel of Fig. 7.6: the polymer chain, the linker region and the dye moiety. The GS and EES FFS employed for the dye are the same as before (see Tab. 7.1), while the parameters concerning the dihedrals located in the dye-polymer junction were specifically parametrised. A least square minimisation procedure was used on purposely computed QM torsional energy scans, performed on small portions of the system that include the TEMPO moiety and the first aliphatic monomers present on the grafting site, as shown in Fig. 7.6, bottom panel.

With the mentioned entirely intra-molecular FF, a final simulation was performed for each of the two states of the dye. To avoid highly repulsive interactions, the NfO-TEMPO dye was linked onto a non equilibrated strained conformation of the polymer, where most of the aliphatic chain dihedrals were in their *trans* configuration. Thereafter, an effective entangling of the chain was achieved with short equilibration runs in the NVT ensemble, where the NfO-TEMPO dye was frozen at its equilibrium conformation. During these runs, the volume of the simulation box was slowly decreased and the torsional barriers of the chain dihedrals were artificially lowered to increase the rate of conversion from the *gauche* to the *trans* conformation. After 10 ns, a pressure coupling was introduced and all the constraints on the dye substituent were removed. The final NPT simulation run 10 ns long was performed at 300 K and 1 atm, achieving again the final density of 890 kg/m^3. A visualisation of the system is provided in Fig. 1.2.

It may be worth mentioning that the experimental concentration of the dye inside the polymer is so low that the grafted molecule cannot be found in close proximity to another probe, hence the choice of modeling just one molecule per polymer chain should best represent this scenario.

7.3.3 Structural properties

Pair distribution functions

From the two MD simulations, the dye - polymer pair distribution function were calculated and plotted together with the dye - toluene equivalents as reported in Fig. 7.7. For a better comparison between the two embeddings, the $g(r)$s were considered for

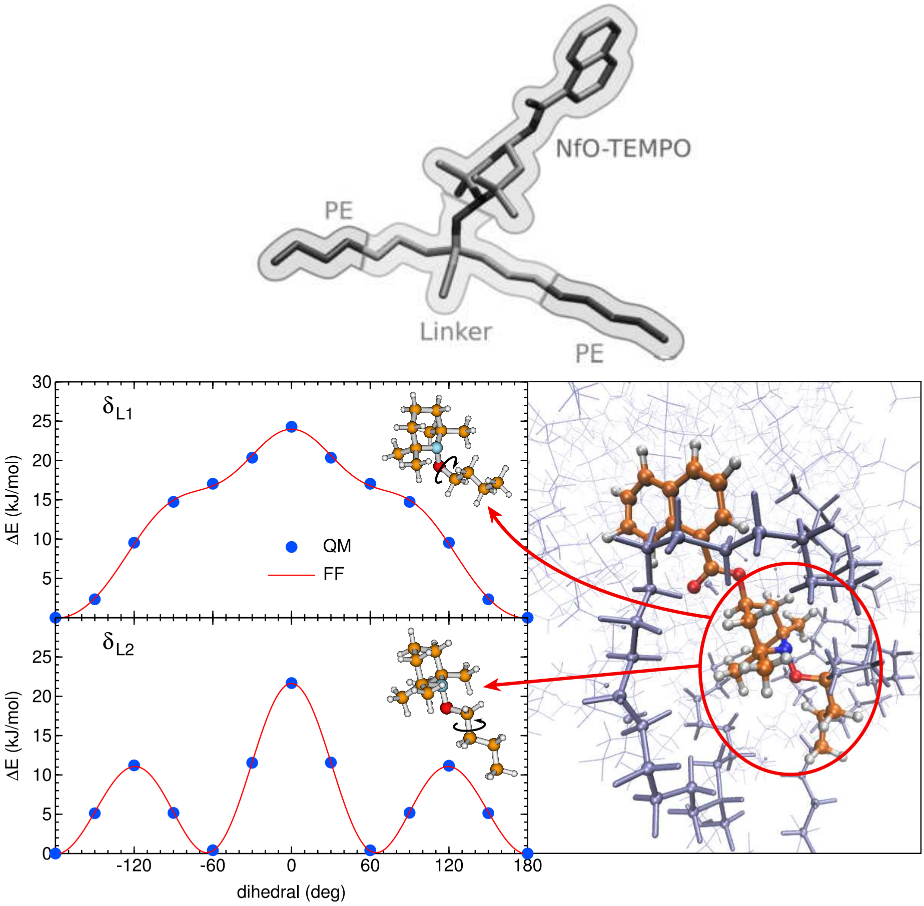


Figure 7.6: Top: the three different FF regions that make up the macromolecular assembly dye + PE.
Bottom, left panel: parametrisation of the torsional profiles in the grafting TEMPO-polymer region. The small molecular targets employed as models for the grafting region are shown in the insets. QM data, computed at the PBE/N07D level, are reported with circles, while the parametrised FF torsional profile is shown with a solid line. Right panel: TEMPO-polymer grafting region and selection of the small molecular targets.

the NfO-TEMPO centre of mass with respect to any of the aromatic carbon atoms of toluene, or any of the carbon atoms of PE: for this reason the curves in Fig. 7.4 and Fig. 7.7 appear different.

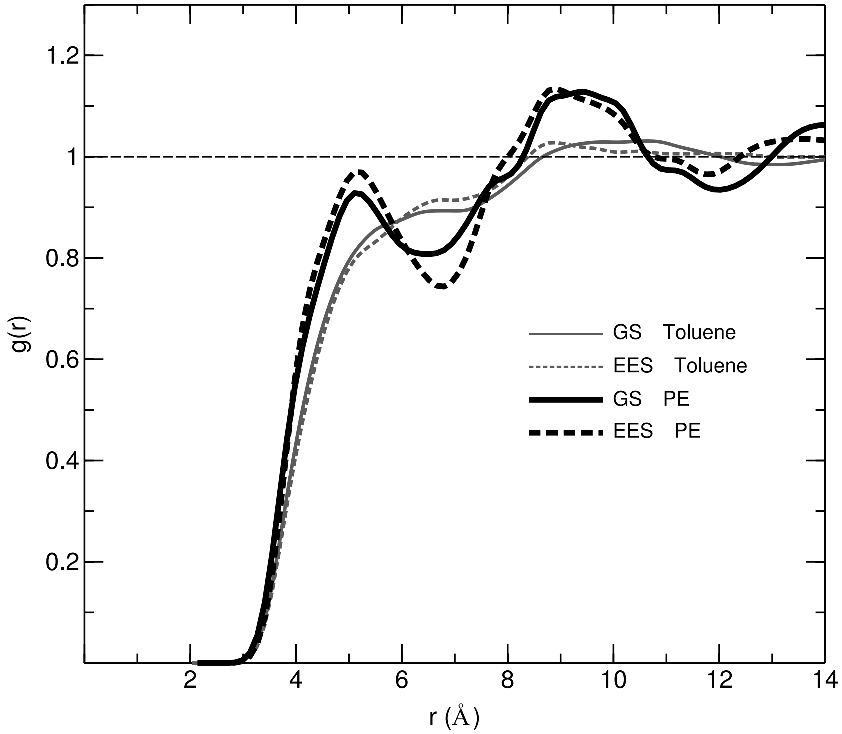


Figure 7.7: Dye – embedding pair distribution functions, for both the embedding (toluene – thin grey lines – and PE – thick black lines) and the electronic state of the solute. GS are represented with solid lines, EES with dashed lines.

By comparing the behaviour in polymer and in toluene, despite all functions present a similar starting behaviour, the influence of the surroundings rapidly differentiates their shapes, with the $g(r)$'s computed in polymer much more structured than those computed in solution. On the other hand, the correlation functions of different electronic states (GS and EES) embedded in the same *medium* are less distinguishable, even if the excited state in polymer allows more embedding carbon atoms in the first shell.

Comparison of the first peaks in toluene and polymer ($\sim$ 5.8 Å and 4.8 Å, respectively) shows that, for both GS and EES, the polymer units are allowed to get sensibly closer to the dye, suggesting that a tighter cage is formed around the naphthalene moiety by the entangled polymer. The well defined minimum of the first peak in polymer and the absence of a net decay in toluene suggest a longer lifetime of the polymer cage and a more "dynamic" character of the toluene first neighbour shell. Moving to longer distances, a second peak is more clearly definite in the polymer – more structured – case; anyway, also here the pair distribution function reaches bulk behaviour ($g(r) \simeq 1$) for $r \simeq 14$.

Torsions

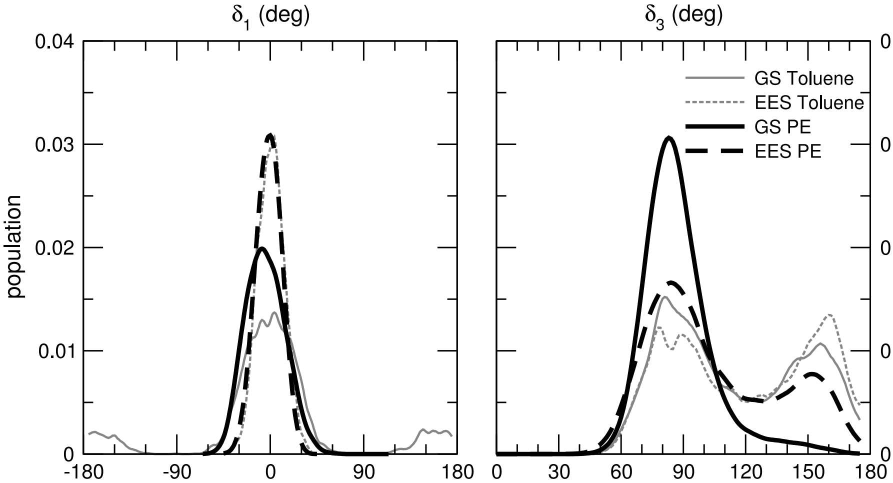


Figure 7.8: Population distribution of the two flexible dihedrals in the four embedded simulation pictures.

A deeper insight on the surrounding environment and its correlation with the molecular conformations adopted by the dye can be gained by investigating the behaviour of the NfO-TEMPO δ_1 and δ_3 dihedral angles. In the two panels of Fig.7.8 the distributions of the dihedral angles in the dye's GS and EES are shown for both environments.

As already noted in the solvated case, δ_1 distribution is significantly narrower for the EES, and its maximum is centred at about 0°, which corresponds to one of the possible planar conformations. Since this reduced flexibility has been ascribed to the increased delocalisation of the π cloud towards the regions in proximity of the TEMPO moiety (which is an intrinsic feature of the dye) it is not surprising that the same behaviour is found when the dye in its EES is embedded in a different environment like PE. In fact, the two δ_1 distributions (EES in toluene and in PE) are almost identical.

On the contrary, a major difference, rooted in the different embedding, is evident when comparing the GS δ_1 distributions obtained for the solvated and grafted systems. In fact, only in the former solvent the dye is able to adopt the two planar conformations ($\delta_1 \sim 0°$ and $\delta_1 \sim \pm 180°$) expected from the inspection of the corresponding torsional energy profile, whereas only the $\delta_1 \sim 0°$ conformer is populated in the polymer matrix (like it was *vacuo*, for a different reason). This observation is in agreement with the hypothesis drawn from the $g(r)$ analysis, where the polymeric chain entangles the dye, thus reducing its internal flexibility.

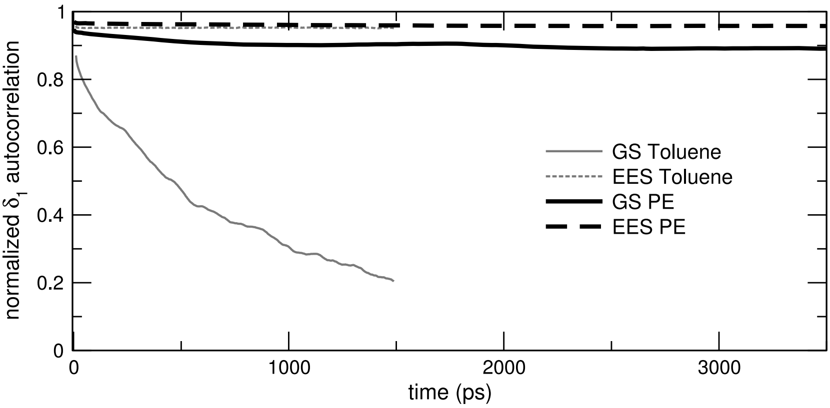


Figure 7.9: Normalised auto-correlation function of the δ_1 torsion angle for the four embedded simulations.

This is also confirmed by looking at the δ_1 time ACF (see. section 2.3)

$$\Delta(t) =< cos[\delta_1(t_0 + t)]cos[\delta_1(t_0)] >$$

registered for all systems and displayed in Figure 7.9. Indeed, only in toluene solution the dye in its GS is free to explore all the accessible conformations, being $\Delta(t)$ almost zero after 1.5 ns, meaning that the dihedral value is retained, on average, for shorter times. On the contrary, the ACF in polymer remains highly correlated for more than 3 ns, indicating that the carboxy moiety is forced by the embedding *medium* to preserve its co-planarity with the naphthalene unit.

The auto-correlations of both the EES simulations are even slower to decrease, as expected from the narrower distributions and the stiffer character of the subtending bond.

Analogous data computed for δ_3 and shown in the right panels of Figure 7.8 are apparently in contrast with the above analysis. Indeed, the described "caging" effect of the polymer is only evident in the GS, whereas the dye in its EES seems able adopt conformations similar to those observed in toluene solution. However, the similar behaviour of the δ_3 distributions of the EES in toluene and polymer can still find a rationale without contradicting the "caging" hypothesis.

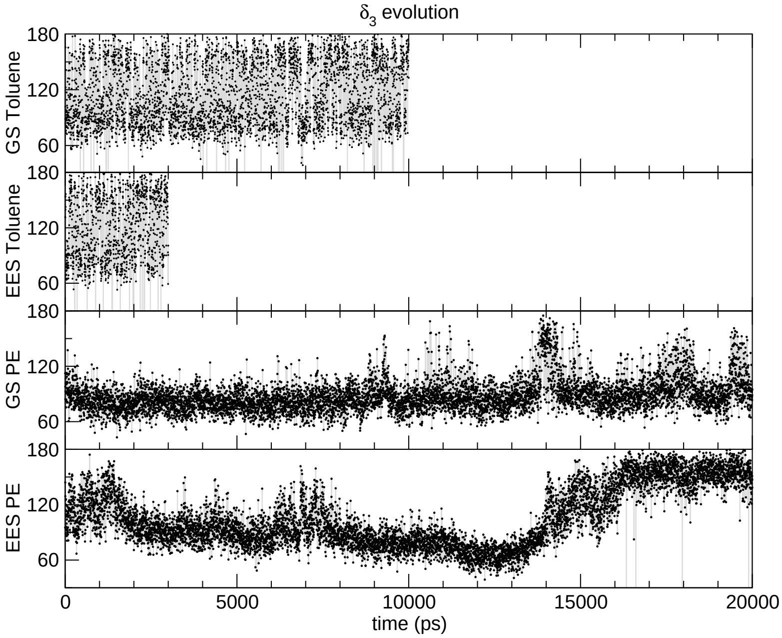


Figure 7.10: Time evolution of the δ_3 dihedral with a resolution of 2.5 ps, for the whole production time of the four simulations

Indeed, if the value of δ_3 is plotted over the simulation time like in Fig. 7.10 a remarkable difference appears: in toluene fast inter-conversions between the two conformers are allowed, whereas in the more hindered polymer matrix only one very slow ($\sim$ 12.5–17 ns) transition is found from one conformer to the other, whereas only small variations of the dihedral angle are allowed by the cage for shorter time scales. The resulting distributions (Fig. 7.8, right, dashed lines) are thus similarly wide, but the exploration of the secondary well happens with remarkably different dynamics and time evolutions.

Going back to the δ_1 dihedral dynamics, a difference can be drawn between the solvated and grafted system (only GS) when it comes to the width of the main ($\delta_1 \simeq 0°$) well.

It should be noted that the same FF term $V_{\mathrm{tors}}(\delta_1)$ (Eq. 2.1) is used for the intrinsic potential that describes the intramolecular torsion, with the same parameters giving rise to the profile in Fig. 7.2, top left panel. As a consequence, the different population distributions must be ascribed to a "mean field" due to the surroundings, that can be evaluated from the differences between the intramolecular potential energy and the actual resulting distribution.

In detail, in the hypothesis that the effective potential energy curve governing δ_μ dynamics can be partitioned into an intra molecular component (*i.e.* V_{tors}) and the mean field $W(\delta_\mu)$, $P(\delta_\mu)$ can be expressed as

$$P(\delta_\mu) \propto e^{-\frac{V_{tors}(\delta_\mu)+W(\delta_\mu)}{k_B T}} \tag{7.1}$$

and therefore, up to an additive constant,

$$W(\delta_\mu) = -(E_{tors}(\delta_\mu) + k_B T ln[P(\delta_\mu)]) \tag{7.2}$$

The above equations are applied to δ_1, and the resulting mean field $W(\delta_1)$ is displayed in Fig. 7.11.

The resulting mean field W(δ_1) in the polymer and in toluene solution presents a marked difference between the two surrounding media being much more flat in toluene than in the other environment, and close to zero (*i.e.* toluene slightly affects the dye's inherent conformational freedom). It is also worth noticing that the entanglement of the polymer around the dye creates a supplementary well, centred at about δ_1 = 0, that constrains this angle to *librate* within a limited interval (roughly -50° to 30°) rather than exploring the complete range of values.

7.3.4 Dynamic properties

A more direct way to calculate the mobility of the dye inside the embedding is to consider the translational and rotational diffusion and the inertia axes reorientation times. For their definitions, see section 2.3.

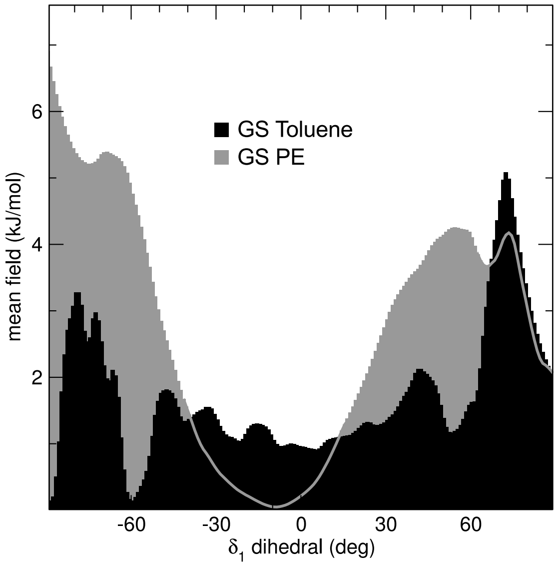


Figure 7.11: Mean field W(δ_1) (Eq. 7.2) due to the different embeddings (grey: polymer, black: toluene) acting on the NfO-TEMPO conformational behaviour

The mean square displacement of the fluorophore's centre of mass is plotted in fig. 7.12 for the GS dynamics in toluene and in polymer.

As regards the solvated dye, the MSD shows the usual diffusive behaviour of a solute in a liquid medium. Indeed the MSD(t) log-log curve is fitted by a straight line with a slope of 1.03 ± 0.04, indicating that the MSD obeys the power law $MSD(t) = 2D \cdot t^1$. The proportionality factor is obtained by fitting the curve in linear scale, and it gives the translational diffusion coefficient of NfO-TEMPO-Me in toluene

$$D^{\mathrm{tr}}_{\mathrm{tol}} = (0.66 \pm 0.04) \cdot 10^{-9}\mathrm{m}^2\mathrm{s}^{-1}$$

As of the grafted system, a sub-diffusive behaviour is obtained as a consequence of the polymer entanglement. Thus, for integration times up to 1.4 ns, the asymptotic diffusive regime is not reached and it is impossible to define a translational diffusion coefficient.

Proceeding to the rotations, a summary of the results is reported in 7.13 The angular velocity auto-correlation functions (AVAFS) are plotted in the top panel, concerning the projections ω_a, ω_b, ω_c onto the three principal inertia axes. In the stationary point conformation, the inertia axes are roughly oriented as follows (see fig. 7.1 for the defi-

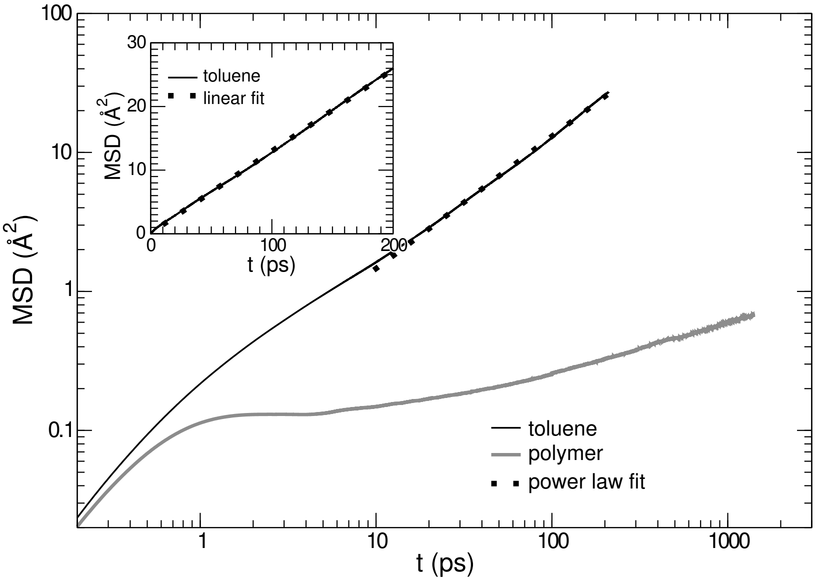


Figure 7.12: Big panel: comparison in logarithmic scale between MSD(t) of the dye in the two different environments: toluene solution and grafted to polymer. Inset: The MSD curve for the system in solution is shown in linear scale. Analytic fits of the curve in the t > 80 ps region are reported.

nitions):

a. "Long axis": connecting the TEMPO Nitrogen atom (NT) to the C2 atom in the first naphthalene ring

b. Connecting the C3 atoms of the two naphthalene rings

c. Perpendicular to the previous ones

The asymptotic value of the AVAFs integrals (see Eq. 2.11) are used as values for the rotational diffusion coefficients, and are summarised in the table below:

Property	Toluene	Polymer
D_a^{rot} (s^{-1})	1.28×10^{10}	0.07×10^{10}
D_b^{rot} (s^{-1})	0.05×10^{10}	0.02×10^{10}
D_c^{rot} (s^{-1})	0.02×10^{10}	0.02×10^{10}

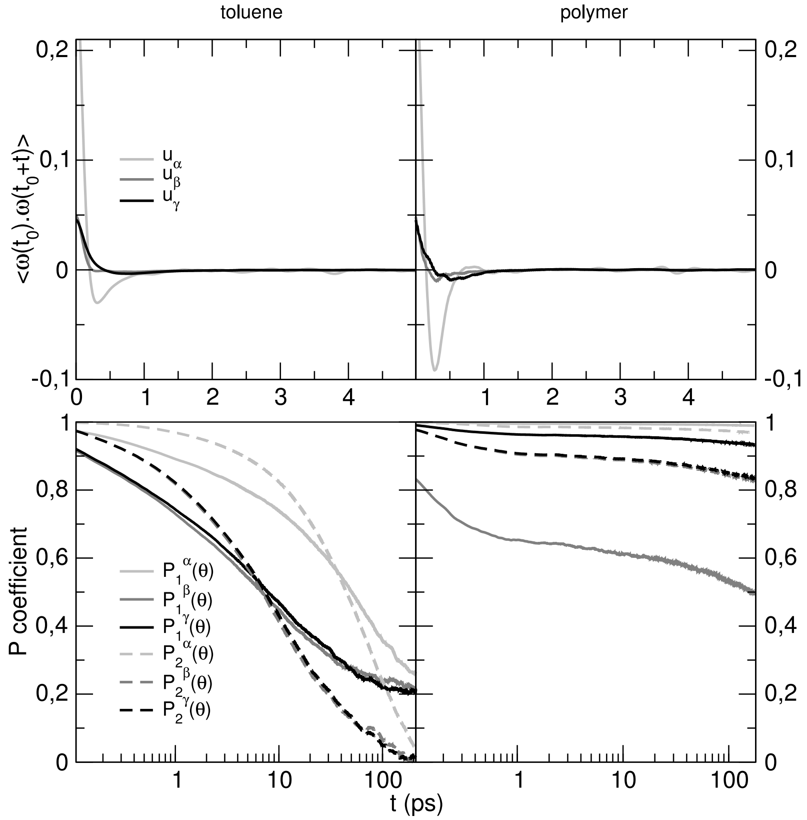


Figure 7.13: Rotational behaviour of the NfO-TEMPO GS in the two considered environments. Top: angular velocity aurocorrelation function along the three principal inertia axes; bottom: orientational first and second order parameters $P_1(\theta)$ (solid lines) and $P_2(\theta)$ (dashed lines) concerning the three axes.

From the plot in the bottom panel of fig. 7.13 an interesting result can be inferred about the polymer cage. By looking at the first rank orientational parameter $P_k^1(\tau) = \cos(\phi_k(\tau))$, $\phi_k(\tau)$ being the angle between the 'k' axis at times t_0 and $t_0 + \tau$, it can be seen that the definition of the "long" axis a is more stable, over time, than b and c that can vary and mix as a result of the molecule conformational changes. Moreover, the orientational parameters drop far more rapidly in the case of toluene, indicating the possibility for a complete reorientation in hundreds of ps, than in polymer. In particular, over a integration time of $\tau \simeq 150$ ps, the principal inertia axis of NfO-TEMPO grafted onto PE remains almost unchanged.

7.4 Vibrational analysis

Before conducting the statistical determination of optical spectra, a discussion about the vibrational features of the dye was done.

From a simulation of the grafted NfO-TEMPO in the NVE ensemble, performed with a reduced timestep of 0.2 fs and a print frequency of 0.4 fs, a vibrational power spectrum was calculated, performing a Fourier transform of the velocity auto-correlation function. As stated in section 2.3.3, the spectral analysis of a ACF is expected to give a reliable set of normal modes. Nevertheless, the actual intensities of a IR spectrum must be calculated by means of a good classical approximation to the molecular dipole moment ACF.

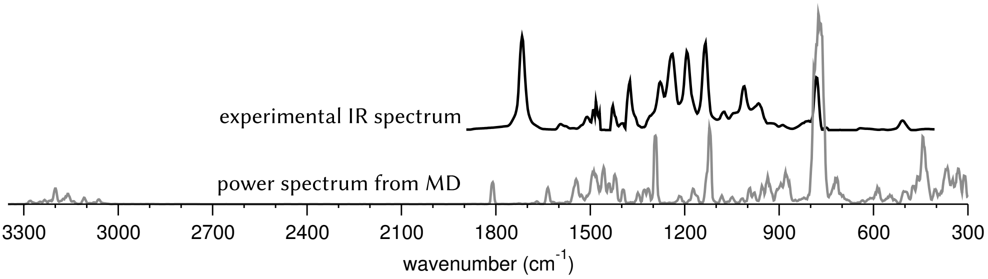


Figure 7.14: Vibrational spectra of NfO-TEMPO grafted to PE (arbitrary units). The experimental curve (black) is obtained as a difference of the IR spectra of NfO-TEMPO+PE and PE alone. The calculated spectrum is obtained as a Fourier transform of the VACF of NfO-TEMPO from a simulation in the polymer environment. The high-frequency region is not reported in the experimental reference[63].

Firstly, the VACF power spectrum of the NfO-TEMPO grafted molecule is plotted in fig. 7.14 and compared to the experimental data. The experimental curve is obtained as a difference between the IR spectrum of the doped polyolefin, and the polyolefin alone.

Restating that the intensities from the two curves must not be compared at this stage, the main features of the vibrational spectra are highlighted in the simulated one. See, for example, the modes placed at $\simeq$1800 cm^{-1} and 800 cm^{-1}.

In order to allow a direct comparison with experiments, the actual "semi-classical" IR absorption spectrum was calculated for the isolated and the grafted dye, by using the quantum corrected Fourier transform of the electric dipole ACF, according to Equation (2.12). The experimental dye's signal was again extracted from the provided data by taking the difference of the grafted and the isolated polymer spectrum (see Figure 7.15, dashed and solid turquoise lines). In the investigated window (400 - 1900 cm^{-1}), the experimental spectrum presents a faint peak around 510 cm^{-1}, followed by several peaks delimited by the already mentioned strong ones near 800 and 1800 cm^{-1}, respectively. Actually, modes with higher frequencies are present, though not plotted, in the experimental spectrum, with a peak reported at 3052 cm^{-1}. On the contrary these modes, related to the C–H stretching, are totally absent in the velocity power-spectrum calculated at the oxygen site, since they do not affect this particular atom (Fig 7.14).

To make the crude assignment of the IR absorption lines more accurate, also the harmonic DFT-level calculation was used as a reference, that was already performed to parametrise the FF. The experimental IR spectrum, the spectra computed from the MD simulation in polymer and in gas phase, and the QM reference are plotted Figure 7.15(a-d). As regards the most evident peaks of the QM reference, an inspection performed using visualisation tools revealed that the two lines at 802.7 and 1797.1 cm^{-1} correspond to the out-of-plane bending of the naphthalene hydrogens and to the C=O stretching; the mode at 1309.1 cm^{-1} involves an in-plane motion of the naphthalene hydrogens.

The fair agreement, in term of wavenumbers, of the harmonic DFT-level spectrum and the MD based ones, could be easily predicted, since the simulations use a FF based on the DFT hessian matrix. In this case, also the agreement with the experiment is good, though the lines may be shifted (*e.g.* the C=O band is translated by 80 cm^{-1}). On the other hand, as regards the peak intensities, the vibrational spectra calculated with the quantum corrected dipole ACF are in fairly good agreement with their experimental counterpart.

It should be pointed out that the inclusion of anharmonicity in the FF definition, that is not present in the model employed here, could enhance the vibrational description of the system and lead to a more quantitative correspondence. Having said that, it can be underlined that the adoption of a specifically tailored FF allows for a rather reliable representation of the vibrational behaviour of the dye. Since the present statistical determination of the dye's electronic spectra is based on configurations extracted from classical MD simulations, the resulting spectrum will take into account the vibrational broadening with reasonable accuracy.

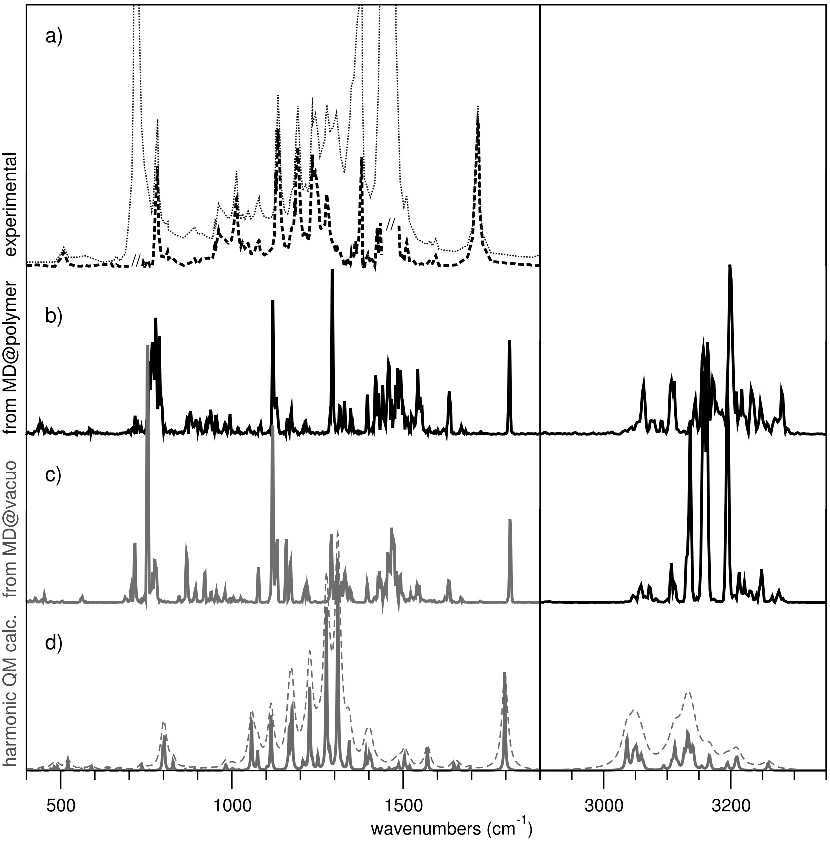


Figure 7.15: Experimental and computed IR spectra; from top to bottom: a) Experimental IR spectrum of the dye+polymer system (dashed line) [63], and the same spectrum removing the signals from the polymer itself (solid line). b)) Calculated IR spectrum the dye, obtained from MD simulations of the grafted system. c) Calculated spectrum from MD for the NfO-TEMPO-Me in vacuo. d) Harmonic IR spectrum calculated at the DFT level on the isolated molecule, broadened with Lorentzian functions with HWHM 2 cm^{-1} (solid line) and 10 cm^{-1} (dashed line).

7.5 Optical spectra

Finally, for all the dye IN VACUO, solvated and grafted systems, the statistical lineshapes were calculated with the statistical method.

7.5.1 Absorption spectra and choice of the DFT functional

A first study, which was reported in Ref. [65], took advantage of only the experiments in gas and liquid phase to determine the most effective DFT method to address the TD calculations.

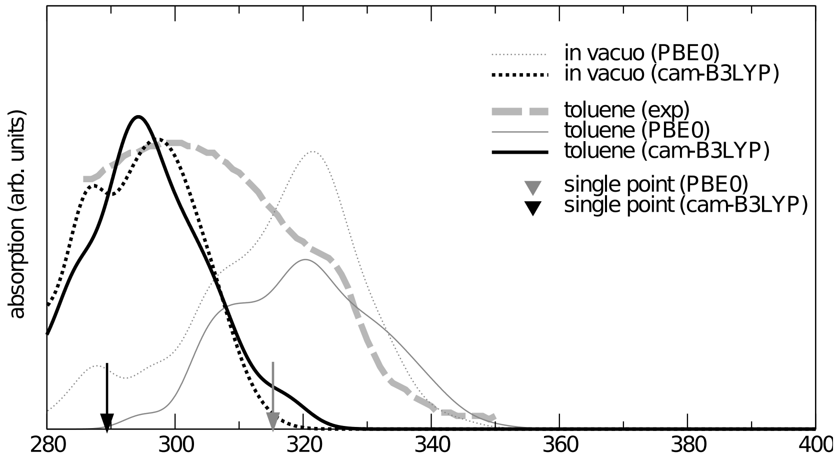


Figure 7.16: Absorption spectra in vacuo (dotted) and in toluene solution (solid line) calculated with different DFT functionals (PBE0: thin grey, CAM-B3LYP: thick black) in the statistical approach, compared to the static approach (placemarks) and to the experimental spectrum. The same normalisation was used for all the computed spectra.

At first, PBE0 XC functional was used for the TDDFT "electronic embedding" calculations, and the results are shown with thin grey lines in Fig. 7.16. The dynamic approach red-shifts the absorption wavelength, $\lambda_{\max}^{\mathrm{abs}}$, by about 2 nm, with respect to the static approach employed in the previous works of Ref. [64] (grey arrow). Unfortunately, the static value (315 nm) did already overestimate the experimental $\lambda_{\max}^{\mathrm{abs}}$ by 16 nm, so that dynamical effects seem to worsen agreement with experiment.

Since the choice of the functional / basis set could play a role in determining this trend, other functional/basis set pairs have been explored, eventually obtaining remarkably accurate results by combining the CAM-B3LYP functional with the cc-pvDz basis set. In this case, the static approach (black placemark in the graph) underestimates the experimental peak wavelength by 10 nm, but this error is corrected by the MD treatment with a 3 nm red-shift. The "thermal" broadening of the absorption line, obtained by including different snapshots, allows a more direct comparison of the spectra, resulting in a good agreement between the computed and the experimental [64] peaks of absorption (292 nm and 299 nm, respectively). Concerning solvent effects, it should be noted that solvatochromic shifts are quite negligible, leading to very similar spectra in the gas phase and in solution.

Once the DFT functional of choice has been assessed, the following calculations all benefit from the CAM-B3LYP/cc-pvDz treatment.

At first, the same four simulations (GS, EES *in vacuo* and in toluene) were used for the simulation of the emission lines.

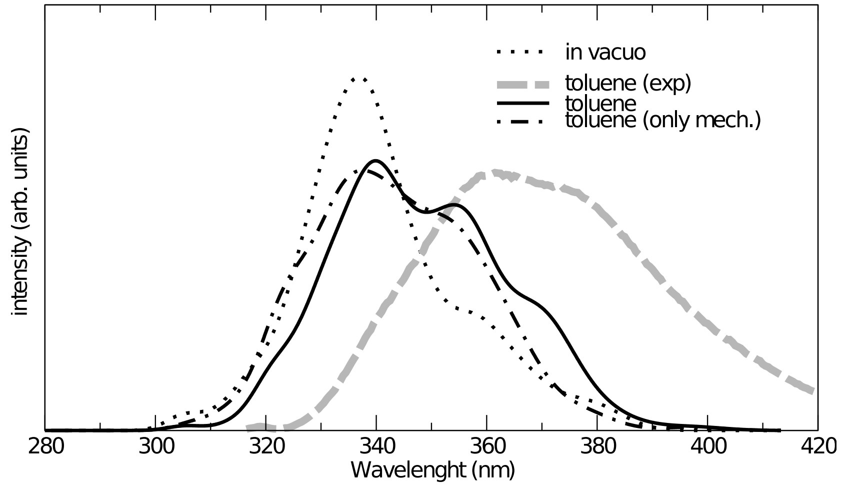


Figure 7.17: Emission spectra of the NfO-TEMPO dye, calculated in vacuo (dotted) and toluene solution (solid line), considering also the "mechanical embedding" alone (dash-dotted line).

Inspection of Fig. 7.17 shows that the emission profile of NfO-TEMPO-Me in toluene solution reproduces fairly well the experimental data, especially the line-shape and its decay, although the peak is blue-shifted by about 15 nm. At variance with absorption, the effect of the solvent on the spectral shape is quite significant.

In this case the spectrum in the gas phase is less broadened and centred at a shorter wavelength (by about 17 nm) so that a non negligible solvatochromic shift is observed. It is worth noticing that the increased sensitivity to the solvent, found for EES with respect to GS, is consistent with the augmented flexibility of the naphthalene skeleton (*i.e.* where HOMO and LUMO orbitals are essentially localised) evidenced by the FF and with the larger number of solvent molecule found in the first solvation shell by pair correlation function analysis.

Finally, the main contribution to the broadening is accounted for already at ME level (section 1.1, whereas the inclusion of the direct electrostatic interaction in the TDDFT calculations, achieved with EE, only tunes the signal, evidencing two shoulders at about 355 and 370 nm.

7.5.2 Spectra in polymer matrix

In conclusion, all the spectra were calculated, with the same method and exploiting the MM and ME schemes, for the dye grafted onto a polymer matrix.

The resulting lines are reported, together with the ones in toluene, in figure 7.18. Table 7.2 reports a summary of the peaks and Stokes shifts observed.

System	Method	Absorption peak (nm)	Stokes shift (nm)
in vacuo	Computed	292	37
in toluene	Computed	292	54
	Experim.	299	62
in polymer	Computed	294	49
	Experim.	296	62

Table 7.2: Values of the absorption wavelengths and of the Stokes shift, calculated at the line peaks, for the NfO-TEMPO dye in the considered environments.

By looking at Table 7.2, the results, as concerns the peak wavelengths, are as good as found in the toluene case, when compared to the experimental findings. In particular, the absorption peak is reproduced within 2 nm, while the error on the Stokes shift is of about 13 nm.

By inspection of the line shapes, it can be observed that they well compare to the experimental ones, particularly when the broadening and the decay slope are considered. The experimental bulged decay of the absorption line at about 320 nm corresponds to a secondary peak, although overestimated and slightly shifted, in the computed spectrum ($\simeq$ 310 nm). The computed emission line is found to be as broad as the reference one; moreover, the decay toward red loses the *plateau* observed in the solvated case ($\simeq$ 360 – 380 nm exp.), as expected from the experimental spectra. Eventually, the experimental curve for longer wavelengths exhibits a shoulder at about 395 nm, that can be found

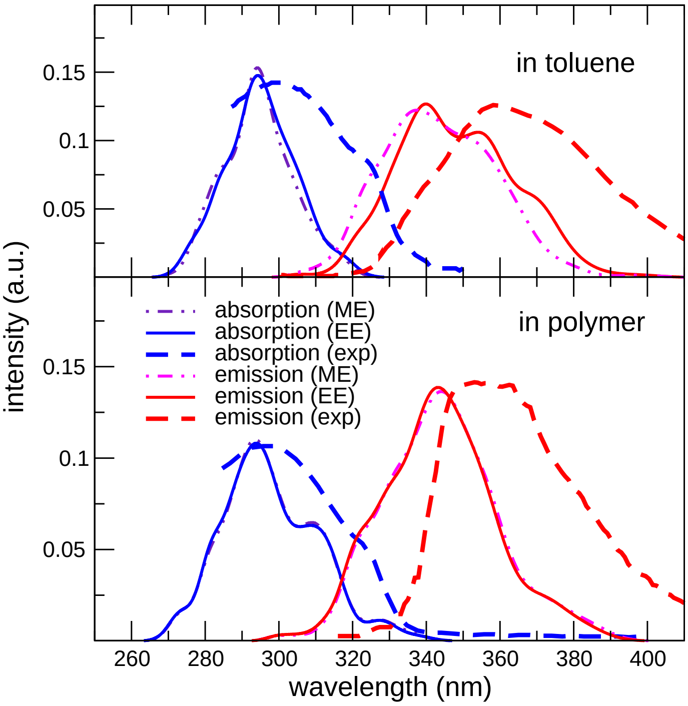


Figure 7.18: Summary of the absorption and emission spectra of NfO-TEMPO-Me in toluene solution and NfO-TEMPO grafted onto polymer matrix. The calculated lines are reported using both the EE and ME schemes. The experimental curves are plotted for comparison as dashed lines.

also in the simulated spectrum at 370 nm.

Moreover, ME and EE schemes converge to almost exactly the same results; it may be worth noticing that this is a feature shared with the spectra computed in toluene, indicating that in both cases rather weak electrostatic interactions take place between the dye and the considered embeddings, that are not able to sensibly alter the absorption signal.

Chapter 8

Conclusions

In the presented work, a general approach has been illustrated that takes advantage of state of the art QM/MM techniques for the computational treatment of several features of medium sized molecules. The investigated properties are expected to depend on thermally accessible configurations and on the interactions with the environment, and the described protocol is able to deal with both these features.

As regards the former, the balance between twisted conformers due to the flexibility of the inner dihedrals has been addressed through the employment of MD simulations in the condensed phase. To achieve reliable results at the MM level, the simulations were performed on the basis of accurate force fields obtained with the JOYCE method. The latter makes uses of DFT-level electronic calculations to correctly reproduce the vibrational and torsional features of the molecule's internal motion. The versatile potential expression of the employed functional allows to adjust several properties of the molecular model: namely, the description of the torsions (periodic or harmonic), the inclusion of intramolecular non bonded interactions, the employment of coupling terms. As demonstrated, the protocol has offered the possibility to build the most accurate classical models existing in literature for a group of molecules, specifically suitable for the QM level precision which is required for spectroscopic application. The developed models have shown better results than standard literature force fields in predicting conformational changes of the compounds in both their ground and electronically excited states.

The attention to the intramolecular description of the investigated compounds has been combined to a reasoned description of their interaction with the environment and, in the case of the application to a polymer matrix, to the conception of an entire new model from scratch.

Specifically, the peculiar interplay between electronegative atoms – carbonyl oxygen in acrolein and pyridine nitrogen in nicotine – and hydrogens has been mimicked through a rational choice of explicit lone pairs; this has proven to be beneficial for the final spectroscopic description of acrolein, and for the excellent QM - MM agreement for what concerns the hydrogen bond dynamics in nicotine.

Moreover, the model for an apolar polyolefin has been set up, as a prototype for the description of a complex environment. The examination of the resulting trajectories for the probe + polymer has revealed a reliable prediction of physical chemical properties of the polymer chain, with particular regard to time-dependent ones (translational and rotational diffusion) which succeeded to describe the peculiar β-relaxation regime of this macromolecule.

Finally, exploiting the resulting MD trajectories, a statistical approach has been set up and used for the generation of UV-vis spectra of absorption (for acrolein and the NfO-TEMPO compound) and emission (for the latter only, both in toluene and grafted to the polyolefin).

In fact, the accurate MD simulations carried out with the DFT-derived force fields can be used to generate sets of statistically uncorrelated conformations of the target molecule and its neighbours. Computational spectroscopy techniques, based on TDDFT are then carried out on them. At this stage, different procedures for the inclusion of the solvent have been discussed, showing that an high level description is required for the first embedding shell, particularly when dealing with solvents as much polarizable as water.

For such cases, the role of polarization is analysed, by means of the fluctuating charge approach. At first, a recent implementation was tested for solvated systems and biological embeddings, especially to assess the dependence of the predicted properties on the thickness of the surrounding layer; it was eventually applied to the QM/FQ/PCM computation of the acrolein molecule's properties, demonstrating its crucial role.

The final "statistical spectra" obtained for the most extensively analysed case, NfO-TEMPO in different environments and electronic states, has all in all shown the main advantage of the presented protocol. Indeed, the simulation was validated by the comparison of the computed UV-visible spectra with their experimental counterparts. Calculated spectra were able to recognise the different effects, experimentally registered, on the broadening of the absorption and emission lines, namely their width, asymmetry and secondary peaks, with also good numerical comparability of the peak wavelengths and Stokes shifts. The differences between the environments have been highlighted and their origin connected with the hindered dynamics within the solvent and the polymer cages.

This work is presented as a further advance in setting up a sound protocol, able on the one hand to efficiently flank the experiment in the design of novel materials and, on the other hand, to help the identification of spectroscopically revealed properties of compounds with peculiar fingerprints.

Future applications of the proposed procedure could benefit from a better understanding of the different mechanisms that can alter the properties of organic dyes and similar classes of medium sized molecules. Indeed, The protocol was proven able to describe and correlate the evolution of their conformations, their vibrational and optical properties and their interaction with embeddings of increasing structural complexity and polarity.

Chapter 9

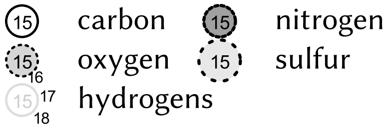
Appendix: force field parameters

All the force field parameters used in this thesis, that are not entirely taken from literature references, are enumerated in this appendix together with the indication of the source or of the method used to obtain them.

9.1 Legend

9.1.1 Chemical elements

The considered molecules will be sketched throughout the chapter, to map the atom numbers – used in the force fields – onto the molecular structure. The following color code is used to distinguish between the encountered elements:



9.1.2 Common sources

Parameters sources that have been used are referenced within each table. Sources used for many FF are reported by means of acronyms: their references are enumerated here.

Non bonded parameters

The source for the Van der Waals Lennard-Jones parameters is usually the OPLS force field. In this case, they are marked with OPLS: see [34, 66].

Charges used for Coulomb interactions can be taken from the same sources or generated directly from QM: the population analysis method is reported, either MULLI (Mulliken) or CM5 [57]. The population analysis are performed with Gaussian 09, development version [33], and may benefit from the PCM treatment [13]: the solvent is specified in this case.

Fluctuating charges parameters

The fluctuating charge parameters, namely η_i and χ_i, are either taken from the water models by Rick *et al.* (RICK [47]) or from CHARMM [46]. Even if off-diagonal η_{ij} electronegativities are provided from the source, they've been substituted with the arithmetic mean of η_i and η_j, for convenient use with earlier implementations of FQ in Gaussian 09.

Bonded parameters

The intramolecular parameters are always obtained through the Joyce procedure, profusely described along the thesis. The JOYCE label is used only when it's unclear.

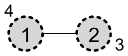
Refinements

The employed parameters may differ from the ones in the referenced papers, if a refinement was needed in order to match some physical property in condensed phase. In this case, a $\underline{\ }^{R}$ superscript is placed next to them.

9.2 Joyce validation set

The following parameters have been set up and used in chapter 4. Except for the last one (nicotine), none of them has been used for MD simulations; in facts, when intramolecular non bonded interactions have been neglected – or fall below the 1-4 bond distance, like in H-Peroxide –, charges and Lennard-Jones parameters are purely simbolic.

9.2.1 H-Peroxide

Atom numbering	Atom		Charge		Lennard-Jones		
	#	type	q (e)	source	ϵ (kJ/mol)	σ (nm)	source
	1	Op	-0.468	MULLI	0.290	0.586	OPLS
	2	Op	-0.468		0.290	0.586	
	3	Hp	0.468		0.250	0.126	
	4	Hp	0.468		0.250	0.126	

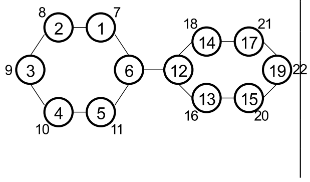
Harmonic term	Q^0_μ	k_μ
Stretching	(nm)	(kJ/mol nm^2)
1 2	0.1396	343427
2 3	0.0949	563241
1 4	0.0949	563241

Harmonic term	Q^0_μ	k_μ
Bending	(°)	(kJ/mol)
1 2 3	102.08	703.781
2 1 4	102.08	703.781

Periodic dihedral	n^μ_1	$k^d_{1\mu}$ (kJ/mol)	γ^μ_1 (°)
4 1 2 3	2	5.178	52.02

Non bonded intramolec.: all excluded (< 1-4)

9.2.2 Biphenyl

Atom numbering	Atom		Charge		Lennard-Jones		
	#	type	q (e)	source	ϵ (kJ/mol)	σ (nm)	source
	1, 5, 13, 14	C3	-0.218	MULLI	0.355	0.293	OPLS
	2, 4, 15, 17	C2	-0.202		0.355	0.293	
	6, 12	C4	0.019		0.355	0.293	
	3, 19	C1	-0.204		0.355	0.293	
	7, 11,16,18	H3	0.210		0.220	0.126	
	8, 10,20,21	H2	0.203		0.220	0.126	
	9, 22	H1	0.201		0.220	0.126	

Harmonic term	Q^0_μ	k_μ	Harmonic term	Q^0_μ	k_μ
Stretching	(nm)	(kJ/mol nm^2)	Bending	(°)	(kJ/mol)
C3 C2	0.1385	385892	C3 C2 C1	120.24	836.3809
C2 C1	0.1385	371801	C2 C3 C4	120.86	757.2595
C3 C4	0.1393	311679	C2 C3 H3	119.58	420.0504
C3 H3	0.1075	373653	C3 C2 H2	119.67	393.7594
C2 H2	0.1076	370952	C2 C1 C2	119.47	706.0529
C1 H1	0.1075	372543	C1 C2 H2	120.08	346.5010
C4 C4	0.1491	297697	C2 C1 H1	120.27	336.8249
			C3 C4 C3	118.33	650.7777
			C4 C3 H3	119.55	350.4985
			C3 C4 C4	120.84	666.9022

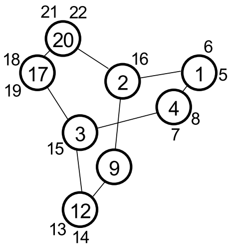
Periodic dihedral	n_1^μ	$k^d_{1\mu}$ (kJ/mol)	γ_1^μ (°)	Periodic dihedral	n_1^μ	$k^d_{1\mu}$ (kJ/mol)	γ_1^μ (°)
C3 C2 C1 C2	2	22.474	180.00	H3 C3 C4 C3	2	16.989	180.0
C4 C3 C2 C1	2	14.698	180.00	H3 C3 C4 C4	2	9.642	180.0
H3 C3 C2 C1	2	19.161	180.00	H3 C2 C4 C3	2	34.764	180.0
C3 C2 C1 H1	2	18.283	180.00	H2 C3 C1 C2	2	31.171	180.0
C2 C3 C4 C3	2	24.203	180.00	H1 C2 C2 C1	2	29.533	180.0
C4 C3 C2 H2	2	18.458	180.00	H2 C1 C3 C2	2	31.167	180.0
C2 C3 C4 C4	2	15.393	180.00	C4 C3 C3 C4	2	7.181	180.0
H3 C3 C2 H2	2	9.680	180.00	C4 C4 C3 C3	2	7.124	180.0
H2 C2 C1 C2	2	19.189	180.00	H3 C4 C2 C3	2	34.767	180.0
H2 C2 C1 H1	2	9.747	180.00				

(all the bonded pairs, triplets and quadruplets above are included and summarized by atomtypes)

Periodic dihedral	n_1^μ	$k^d_{1\mu}$ (kJ/mol)	γ_1^μ (°)	n_2^μ	$k^d_{2\mu}$	γ_2^μ
1 6 12 13 "main"	2	-4.064	0.0	4	0.693	0.0
5 6 12 14	2	-4.064	0.0	4	0.693	0.0

Non bonded intramolec.: 1-4 excluded, 1-5+ included, no shielding

9.2.3 Bicyclooctane

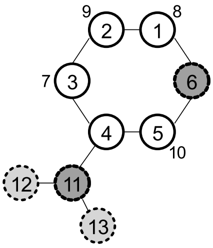
Atom numbering	Atom		Charge		Lennard-Jones		
	#	type	q (e)	source	ϵ (kJ/mol)	σ (nm)	source
	1, 4, 9, 12, 17, 20	C1	-0.31	MULLI	0.350	0.276	OPLS
	2	C2	-0.18		0.350	0.276	
	3	C2	-0.18		0.350	0.276	
	5, 6, 7, 8, 10, 11, 13, 14, 18, 19, 21, 22	H1	0.16		0.250	0.126	
	15	H2	0.16		0.250	0.126	
	16	H2	0.16		0.250	0.126	

Harmonic term	Q_μ^0	k_μ
Stretching	(nm)	(kJ/mol nm^2)
1 2	0.1536	223306
1 4	0.1552	222983
3 4	0.1536	223306
1 5	0.1086	339993
1 6	0.1086	339993
4 7	0.1087	339993
4 8	0.1086	339993
2 9	0.1536	223306
9 10	0.1086	339993
9 11	0.1086	339993
3 12	0.1536	223306
9 12	0.1552	222983
12 13	0.1087	339993
12 14	0.1086	339993
3 15	0.1086	342392
2 16	0.1086	342392
3 17	0.1536	223306
17 18	0.1086	339993
17 19	0.1086	339993
2 20	0.1536	223306
17 20	0.1552	222983
20 21	0.1086	339993
20 22	0.1086	339993
Bending	(°)	(kJ/mol)
2 1 4	109.68	706.538
2 1 5	109.81	409.664
2 1 6	109.54	409.664
1 2 9	109.18	774.888
1 2 16	109.75	440.428
1 2 20	109.20	774.888
1 4 3	109.68	706.538
4 1 5	110.66	471.249
4 1 6	110.64	471.249
1 4 7	110.63	471.249
1 4 8	110.65	471.249
3 4 7	109.55	409.664
3 4 8	109.81	409.664
4 3 12	109.19	774.888
4 3 15	109.76	440.428
4 3 17	109.18	774.888
5 1 6	106.46	381.522
7 4 8	106.46	381.522
2 9 10	109.80	409.664
2 9 11	109.54	409.664
2 9 12	109.68	706.538
9 2 16	109.76	440.428
9 2 20	109.17	774.888
10 9 11	106.46	381.522
10 9 12	110.65	471.249
11 9 12	110.64	471.249
3 12 9	109.68	706.538
3 12 13	109.55	409.664
3 12 14	109.80	409.664
12 3 15	109.76	440.428
12 3 17	109.19	774.888
9 12 13	110.64	471.249
9 12 14	110.66	471.249
13 12 14	106.46	381.522
15 3 17	109.75	440.428
16 2 20	109.76	440.428
3 17 18	109.82	409.664
3 17 19	109.52	409.664
3 17 20	109.68	706.538
18 17 19	106.46	381.522
18 17 20	110.66	471.249
19 17 20	110.64	471.249
2 20 17	109.68	706.538
2 20 21	109.53	409.664
2 20 22	109.82	409.664
17 20 21	110.63	471.249
17 20 22	110.66	471.249
21 20 22	106.46	381.522

Periodic dihedral	n_1^μ	$k_{1\mu}^d$ (kJ/mol)	γ_1^μ (°)	Periodic dihedral	n_1^μ	$k_{1\mu}^d$ (kJ/mol)	γ_1^μ (°)
3 1 4 3	3	12.844	0.0	16 2 9 10,11	3	0.040	0.0
2 1 4 7, 8	3	-1.742	0.0	20 2 9 10,11	3	1.572	0.0
4 1 2 9	3	7.401	0.0	2 9 12 3	3	12.844	0.0
4 1 2 16	3	2.692	0.0	2 9 12 13,14	3	-1.742	0.0
4 1 2 20	3	7.401	0.0	16 2 9 12	3	2.692	0.0
5, 6 1 2 9	3	1.572	0.0	20 2 9 12	3	7.401	0.0
5, 6 1 2 16	3	0.040	0.0	9 2 20 17	3	7.401	0.0
5, 6 1 2 20	3	1.572	0.0	9 2 20 21,22	3	1.572	0.0
1 2 9 10,11	3	1.572	0.0	10,11 9 12 3	3	-1.742	0.0
1 2 9 12	3	7.401	0.0	10,11 9 12 13,14	3	-1.394	0.0
1 2 20 17	3	7.401	0.0	15 3 12 9	3	2.692	0.0
1 2 20 21,22	3	1.572	0.0	17 3 12 9	3	7.401	0.0
5, 6 1 4 3	3	-1.742	0.0	15 3 12 13,14	3	0.040	0.0
12 3 4 1	3	7.401	0.0	17 3 12 13,14	3	1.572	0.0
15 3 4 1	3	2.692	0.0	12 3 17 18,19	3	1.572	0.0
17 3 4 1	3	7.401	0.0	12 3 17 20	3	7.401	0.0
5, 6 1 4 7, 8	3	-1.394	0.0	15 3 17 18,19	3	0.040	0.0
12 3 4 7, 8	3	1.572	0.0	15 3 17 20	3	2.692	0.0
15 3 4 7, 8	3	0.040	0.0	16 2 20 17	3	2.692	0.0
17 3 4 7, 8	3	1.572	0.0	16 2 20 21,22	3	0.040	0.0
4 3 12 9	3	7.401	0.0	3 17 20 2	3	12.844	0.0
4 3 12 13,14	3	1.572	0.0	3 17 20 21,22	3	-1.742	0.0
4 3 17 18,19	3	1.572	0.0	18,19 17 20 2	3	-1.742	0.0
4 3 17 20	3	7.401	0.0	18,19 17 20 21,22	3	-1.394	0.0

Non bonded intramolec.: all excluded

9.2.4 Nitropyridine

Atom numbering	Atom		Charge		Lennard-Jones		
	#	type	q (e)	source	ϵ (kJ/mol)	σ (nm)	source
	1	C1	0.095	MULLI	0.355	0.292	OPLS
	2	C2	-0.275		0.355	0.292	
	3	C3	-0.102		0.355	0.292	
	4	C2	0.093		0.355	0.292	
	5	C1	0.109		0.355	0.292	
	6	Np	-0.522		0.325	0.711	
	7	H	0.282		0.242	0.125	
	8	H	0.225		0.242	0.125	
	9	H	0.233		0.242	0.125	
	10	H	0.271		0.242	0.125	
	11	No	0.520		0.325	0.711	
	12	On	-0.468		0.290	0.585	
	13	On	-0.463		0.290	0.585	

Harmonic term	Q^0_μ	k_μ
Stretching	(nm)	(kJ/mol nm^2)
1 2	0.1387	381334
2 3	0.1380	366175
3 4	0.1382	366175
4 5	0.1385	381334
1 6	0.1322	416151
5 6	0.1318	416151
3 7	0.1072	386444
1 8	0.1075	380381
2 9	0.1073	379034
5 10	0.1072	380381
4 11	0.1452	202652
11 12	0.1194	690690
11 13	0.1193	690690

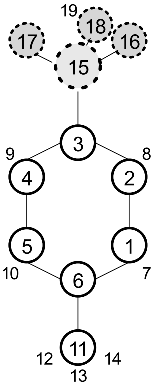
Harmonic term	Q^0_μ	k_μ
Bending	(°)	(kJ/mol)
1 2 3	118.33	484.504
2 1 6	123.94	1065.812
2 1 8	120.14	324.675
1 2 9	120.55	452.424
2 3 4	117.25	805.174
2 3 7	122.27	330.887
3 2 9	121.12	308.914
3 4 5	120.52	484.504
4 3 7	120.49	330.887
3 4 11	119.79	900.377
4 5 6	121.80	1065.812
4 5 10	120.57	324.675
5 4 11	119.69	739.463
1 6 5	118.16	536.838
6 1 8	115.92	479.173
6 5 10	117.63	479.173
4 11 12	117.43	209.731
4 11 13	117.61	209.731
12 11 13	124.95	1662.133

Periodic dihedral	n^μ_1	$k^d_{1\mu}$ (kJ/mol)	γ^μ_1 (°)	Periodic dihedral	n^μ_1	$k^d_{1\mu}$ (kJ/mol)	γ^μ_1 (°)
1 2 3 4	2	5.596	180.00	9 2 3 7	2	12.967	180.00
6 1 2 3	2	18.475	180.00	3 4 5 6	2	18.475	180.00
1 2 3 7	2	20.688	180.00	7 3 4 5	2	20.688	180.00
8 1 2 3	2	26.639	180.00	3 4 5 10	2	26.639	180.00
2 1 6 5	2	21.587	180.00	7 3 4 11	2	18.174	180.00
6 1 2 9	2	30.441	180.00	4 5 6 1	2	21.587	180.00
8 1 2 9	2	17.412	180.00	11 4 5 6	2	76.277	180.00
2 3 4 5	2	5.596	180.00	11 4 5 10	2	23.080	180.00
9 2 3 4	2	8.212	180.00	8 1 6 5	2	27.624	180.00
2 3 4 11	2	47.555	180.00	10 5 6 1	2	27.624	180.00

Periodic dihedral	n^μ_1	$k^d_{1\mu}$ (kJ/mol)	γ^μ_1 (°)	n^μ_2	$k^d_{2\mu}$	γ^μ_2
5 4 11 13 “main”	2	-12.035	0.0	4	0.126	0.0
3 4 11 12	2	-12.035	0.0	4	0.126	0.0

Non bonded intramolec.: 1-4+ included for atomtypes H, No, On

9.2.5 Toluensulfonic acid

Atom numbering	Atom		Charge		Lennard-Jones		
	#	type	q (e)	source	ϵ (kJ/mol)	σ (nm)	source
	1	C3	-0.233	MULLI	0.3550	0.29288	OPLS
	2	C2	-0.164		0.3550	0.29288	
	3	C1	-0.329		0.3550	0.29288	
	4	C2	-0.155		0.3550	0.29288	
	5	C3	-0.224		0.3550	0.29288	
	6	C4	0.059		0.3550	0.29288	
	7	H	0.214		0.2420	0.12555	
	8	H	0.258		0.2420	0.12555	
	9	H	0.260		0.2420	0.12555	
	10	H	0.215		0.2420	0.12555	
	11	CT	-0.513		0.3550	0.29288	
	12	HT	0.186		0.2420	0.12555	
	13	HT	0.192		0.2420	0.12555	
	14	HT	0.182		0.2420	0.12555	
	15	So	1.629		0.3550	0.10460	
	16	Os	-0.674		0.2960	0.87864	
	17	Os	-0.633		0.2960	0.87864	
	18	Oh	-0.766		0.2960	0.87864	
	19	Ho	0.496		0.0000	0.00000	

Harmonic term	Q^0_μ	k_μ
Stretching	(nm)	(kJ/mol nm^2)
1 2	0.1385	375987
2 3	0.1383	357513
3 4	0.1389	357513
4 5	0.1379	375987
1 6	0.1388	361182
5 6	0.1394	361182
1 7	0.1075	371947
2 8	0.1073	381470
4 9	0.1073	381470
5 10	0.1076	371947
6 11	0.1510	242518
11 12	0.1085	348746
11 13	0.1086	348746
11 14	0.1083	348746
3 15	0.1759	258548
15 16	0.1428	694482
15 17	0.1419	694482
15 18	0.1591	304804
18 19	0.0954	550906
Bending	(°)	(kJ/mol)
1 2 3	119.22	722.160
2 1 6	120.97	798.764
2 1 7	119.22	325.948
1 2 8	120.61	436.086
2 3 4	120.89	418.705

Harmonic term	Q^0_μ	k_μ
Bending	(°)	(kJ/mol)
3 2 8	120.16	340.722
2 3 15	119.74	776.059
3 4 5	119.17	722.160
3 4 9	120.11	340.722
4 3 15	119.37	776.059
4 5 6	121.04	798.764
5 4 9	120.72	436.086
4 5 10	119.32	325.948
1 6 5	118.71	537.138
6 1 7	119.80	409.359
1 6 11	121.12	567.415
6 5 10	119.64	409.359
5 6 11	120.17	567.415
6 11 12	111.01	429.848
6 11 13	110.65	429.848
6 11 14	111.42	429.848
12 11 13	107.56	397.794
12 11 14	108.14	397.794
13 11 14	107.91	397.794
3 15 16	109.23	923.707
3 15 17	108.90	923.707
3 15 18	103.06	771.251
16 15 17	121.57	1129.163
16 15 18	106.00	997.713
17 15 18	106.46	997.713
15 18 19	109.46	473.295

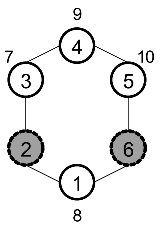
Periodic dihedral	n_1^μ	$k_{1\mu}^d$ (kJ/mol)	γ_1^μ (°)	Periodic dihedral	n_1^μ	$k_{1\mu}^d$ (kJ/mol)	γ_1^μ (°)
1 2 3 4	2	22.344	180.0	3 4 5 6	2	11.117	180.0
6 1 2 3	2	11.117	180.0	3 4 5 10	2	12.218	180.0
7 1 2 3	2	12.218	180.0	15 3 4 5	2	29.654	180.0
1 2 3 15	2	29.654	180.0	15 3 4 9	2	21.418	180.0
2 1 6 5	2	36.603	180.0	4 5 6 1	2	36.603	180.0
6 1 2 8	2	15.063	180.0	9 4 5 6	2	15.063	180.0
2 1 6 11	2	31.411	180.0	4 5 6 11	2	31.411	180.0
7 1 2 8	2	11.506	180.0	9 4 5 10	2	11.506	180.0
2 3 4 5	2	22.344	180.0	7 1 6 5	2	20.114	180.0
8 2 3 4	2	20.246	180.0	10 5 6 1	2	20.114	180.0
2 3 4 9	2	20.246	180.0	7 1 6 11	2	22.423	180.0
8 2 3 15	2	21.418	180.0	10 5 6 11	2	22.423	180.0

Periodic dihedral	n_1^μ	$k_{1\mu}^d$ (kJ/mol)	γ_1^μ (°)	n_2^μ	$k_{2\mu}^d$	γ_2^μ	n_3^μ	$k_{3\mu}^d$	γ_3^μ	n_4^μ	$k_{4\mu}^d$	γ_4^μ
2 3 15 16 “δ_1”	1	0.088	0.0	2	-5.375	0.0	3	0.061	0.0	4	-54.68	0.0
4 3 15 17	1	0.088	0.0	2	-5.375	0.0	3	0.061	0.0	4	-54.68	0.0
3 15 18 19 “δ_2”	1	8.657	0.0	2	5.339	0.0	3	1.580	0.0	4	0.235	0.0
1 6 1 2 “δ_3”	3	0.309	0.0									
1 6 1 3	3	0.309	0.0									
1 6 1 4	3	0.309	0.0									
5 6 1 2	3	0.309	0.0									
5 6 1 3	3	0.309	0.0									
5 6 1 4	3	0.309	0.0									

Non bonded intramolec.: 1-4+ included for atomtypes H, So, Oh, Ho

9.3 FFPol for pyrimidine

Pyrimidine was described by means of a “FFPol” force field, that exploits Joyce level valence terms and FQ charges: see section 2.5.2. Instead of charges and Lennard-Jones parameters, FQ parameters are reported.

Atom numbering	Atom		FQ		
	#	type	χ (kcal/mol e)	η (kcal/mol e)	source
	1	C1	332.37	225.48	CHARMM
	2	N	396.73	246.94	
	3	C2	332.37	225.48	
	4	C3	332.37	225.48	
	5	C2	332.37	225.48	
	6	N	396.73	246.94	
	7	H2	319.46	454.14	
	8	H1	319.46	454.14	
	9	H3	319.46	454.14	
	10	H2	319.46	454.14	

Harmonic term	Q^0_μ	k_μ	Harmonic term	Q^0_μ	k_μ	source
Stretching	(nm)	(kJ/mol nm^2)	Bending	(°)	(kJ/mol)	
1 2	0.1339	293972	1 2 3	115.56	290.453	Joyce
2 3	0.1340	364206	2 1 6	127.57	1394.318	
3 4	0.1395	345097	2 1 8	116.21	336.879	
4 5	0.1395	345097	2 3 4	122.43	968.791	
1 6	0.1339	293972	2 3 7	116.46	503.928	
5 6	0.1340	364206	3 4 5	116.44	424.568	
3 7	0.1096	327510	4 3 7	121.11	247.427	
1 8	0.1096	338720	3 4 9	121.78	306.287	
4 9	0.1092	337901	4 5 6	122.44	968.791	
5 10	0.1096	327510	5 4 9	121.78	306.287	
			4 5 10	121.11	247.427	
Impr. dihedral	(°)	(kJ/mol)	1 6 5	115.56	290.453	
8 2 6 1	0.0	231.174	6 1 8	116.21	336.879	
7 2 4 3	0.0	219.442	6 5 10	116.46	503.928	
9 3 5 4	0.0	124.890				
10 4 6 5	0.0	219.438				

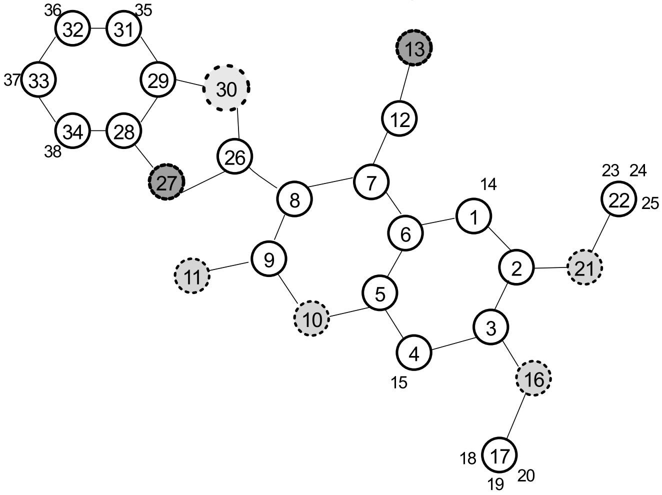
Periodic dihedral	n^μ_1	$k^d_{1\mu}$ (kJ/mol)	γ^μ_1 (°)	Periodic dihedral	n^μ_1	$k^d_{1\mu}$ (kJ/mol)	γ^μ_1 (°)
1 2 3 4	2	16.733	180.0	7 3 4 5	2	16.019	180.0
6 1 2 3	2	8.424	180.0	3 4 5 10	2	16.019	180.0
1 2 3 7	2	22.253	180.0	7 3 4 9	2	6.135	180.0
8 1 2 3	2	20.554	180.0	4 5 6 1	2	16.733	180.0
2 1 6 5	2	8.424	180.0	9 4 5 6	2	13.833	180.0
2 3 4 5	2	15.648	180.0	9 4 5 10	2	6.135	180.0
2 3 4 9	2	13.833	180.0	8 1 6 5	2	20.554	180.0
3 4 5 6	2	15.648	180.0	10 5 6 1	2	22.253	180.0

All the intramolecular electrostatic treated through the FQ approach.
No LJ pairs generated.

9.4 Coumarin-based dye

Two force fields were generated, for just as many considered electronic states (S0, S1), for the coumarin-based dye. They share the definition of atom names and LJ parameters, while charges and valence intramolecular terms vary.

Atom numbering:



Atom		Charges			Lennard-Jones		
#	type	q$_{S0}$ (e)	q$_{S1}$ (e)	source	ϵ (kJ/mol)	σ (nm)	source
1	C2	-0.476	-0.399		0.355	0.292	
2	C1	0.520	0.511		0.355	0.292	
3	C1	0.390	0.377		0.355	0.292	
4	C2	-0.090	-0.083		0.355	0.292	
5	Cx	-0.022	-0.053		0.355	0.292	
6	C3	0.160	0.199		0.355	0.292	
7	C4	0.720	0.712		0.355	0.317	
8	C5	0.381	0.311		0.355	0.317	
9	Co	0.744	0.726	MULLI	0.375	0.439	OPLS
10	O6	-0.908	-0.935		0.307	0.711	
11	Oc	-0.546	-0.547		0.296	0.878	
12	Cn	0.507	0.587		0.355	0.292	
13	N	-0.965	-1.095		0.325	0.711	
14	H2	0.034	0.040		0.242	0.125	
15	H2	0.091	0.092		0.242	0.125	
16	O	-0.878	-0.879		0.290	0.585	
17	Ct	0.099	0.100		0.350	0.276	

Atom		Charges			Lennard-Jones		
#	type	q$_{S0}$ (e)	q$_{S1}$ (e)	source	ϵ (kJ/mol)	σ (nm)	source
18	Ht	0.119	0.122	MULLI	0.250	0.125	OPLS
19	Ht	0.119	0.122		0.250	0.125	
20	Ht	0.123	0.124		0.250	0.125	
21	O	-0.926	-0.896		0.290	0.585	
22	Ct	0.124	0.136		0.350	0.276	
23	Ht	0.125	0.135		0.250	0.125	
24	Ht	0.125	0.135		0.250	0.125	
25	Ht	0.118	0.123		0.250	0.125	
26	C7	0.235	0.246		0.350	0.276	
27	N5	-0.364	-0.388		0.325	0.711	
28	C8	0.833	0.824		0.355	0.292	
29	C9	-0.477	-0.521		0.355	0.292	
30	S	0.174	0.192		0.355	1.046	
31	Ca	-0.241	-0.222		0.355	0.292	
32	Ca	-0.111	-0.121		0.355	0.292	
33	Ca	-0.089	-0.064		0.355	0.292	
34	Ca	-0.002	0.011		0.355	0.292	
35	Ha	0.089	0.094		0.242	0.125	
36	Ha	0.081	0.087		0.242	0.125	
37	Ha	0.076	0.080		0.242	0.125	
38	Ha	0.102	0.111		0.242	0.125	

Stretching		S0 parameters		S1 parameters	
		r^0 (nm)	k^s (kJ/mol nm^2)	r^0 (nm)	k^s (kJ/mol nm^2)
1	2	0.1380	333651	0.1390	316021
2	3	0.1431	256187	0.1437	221778
3	4	0.1390	333651	0.1387	316021
4	5	0.1395	321706	0.1398	300927
1	6	0.1417	299309	0.1406	274203
5	6	0.1397	295487	0.1414	270070
6	7	0.1442	244664	0.1446	177878
7	8	0.1384	350057	0.1431	271035
8	9	0.1469	186524	0.1444	203588
5	10	0.1362	298360	0.1358	301277
9	10	0.1377	201490	0.1403	150353
9	11	0.1211	665806	0.1216	629345
7	12	0.1435	291632	0.1408	309784
12	13	0.1162	1137905	0.1172	1023042
1	14	0.1083	342563	0.1083	341591
4	15	0.1082	342563	0.1083	341591
3	16	0.1346	324115	0.1345	326968
16	17	0.1426	256262	0.1427	239517

Stretching		S0 parameters		S1 parameters	
		r^0 (nm)	k^s (kJ/mol nm^2)	r^0 (nm)	k^s (kJ/mol nm^2)
17	18	0.1097	313711	0.1097	315469
17	19	0.1097	313711	0.1097	315469
17	20	0.1091	313711	0.1090	315469
2	21	0.1355	324115	0.1345	326968
21	22	0.1424	256262	0.1430	239517
22	23	0.1097	313711	0.1097	315469
22	24	0.1097	313711	0.1096	315469
22	25	0.1091	313711	0.1090	315469
8	26	0.1460	245013	0.1432	278676
26	27	0.1303	406852	0.1326	341530
27	28	0.1374	300370	0.1358	316363
28	29	0.1415	194342	0.1423	207208
26	30	0.1787	147945	0.1788	142032
29	30	0.1752	210414	0.1762	213260
29	31	0.1398	282563	0.1388	326356
31	32	0.1390	322824	0.1403	283078
32	33	0.1409	322824	0.1406	283078
28	34	0.1405	301335	0.1414	310957
33	34	0.1386	322824	0.1387	283078
31	35	0.1086	333869	0.1086	334376
32	36	0.1087	333869	0.1087	334376
33	37	0.1086	333869	0.1086	334376
34	38	0.1086	333869	0.1086	334376

Bending			S0 parameters		S1 parameters	
			θ^0 (°)	k^b (kJ/mol nm^2)	θ^0 (°)	k^b (kJ/mol nm^2)
1	2	3	119.63	327.621	119.89	314.1379
2	1	6	121.08	539.925	121.58	644.4307
2	1	14	120.49	305.600	120.35	304.2934
1	2	21	125.18	701.010	124.80	830.8288
2	3	4	119.83	327.621	119.08	314.1379
2	3	16	115.57	493.664	115.68	464.7691
3	2	21	115.19	493.664	115.30	464.7691
3	4	5	119.37	519.931	119.97	511.5862
3	4	15	122.15	305.600	122.07	304.2934
4	3	16	124.60	701.010	125.25	830.8288
4	5	6	122.11	165.825	122.16	315.9820
4	5	10	116.67	865.794	115.70	897.4059
5	4	15	118.48	337.206	117.96	327.8010
1	6	5	117.99	524.404	117.32	404.7077
1	6	7	124.20	642.508	123.25	668.9554
6	1	14	118.43	346.141	118.07	325.7872
5	6	7	117.81	643.989	119.43	711.8025
6	5	10	121.22	522.473	122.14	537.5053
6	7	8	121.10	229.639	117.69	199.4782
6	7	12	115.28	554.763	116.30	420.5716
9	8	26	118.05	800.819	117.96	530.6359

Bending			S0 parameters		S1 parameters	
			θ^0 (°)	k^b (kJ/mol nm^2)	θ^0 (°)	k^b (kJ/mol nm^2)
7	8	9	118.98	544.095	120.65	511.5534
8	7	12	123.61	380.539	126.00	451.2584
7	8	26	122.97	669.601	121.39	558.6880
8	9	10	117.65	216.185	118.78	311.2603
8	9	11	125.21	172.049	126.31	178.8570
5	10	9	123.24	570.350	121.32	513.4090
10	9	11	117.14	998.799	114.92	958.9179
7	12	13	169.28	352.381	168.91	371.7686
3	16	17	118.71	677.997	118.62	604.7823
16	17	18	111.06	494.374	110.93	495.4111
16	17	19	111.06	494.374	110.92	495.4111
16	17	20	105.60	494.374	105.61	495.4111
18	17	19	109.79	312.742	109.79	313.9696
18	17	20	109.62	312.742	109.75	313.9696
19	17	20	109.62	312.742	109.76	313.9696
2	21	22	118.08	677.997	118.99	604.7823
21	22	23	111.20	494.374	110.78	495.4111
21	22	24	111.21	494.374	110.84	495.4111
21	22	25	105.67	494.374	105.44	495.4111
23	22	24	109.51	312.742	109.73	313.9696
23	22	25	109.59	312.742	109.96	313.9696
24	22	25	109.59	312.742	110.02	313.9696
8	26	27	121.98	160.344	121.81	362.1984
8	26	30	123.06	200.270	123.49	202.9364
26	27	28	112.13	378.196	112.03	496.8849
27	26	30	114.96	1102.827	114.70	964.3459
27	28	29	114.97	589.692	115.56	688.6182
27	28	34	125.13	777.355	124.78	658.7802
28	29	30	109.76	1104.236	109.54	758.4330
28	29	31	121.22	273.143	121.17	73.3091
29	28	34	119.91	270.262	119.66	436.0720
26	30	29	88.18	697.083	88.18	829.7803
30	29	31	129.02	789.753	129.29	927.1016
29	31	32	118.06	600.793	118.30	573.2685
29	31	35	121.15	309.734	121.34	287.4537
31	32	33	121.15	678.380	121.20	669.4015
32	31	35	120.79	326.140	120.36	327.1108
31	32	36	119.38	326.140	119.19	327.1108
32	33	34	120.96	678.380	120.73	669.4015
33	32	36	119.47	326.140	119.61	327.1108
32	33	37	119.32	326.140	119.43	327.1108
28	34	33	118.70	670.242	118.94	568.6176
28	34	38	119.38	305.483	118.99	286.3264
34	33	37	119.72	326.140	119.84	327.1108
33	34	38	121.92	326.140	122.08	327.1108

Equilibrium values for harmonic dihedrals (ϕ^0 in Eq. 2.1) and phase values for periodic ones (γ_j) are zero for all the following data.

Harm. tors.	S0 k^t (kJ/mol)	S1
1 2 3 4	89.940	90.816
2 3 4 5	89.940	90.816
3 4 5 6	89.940	90.816
4 5 6 1	89.940	90.816
2 1 6 5	89.940	90.816
6 1 2 3	89.940	90.816
14 1 2 21	64.101	60.535
16 3 4 15	64.101	60.535
21 2 3 16	186.456	185.890
15 4 5 10	62.313	71.937
14 1 6 7	81.276	92.242
5 6 7 8	86.564	74.613
6 7 8 9	86.564	74.613
7 8 9 10	23.344	30.616
8 9 10 5	48.789	31.461
6 5 10 9	48.789	31.461
10 5 6 7	23.344	30.616
1 6 7 12	86.564	74.613
12 7 8 26	86.564	74.613
26 7 9 8	161.885	156.282
9 8 11 10	600.253	459.932
7 6 8 13	80.236	37.671
26 27 28 29	129.184	149.429
27 28 29 30	249.725	259.868
28 29 30 26	232.120	207.175
27 26 30 29	237.190	187.305
30 26 27 28	179.808	148.197
26 8 27 30	161.885	156.282
28 29 31 32	96.495	96.432
29 31 32 33	96.495	96.432
31 32 33 34	96.495	96.432
32 33 34 28	96.495	96.432
29 28 34 33	96.495	96.432
34 28 29 31	96.495	96.432
27 28 34 38	118.280	114.472
37 33 34 38	62.345	61.701
30 29 31 35	118.280	114.472
35 31 32 36	62.345	61.701
36 32 33 37	62.345	61.701

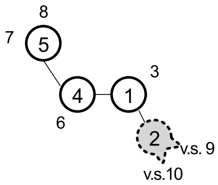
Periodic dihedral	n_j^μ	S0 $k_{j\mu}^d$	S1
$\binom{9}{7}$ 8 26 $\binom{27}{30}$	1	6.829	5.187
	2	-0.650	-6.795
	3	1.648	2.092
	4	0.615	0.777
1 2 21 22	1	-6.844	-4.774
	2	-1.128	-2.916
	3	-3.731	-2.748
	4	-0.109	-1.751
4 3 16 17	1	-6.844	-4.774
	2	-1.128	-2.916
	3	-3.731	-2.748
	4	-0.109	-1.751
3 16 17 18/19/20	3	4.855	4.173
2 21 22 23/24/25	3	4.855	4.173

Non bonded intramolec.: only Ct atoms.

9.5 Vacff and Pcmff models for acrolein

The two force fields are reported for the flexible description of the acrolein molecule, based respectively on the hessian matrix *in vacuo* and in pcm. The hessian matrices were calculated at dft level as reported in chapter 5. However, the same charges were employed for the two simulations, as well as the same dihedral description.

The spc water model, used for the solvated simulations, is not reported here.

Atom numbering	Atom		Charge		Lennard-Jones		
	#	type	q (e)	source	ϵ (kJ/mol)	σ (nm)	source
	1	CO	0.114	mulli	0.375	0.439	opls
	2	O	0.0		0.296	0.878	
	3	H	0.105		0.242	0.062	
	4	C1	-0.095		0.355	0.317	
	5	C2	-0.147		0.355	0.317	
	6	H1	0.118		0.242	0.062	
	7	H2	0.111		0.242	0.125	
	8	H2	0.107		0.242	0.125	
	9	XO	-0.157		0.0	0.0	
	10	XO	-0.157		0.0	0.0	

Stretching		vacff parameters		pcmff parameters	
		r^0 (nm)	k^s (kJ/mol nm^2)	r^0 (nm)	k^s (kJ/mol nm^2)
1	2	0.1211	659757	0.1219	605040
1	3	0.1111	271860	0.1107	280048
1	4	0.1474	226320	0.1466	239223
4	5	0.1335	489530	0.1338	490343
4	6	0.1086	328492	0.1085	329230
5	7	0.1084	331551	0.1083	332562
5	8	0.1086	331551	0.1086	332562

Bending			vacff parameters		pcmff parameters	
			θ^0 (nm)	k^b (kJ/mol nm^2)	θ^0 (nm)	k^b (kJ/mol nm^2)
2	1	3	120.84	503.5880	120.41	500.9884
2	1	4	124.33	734.4134	124.39	765.0927
3	1	4	114.83	246.6989	115.20	233.1417
1	4	5	121.09	595.0939	120.57	598.7081
1	4	6	116.58	309.0747	117.18	305.6475
5	4	6	122.33	329.6074	122.25	332.8363
4	5	7	122.19	339.2822	121.94	337.2557
4	5	8	121.00	339.2822	121.03	337.2557
7	5	8	116.81	249.1921	117.02	251.7031

Impr.dih.	vacff parameters		pcmff parameters	
	ϕ^0 (nm)	k^b (kJ/mol nm^2)	ϕ^0 (nm)	k^b (kJ/mol nm^2)
4 2 3 1	0.0	501.800	*same*	
4 1 5 6	0.0	403.159		
8 4 7 5	0.0	200.449		

Periodic dihedral	n_1^μ	$k_{1\mu}^d$ (kJ/mol)	γ_1^μ (°)	n_2^μ	$k_{2\mu}^d$	γ_2^μ	n_3^μ	$k_{3\mu}^d$	γ_3^μ	n_4^μ	$k_{4\mu}^d$	γ_4^μ
$\binom{1}{6}$ 4 5 $\binom{7}{8}$	2	26.515	180.0									
$\binom{2}{3}$ 1 4 $\binom{5}{6}$	1	1.282	180.0	2	7.800	180.0	3	0.982	180.00	4	-0.410	180.00

Non bonded intramolec: all excluded.

Bibliography

[1] Barone, V.; Improta, R.; Rega, N. *Acc. Chem. Res.* **2008**, *41*(5), 605–616.

[2] Dykstra, C.; Frenking, G.; Kim, K.; Scuseria, G. *Theory and applications of computational chemistry: the first forty years;* Elsevier, 2011.

[3] Neese, F. *Coord. Chem. Rev.* **2009**, *253*(5), 526–563.

[4] Adamo, C.; Jacquemin, D. *Chem. Soc. Rev.* **2013**, *42*(3), 845–56.

[5] Autschbach, J.; Ziegler, T.; van Gisbergen, S. J.; Baerends, E. J. *J. Chem. Phys.* **2002**, *116*(16), 6930–6940.

[6] Crawford, T. D. *Theor. Chem. Acc.* **2006**, *115*(4), 227–245.

[7] Van Dishoeck, E. F. *arXiv preprint astro-ph/0403061* **2004**.

[8] Szymanski, W.; Beierle, J. M.; Kistemaker, H. A.; Velema, W. A.; Feringa, B. L. *Chem. Rev.* **2013**, *113*(8), 6114–6178.

[9] Canuto, S. *Solvation effects on molecules and biomolecules: computational methods and applications*, Vol. 6; Springer Science & Business Media, 2010.

[10] Ow, H.; Larson, D. R.; Srivastava, M.; Baird, B. A.; Webb, W. W.; Wiesner, U. *Nano Lett.* **2005**, *5*(1), 113–117.

[11] Demaison, J.; Boggs, J. E.; Császár, A. G. *Equilibrium molecular structures: From spectroscopy to quantum chemistry;* CRC Press, 2010.

[12] Manzoni, V.; Lyra, M. L.; Gester, R. M.; Coutinho, K.; Canuto, S. *Physical Chemistry Chemical Physics* **2010**, *12*(42), 14023–14033.

[13] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 4708.

[14] Malcıoglu, O. B.; Calzolari, A.; Gebauer, R.; Varsano, D.; Baroni, S. *J. Am. Chem. Soc.* **2011**, *133*(39), 15425–15433.

[15] Hoffmann, M.; Wanko, M.; Strodel, P.; König, P. H.; Frauenheim, T.; Schulten, K.; Thiel, W.; Tajkhorshid, E.; Elstner, M. *J. Am. Chem. Soc.* **2006**, *128*(33), 10808–10818.

[16] Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*(7), 1198–1229.

[17] Rega, N.; Cossi, M.; Barone, V. *J. Am. Chem. Soc.* **1998**, *120*(23), 5723–5732.

[18] Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. *J. Phys. Chem. B* **2011**, *115*(12), 3027–3037.

[19] Brancato, G.; Barone, V.; Rega, N. *Theor. Chem. Acc.* **2007**, *117*(5-6), 1001–1015.

[20] Rappe, A. K.; III, W. A. G. *J. Phys. Chem.* **1991**, *95*, 3358.

[21] Lipparini, F. PhD thesis, **2013**.

[22] Lipparini, F.; Cappelli, C.; Barone, V. *Journal of Chemical Theory and Computation* **2012**, *8*(11), 4153–4165.

[23] Lipparini, F.; Cappelli, C.; Scalmani, G.; De Mitri, N.; Barone, V. *Journal of Chemical Theory and Computation* **2012**, *8*(11), 4270–4278.

[24] Barone, V.; Bloino, J.; Biczysko, M.; Santoro, F. *J. Chem. Theory Comput.* **2009**, *5*, 540.

[25] Bloino, J.; Biczysko, M.; Santoro, F.; Barone, V. *J. Chem. Theory Comput.* **2010**, *6*, 1256.

[26] Prampolini, G.; Bellina, F.; Biczysko, M.; Cappelli, C.; Carta, L.; Lessi, M.; Pucci, A.; Ruggeri, G.; Barone, V. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2013**, *19*(6), 1996–2004.

[27] Cacelli, I.; Prampolini, G. *J. Chem. Theory Comput.* **2007**, *3*, 1803.

[28] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*, Oxford science publications; Oxford University Press, USA, 1989.

[29] McQuarrie, D. *Statistical Thermodynamics;* University Science Books: Mill Valley, CA, 1973.

[30] Mackerell, A. D. *J. Comp. Chem.* **2004**, *25*(13), 1584–1604.

[31] Burger, S.; Lacasse, M.; Verstraelen, T.; Drewry, J.; Gunning, P.; Ayers, P. *J. Chem. Theory Comput.* **2012**, *8*, 554.

[32] Cacelli, I.; Prampolini, G. *J. Chem. Theory Comput.* **2007**, *3*, 1803.

[33] Gaussian09, revision c.01. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.; Iyengar, S. S.; Tomasi, J.; .; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Parandekar, P. V.; Mayhall, N. J.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fo, D. J.; Gaussian, Inc., Wallingford CT, **2009**.

[34] Jorgensen, W.; Maxwell, D.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

[35] van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P.; Tieleman, D.; Sijbers, A.; Feenstra, K.; van Drunen, R.; Berendsen, H. *GROMACS4.5;* Gromacs User Manual version 4.5.4: ww.gromacs.org, 2010.

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

[37] Salvadori, A.; Licari, D.; Mancini, G.; Brogni, A.; De Mitri, N.; Barone, V. In *"Reference Module in Chemistry, Molecular Sciences and Chemical Engineering"*; Reedijk, J., Ed.; Springer: Amsterdam, Netherlands, 2014.

[38] Maple, J. R.; Dinur, U.; Hagler, A. T. *Proc. Natl. Acad. Sci. USA* **1988**, *85*(15), 5350–5354.

[39] Jorgensen, W. L. *J. Am. Chem. Soc.* **1981**, *103*, 335–350.

[40] Wang, L.-P.; Martinez, T. J.; Pande, V. S. *J. Phys. Chem. Lett.* **2014**, *5*(11), 1885–1891.

[41] Ramirez, R.; Lopez-Ciudad, T.; Kumar P, P.; Marx, D. *The Journal of Chemical Physics* **2004**, *121*(9), 3973–3983.

[42] Steiner, T. *Angew. Chem. Int. Ed.* **2002**, *41*(1), 48–76.

[43] Rapaport, D. *Mol. Phys.* **1983**, *50*(5), 1151–1162.

[44] Pagliai, M.; Cardini, G.; Righini, R.; Schettino, V. *J. Chem. Phys.* **2003**, *119*(13), 6655–6662.

[45] Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. *J. Phys. Condens. Matter* **2009**, *21*(33), 333102.

[46] Patel, S.; Brooks, C. L. *J. Comp. Chem.* **2004**, *25*, 1.

[47] Rick, S. W.; Stuart, S. J.; Berne, B. J. *The Journal of Chemical Physics* **1994**, *101*(7), 6141–6156.

[48] Ohno, K. *Theor. Chem. Acc.* **1964**, *2*, 219–227.

[49] Lipparini, F.; Cappelli, C.; Scalmani, G.; De Mitri, N.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*(11), 4270–4278.

[50] Lipparini, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*(11), 4153–4165.

[51] Biczysko, M.; Bloino, J.; Brancato, G.; Cacelli, I.; Cappelli, C.; Ferretti, A.; Lami, A.; Monti, S.; Pedone, A.; Prampolini, G.; Puzzarini, C.; Santoro, F.; Trani, F.; Villani, G. *Theor. Chem. Acc.* **2012**, *131*, 1201.

[52] Thom H. Dunning, J. *J. Chem. Phys.* **1989**, *90*(2), 1007–1023.

[53] Jorgensen, W.; Maxwell, D.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

[54] Case, D. A.; Cheatham, T.; Darden, T.; Gohlke, H.; Luo, R.; Jr.", K. M. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. *J. Computat. Chem.* **2005**, *26*, 1668.

[55] Seminario, J. *Int. J. Quant. Chem.* **1996**, *60*, 1271.

[56] Rappé, A.; Casewit, C.; Colwell, K.; .; Goddard, W.; Skiff, W. *J. Am. Chem. Soc.* **1992**, *114*, 10024.

[57] Marenich, A.; Jerome, S.; Cramer, C.; Truhlar, D. *J. Chem. Theory Comput.* **2012**, *8*, 575.

[58] Signore, G.; Nifosí, R.; Albertazzi, L.; Storti, B.; Bizzarri, R. *J. Am. Chem. Soc.* **2010**, *132*, 1276.

[59] Brancato, G.; Rega, N.; Barone, V. *The Journal of Chemical Physics* **2006**, *125*(16), –.

[60] Barone, V.; Baiardi, A.; Bloino, J. *Chirality* **2014**, *26*(9), 228–240.

[61] Guillot, B. **2002**, *101*(1), 219–260.

[62] Cicogna, F.; Coiai, S.; Passaglia, E.; Tucci, I.; Ricci, L.; Ciardelli, F.; Batistini, A. *J. Polym. Science. A* **2011**, *49*, 781.

[63] Cicogna, F.; Coiai, S.; Pinzino, C.; Ciardelli, F.; Passaglia, E. *React. Func. Polym.* **2012**, *72*, 695.

[64] Monti, S.; Cicogna, F.; Passaglia, E.; Prampolini, G.; Barone, V. *Phys. Chem. Chem. Phys.* **2011**, *13*, 21471–8.

[65] De Mitri, N.; Monti, S.; Prampolini, G.; Barone, V. *J. Chem. Theory Comput.* **2013**, *9*(10), 4507–4516.

[66] Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. *J. Comp. Chem.* **1997**, *18*, 1955.


260
280
300
320