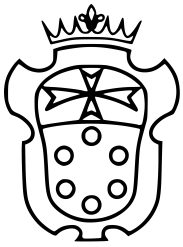


SCUOLA
NORMALE
SUPERIORE

Scuola Normale Superiore

Classe di Scienze

Tesi di Perfezionamento in Methods and Models for Molecular Sciences

XXXIV CICLO

Modeling Drugs in Complex Environments: Solution, Cell Membranes, and DNA

Supervisor:
Prof. Chiara Cappelli

Sara Luz GÓMEZ MAYA

Academic Year: 2021/2022

To my family

.

Acknowledgements

First I would like to thank my supervisor Professor Chiara Cappelli for giving me the opportunity to work on the COSINE network and to be part of her research group at SNS. Thank you for the confidence, for the patience, for sharing with me her knowledge and expertise, and for giving me the freedom to be involved in other projects I found interesting as well.

I extend my thanks to all the current and past members of the EmbedLab@SNS group: Gioia Marrazzini, Franco Egidi, Alessandra Puglisi, Chiara Sepali, Pablo Grobras, Piero Lafiosca, Sveva Sodomaco, Sulejman Skoko, Luca Nicoli, Luca Bonatti, Linda Goletto and with a special emphasis to the brilliant Tommaso Giovannini, who has guided me through the entire Ph.D. process as my "intermediate supervisor". I also want to acknowledge Barbara Rossi for hosting me in Elettra Synchrotron and having made my visit very productive. I am extraordinarily thankful to Prof. Albeiro Restrepo for his invaluable advice, and to Nathy for all the collaborative work we have carried out together.

I want to thank my family for their support, especially my mother and my sister for their love and for believing in me. Thanks also to Manu for always supporting and helping me. Every achievement is yours!!!

List of Publications

Publications included in this thesis

[1] **Sara Gómez**, Tommaso Giovannini, and Chiara Cappelli. *ACS Physical Chemistry Au*, 3(1):1–16, 2023. doi: 10.1021/acsphyschemau.2c00050. URL https://doi.org/10.1021/acsphyschemau.2c00050.

[2] **Sara Gómez**, Tommaso Giovannini, and Chiara Cappelli. Absorption spectra of xanthines in aqueous solution: a computational study. *Phys. Chem. Chem. Phys.*, 22:5929–5941, 2020. doi: 10.1039/C9CP05420K. URL http://dx.doi.org/10.1039/C9CP05420K.

[3] Ring vibrations to sense anionic Ibuprofen in aqueous solution as revealed by resonance raman. *Molecules*, 27(2), 2022. ISSN 1420-3049. doi: 10.3390/molecules27020442. URL https://www.mdpi.com/1420-3049/27/2/442.

[4] **Sara Gómez**, Franco Egidi, Alessandra Puglisi, Tommaso Giovannini, Barbara Rossi, and Chiara Cappelli. Unlocking the power of resonance Raman spectroscopy: The case of amides in aqueous solution. *J. Mol. Liq.*, 346:117841, 2021. ISSN 0167-7322. doi: https://doi.org/10.1016/j.molliq.2021.117841. URL https://www.sciencedirect.com/science/article/pii/S0167732221025666.

[5] **Sara Gómez**, Cettina Bottari, Franco Egidi, Tommaso Giovannini, Barbara Rossi, and Chiara Cappelli. Amide spectral fingerprints are hydrogen bonding-mediated. *J. Phys. Chem. Lett.*, 13(26):6200–6207, 2022. doi: 10.1021/acs.jpclett.2c01277. URL https://doi.org/10.1021/acs.jpclett.2c01277.

[6] Natalia Rojas-Valencia, **Sara Gómez**, Sebastian Montillo, Marcela Manrique-Moreno, Chiara Cappelli, Cacier Hadad, and Albeiro Restrepo. Evolution of bonding during the insertion of anionic Ibuprofen into model cell membranes. *J. Phys. Chem. B*, 124(1):79–90, 2020. doi: 10.1021/acs.jpcb.9b09705. URL https://doi.org/10.1021/acs.jpcb.9b09705.

[7] Natalia Rojas-Valencia, **Sara Gómez**, Francisco Núñez-Zarur, Chiara Cappelli, Cacier Hadad, and Albeiro Restrepo. Thermodynamics and intermolecular interactions during the insertion of anionic Naproxen into model cell membranes. *J. Phys. Chem. B*, 125(36):10383–10391, 2021. doi: 10.1021/acs.jpcb.1c06766. URL https://doi.org/10.1021/acs.jpcb.1c06766.

[8] Piero Lafiosca, **Sara Gómez**, Tommaso Giovannini, and Chiara Cappelli. Absorption properties of large complex molecular systems: The DFTB/Fluctuating Charge approach. *J. Chem. Theory Comput.*, 18(3):1765–1779, 2022. doi: 10.1021/acs.jctc.1c01066. URL https://doi.org/10.1021/acs.jctc.1c01066.

[9] Sara Gómez, Piero Lafiosca, Franco Egidi, Tommaso Giovannini, and Chiara Cappelli. UV-Resonance Raman Spectra of Systems in Complex Environments: A Multiscale Modeling applied to Doxorubicin intercalated into DNA. *Just Accepted in J. Chem. Inf. Model.*, 2023.

[10] Natalia Rojas-Valencia, Sara Gómez, Tommaso Giovannini, Chiara Cappelli, Albeiro Restrepo, and Francisco Núñez Zarur. Water maintains the UV-Vis spectral features during the insertion of anionic Naproxen and Ibuprofen into model cell membranes. *Submitted*, 2023.

Publications not included in this thesis

[11] **Sara Gómez**, Natalia Rojas-Valencia, Santiago A. Gómez, Chiara Cappelli, Gabriel Merino, and Albeiro Restrepo. A molecular twist on hydrophobicity. *Chem. Sci.*, 12:9233–9245, 2021. doi: 10.1039/D1SC02673A. URL `http://dx.doi.org/10.1039/D1SC02673A`.

[12] Linda Goletto, **Sara Gómez**, Josefine Andersen, Henrik Koch, and Tommaso Giovannini. Linear response properties of solvated systems: A computational study. *Phys. Chem. Chem. Phys.*, 24:27866–27878, 2022. doi: 10.1039/D2CP04512E. URL `http://dx.doi.org/10.1039/D2CP04512E`.

[13] Santiago A. Gómez, Natalia Rojas-Valencia, **Sara Gómez**, Franco Egidi, Chiara Cappelli, and Albeiro Restrepo. Binding of SARS-CoV-2 to cell receptors: A tale of molecular evolution. *ChemBioChem*, 22 (4):724–732, 2021. doi: https://doi.org/10.1002/cbic.202000618. URL `https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/cbic.202000618`.

[14] Santiago A Gomez, Natalia Rojas-Valencia, **Sara Gómez**, Chiara Cappelli, and Albeiro Restrepo. The role of spike protein mutations in the infectious power of SARS-COV-2 variants: A molecular interaction perspective. *ChemBioChem*, 23(7):e202100393, 2021. doi: https://doi.org/10.1002/cbic.202100393. URL `https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/cbic.202100393`.

[15] Lina Uribe, **Sara Gómez**, Tommaso Giovannini, Franco Egidi, and Albeiro Restrepo. An efficient and robust procedure to calculate absorption spectra of aqueous charged species applied to NO_2^-. *Phys. Chem. Chem. Phys.*, 23:14857–14872, 2021. doi: 10.1039/D1CP00652E. URL `http://dx.doi.org/10.1039/D1CP00652E`.

[16] Lina Uribe, **Sara Gómez**, Franco Egidi, Tommaso Giovannini, and Albeiro Restrepo. Computational hints for the simultaneous spectroscopic detection of common contaminants in water. *J. Mol. Liq.*, 355:118908, 2022. ISSN 0167-7322. doi: https://doi.org/10.1016/j.molliq.2022.118908. URL `https://www.sciencedirect.com/science/article/pii/S0167732222004469`.

[17] Alejandro Arias, **Sara Gómez**, Natalia Rojas-Valencia, Francisco Núñez-Zarur, Chiara Cappelli, Juliana A. Murillo-López, and Albeiro Restrepo. Formation and evolution of C–C, C–O, C=O and C–N bonds in chemical reactions of prebiotic interest. *RSC Adv.*, 12:28804–28817, 2022. doi: 10.1039/D2RA06000K. URL `http://dx.doi.org/10.1039/D2RA06000K`.

[18] Santiago Gómez, **Sara Gómez**, Jorge David, Doris Guerra, Chiara Cappelli, and Albeiro Restrepo. Dissecting bonding interactions in cysteine dimers. *Molecules*, 27(24), 2022. ISSN 1420-3049. doi: 10.3390/molecules27248665. URL `https://www.mdpi.com/1420-3049/27/24/8665`.

[19] Chiara Sepali, Piero Lafiosca, **Sara Gómez**, Tommaso Giovannini, and Chiara Cappelli. Computational Protocol to simulate solvent effects on Raman and Raman Optical Activity spectra of systems in aqueous solution. *Submitted*, 2023.

[20] **Sara Gómez**, Matteo Ambrosetti, Tommaso Giovannini, and Chiara Cappelli. A close-up look at the electronic spectroscopic signatures of common pharmaceuticals in solution. *Submitted*, 2023.

Preface

This thesis has been submitted in partial fulfillment of the requirements for obtaining a Ph.D. degree in Methods and Models for Molecular Sciences at Scuola Normale Superiore di Pisa, Italy, Classe di Scienze. The work was carried out under the supervision of Professor Chiara Cappelli in a four-year period: November 2018 – November 2022. During the first three years of the doctoral studies, the student was enrolled as a fellow in the H2020–MSCA–ITN–2017, Marie Skłodowska Curie Innovative Training Network "COmputational Spectroscopy In Natural sciences and Engineering (COSINE, grant number 765739).

The thesis is based on ten publications (Papers 1–10). At the time of writing, Paper 10 is in peer review, Paper 9 is just accepted, while Papers 1–8 are published in peer-reviewed journals. During the course of the Ph.D., a few other manuscripts have also been prepared (Papers 11–20) which have been focused on applications of the proposed protocols. Two of those derived works are also under review (Papers 19,20).

The thesis is structured into seven main chapters. Chapter 1 provides the overview and the motivation behind the creation of a computational protocol for investigating complex systems. Chapter 2 explains the steps of the protocol, including a brief theoretical background of embedding approaches and the spectral signatures studied. Chapter 3 gives a summary of the key scientific findings from the author's publications addressed in the thesis. Applications of the protocol to systems in aqueous solution, inserted in cell membranes and intercalated into DNA are described in Chapters 4, 5 and 6, respectively. Finally, Chapter 7 draws some conclusions and future perspectives of the work.

Contents

List of Figures

List of Tables

List of Abbreviations

AH Adiabatic Hessian 205, 220, 221

AS Adiabatic Shift 205, 218, 220, 221

BOA Born-Oppenheimer Approximation 206, 208

cLR corrected Linear Response 18–20

CT Charge Transfer 14, 15, 24

DFT Density Functional Theory 15, 17, 19, 20, 25, 153

DFTB Density Functional Tight Binding 1, 25, 153

DMA Dimethyl Acetamide 27

DMPC Dimyristoylphosphatidylcholine 25

DNA Deoxyribonucleic acid vi, 1, 3–5, 7, 25, 153

DOX Doxorubicin 1, 7, 25, 153

ECD Electronic Circular Dichroism 11, 17, 203

EE Electrostatic Embedding 2, 3, 14, 19

FC Franck-Condon 20, 21, 205, 215, 216, 218, 219, 222

FF Force Field 2, 8–10, 22, 203

FQ Fluctuating Charges 1, 2, 4, 5, 14–16, 18, 19, 21, 23, 25, 27, 153, 203, 204

FQFμ Fluctuating Charges and Fluctuating Dipoles 1, 2, 14–16, 18, 19, 203

GS Ground State 18, 19

HT Herzberg-Teller 215, 216, 222

IMDHO Independent Mode Displaced Harmonic Oscillator ix, 205, 217, 218, 220, 222

LR Linear Response 17–19

MD Molecular Dynamics 1, 3, 4, 6–11, 13, 22–25, 27, 99, 153, 203, 204

ME Mechanical Embedding 14

MM Molecular Mechanics 2, 5, 6, 10, 13–16, 19, 153, 203, 204

MO Molecular Orbitals 18, 19

NAGMA N-acetylglycine-N-methylamide 7, 24, 27

NALMA N-acetyl-leucine-methylamide 24, 27

NBO Natural Bond Orbitals 1, 11, 12, 23–25, 99

NCI Non-covalent interaction 1, 11, 12, 24, 25, 99

NMA N-Methyl Acetamide 27

NPT Isothermal–isobaric ensemble (constant temperature and constant pressure ensemble) 9

NVT Canonical Ensemble, N: number of particles, V: Volume, T: Temperature 9

PBC Periodic Boundary Conditions 9, 10

PCM Polarizable Continuum Model 13

PDB Protein Data Bank 8

PE Polarizable Embedding 2, 14

PES Potential Energy Surface 20, 21, 210, 214, 217–221

PHVA Partial Hessian Vibrational Approach 4

QM Quantum Mechanics 1, 2, 4–6, 9, 10, 13–19, 21–23, 27, 153, 203, 204

QM/MM Quantum Mechanics/Molecular Mechanics x, 1, 2, 4–6, 10, 13–16, 19, 22–25, 27, 153, 203, 204

QTAIM Quantum Theory of Atoms in Molecules 1, 11, 12, 24, 25, 99

RMSD Root Mean Square Deviation 11

RR Resonance Raman ix, x, 1–5, 11, 17, 18, 20, 21, 23–25, 27, 153, 203, 205, 213–222

RREP Resonance Raman Excitation Profile 218, 221

SCF Self-Consistent Field 15–19

SOS Sum Over States 18, 222

TD Time-Dependent 17, 18, 21, 25, 217, 218

TI Time-Independent 18, 21, 217

UV-Vis Ultraviolet/visible 1, 3–5, 17, 23, 25, 26, 99, 203

VG Vertical Gradient 20, 21, 205, 217, 218, 221

VH Vertical Hessian 205, 220, 221

Abstract

This Thesis focuses on developing and applying protocols, based on multiscale Quantum Mechanics/Molecular Mechanics (QM/MM) simulations, for studying systems in complex environments, with a special emphasis on spectroscopic and molecular response properties. When the systems are embedded in aqueous environments, the atomistic mutually polarizable QM/Fluctuating Charges (FQ) and QM/Fluctuating Charges and Fluctuating Dipoles (FQFμ) embedding schemes are applied to the calculation of Ultraviolet/visible (UV-Vis), Raman, and Resonance Raman (RR) spectra. Biologically important molecules, like small amides, small peptides, and a series of common drugs covering xanthines (caffeine, and others), and anionic Ibuprofen and Naproxen, are part of the test cases.

As an extension of the designed protocols, a novel strategy to study the nature of the interactions taking place in complex environments is proposed. The method, similar to that used in the modeling of spectroscopies, is based on the extraction of multiple configurations from Molecular Dynamics (MD) trajectories and then computing quantum mechanical descriptors, as those derived from Natural Bond Orbitals (NBO), Quantum Theory of Atoms in Molecules (QTAIM), and Non-covalent interaction (NCI) formalisms. The procedure is applied to the analysis of the evolution of intermolecular interactions during the insertion of non-steroidal anti-inflammatory drugs, namely anionic Ibuprofen and Naproxen, into model cell membranes. QM/MM-based UV-Vis absorption spectra of both drugs are also modeled as a function of their position from the purely aqueous phase to the approximate center of the membrane. The reasons for the small changes in experimental UV-Vis spectra are then revealed.

A further extension of the model to deal with electronic and vibrational spectroscopies of targets embedded in other large biomolecular environments is developed and applied to the intercalation complex between Doxorubicin (DOX), a well-known anticancer drug, and DNA. The coupling between the Density Functional Tight Binding (DFTB) approach and the FQ approach is tested to simulate UV-Vis spectra, and innovative methodologies to compute normal modes are presented.

To demonstrate the reliability of the developed protocols, the versatile methodology is validated by comparison with experimental data. The excellent agreement attained makes it possible to confidently use our simulations to interpret/predict properties and spectra of complex biosystems.

Chapter 1

Introduction

Accounting for environmental effects when modeling system properties has been considered a challenge for years, especially in the spectroscopic field.[22–25] Different computational protocols have been developed to accurately describe both the chromophore's response to the light and the perturbation provided by its surrounding medium. Most of them belong to the family of the so-called focused models,[22,26,27] in which the system is usually divided into two portions: the target, which inherits the measured property, generally described at the QM level, and the environment, often treated by means of classical physics to reduce the overall computational cost.[26,28–32] If an atomistic description of the embedding is retained, MM Force Field (FF)s are exploited in the resulting QM/MM approaches.[28,33] QM/MM schemes differ from each other in how they describe the interaction between the target fragment and the environment. This interaction might be restricted to electrostatics (including polarization) or might also contain other contributing forces of quantum nature, like Pauli repulsion and dispersion.[33] For instance, in the Electrostatic Embedding (EE)[34,35] method, the environment charge distribution is included as point charges in the QM Hamiltonian, whereas in the Polarizable Embedding (PE)[36–38] approaches the mutual polarization between the two regions is taken into account, thus providing the most physically consistent picture of the interacting system.

There are various techniques to simulate the mutual polarization between the quantum region and the surroundings. Modeling the polarization via the induced-dipole model is the strategy exploited in the Polarizable Embedding model[36], the Effective Fragment Potential (EFP) models[39–41] and in QM/AMOEBA[42]. In contrast, QM/FQ[25] and QM/FQFμ,[43] use the electronegativity equalization principle to allow variations in the charges (or in both charges and dipoles) assigned to MM atoms. In this thesis, the focus is on QM/FQ and QM/FQFμ approaches.

Many of the reported methodologies to perform hybrid, polarizable QM/MM calculations can be summarized by a series of steps that start with a configurational sampling, and from this, a number of structures are retrieved and used in subsequent quantum-classical calculations.[22,37,44–47] When the embedded system under investigation is represented by a solvated molecule, the partitioning between the QM and MM portions is generally straightforward since non-covalent bonds are established between the chromophore (the solute) and the solvent. For such systems, well-established and robust computational protocols have been developed in recent years[22,44,45] and QM/FQ and QM/FQFμ models have been proven to be particularly successful for simulating standard spectroscopies[22] and more challenging spectroscopies, such as RR of molecules in aqueous solutions. In

this respect, a protocol to interpret and reliably predict RR spectra of molecules in solution was developed and applied to diverse systems[3–5] as part of this thesis.

Resonance Raman and especially UVRR has a very large potential to unveil the structural and electronic properties of systems embedded in different environments. RR provides selectivity and sensitivity through the enhancement of certain vibrations associated with particular chromophores when the incident wavelength is tuned to match a specific electronic transition of the system.[24,48] Experimentally, many biological systems are studied by using this technique. However, the modeling of such spectroscopy is far from trivial for molecules in solution[4,49–51] and becomes almost prohibitively expensive for medium-to-large-sized targets embedded in more complex environments.

Given the combined nature of the RR signal, electronic and vibrational effects must be included in the model when computing the property in any environment. Indeed, the calculation of the UV-Vis spectrum is a prerequisite of any RR simulation, because the identification of the transition to which the resonance occurs is made possible by the analysis of the electronic excited states, and the selection of the excitation energy is enabled by comparing calculated and experimental spectra. In addition, it is well known that the ingredients of those spectroscopies, namely, electronic transitions, normal modes, and polarizabilities are all influenced by the environment.[1,50]

As mentioned before, for systems in solution, basically all kinds of spectroscopies have been covered in recently reported protocols[4,22] but when it comes to complex environments such as proteins, DNA, and membranes, simulations usually focus on electronic spectroscopies[52,53] and just a few reports consider vibrational properties[54–56]. Clearly, the current methods for isolated or solvated molecules must be modified if the system complexity grows (for example, for heterogeneous systems) in order to adequately cover the new aspects and interactions. As a matter of fact, it is possible to investigate in detail the fundamental interactions, at the molecular level, that take place between the target and its surroundings adopting the common practice of subjecting individual configurations afforded by MD simulations to quantum mechanical analysis.[6]

Although the environment's polarization is not taken into consideration by the EE, it is nevertheless helpful in many situations, as long as the methodology can be validated and the results are in agreement with experimental reports. In fact, in some cases, the less accurate, but computationally inexpensive EE permits a good reproduction of spectral measurements and together with quantum mechanical descriptors of bonding help to provide explanations of experimental behaviors.[10] UV-Vis of drugs inserting into membranes[10] and UV-Vis and subsequent RR of drugs intercalated into DNA[9] are examples where the EE scheme is generally accurate.

Going back to RR simulations, it is well known that the computation of normal

modes is one of the key components. Since in the sequential classical↦QM approaches, the sampling is often obtained through MD simulations, most of the configurations would mimic a situation out-of-equilibrium (in terms of quantum mechanics) and any vibrational analysis could give rise to imaginary frequencies, representing oscillations around an equilibrium position. The alternative solution of refining the geometry of the snapshots through partial optimizations, in other words optimizing the quantum region via QM/MM methods while keeping the environment frozen (Partial Hessian Vibrational Approach (PHVA)), could have a negative effect on the dynamics/conformations of the quantum region. Also, the acquisition of the harmonic modes is time-consuming, and dealing with large systems implies including hundreds of vibrations in the calculation of the final spectra. The development of a protocol to calculate RR spectroscopies of complex systems, by proposing a series of strategies to compute the normal modes aiming at circumventing, on one hand, the partial optimization of the structures sampled through MD runs, and on the other hand, the cost of performing vibrational analysis on each configuration, is also part of this thesis.

Finally, it is clear that the methodologies are flexible and thus suitable for promising applications to complex biosystems. The general protocol is discussed in Chapter 2 as well as the theoretical basis of the FQ models and their extension to spectral signals under investigation. Then, the practical performance of the different variants of the protocol when calculating the properties of diverse systems is illustrated. Systems in solution are discussed in Chapter 4 and are the topics of the Perspective[1], and Papers 2, 3, 4, 5. In particular, the perspective lists the key physicochemical factors that must be taken into account while developing a reliable method to describe the absorption properties of solvated systems. Hence, it is important to pay special attention to dynamical elements and strong solute-solvent interactions, in order to accurately represent the intricacies of the solvation phenomena. Notwithstanding, all those elements permeate other spectroscopies and must be integrated into protocols for the simulation not only of absorption properties but also of RR spectroscopies. Chapter 5 shows how the combination of QM descriptors (calculated in complex systems[6,7]) and spectra (Paper 10) allows elucidating the root causes of experimental behaviors, such as negligible changes in UV-Vis spectra of drugs when going from aqueous solution to the aqueous lipidic environment. The extension of the atomistic multiscale computational protocol to more complex systems, such as chromophores embedded in biological matrices is discussed in Chapter 6. Doxorubicin, a widely used chemotherapy agent, intercalated into DNA (Papers 8,9) was chosen as a test case.

Chapter 2

Protocol for studying systems in complex environments

Understanding spectroscopic signatures of chromophores in complex environments always constitutes an enormous challenge that can be aided by computational chemistry. We have developed a multiscale protocol based on both polarizable and non-polarizable QM/MM approaches, primarily for the calculation of electronic absorption and RR spectroscopies of systems embedded in environments ranging from aqueous solution to DNA. For systems dissolved in water, we exploited the potentialities of the polarizable QM/FQ model, while for probes intercalated into DNA we used fixed charges in the MM region.

This chapter contains a brief outline of the steps involved in the workflow of a QM/MM spectral calculation, specifically oriented to UV-Vis and RR spectroscopies. It applies to the study of systems ranging from small solute-solvent to large, biologically relevant systems, such as proteins and membranes.

The protocol is quite versatile and there is a number of computational choices also named *parameters* that need to be carefully evaluated and selected in each one of the following steps (see Figure 2.1), to ensure the quality of the final results:

(i) Definition of the QM/MM system

(ii) Conformational/Configurational Sampling

(iii) Extraction of structures

(iv) QM/MM calculations of the spectral property/signal

(v) Analysis and refinement

In the next sections, each step is discussed.

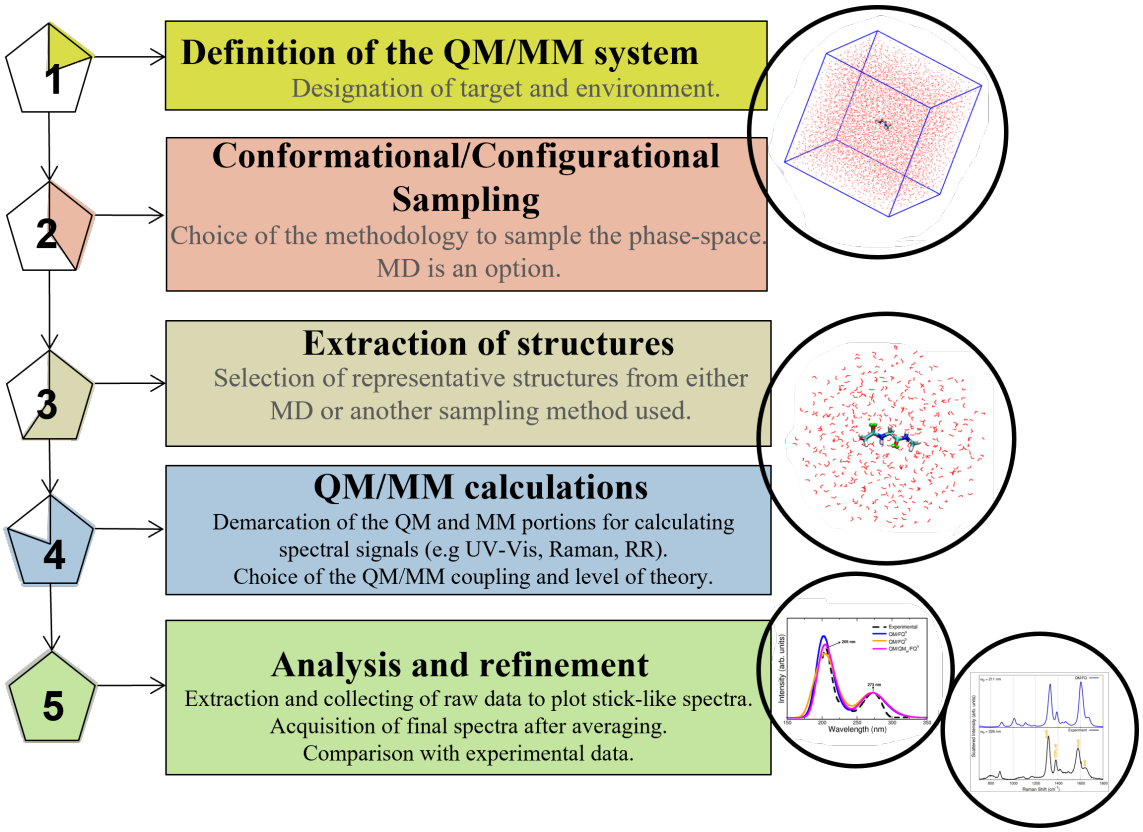


Figure 2.1. Flowchart of the computational protocol

2.1. Definition of the QM/MM system

Due to the prohibitively high computational cost associated with an entire QM treatment when simulating a spectral property of a complex system, keeping a fully atomistic description of the environment implies the necessity of using a hybrid QM/MM approach (see Section 2.4.1). Prior to the generation of configurations and before any calculations are carried out, the system must be defined, demarcating the QM (generally the solute/chromophore) and MM (most often the solvent/environment) portions, the first being responsible for the spectral response. This definition is system- and property-dependent and is useful in the subsequent steps, for instance in the parametrization of the force field previous to MD runs (see Section 2.2). Figure 2.2a shows how the regions for target and environment are chosen in the systems under investigation.

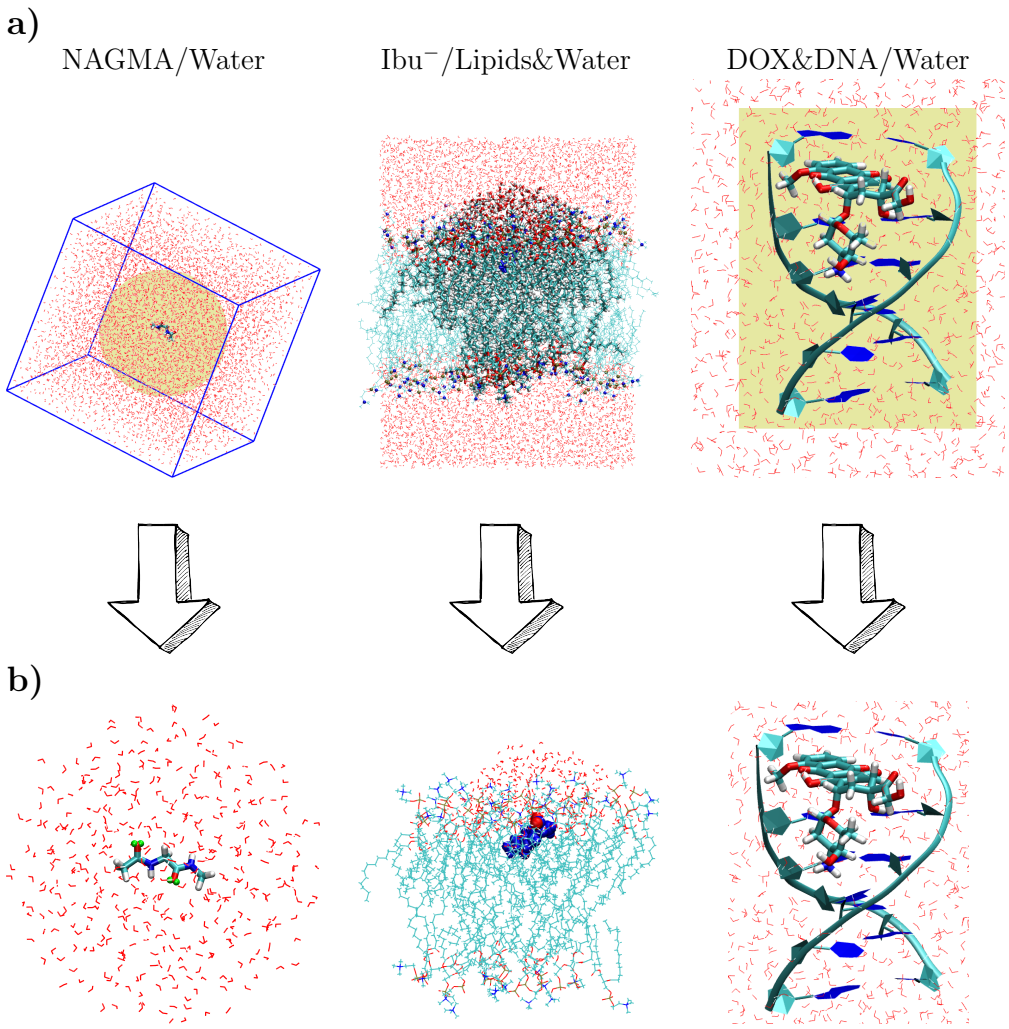


Figure 2.2. Schematic representation of the different systems studied in this thesis: NAGMA, a dipeptide in aqueous solution[5], anionic Ibuprofen inserted in a model cell membrane[6,10] and Doxorubicin intercalated into DNA.[8,9] a) Example of a complete snapshot extracted from MD runs. b) Cut snapshot (see the shaded area in panel a)) maintaining the environment molecules that are expected to establish strong and specific interactions with the target.

2.2. Molecular Dynamics as a tool for sampling

Including environmental effects within the modeling of spectroscopic properties is a critical task since spectral measurements are often conducted on non-isolated chemical systems. The primary goal of this step is to explore the phase space of the target-environment system in order to obtain diverse arrangements or the so-called configurations in which final spectra are calculated. The choice of a proper sampling strategy is crucial for the success of the protocol since it influences the remaining steps. There is a large variety of methodologies, employing both classical and quantum mechanics,[57] that are useful to sample conformations/configurations of systems in solution and in more complex environments. Among the most used techniques are conformer generators, stochastic and genetic algorithms, Molecular Docking, Monte-Carlo and MD

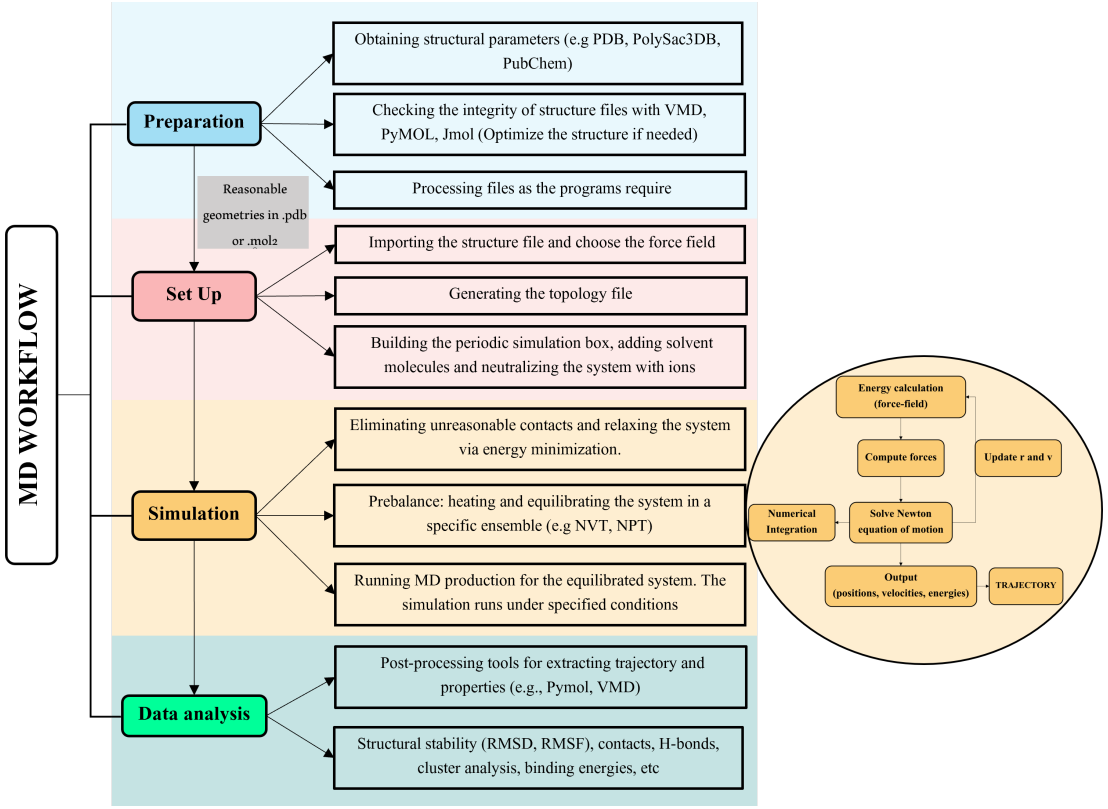


Figure 2.3. MD workflow adapted from Ref. 68 taking as an example the steps to follow in the GROMACS program.

simulations.[58–64] Indeed, MD simulations are well-established sampling strategies in the literature. Next, a concise description of MD simulations is presented but for a more complete discussion, the reader is referred to some of the extensive explanations from the existing literature.[57,65–67]

MD simulation is a method for studying how atoms and molecules move, behave and interact over a period of time, thus providing a view of the dynamic evolution of the system. It is based on Newtonian mechanics by interatomic potentials or molecular mechanics FF. The latter can be seen as a combination of two components: (i) a set of equations used to generate the potential energies and their derivatives, which ultimately lead to molecular forces, and (2) suitable parameters that are used in this set of equations. Such parameters define the atoms/residues, and their masses, charges, types, etc., involved in a system under study. FFs also contain bonded and nonbonded parameters related to bonded and nonbonded interactions.

Regardless of the system, the MD simulation procedure requires a series of steps, illustrated in Figure 2.3: preparation of necessary material, setting up, simulation, and data analysis or post-processing. What is own of each system is the availability of atomic coordinates since not each initial molecular configuration can be found in a database like the Protein Data Bank (PDB) website.[69,70] In that case, researchers have to build up the geometric information of the molecules of

interest using diverse computational packages and make sure about their reliability and validity. For instance, most of the compounds studied in the works of this thesis are non-standard ligands, then early modeling of the lowest-energy conformers and pretreatment of the molecular files was needed.

The second big step involves the setting up of the simulation. Molecules and atoms are endowed with a FF and according to the selection, the topology –molecular description– is constructed. Plenty of FFs are available in the literature[71] varying in the scope of their application: AMBER, GAFF, GROMOS, OPLS, CHARMM, UFF, MMFF, ReaxFF, MARTINI, among others, form the group of the most widely used. Keeping in mind the ultimate goal of the protocol which is the spectral calculation, the choice of the FF is a critical part since the combination of classical MD simulations and quantum-chemical calculations could lead to the so-called "geometry mismatch" problem.[72–78] The inconsistency arises from the fact that structures derived from general available FFs could be unreliable for subsequent embedding calculations, thus influencing the accuracy and the final shape of the spectra. To overcome this drawback, the use of tailored FFs (reparametrized so that they reproduce the QM description) has been proposed.[79–82] Once the FF has been chosen and/or adapted, the simulation box is built up and solvent molecules and counterions are added to neutralize the system. At this point, Periodic Boundary Conditions (PBC) are often induced to mimic a large (infinite) system by treating a relatively small part of it in a *unit cell*, thus maintaining the number of particles in the simulation system and eliminating boundary effects.[83]

Afterward, the central piece of the MD method comes into play. In the simulation step, the parameters for running the MD stages should be adjusted properly according to the simulation's purposes. Stages encompass an energy minimization with the goal of excluding unreasonable contacts (e.g. solvent clashes) from the system, a pre-balancing or equilibration stage for a short time scale (about 1 ns) of the system via NPT or NVT ensembles, and finally a production stage under specified conditions. The repeated (n steps) acquisition of the calculated forces, which are then used to solve Newton's laws of motion through numerical integration, is at the heart of the simulation. In each cycle of the process, positions and velocities are updated and the potential energy is recalculated. In this way, the outcome of an MD simulation is a trajectory that describes the variation in time of position, velocities, and energies of a system. The post-processing step is devoted to analyzing the results: H-bonding, number of contacts, binding energies, etc.

In the context of computational spectroscopy, MD gives the conformational, functional, and dynamic changes of target–environment systems under different conditions (e.g. P, T), desirably those used in the experiments. To meet the conditions of a good statistical description of the experimental system, a large number of simulations (also named *replicas*) could be required. Thus, it is advised to change MD starting points (for instance test different orientations) and perform

extra runs. MD simulations of the systems studied in this thesis are performed at the purely classical level and by imposing PBC on cubic boxes. However, it is worth mentioning that MD simulations are totally flexible and can use either QM or MM, or a mixture of both to calculate forces, a feature leading to *ab initio* or combined QM/MM MD.[57,84–86] As a matter of fact, those alternative MD simulations are other ways to circumvent the problem with the FF choice, mentioned above. Still, the associated computational cost can be a limitation for many systems.

In addition, apart from the classical all-atom MD simulation, other methods have been proposed, among them, coarse-grained (CG) MD simulation[87–90], steered MD simulation[91–94], and accelerated (AC) MD simulations[83] that include essential dynamics[95], replica exchange MD simulation[96,97], hyperdynamics[98,99], metadynamics[100,101], and temperature-accelerated dynamics.[102] Notice that the simulation can be performed via different programs according to the purpose.

To conclude, subsequent QM/MM calculations (step *iv* of the protocol, section 2.4.1) are regularly performed on structures extracted from MD trajectories, but the procedure can also be extended to configurations resulting from other sampling strategies.[15]

2.2.1. Description of the nature of the interactions

After running the production stage of an MD simulation, the result at hand is a trajectory made out of multiple configurations. To analyze target↔environment interactions and also help in the definition of the boundaries to calculate spectral properties in the QM/MM framework, the pair correlation functions, $g(r)$, are very useful. The general $g_{\text{target}-\text{environment}}(r)$ between the target's center of mass and the centers of mass of all the solvent/environment molecules included in the simulation box or the specific atomic $g_{\text{X}-\text{Y}}(r)$ (X=hetero-atoms O, N, S, etc or H in the target, Y= atoms from the environment, e.g. water protons or water oxygen atoms), pair correlation functions can be plotted, and from them, the closest interaction between solute and solvent atoms can be extracted. Overall, hydrogen bonds are found to be the most stabilizing interactions in the different environments, but in no-water solutions, more complex contacts emerge. A $g(r)$ makes evident how a solvent or an environment is structured in different shells or layers. Also, from its asymptotic value (plateau), it is possible to choose a cutting distance to make sure that all the most interacting surrounding molecules are included in the QM/MM scheme to compute the property of interest.

Another advantage of the MD trajectories is that they can recover bulk properties of systems. Therefore, based on the quantum mechanics postulate that says that macroscopic properties are the statistical averages of microstates, it could be argued that taking snapshots from an equilibrated MD is a proper procedure to analyze the quantum interactions responsible for intermolecular bonding and in turn, for several effects on the spectral properties.

Together with the pair correlation functions, a number of descriptors of chemical bonding derived from the QTAIM, from the NCI, and from NBO methods, can be computed to identify and characterized the main interactions. Table 2.1 provides a summary of the standard set of global and local descriptors that have been used in the thesis and how they are related to bonding. For rigorous theoretical developments of the corresponding formalisms, the interested reader is referred to Refs. 103–115. As for the software, the analysis of the hydration patterns is carried out with either the GROMACS[116] package, TRAVIS[117,118], MDAnalysis[119,120] or VMD[121] tools, whereas intermolecular interactions are dissected with a combination of different codes, namely, AIMall[122], NBO[123], and NCIPLOT[124].

2.3. Extraction of structures

Once the target-environment phase space is explored, a number of snapshots are extracted and used for the following quantum-classical calculations. Snapshots are always chosen from the last nanoseconds of the MD production runs, with a large enough separation in time to ensure uncorrelated conformations. The choice of the size and number of snapshots is not univocal. Therefore, it is always recommended to check the convergence of the spectra obtained as varying both the size of the environment and the number of configurations in a reasonable range.

Concerning the size of the configuration, a good suggestion is to investigate the pair correlation functions (see Section 2.2.1) in order to determine to what extent the target "feels" more strongly the surrounding environment. For solute-solvent systems, snapshots are often cut in sphere-shaped droplets and a radius from 10 to 20 Å has been demonstrated as a good choice, yet subjected to the system size.[129] Undoubtedly, other cuts can be also done for example following the shape of the target, as long as the configuration can account for the local and specific interactions with the neighbor molecules. Final cut snapshots of the systems examined in this thesis are displayed in Figure 2.2b).

As for the convergence of the final spectra, it always depends on the desired property/spectroscopy as well as on the system's degrees of freedom. In the case of electronic absorption[2] and RR spectra[4], just a few hundred snapshots have proven to give excellent results, but properties such as Electronic Circular Dichroism (ECD) may require a larger number[20].

Furthermore, the computational sample can be also prepared by retrieving just some representative structures. For that purpose, procedures and algorithms grouping together snapshots based on their structural characteristics are beneficial. The GROMOS clustering method proposed in Ref. 130 to group microstates falling within a structural Root Mean Square Deviation (RMSD) has been one of the most widespread tools. The number of members in each family group defines its relative weight. The central structure of each cluster, that one having the smallest average RMSD from all other structures of the cluster, is taken as the

Table 2.1. Inventory of QTAIM, NBO, NCI derived descriptors of bonding used in different works along this thesis[2,3,6,7] to dissect formal bonds as well as explicit solute↔solvent, and solvent↔solvent ($H_2O\cdots H$–O–H) intermolecular interactions. Closed shell is a generic term that includes long-range, van der Waals, ionic, and hydrogen bonding interactions. $\mathbf{r}_c$ explicitly indicates that QTAIM properties were evaluated at bond critical points. For the $E^{(2)}_{d\to a}$ term ϕ_d, ϕ_a are the donor and acceptor orbitals, $\varepsilon_d, \varepsilon_a$ are the corresponding energies, and $\hat{\mathcal{F}}$ is the one–particle Fock operator. Table adapted from Refs. 125–127.

Descriptor	Relationship to bonding
Global	
Density differences	Redistribution of charge
Non–covalent interactions	Green surfaces → weak to moderate strength attractive interactions Blue surfaces → strong attractive interactions
Local	
$\rho(\mathbf{r}_c)$	Large values are indicative of covalency Small values are indicative of closed shell interactions
$\nabla^2\rho(\mathbf{r}_c)$	$< 0 \Rightarrow$ local maxima (local concentration of charge): covalent interactions $> 0 \Rightarrow$ local minima (local depletion of charge): closed shell interactions
$\mathcal{H}(\mathbf{r}_c) = \mathcal{G}(\mathbf{r}_c) + \mathcal{V}(\mathbf{r}_c)$	$< 0 \Rightarrow$ local dominance of attractive potential energy: covalent interactions $> 0 \Rightarrow$ local dominance of repulsive kinetic energy: closed shell interactions
$\lvert\mathcal{V}(\mathbf{r}_c)\rvert / \mathcal{G}(\mathbf{r}_c)$	$> 2 \Rightarrow$ Covalent $\in [1,2] \Rightarrow$ Intermediate character $\in [0,1] \Rightarrow$ Closed shell
$E^{(2)}_{d\to a} = -q_d \frac{\lvert\langle\phi_d\rvert\hat{\mathcal{F}}\lvert\phi_a\rangle\rvert^2}{\varepsilon_a - \varepsilon_d}$	Large values: strong interactions Small values: weak interactions
Wiberg bond indices (WBI)[128]	degree of sharing of the electron density between two atoms Strength of the interaction
NBO charges	Accumulation/depletion of charge

representative configuration and then used in the embedding calculations.

2.4. Embedding Methods

Since the majority of spectroscopies are recorded in the condensed phase, adequate theoretical models that can take into account how the environment affects the spectral signals are required. A full QM treatment of the whole system could be appealing, but the problem would become intractable from the computational point of view.[131–133] To get expressions that are computationally manageable, it is frequently necessary to reduce the problem's high dimension. Depending on the type of problem at hand, simulations at various theoretical levels are used, from full quantum mechanics to mixed QM/classical approaches. In the latter, the chemical system is divided into a TARGET REGION (treated using quantum mechanics) and its ENVIRONMENT (treated classically).

Depending on how the classical part is simulated, QM/classical models can be split into major classes. The most common method is the Polarizable Continuum Model (PCM), which uses a continuum, polarizable dielectric to describe the classical portion.[26,27,30–32,134–152] In PCM, the target is placed within a molecule-shaped cavity in the embedding medium (built as the envelope of spheres centered on the target atoms) and the electrostatic problem is solved using a boundary element approach. PCM implicitly includes the statistical average of the possible configurations of the environment, but lacking any atomistic description of the environment, PCM cannot properly treat specific target↔environment interactions, which are often essential for getting accurate results. The inclusion of a limited number of target-surrounding molecules in the QM layer (so-called QM/QM$_w$/PCM) usually helps but this procedure raises two important questions. First, how many molecules are needed to well reproduce the spectral behavior, and second, where to put them in order to account for a dynamical treatment averaging all the possible configurations. Mixed quantum mechanical/molecular mechanical (QM/MM) treatments[28,29,33,153] offer an alternative solution because they maintain the atomistic nature of the environment, overcoming the typical limits of continuum approaches.

2.4.1. Quantum Mechanics/Molecular Mechanics Models

As mentioned above, the strength of QM/MM methods resides in the fully atomistic description of the classical MM portion, which allows for studying QM–environment specific interactions. To account for the dynamical aspects of embedding phenomena, QM/MM approaches require an accurate sampling of the system phase-space. In a previous section, it has been shown that MD simulations constitute an excellent source of configurations. The issue is that the property of interest must be calculated on each snapshot until reaching convergence, and it increases the computational cost associated with a QM/MM simulation.[42,129,154–163]

The total energy of the QM/MM system can be written as:[33]

$$E = E_{\mathrm{QM}} + E_{\mathrm{MM}} + E^{\mathrm{int}}_{\mathrm{QM/MM}} \tag{2.1}$$

where E_{QM} and E_{MM} are QM and MM energies, and $E^{\mathrm{int}}_{\mathrm{QM/MM}}$ is the interaction energy, which may be decomposed as follows:

$$E^{\mathrm{int}}_{\mathrm{QM/MM}} = E^{\mathrm{ele}}_{\mathrm{QM/MM}} + E^{\mathrm{pol}}_{\mathrm{QM/MM}} + E^{\mathrm{rep}}_{\mathrm{QM/MM}} + E^{\mathrm{dis}}_{\mathrm{QM/MM}} + E^{\mathrm{CT}}_{\mathrm{QM/MM}} \tag{2.2}$$

where electrostatic ($E^{\mathrm{ele}}_{\mathrm{QM/MM}}$), polarization ($E^{\mathrm{pol}}_{\mathrm{QM/MM}}$), repulsion ($E^{\mathrm{rep}}_{\mathrm{QM/MM}}$), dispersion ($E^{\mathrm{dis}}_{\mathrm{QM/MM}}$) and Charge Transfer (CT), $E^{\mathrm{CT}}_{\mathrm{QM/MM}}$) contributions are highlighted.

Different QM/MM approaches can be formulated depending on how the QM and MM portions are coupled,[164] which is translated into the way $E^{\mathrm{int}}_{\mathrm{QM/MM}}$ is defined. The variants include models based on Mechanical Embedding (ME), EE, PE. In QM/ME the electronic structure calculation comprises the isolated QM region, and the effect of the environment is computed entirely at the MM level. Regardless of the region they belong to, ME demands the assignment of point charges and empirical parameters to all of the atoms.[165] Conversely, the QM/EE model does not require the parametrization of partial charges for the QM atoms. In QM/EE, the MM partial charges are explicitly incorporated in the QM Hamiltonian, allowing the electron density to become polarized as a result of the electrostatic field produced by the environment. When the polarization of the MM region in response to the electron density is modeled, it leads to the polarizable QM/PE approaches. Therefore, QM/PE is a more realistic description of the QM/MM interaction since the mutual target-environment polarization effects are recovered. Polarizable embedding schemes are unquestionably the most advanced QM/MM methods.

2.4.2. Polarizable QM/MM models

Implementing effective polarization schemes for the MM portion has taken up a significant amount of recent effort in the development of QM/MM approaches for different applications, especially in the spectroscopic field. Some of the polarizable strategies proposed in the literature are summarized in Table 2.2 along with the corresponding references for their formalisms and selected applications. A wide range of semiempirical and ab initio electronic structure formalisms have been interfaced with these various polarizable schemes. In the following subsection FQ and FQFμ approaches will be described in more detail, since these were the Polarizable QM/MM models used in the thesis. Both methods have been developed and extended to calculate spectroscopic and response properties of molecules in solution,[22] and for QM/FQ there also exists an extension to non-aqueous media.[166]

2.4.2.1. Fluctuating Charges and Fluctuating Charges and Dipoles

In QM/FQ, each MM atom is assigned a charge (q) which can vary according to the electronegativity equalization principle (EEP),[216–218] i.e. a charge flow occurs

Table 2.2. Strategies available in the literature to deal with mutual QM/MM polarization

Model	References
Drude oscillator	167–174
Induced dipoles	163,175–184
AMOEBA	42,162,185–191
QM/MMPol	41,192,193
QM/Discrete Reaction Field (DRF)	194–199
QM/Polarizable Embedding (PE)	200–202
Effective fragment potential (EFP)	203
Fluctuating Charges (FQ)	22,25,129,154–161,204–212
Fluctuating Charges and Fluctuating Dipoles (FQFμ)	13,43,212–215

when two atoms have different chemical potential.[25] The FQ force field is defined in terms of two atomic parameters, namely the electronegativity (χ) and the chemical hardness (η)[204], both rigorously defined within the conceptual Density Functional Theory (DFT) framework.[219–224] Differently, in QM/FQFμ, an additional polarization source is introduced to model the anisotropic nature of non-covalent interactions. It is incorporated in terms of a set of fluctuating dipoles μ, which are assigned to MM atoms and are expressed through the atomic polarizability α.[43].

If the QM portion is described at the Self-Consistent Field (SCF) level, the total QM/FQFμ energy reads (QM/FQ is recovered by discarding all terms depending on μ):[43]

$$\begin{aligned}\mathcal{E}(\mathbf{D},\mathbf{q},\boldsymbol{\mu},\boldsymbol{\lambda}) = \mathrm{tr}\mathbf{h}\mathbf{D} + \frac{\mathbf{1}}{\mathbf{2}}\mathrm{tr}\mathbf{D}\mathbf{G}(\mathbf{D}) + \frac{\mathbf{1}}{\mathbf{2}}\mathbf{q}^\dagger\mathbf{T}^{\mathbf{qq}}\mathbf{q} + \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \\ + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q} + \mathbf{q}^\dagger\mathbf{V}(\mathbf{D}) - \boldsymbol{\mu}^\dagger\mathbf{E}(\mathbf{D})\end{aligned} \tag{2.3}$$

where $\mathbf{h}$ and $\mathbf{G}$ are the usual one- and two-electron matrices, and $\mathbf{D}$ is the density matrix. $\boldsymbol{\chi}$ collects atomic electronegativities, whereas T_{ij}^{qq}, $\mathbf{T}_{ij}^{q\mu}$ and $\mathbf{T}_{ij}^{\mu\mu}$ are charge-charge, charge-dipole and dipole-dipole interaction kernels, respectively. Their expressions can be found in Refs. 43,225. $\mathbf{q}^\dagger\mathbf{V}(\mathbf{D})$ and $\boldsymbol{\mu}^\dagger\mathbf{E}(\mathbf{D})$ describe the electrostatic interactions between the QM density and the FQs and Fμs, respectively. $\boldsymbol{\lambda}$ is a set of Lagrangian multipliers that impose specific charge constraints, which may be: (i) the entire MM system is constrained to a fixed charge value, thus allowing CT between solvent molecules (the resulting approaches are named QM/FQ$_{\mathrm{CT}}$ and QM/FQFμ_{CT}); (ii) the charge constrain is imposed to each MM molecule, thus no CT can occur between MM molecules.[43] Independently of the charge constraints that are exploited, the effective Fock matrix in the atomic orbitals (AO) basis set $\{\chi_\mu\}$ is:[43]

$$\tilde{F}_{\mu\nu} = \frac{\partial\mathcal{E}}{\partial D_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{D}) + \mathbf{V}_{\mu\nu}^\dagger\mathbf{q} - \mathbf{E}_{\mu\nu}^\dagger\boldsymbol{\mu} \tag{2.4}$$

Notice that in the non-polarizable QM/MM, MM charges are fixed, therefore the QM/MM contribution to the Fock matrix ($\mathbf{V}_{\mu\nu}^\dagger\mathbf{q}$) does not vary along the SCF

procedure. On the contrary, in QM/FQ and QM/FQFμ MM variables (charges and dipoles) explicitly depend on the QM density. As a consequence, their contribution to the Fock matrix has to be computed at each SCF step, thus describing mutual QM/MM polarization effects. Charges and dipoles are computed by imposing the global functional to be stationary with respect to charges, dipoles, and Lagrangian multipliers. This results in the following linear system:[43]

$$\left(\begin{array}{cc|c} \mathbf{T}^{qq} & \mathbf{1}_{\boldsymbol{\lambda}} & \mathbf{T}^{q\mu} \\ \mathbf{1}_{\boldsymbol{\lambda}}^{\dagger} & \mathbf{0} & \mathbf{0} \\ \hline -\mathbf{T}^{q\mu^{\dagger}} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{array}\right) \left(\begin{array}{c} \mathbf{q} \\ \boldsymbol{\lambda} \\ \hline \boldsymbol{\mu} \end{array}\right) = \left(\begin{array}{c} -\boldsymbol{\chi} \\ \mathbf{Q}_{\mathrm{tot}} \\ \hline \mathbf{0} \end{array}\right) + \left(\begin{array}{c} -\mathbf{V}(\mathbf{D}) \\ \mathbf{0} \\ \hline \mathbf{E}(\mathbf{D}) \end{array}\right) \tag{2.5}$$

$$\tag{2.6}$$

$$\mathbf{M}\mathbf{L}_{\lambda} = -\mathbf{C}_{Q} - \mathbf{R}(\mathbf{D}) \tag{2.7}$$

where $\mathbf{1}_{\lambda}$ accounts for the Lagrangian blocks, $\mathbf{C}_{Q}$ collects atomic electronegativities and charge constraints. $\mathbf{L}_{\lambda}$ is the vector containing charges, dipoles, and Lagrangian multipliers, and $\mathbf{R}(\mathbf{D})$ represents the QM potential and field. Again, the FQ linear system can be easily recovered from Eq. 2.5, by simply discarding rows/columns involving μs and their response. At this point, it is important to emphasize that the quality of QM/FQ and QM/FQFμ strongly depends on the values assigned to the parameters that define the model and enter the corresponding equations, for instance, χ and η in the QM/FQ case.

As a side note, although many QM/MM approaches limit the QM/MM interaction to the electrostatics, non-electrostatic interactions may also be included. Indeed, the most physically reliable description of the spectral properties of a complex system may only be achieved if all QM/MM interactions written in Eq. 2.2 are described. Non-electrostatic interactions (repulsion and dispersion) are commonly included through (classical) parametric functions, such as the Lennard-Jones potential[226], which does not depend on the QM density and is a term only added to the total energy of the system. Therefore, it only indirectly influences molecular properties. Actually, it is difficult to describe the non-electrostatic interactions that have a purely quantum nature, by using a MM portion that is classical. Nevertheless, a variety of approaches have been developed in the literature.[200–202,227–235]

In addition, QM/FQ and QM/FQFμ models have been extended to non-electrostatic interactions, namely, Pauli repulsion and quantum dispersion.[236,237] In short, each MM molecule is endowed with a set of s-type Gaussian functions, which simulate the density of the MM portion (ρ^{MM}). If the repulsion energy term is written as the opposite of an exchange integral between the QM density, ρ^{QM} and MM density, $E^{\mathrm{rep}}_{\mathrm{QM/MM}}$ becomes:[238,239]

$$\begin{aligned} E^{\mathrm{rep}}_{\mathrm{QM/MM}} &= \frac{1}{2}\int \frac{\mathrm{d}\mathbf{r_1}\,\mathrm{d}\mathbf{r_2}}{r_{12}}\rho_{\mathrm{QM}}(\mathbf{r_1},\mathbf{r_2})\rho_{\mathrm{MM}}(\mathbf{r_2},\mathbf{r_1}) \\ &= \frac{1}{2}\sum_{\mathbf{R}}\int \frac{\mathrm{d}\mathbf{r_1}\,\mathrm{d}\mathbf{r_2}}{r_{12}}\rho_{\mathrm{QM}}(\mathbf{r_1},\mathbf{r_2})\cdot\left[\xi^{2}_{\mathbf{R}}e^{-\beta_{\mathbf{R}}(\mathbf{r_1}-\mathbf{R})^2}\cdot e^{-\beta_{\mathbf{R}}(\mathbf{r_2}-\mathbf{R})^2}\right] \end{aligned} \tag{2.8}$$

where **R** runs over the centroids of the s-type Gaussian functions, whereas β and ξ are two free parameters of the model, which can be determined by reproducing the exchange-repulsion energy contribution calculated by some energy decomposition analysis (EDA).[240–242] This method has been firstly applied to the calculation of spectral properties of molecular systems in aqueous solution.[236].

To conclude, it is worth mentioning that to overcome the limitations associated with QM/classical approaches, the entire system might be treated using a QM description, where all interactions are described at the QM level.[40,243–245,245–256] In those cases, the system is partitioned into at least one active and one inactive part, that allows reducing the computational cost with respect to a full QM calculation at the same level. Similar to what is done in focused models, only the active wavefunction/density is usually optimized, whereas the inactive wavefunction/density remains frozen. Most quantum embedding methods are developed within the framework of DFT.[250,257,258]

2.5. Spectroscopical Signatures under investigation

In this work, various kinds of spectroscopies, including UV-Vis, ECD, Raman, and RR of systems in aqueous solution and in more complex environments, have been investigated using the multiscale models discussed in the preceding sections.

The simplest and least expensive analytical method for examining a system's electrical characteristics is frequently one-photon absorption spectroscopy in the UV-Visible range. Highly correlated approaches (Configuration interaction, CI and Multi-Configurational SCF, MCSCF methods) have historically been used to calculate electronic transitions and excited states. However, due to a favorable balance between cost and accuracy, TD-DFT has emerged as the most used method for calculating absorption properties and spectra of medium-sized molecules.[129,164] There are two formalisms to describe electronic excitations and spectroscopy in TD-DFT, namely, real-time and Linear Response (LR), denoted in what follows as RT-TDDFT and LR-TDDFT, respectively. In theory, to generate a spectrum, RT-TDDFT is more computationally expensive because it necessitates the time propagation of the electron density during extensive simulations lasting several tens of femtoseconds. For that reason, the LR approach is the most popular class of TD-DFT implementations to compute electronic transitions.

ECD is the simplest chiroptical spectroscopy that can be calculated. The assessment of excited-state energies and transition densities is necessary for ECD. After obtaining transition densities, dipole strengths and rotatory strengths can be used to compute basic absorption and ECD cross-sections.[44]

In vibrational Raman scattering, transitions occur between two vibronic levels of the molecule. Quantum mechanics provides a formal expression from which

Raman intensities (or Raman cross sections) can be calculated, based on the Raman polarizability tensor.[259] Two forms of scattering –normal Raman scattering and RR scattering– can be identified depending on the frequency of the incident light and the characteristics of the system.[260]

RR spectroscopy exploits the fact that during Raman measurements the incident frequency is tuned into an electronic absorption band, enhancing selected vibrational modes.[24,48] RR offers unique selectivity as well as a high sensitivity to experimentally detect even traces of compounds, and thus it finds analytical applications in agriculture, life sciences, explosive detection, art, archaeology, and forensics, with additional current research in carbon nanotubes.[261] The key ingredient in the simulation of RR spectra is the calculation of the transition polarizability tensor,[262–265] for which different methods have been proposed in both the Time-Independent (TI)[3–5,16,49,266] –based on the Sum Over States (SOS), expressions–, and the Time-Dependent (TD) –based on wave packet dynamics– approaches,[267–272] with the inclusion of Duschinsky, Herzberg-Teller, anharmonicity, and solvation effects.[50,273] An alternative method that relies on a short-time dynamics approximation is to calculate RR intensities directly from the geometrical derivatives of the frequency-dependent complex polarizability with respect to the normal coordinates.[274–277] The main advantage of that strategy is that all the electronic states are included in the polarizability, being also well-suited for dealing with large molecules or small molecules in complex environments.[274,278] The reader is referred to Appendix A for further details on Raman and RR spectroscopies.

QM/FQ and QM/FQFμ terms (see Eqs. 2.7 and 2.5) propagate to the solute's response equations when calculating response or spectroscopic properties, ensuring that polarization effects are completely taken into account in the computed final spectral data. This is the topic of the next section.

2.6. QM/FQ and QM/FQFμ Approaches for Computing Absorption and RR Spectra

2.6.1. Absorption Spectra

To further extend QM/FQ and QM/FQFμ to the calculation of absorption properties of solvated systems, the modification of the Ground State (GS) Molecular Orbitals (MO) which results from the SCF procedure is not sufficient.
Due to the presence of non-linear terms in the Hamiltonian as a result of the use of polarizable approaches (see Eq. 2.4), two alternative formalisms can be utilized, namely: (i) all polarization sources linearly respond to the transition density (LR);[214,279] (ii) they adjust to the excited-state electronic configuration, in a state-specific fashion. The latter treatment may be limited to a first-order correction, giving rise to the so-called corrected Linear Response (cLR) approach.[214,279]

Both LR and cLR formalisms have already been discussed in the literature for polarizable QM/MM approaches.[25,36,42,176,214,280–283] Here, it is briefly recalled how Casida's equations[284] for the calculation of electronic excitation energies ω are modified:

$$\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^* & \tilde{\mathbf{A}}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{2.9}$$

where $\tilde{A}$ and $\tilde{B}$ matrices read:[214]

$$\tilde{A}_{ai,bj} = (\epsilon_a - \epsilon_i)\delta_{ab}\delta_{ij} + (ai|bj) - c_x(ab|ij) + c_l f^{xc}_{ai,bj} + C^{pol}_{ai,bj} \tag{2.10}$$

$$\tilde{B}_{ai,bj} = (ai|bj) - c_x(aj|ib) + C^{pol}_{ai,bj} \tag{2.11}$$

ε indicates MO energies (with the common notation: virtual MO $a, b, ...$; occupied MO $i, j, ...$), $(pq|rs)$ are two-electron integrals, c_x and c_l are coefficients defining the SCF level (HF: c_x =1, c_l = 0; pure DFT: c_x =0, c_l = 1). As specified in Eqs. 2.10 and 2.11, additional terms with respect to the in-vacuo formulation are present for polarizable embedding approaches. In particular, both direct contributions (the C^{pol} term in Eqs. 2.10 and 2.11) and indirect effects (modifications of GS MO coefficients and energies) appear. C^{pol} is specified according to the polarizable embedding approach. In case of QM/FQFμ:[214]

$$\begin{aligned} C^{FQF\mu}_{ai,bj} = &\sum_p^{N_q} \left(\int_{\mathbb{R}^3} \phi_a(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \phi_i(\mathbf{r})\, \mathrm{d}\mathbf{r} \right) \cdot q_p^T(\phi_b, \phi_i) + \\ &- \sum_p^{N_\mu} \left(\int_{\mathbb{R}^3} \phi_a(\mathbf{r}) \frac{(\mathbf{r} - \mathbf{r}_p)}{|\mathbf{r} - \mathbf{r}_p|^3} \phi_i(\mathbf{r})\, \mathrm{d}\mathbf{r} \right) \cdot \boldsymbol{\mu}_p^T(\phi_b, \phi_i) \end{aligned} \tag{2.12}$$

Here, q^T and $\boldsymbol{\mu}^T$ are perturbed FQs and Fμs (placed at positions $\mathbf{r}_p$), which are adjusted to the transition density $\mathbf{D}^T_K = \mathbf{X}_K + \mathbf{Y}_K$.[214] They are calculated by solving a modified set of linear equations, explicitly depending on the electric potential and field due to the QM transition density:[214]

$$\mathbf{M}\mathbf{L}^T_\lambda = -\mathbf{R}(\mathbf{D}^T_K) \tag{2.13}$$

Notice that, in the case of the EE approach only the GS MO coefficients and energies are modified, i.e. no direct contributions are included in the equations defining excited state energies.

The first conceptual step of both LR and cLR is the definition o the K-th solute electronic excitation, by keeping the solvent response frozen, i.e. by imposing C^{pol} = 0 in Eqs. 2.10 and 2.11. Then, MM polarizable variables are adjusted to the K density. In the LR regime, the response to the whole transition densities is

considered, so that only the dynamic solute-solvent interactions (some dispersion interactions) are taken into account. In contrast, energy differences due to the relaxation of the solute density are not taken into consideration. The latter is considered by the cLR approach, which instead discards the dynamic aspects of solute-solvent interactions. Clearly, the two contributions describe two different physico-chemical phenomena, which can be seen as complementary. A model to account for both contributions at the same time, the so-called cLR^2 approach, has been recently proposed.[166,285]

2.6.2. Raman and RR spectroscopy

The spontaneous Raman scattering cross-section is typically calculated at the DFT level using response theory by differentiating the dynamic electric polarizability with respect to the normal mode displacements, calculated for a perturbation with angular frequency ω corresponding to that of the light source (for example a laser or synchrotron light). By means of the vibrational transition polarizability α_i associated with an excitation of the i-th normal mode, it is possible to describe the cross-section σ_i in terms of the Raman rotational invariants:

$$a_i^2 = \frac{1}{9}\sum_{ab}\alpha^*_{aa,i}\alpha_{bb,i} = \frac{1}{9}|\alpha_{xx,i} + \alpha_{yy,i} + \alpha_{zz,i}|^2 \tag{2.14}$$

$$g_i^2 = \frac{1}{2}\sum_{ab}(3\alpha^*_{ab,i}\alpha_{ab,i} - \alpha^{i*}_{aa,i}\alpha_{bb,i}) \tag{2.15}$$

$$\sigma_i = \left(\frac{\omega - \omega_i}{c}\right)^4 \frac{45a_i^2 + 7g_i^2}{45} \tag{2.16}$$

To characterize the Potential Energy Surface (PES) of a molecule, the harmonic approximation is usually invoked. In the non-resonant regime, the vibrational transition polarizability is expanded in a Taylor series to first order, and only the first derivative term is typically kept and calculated either numerically or analytically.[286] In the method, the imaginary part of the polarizability is assumed to be negligible, but this is true only when the incident radiation is far-from-resonance. Conversely, in the resonant case, the following full sum-over-state expression must be considered

$$\alpha_{ab,i} = \frac{1}{\hbar}\sum_{m'}\frac{\langle i|\mu_a|m'\rangle\langle m'|\mu_b|0\rangle}{\omega_{m'} - \omega_i - \omega - i\gamma} \tag{2.17}$$

where the summation runs over all vibronic states belonging to the PES of the resonant electronic state, while γ is the excited state's phenomenological damping constant. Hence, RR (and Raman with the corresponding tensor) intensities are obtained in terms of geometrical derivatives of electric dipole-electric dipole polarizability $\alpha_{ab,i}$, as it is done in Papers 4, 3, 5, following the details of Ref. 50. In particular, the Vertical Gradient (VG), Franck-Condon (FC) approximation is

employed, where the vibrational frequencies and normal modes of the excited state are assumed to be the same as the ground state, and the transition dipole moments are considered to be independent of the molecular geometry.[50] VG|FC variant is part of the TI[49,50] method, though an equivalent TD[267,269,287] formulation has also been reported in the literature (see Appendix A). Calculations of vibrational RR spectra of isolated and solvated molecules have been performed by resorting to these frameworks, combining different strategies for the description of the excited states[264,288–294].

The method employed to calculate the RR spectrum is considerably more involved compared to the one for spontaneous Raman. However, it should be emphasized that attempting to simulate the RR spectrum using the same methodology employed in the non-resonant case by simply altering the incident frequency so that it is close to that of the electronic transition would lead to completely erroneous results, unless a method that explicitly includes the imaginary part of polarizability is used.[274] The latter approach that is also called "*Resonance Polarizability Derivatives*" is employed in Paper 9.

In order to properly simulate the RR spectrum of a system in solution, it is necessary to have a flexible solvation method being able to model a wide array of molecular properties, involving both electronic and vibrational degrees of freedom, as well as excited states. As mentioned in section 2.4.2.1, solvent effects can be described by means of the FQ force field, in which each atom of the classical layer is endowed with a charge, whose value is not fixed, but is allowed to vary as a response to the electric potential produced by the QM density.[22] In recent years, the QM/FQ method has been extended to the calculation of analytical energy third derivatives, which allow for the calculation of spontaneous Raman spectra,[159] excitation energies,[154] and excited state gradients.[209]

In paper 4, the polarizable QM/FQ model was used for the first time to include solvent effects in all terms within equation 2.17. In fact, the presence of the solvent must be carefully considered and included at the QM/FQ level of theory in all steps of the simulation: the ground state geometry of the molecule is first optimized in the presence of the solvent shell, then the electronic density and harmonic PES are modeled by taking the reaction field due to the water molecules into account. The contribution due to the water molecules also enters the response equations that are solved to calculate excitation energies and to model the excited state PES. For instance, in the far-from-resonance case, the QM/FQ contributions to the electric dipole-electric dipole polarizability α_x are reported in the following equation:

$$\alpha^x_{QM/FQ} = \sum_{\mu\nu} \left[\mathbf{q}^\dagger(\mathbf{D}^e(\omega))\mathbf{V}^x_{\mu\nu}D^e_{\mu\nu}(\omega)\right] + \sum_{\mu\nu} \left[\mathbf{q}^\dagger(\mathbf{D}^e(\omega))\mathbf{V}^x_{\mu\nu}D^e_{\mu\nu}(\omega)\right] \tag{2.18}$$

where e indicates electric perturbation, and ω is the frequency of the external radiation. In addition, FQ contributions are expressed in terms of perturbed FQ charges, i.e. generated by a perturbed QM potential. More details on the

derivation can be found in Ref. 159.

2.7. Analysis and refinement

Once the individual QM/MM calculations have been done on each snapshot, the post-processing stage starts, and the raw outputs are extracted. This is the equivalent of doing an orientational sampling of the response function or of the investigated property. Then, the separate spectra are plotted/convoluted according to the corresponding equations for intensities, and averaged across the snapshots to obtain a final spectrum that is compared with experimental data. Regarding the question about the number of snapshots needed to obtain converged properties, the general recommendation is to assess the convergence of the particular property/spectra when varying the number of configurations.

As explained in Ref. 44 several drawbacks of the entire procedure can be identified in this step:

- insufficient number of snapshots
- insufficiently long MD
- a poor choice of the classical FF used during the MD runs
- inadequacies in the electronic structure method employed to calculate the property

To overcome those shortcomings and refine results some strategies have been suggested. Among them, increasing the number of snapshots to evaluate convergence, lengthening MD simulations, testing enhanced sampling techniques, re-parametrizing the force fields, and checking the QM level (Hamiltonian, basis set) used, are the most useful tactics.[22]

Chapter 3

Overview of the Attached Papers

In this Chapter, a synopsis of each one of the papers included in the Thesis is presented. The author's contribution to these publications is summarized in Table 3.1.

In the **Perspective**[1], the essential physico-chemical aspects that need to be considered when building a reliable approach to describe absorption properties of solvated systems are outlined. The particular focus is on how to properly model the complexity of the solvation phenomenon, arising from dynamical aspects and specific, strong solute-solvent interactions. To this end, conformational and configurational sampling techniques, such as MD, have to be coupled to accurate fully atomistic QM/MM methodologies. By exploiting different illustrative applications, the perspective shows that an effective reproduction of experimental spectral signals can be achieved by delicately balancing exhaustive sampling, hydrogen bonding, mutual polarization, and non-electrostatic effects.

Paper 2 presents a detailed computational analysis of UV-Vis spectra of caffeine, paraxanthine, and theophylline in aqueous solution. A hierarchy of solvation approaches for modeling the aqueous environment is tested, ranging from continuum model to non-polarizable and polarizable QM/MM models, with and without the explicit inclusion of water molecules in the QM portion. Computed results are directly compared with experimental data, so to highlight the role of electrostatic, polarization, and hydrogen bonding solute-solvent interactions.

In **Paper 3**, the potentialities of RR spectroscopy to detect Ibuprofen in diluted aqueous solutions are unraveled. A fully polarizable QM/MM methodology based on FQ coupled to MD is exploited to take into account the dynamical aspects of the solvation phenomenon. The work's findings, which are discussed in light of a NBO analysis, reveal that a selective enhancement of the Raman signal due to the normal mode associated with the C-C stretching in the ring, $\nu_{C=C}$, can be achieved by properly tuning the incident wavelength, thus facilitating the recognition of Ibuprofen in water samples.

Paper 4 reports a joined experimental and computational study of Raman and RR spectra of amides in aqueous solution. By employing state-of-the-art QM/MM methods combined with synchrotron-based UVRR spectroscopy, a protocol to interpret and reliably predict RR spectra of amide systems in water, which are prototypical systems for the peptide bond, is proposed. Along the paper, it is demonstrated that the main experimental spectral features can be correctly reproduced by simultaneously taking into account the dynamical aspects of the solvation phenomenon, specific solute-solvent hydrogen bond interactions, and

Activity	Paper									
	1	2	3	4	5	6	7	8	9	10
Original idea		✓	✓		✓	✓	✓			✓
Conceptualization		✓	✓						✓	✓
Investigation	✓	✓	✓	✓	✓			✓	✓	
Methodology - Implementation				✓		✓			✓	
Methodology - Calculations		✓	✓	✓	✓	✓	✓	✓	✓	✓
Methodology - Data extraction		✓	✓	✓	✓	✓	✓	✓	✓	✓
Methodology - Experiments				✓	✓					
Data curation	✓								✓	
Formal analysis	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Validation		✓	✓	✓	✓				✓	
Visualization	✓	✓	✓	✓	✓			✓	✓	
Writing – original draft	✓	✓	✓	✓	✓			✓	✓	✓
Writing – review & editing						✓	✓			
Response to reviewers' concerns	✓	✓	✓	✓	✓	✓	✓	✓		

Table 3.1. Contributions of this thesis' author to the different works presented herein.

mutual solute-solvent polarization effects.

In **Paper 5**, the origin of the peculiar amide spectral features of proteins in aqueous solution is investigated by exploiting a combined theoretical and experimental approach to study UVRR spectra of peptide molecular models, namely NAGMA and NALMA. UVRR spectra are recorded by tuning Synchrotron Radiation at several excitation wavelengths and modeled by using a multiscale protocol based on a polarizable QM/MM approach. Thanks to the unparalleled agreement between theory and experiment, it is confirmed that specific hydrogen bond interactions, which dominate hydration dynamics around these solutes, play a crucial role in the selective enhancement of amide signals. The results further argue the capability of vibrational spectroscopy methods as valuable tools for refined structural analysis of peptides and proteins in aqueous solution.

Paper 6 shows that despite their purely classical origin, randomly chosen configurations from MD simulations provide deep insight into the purely quantum nature of bonding interactions. Descriptors of chemical bonding derived from five different analysis tools based on quantum mechanics (natural charges, electron density differences, QTAIM, NBO, and NCI index) consistently afford a picture of a wall of weak, non–covalent intermolecular interactions separating anionic Ibuprofen from the environment. This wall, arising from the cumulative effect of a multitude of individual weak CT interactions to the interstitial region between fragments, stabilizes the drug at all equilibrium positions in the free energy profile for its insertion into model cell membranes. The formal charge in anionic Ibuprofen strengthens all intermolecular interactions, having a particularly strong effect on the network of water-to-water hydrogen bonds in the solvent. Electron redistribution during the insertion process leads to a sensible reduction of electron

delocalization in both the $-CO_2^-$ group and in the aromatic ring of Ibuprofen.

In **Paper 7** the insertion process of Naproxen into model Dimyristoylphosphatidylcholine (DMPC) membranes is studied by resorting to state-of-the-art classical and quantum mechanical atomistic computational approaches. MD simulations indicate that anionic Naproxen finds an equilibrium position right at the polar/non–polar interphase when the process takes place in aqueous environments. With respect to the reference aqueous phase, the insertion process faces a small energy barrier of ≈ 5 kJ mol^{-1} and yields a net stabilization of also ≈ 5 kJ mol^{-1}. Entropy changes along the insertion path, mainly due to a growing number of realizable microstates because of structural reorganization, are the main factors driving the insertion. An attractive fluxional wall of non–covalent interactions is characterized by all quantum descriptors of chemical bonding (NBO, QTAIM, NCI, density differences and natural charges). This attractive wall originates in the accumulation of tiny transfers of electron densities to the interstitial region between the fragments from a multitude of individual intermolecular contacts stabilizing the tertiary drug/water/membrane system.

Paper 8 reports on the first formulation of a novel polarizable QM/MM approach, where the DFTB is coupled with the FQ force field. The resulting method (DFTB/FQ) is then extended to linear response within the TD-DFTB framework and challenged to study absorption spectra of large condensed-phase systems.

As already mentioned, UVRR spectroscopy is a valuable tool to study the binding of drugs to biomolecular receptors. The extraction of information at the molecular level from experimental RR spectra is made much easier and more complete thanks to the use of computational approaches, specifically tuned to deal with the complexity of the supramolecular system. In **Paper 9** a protocol to simulate RR spectra of complex systems at different levels of sophistication is proposed, by exploiting a QM/MM approach. The method is challenged to investigate RR spectra of a widely used chemotherapy drug, DOX intercalated into a DNA double strand. The computed results show a good agreement with experimental data, thus confirming the reliability of the computational protocol and its potential extension to other complex systems.

In **Paper 10** UV-Vis spectra of anionic Ibuprofen and Naproxen in a model lipid bilayer of the cell membrane are investigated using computational techniques in combination with a comparative analysis of drug spectra in purely aqueous environments. The simulations aim at elucidating the intricacies behind the negligible changes in the maximum absorption wavelength in the experimental spectra. A set of configurations of the systems constituted by lipid, water, and drugs or just water and drugs are obtained from classical MD simulations and UV-Vis spectra are computed in the framework of atomistic QM/MM approaches together with TD/DFT. The results suggest that the molecular orbitals involved in the electronic transitions are the same regardless of the chemical environment. A thorough analysis of the contacts between the drug and water molecules reveals

that no significant changes in UV-Vis spectra arise as a consequence of anionic Ibuprofen and Naproxen molecules being permanently microsolvated by water molecules despite the presence of lipid molecules.

Chapter 4

Biologically important molecules in aqueous environments

How to create an accurate atomistic model to simulate the absorption spectra of systems in solution? As demonstrated in the perspective[1], computational models can effectively replicate experimental spectral signals and have predictive power by carefully balancing exhaustive sampling, hydrogen bonding, mutual polarization, and nonelectrostatic effects. But those key ingredients go beyond absorption spectra and permeate other kinds of spectroscopies.

Models integrating QM/FQ simulations have been useful in the interpretation of the UV–Vis absorption spectra of chromophores like caffeine and other xanthines[2] embedded in an aqueous environment. Furthermore, it is well known that the simulation of RR spectra is accompanied by the calculation of the absorption spectrum. Through a combined experimental and computational work, a QM/FQ protocol for multiscale simulations of RR spectroscopy[4] is set up, initially studying the RR spectra of prototypical systems of the peptide bond, among them, small amides like Acetamide, N-Methyl Acetamide (NMA) and Dimethyl Acetamide (DMA). Such a protocol can be considered of general application and involves different steps starting with a sampling of configurations from MD trajectories, followed by the extraction of a good enough number of snapshots (until reaching convergence in the desired property) in which the QM/MM calculations are carried out.

By applying the QM/FQ RR protocol and exploiting the experimental facilities at Elettra Sinctrotrone Trieste S.C.p.A. alongside state-of-the-art quantum-chemical models, RR spectra of oligopeptides of biological interest, as N-acetylglycine-N-methylamide (NAGMA) and N-acetyl-leucine-methylamide (NALMA)[5] are measured and interpreted, improving the methodology by testing different parametrizations of the FQ model. The importance of the polarization effects is also quantified by comparing QM/MM results when polarizable and non–polarizable schemes are used. In fact, the selective enhancement of some significant bands in the RR spectra of peptides is only possible if mutual polarization effects are taken into account. Additionally, the application of the protocol allows demonstrating that it is feasible to sense anionic Ibuprofen in aqueous solution by means of RR spectroscopies.[3]

4.1. Modeling Electronic Absorption Spectra of Systems in Solution: A General Overview[1]

ACS PHYSICAL CHEMISTRY Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/physchemau

Perspective

Multiple Facets of Modeling Electronic Absorption Spectra of Systems in Solution

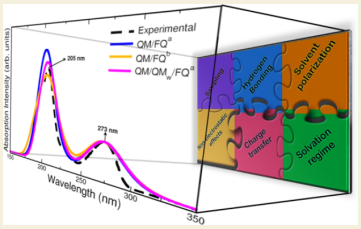
Sara Gómez, Tommaso Giovannini, and Chiara Cappelli*

Cite This: https://doi.org/10.1021/acsphyschemau.2c00050

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: In this Perspective, we outline the essential physicochemical aspects that need to be considered when building a reliable approach to describe absorption properties of solvated systems. In particular, we focus on how to properly model the complexity of the solvation phenomenon, arising from dynamical aspects and specific, strong solute−solvent interactions. To this end, conformational and configurational sampling techniques, such as Molecular Dynamics, have to be coupled to accurate fully atomistic Quantum Mechanical/Molecular Mechanics (QM/MM) methodologies. By exploiting different illustrative applications, we show that an effective reproduction of experimental spectral signals can be achieved by delicately balancing exhaustive sampling, hydrogen bonding, mutual polarization, and nonelectrostatic effects.



KEYWORDS: *aqueous solution, polarization, hydrogen bonding, QM/MM, molecular dynamics*

1. INTRODUCTION

How can we create an accurate atomistic model to simulate absorption spectra of systems in solution? The advances in theoretical and computational chemistry have led to the definition of reliable protocols for the reproduction of many experimental data of molecules in the gas-phase. However, when dealing with solvated molecules, the complexity of the problem increases and the definition of a unique, robust protocol still remains challenging.[1,2] In most of the leading reports, indeed, the simulation of spectral properties is usually performed by using a simplistic representation of the molecular system, where often a single or a limited number of conformations are taken into account for the solute, while the environment is reduced to an unsophisticated nonrealistic view usually mimicked by dielectric constants as in the implicit cases[3] or recreated by a few molecules.[4,5] The reason for such choices is that the complexity of the system constrains its faithful description since that ideal implies high computational costs, becoming almost unattainable for the most part.

However, in the last years, fully atomistic approaches based on a multiscale partitioning of the systems have become more and more diffuse, because they provide a physicochemical consistent portrayal of the solvent, and because their capability to accurately sample the conformational degrees of freedom of the solute−solvent phase-space. All these characteristics are highly desirable features when it comes to spectroscopy.

By retaining the atomistic nature of the whole system, quantum-mechanics/molecular mechanics (QM/MM) approaches[6] for solvation are defined. Depending on the coupling between the QM and MM portions, different approaches can be formulated, ranging from the electrostatic embedding,[7] in which the MM part polarizes the QM part, but not viceversa, to polarizable embedding (PE), where mutual solute−solvent polarization effects are recovered;[1,8,9] clearly, PE gives the most physically consistent picture of the solvation phenomenon. Among the PE approaches that have been developed so far, in this Perspective the focus is on QM/Fluctuating Charges (FQ)[10] and QM/Fluctuating Charges and Fluctuating Dipoles (FQFμ),[11] which have been developed and extended to calculate spectroscopic and response properties of molecules in solution.[12−15]

It has amply been reported in the literature that electronic absorption arising from the solute is affected by the surrounding solvent.[16−19] A typical example are solvatochromic shifts,[14,17,19−21] and more generally the solvent can assist or slow down the common $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ electronic transitions, which results in changes in the appearance, position, intensity, and width of absorption bands.[2,22−24] Hence, the initial question to answer when building a model for solvation is how the solvent interacts with the solute and how these interactions must be meticulously inserted in the model to obtain a detailed description of absorption spectra. To address this problem, PE approaches are the most suitable when coupled to molecular dynamics (MD) simulations, which allow for a correct sampling of the solute solvent phase-space.[1]

Received: September 30, 2022
Revised: November 8, 2022
Accepted: November 8, 2022

In this way, the dynamical aspects of the solvation phenomena, and a reliable description of strong, specific solute−solvent interactions can be achieved.

As it is clear from the previous paragraphs, the definition of a model that successfully reproduces experimental findings would open the door to reliable predictions of spectra for substances or systems whose UV−vis measurements are hard to be performed in the laboratory.[1,25] As in the experiments, many aspects underlie the acquisition of computed absorption spectra and some of them remain critically important. The sampling, the role of the solvent polarization, the hydration patterns commonly via hydrogen bonding, and the inclusion of nonelectrostatic and charge transfer terms are all ingredients of a proper computational modeling and are analyzed and reported in this Perspective. To highlight the role of all the aforementioned elements, we select different established applications, which are used to show the strengths and flaws of the current models. This can also provide options for improving the existing models in order to treat more complex systems and phenomena.

2. THEORETICAL MODELING OF ABSORPTION PHENOMENA

The reliable calculation of absorption properties of molecular systems embedded in an external environment (e.g., a solvent) is challenging. In fact, models need to coherently take into account the various physicochemical aspects of solvation, and how they influence absorption phenomena. Among them, of paramount importance is the accounting for dynamical aspects, which implies an accurate sampling of the solute−solvent phase-space, through the correct identification of all possible conformational minima.[1] Also, depending on the nature of the molecule-environment couple, a reliable model needs to account for the directionality of specific interactions, such as hydrogen bonding.[26] In parallel, a physically consistent description of the spectral signal, i.e., of the electronic properties of the molecular system as perturbed by the environment, is required.[1] Both aspects (i.e., the phase-space sampling and the simulation of the spectral signal) equally contribute to obtain a physically consistent modeling and therefore need to be coherently integrated into the computational protocol. Peculiar aspects related to the phase sampling and specific interactions are discussed in detail in the following sections by resorting to illustrative examples. In this section, the attention is focused on the calculation of the spectral absorption signal.

Multiscale approaches, and QM/MM methods in particular, have recently proven to reliably describe absorption spectra of systems embedded in an external environment.[1] In such models when specified to solutions, the target (i.e., the solute) is described at the QM level, whereas the solvent is atomistically treated at the classical level by means of a parametrized force field (FF).[6] The total energy of the system can therefore be written as[6]

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

where E_{QM} and E_{MM} are QM and MM energies and $E^{int}_{QM/MM}$ is the interaction energy, which may be decomposed as follows:

$$E^{int}_{QM/MM} = E^{ele}_{QM/MM} + E^{pol}_{QM/MM} + E^{rep}_{QM/MM} + E^{dis}_{QM/MM} + E^{CT}_{QM/MM} \tag{2}$$

where electrostatic ($E^{ele}_{QM/MM}$), polarization ($E^{pol}_{QM/MM}$), repulsion ($E^{rep}_{QM/MM}$), dispersion ($E^{dis}_{QM/MM}$), and charge transfer (CT, $E^{CT}_{QM/MM}$) contributions are highlighted.

The various QM/MM approaches differ in the way $E^{int}_{QM/MM}$ is defined. In principle, the most physically reliable description of the electronic properties of the solvated system may only be achieved if all QM/MM interactions are described. However, while electrostatic and polarization effects can be consistently defined within a classical modeling of the solvent layer, repulsion, dispersion, and CT contributions originate from the quantum nature of the electronic degrees of freedom.[27] For this reason, most QM/MM approaches limit the description of QM/MM interactions to electrostatics, yielding the so-called electrostatic embedding,[7] and only in a few cases mutual solute−solvent polarization effects are considered (polarizable embedding).[1] While in electrostatic embedding the interaction term is expressed in terms of a set of MM fixed-value charges, in polarizable QM/MM approaches the interaction term involves a set of electric variables (generally charges and/or dipoles) which are polarized as a response to the QM potential/field, and viceversa.[10] Among the different polarizable QM/MM approaches which have been proposed in the literature, here we concentrate on QM/Fluctuating Charges (QM/FQ) and QM/Fluctuating Charges and Fluctuating Dipoles (QM/FQFμ), which have been developed and amply tested in recent years.[1]

2.1. A Brief Sketch of QM/FQ and QM/FQFμ for Computing Absorption Spectra

In QM/FQ, each MM atom is assigned a charge (q) which can vary according to the electronegativity equalization principle (EEP),[28] i.e. a charge flow occurs when two atoms have different chemical potential.[10] The FQ force field is defined in terms of two atomic parameters, namely the electronegativity (χ) and the chemical hardness (η).[29] Differently, in QM/FQFμ, an additional polarization source is introduced to model the anisotropic nature of noncovalent interactions. It is incorporated in terms of a set of fluctuating dipoles μ, which are assigned to MM atoms and are expressed through the atomic polarizability α.[11]

If the QM portion is described at the SCF level, the total QM/FQFμ energy reads (QM/FQ is recovered by discarding all terms depending on μ):[11]

$$\mathcal{E}(\mathbf{D}, \mathbf{q}, \boldsymbol{\mu}, \lambda) = \mathrm{tr}\,\mathbf{hD} + \frac{1}{2}\mathrm{tr}\,\mathbf{DG(D)} + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^{\dagger}\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + +\mathbf{q}^{\dagger}\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^{\dagger}\mathbf{q} + \boldsymbol{\lambda}^{\dagger}\mathbf{q} + \mathbf{q}^{\dagger}\mathbf{V(D)} - \boldsymbol{\mu}^{\dagger}\mathbf{E(D)} \tag{3}$$

where $\mathbf{h}$ and $\mathbf{G}$ are the usual one- and two-electron matrices and $\mathbf{D}$ is the density matrix. $\boldsymbol{\chi}$ collects atomic electronegativities, whereas $\mathbf{T}^{qq}_{ij}$, $\mathbf{T}^{q\mu}_{ij}$, and $\mathbf{T}^{\mu\mu}_{ij}$ are charge−charge, charge-dipole and dipole−dipole interaction kernels, respectively. Their expressions can be found in refs 11 and 30. $\mathbf{q}^{\dagger}\mathbf{V(D)}$ and $\boldsymbol{\mu}^{\dagger}\mathbf{E(D)}$ describe the electrostatic interactions between the QM density and the FQs and Fμs, respectively. $\boldsymbol{\lambda}$ is a set of Lagrangian multipliers that impose specific charge constraints, which may be (i) the entire MM system is constrained to a fixed charge value, thus allowing CT between solvent molecules (the resulting approaches are named QM/FQ$_{CT}$ and QM/FQFμ_{CT}); (ii) the charge constrain is imposed to each MM molecule and thus no CT can occur between MM molecules.[11]

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

Independently of the charge constraints that are exploited, the effective Fock matrix in the AO basis set $\{\chi_\mu\}$ is[11]

$$\tilde{F}_{\mu\nu} = \frac{\partial \mathcal{E}}{\partial D_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{D}) + \mathbf{V}^{\dagger}_{\mu\nu}\mathbf{q} - \mathbf{E}^{\dagger}_{\mu\nu}\boldsymbol{\mu} \tag{4}$$

Notice that in the nonpolarizable QM/MM, MM charges are fixed, therefore the QM/MM contribution to the Fock matrix ($\mathbf{V}^{\dagger}_{\mu\nu}\mathbf{q}$) does not vary along the SCF procedure. On the contrary, in QM/FQ and QM/FQFμ MM variables (charges and dipoles) explicitly depend on the QM density. As a consequence, their contribution to the Fock matrix has to be computed at each SCF step, thus describing mutual QM/MM polarization effects. Charges and dipoles are computed by imposing the global functional to be stationary with respect to charges, dipoles, and Lagrangian multipliers. This results in the following linear system:[11]

$$\left(\begin{array}{cc|c} \mathbf{T}^{qq} & \mathbf{1}_\lambda & \mathbf{T}^{q\mu} \\ \mathbf{1}^{\dagger}_\lambda & 0 & 0 \\ \hline -\mathbf{T}^{q\mu\dagger} & 0 & \mathbf{T}^{\mu\mu} \end{array}\right)\left(\begin{array}{c} \mathbf{q} \\ \boldsymbol{\lambda} \\ \hline \boldsymbol{\mu} \end{array}\right) = \left(\begin{array}{c} -\boldsymbol{\chi} \\ \mathbf{Q}_{\mathrm{tot}} \\ \hline 0 \end{array}\right) + \left(\begin{array}{c} -\mathbf{V}(\mathbf{D}) \\ 0 \\ \hline \mathbf{E}(\mathbf{D}) \end{array}\right) \tag{5}$$

$$\mathbf{M}\mathbf{L}_\lambda = -\mathbf{C}_{\mathrm{Q}} - \mathbf{R}(\mathbf{D}) \tag{6}$$

where $\mathbf{1}_\lambda$ accounts for the Lagrangian blocks, $\mathbf{C}_{\mathrm{Q}}$ collects atomic electronegativities and charge constraints, $\mathbf{L}_\lambda$ is the vector containing charges, dipoles, and Lagrangian multipliers, and $\mathbf{R}(\mathbf{D})$ represents the QM potential and field. Again, the FQ linear system can be easily recovered from eq 5, by simply discarding rows/columns involving μs and their response. To further extend QM/FQ and QM/FQFμ to the calculation of absorption properties of solvated systems, the modification of the ground state (GS) molecular orbitals (MOs) which results from the SCF procedure is not sufficient. In fact, as a result of the electronic excitation, the solvent degrees of freedom cannot be assumed to be frozen to the solute's GS equilibrium. Indeed, since the time scales associated with electronic excitations are of the order of femtoseconds, it is generally assumed that the solvent degrees of freedom instantaneously readjust to the solute electronic transition, while the vibrational modes, associated with much lower time scales (picoseconds), are frozen to the GS equilibrium. Thus, the solvent enters a so-called electronic "nonequilibrium" regime.[13,31]

Due to presence of nonlinear terms in the Hamiltonian as a result of the use of polarizable approaches (see eq 4), two alternative formalisms can be utilized, namely: (i) all polarization sources linearly respond to the transition density (linear response, LR);[13,32] (ii) they adjust to the excited-state electronic configuration, in a state-specific fashion. The latter treatment may be limited to a first-order correction, giving rise to the so-called corrected linear response (cLR) approach.[13,32]

Both LR and cLR formalisms have already been discussed in the literature for polarizable QM/MM approaches.[10,13,33−39] Here, we briefly recall how Casida's equations[40] for the calculation of electronic excitation energies ω are modified:

$$\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^* & \tilde{\mathbf{A}}^* \end{pmatrix}\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega\begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{7}$$

where $\tilde{A}$ and $\tilde{B}$ matrices read:[13]

$$\tilde{A}_{ai,bj} = (\varepsilon_a - \varepsilon_i)\delta_{ab}\delta_{ij} + (ailbj) - c_x(ablij) + c_l f^{xc}_{ai,bj} + C^{pol}_{ai,bj} \tag{8}$$

$$\tilde{B}_{ai,bj} = (ailbj) - c_x(ajlib) + C^{pol}_{ai,bj} \tag{9}$$

ε indicates MO energies (with the common notation: virtual MOs a, b, ...; occupied MOs i, j, ...), $(pqlrs)$ are two-electron integrals, c_x and c_l are coefficients defining the SCF level (HF: $c_x = 1$, $c_l = 0$; pure DFT: $c_x = 0$, $c_l = 1$). As specified in eqs 8 and 9, additional terms with respect to the in-vacuo formulation are present for polarizable embedding approaches. In particular, both direct contributions (the C^{pol} term in eqs 8 and 9) and indirect effects (modifications of GS MO coefficients and energies) appear. C^{pol} is specified according to the polarizable embedding approach. In case of QM/FQFμ:[13]

$$C^{FQF\mu}_{ai,bj} = \sum_P^{N_q}\left(\int_{\mathbb{R}^3}\phi_a(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}_p|}\phi_i(\mathbf{r})\,\mathrm{d}\mathbf{r}\right)\cdot q^T_p(\phi_b,\phi_i) + -\sum_P^{N_\mu}\left(\int_{\mathbb{R}^3}\phi_a(\mathbf{r})\frac{(\mathbf{r}-\mathbf{r}_p)}{|\mathbf{r}-\mathbf{r}_p|^3}\phi_i(\mathbf{r})\,\mathrm{d}\mathbf{r}\right)\cdot\boldsymbol{\mu}^T_P(\phi_b,\phi_i) \tag{10}$$

Here, q^T and $\boldsymbol{\mu}^T$ are perturbed FQs and Fμs, which are adjusted to the transition density $\mathbf{D}^T_K = \mathbf{X}_K + \mathbf{Y}_K$.[13] They are calculated by solving a modified set of linear equations, explicitly depending on the electric potential and field due to the QM transition density:[13]

$$\mathbf{M}\mathbf{L}^T_\lambda = -\mathbf{R}(\mathbf{D}^T_K) \tag{11}$$

Notice that, in the case of the electrostatic embedding approach, only the GS MO coefficients and energies are modified, i.e., no direct contributions are included in the equations defining excited state energies.

The first conceptual step of both LR and cLR is the definition o the Kth solute electronic excitation, by keeping the solvent response frozen, i.e. by imposing $C^{pol} = 0$ in eqs 8 and 9. Then, MM polarizable variables are adjusted to the K density. In the LR regime, the response to the whole transition densities is considered, so that only the dynamic solute−solvent interactions (some sort of dispersion interactions) are taken into account, while energy differences due to the relaxation of the solute density are not taken into consideration. The latter are considered by the cLR approach, which instead discards the dynamic aspects of solute−solvent interactions. Clearly, the two contributions describe two different physicochemical phenomena, which can be seen as complementary. A model to account for both contributions at the same time, the so-called cLR^2 approach, has been recently proposed.[41,42]

As a side note, to describe the absorption phenomenon, we have presented the extension of polarizable QM/MM approaches to a linear TD-DFT description (see eq 7). It is however known that the choice of the DFT functional might widely affect the electronic response of a molecular system, and should thus be selected based on previous studies reported in the literature.[43−49] To refine such a description, correlated methods, such as Coupled Cluster, can be exploited, though they have not been amply tested in the context of polarizable QM/MM methodologies.[50−53]

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

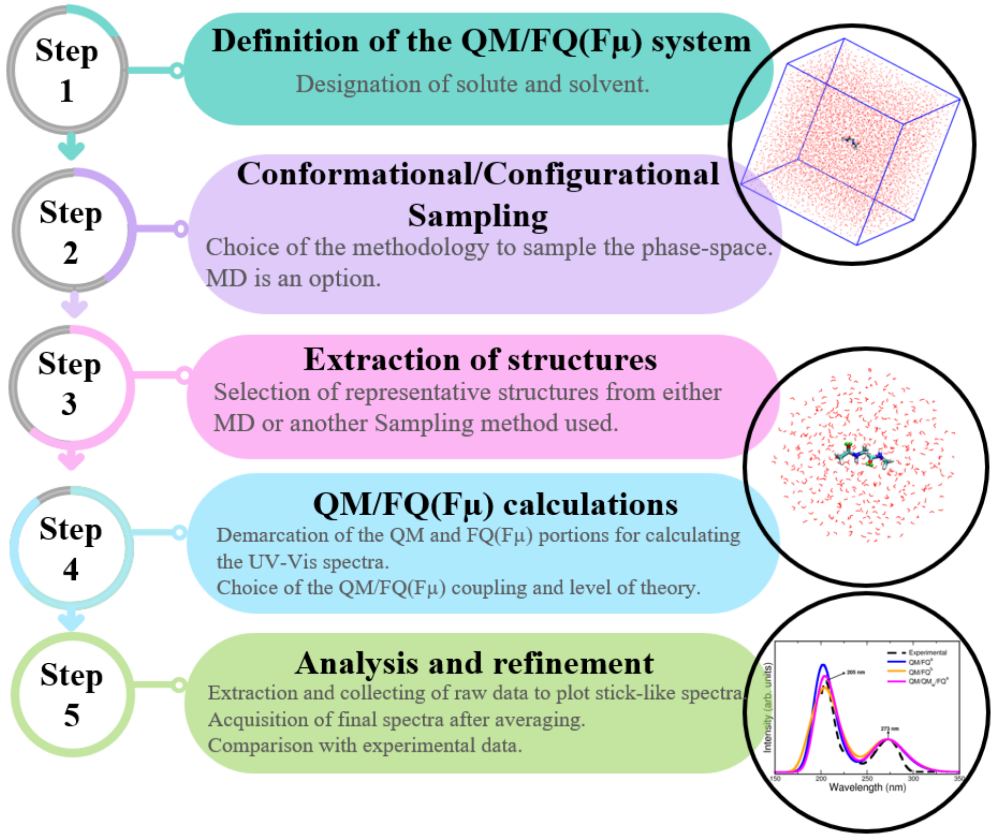


Figure 1. Flowchart of the computational protocol followed in the simulation of absorption spectra.

3. COMPUTATIONAL PROTOCOL

In recent years, we have developed a computational protocol (see Figure 1), which adopts the concepts highlighted in the previous sections and remarkably has successfully been applied to describe absorption properties of solvated systems. Both the direct effect of the environment on UV−vis spectra and the contribution of the spatial arrangement of the solvent to the final spectra are considered when such a procedure is followed. In this Perspective, each one of these steps will be cursorily explained since they have been amply commented elsewhere, along with the "best practices".[1,9] The first step involves (*i*) the definition of the system, demarcating the QM (generally the solute) and MM (most often the solvent) portions, with the first being responsible for the spectral property. Then, (*ii*) a conformational and configurational sampling is performed, by resorting to strategies such as MD simulations, which may also imply a specific reparametrization of the existing FFs. Once the solute−solvent phase space is explored, the computational sample is prepared (*iii*) by extracting some configurations or representative structures, ensuring no correlation between them. Frequently those snapshots are cut in sphere-shaped droplets and, for absorption spectra, a radius less than 20 Å and just hundreds of them have proven to give excellent results.[54−56] Later, (*iv*) QM/FQ calculations of the target property, here electronic absorption spectroscopy, are carried out on the spherical frames obtained at the previous step, at a given QM computational level, which is chosen according to previous studies on similar properties/systems or based on a thorough benchmarking. The two model variants, QM/FQ and QM/FQFμ may be exploited at this step as well as different sets of parameters for water and for nonaqueous solvents. Nonelectrostatic interactions may also be included in the QM/MM modeling.[57] Finally, (*v*) the individual results are extracted, analyzed, and averaged to produce final spectra. At this point, the convergence of the spectra when varying the number of configurations must be assessed. Comparison with experimental data and further refinement of some of the above stages (if needed) sign off the utilization of the protocol.

Table 1 lists all systems whose UV−vis absorption spectra or related quantities (solvatochromic shifts, excitation energies) have been simulated via the computational protocol depicted in Figure 1.

4. ILLUSTRATIVE APPLICATIONS

In this section, we discuss different aspects that need to be considered for constructing a successful model to correctly reproduce electronic absorption spectra of molecules in solution, and in some cases their solvatochromic shifts, by resorting to a set of selected applications. In all cases, according to the original works, MD simulations are performed at the purely classical level and by imposing periodic boundary conditions (PBC) on cubic boxes.

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

Table 1. Record of Different Systems Whose UV−Vis Electronic Absorption Properties or Spectra Have Been Simulated Using the QM/FQ Protocol Depicted in Figure 1[a]

solute	solvent	parametrization	approach	year	ref
formaldehyde	water	A	QM/FQ/PCM	2013	50
nicotine	water	A	QM/FQ/PCM	2015	58
D3 (polythiophene)	water	A	QM/FQ	2018	59
DOX	water	A	QM/FQ	2018	59
	water	A	QM/FQ	2019	55
	water, DNA	A	DFTB/FQ	2022	60
	water	B, C, D	QM/FQ, QM/FQFμ	2022	54
7-methoxycoumarin	water	A	QM/FQ	2019	55
Bodipy 5-Methylcytidine bimane pyridinium dye	water	A	QM/FQ	2019	55
		B, C, D	QM/FQ, QM/FQFμ	2022	54
5-aminophthalimide	water	A	QM/FQ	2019	55
	water	A	QM/FDE/FQ	2021	61
	water	B, C, D	QM/FQ, QM/FQFμ	2022	54
Rhodamine 6G	water	A	QM/FQ	2019	62
Curcumin KK and EK	water	A	QM/FQ	2019	63
PNA	water	B, D	QM/FQ, QM/FQFμ	2019	13
	DIO, THF, ACN	E	QM/FQ	2021	42
	ETH, MET, WTR				
	water	C	CC2-in-MLHF/FQ	2021	51
	water	B, C, D	QM/FQ, QM/FQFμ	2022	54
pyridine	water	B, D	QM/FQ, QM/FQFμ	2019	13
	water	C, D	QM/FQ, QM/FQFμ	2019	14
	water	C	CC2-in-MLHF/FQ	2021	51
pyrimidine	water	B, D	QM/FQ, QM/FQFμ	2019	13
	water	C, D	QM/FQ, QM/FQFμ	2019	14
caffeine	water	A, C	QM/FQ	2020	56
paraxanthine					
theophylline					
luteolin	water	A	QM/FQ	2020	64
kaempferol					
quercetin					
myricetin					
acrolein	water	C, D	QM/FQ, QM/FQFμ	2019	14
	water	C	CC2-in-MLHF/FQ	2021	51
	water	A	QM/FDE/FQ	2021	61
	water	B, C, D	QM/FQ, QM/FQFμ	2022	54
QB	DIO, THF, ACN	E	QM/FQ	2021	42
MER	ETH, MET, WTR				
BET	water	B, C, D	QM/FQ, QM/FQFμ	2022	54
Nitrite	water	A, C	QM/FQ, QM/FQFμ	2021	65
acetamide	water	A	QM/FQ	2022	66
NMA					
DMA					
ubiquitin	water	A	DFTB/FQ	2022	60
a-Ibu	water	A	QM/FQ	2022	67
NAGMA	water	A, C	QM/FQ	2022	68
NALMA					

[a]Parametrizations: A: Rick et al.,[29] B: Carnimeo et al.,[69] C: Giovannini et al.,[70] D (FQFμ): Giovannini et al.,[13] E: Ambrosetti et al.[42]

4.1. Importance of a Comprehensive Sampling

How does the solvent interact with the solute and how are the solute's conformational degrees of freedom affected? In order to answer these questions, an exhaustive sampling of the phase space is a key stage prior to the calculation of the spectral property. Two main factors ensure a high-quality sampling for the whole system, i.e., (i) the conformational sampling of the solute referred to the orientation of the internal dihedral angles within the molecule that originate from intramolecular interactions (such as weak noncovalent interactions, hydrogen bonds, ...), and (ii) the configurational sampling of the solvent or the distribution and orientation of solvent molecules around the solute.[1]

Plenty of methodologies employing both classical and quantum mechanics have shown to be useful for the purpose of sampling conformations/configurations of systems in solution.[71] Conformer generators, stochastic and genetic algorithms, Monte Carlo and MD simulations, among others, form part of the most prosperous techniques proposed in the literature.[72−78] It is worth mentioning that QM/MM

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

calculations (step *iv* of the protocol, see Figure 1) are regularly performed on structures extracted from MD trajectories, but the procedure can also be extended to motifs coming from other sampling strategies.[65] When MD is used, the simulations are quite versatile and they can be conducted in a classical, *ab initio* or combined QM/MM MD way.[71,79−81] The strongest criticism about the choice of a reliable FF might be circumvented by adapting the available FFs to the solute/solvent couple (by means of reparametrization[82−84] as was done in refs 85.86,) Hydration patterns and distribution functions are so useful to perceive the diversity of the sampling and many software come in handy to that end.[87−89]

Regarding spectroscopy, chiral properties (e.g., CD or OR) are usually the most sensitive ones to the system conformation[9] but there are some cases of UV−vis results where the correct solute−solvent sampling has been a determining factor.[58,64] For example, the effects of a conformational sampling obtained from MD runs are evident in the simulation of the UV−vis study of nicotine[58] where due to a proper MD sampling it is revealed the formation of complex specific hydrogen bonding networks of water molecules connecting different parts of the solute, that turn out to be pivotal in the reproduction of the experimental data. Figure 2

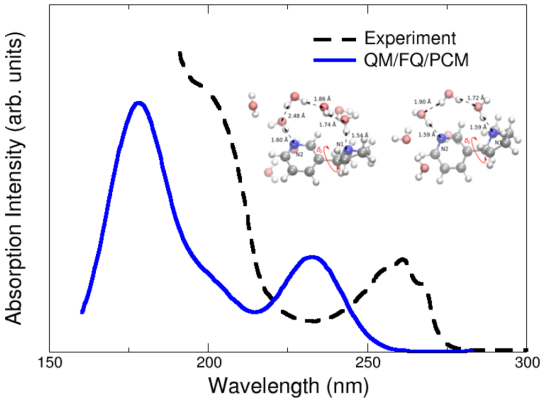


Figure 2. Calculated QM/FQ/PCM absorption spectra of nicotine in aqueous solution. The experimental spectrum (black) taken from ref 58 is given for comparison. Nicotine was freely allowed to move during MD simulations. QM level: CAM-B3LYP/aug-cc-pVDZ. Image adapted from ref 58. Copyright 2015 American Chemical Society.

presents the computed and experimental spectra and representative conformers with one or two bridging water molecules between the nitrogen atoms, that are most frequently sampled during the MD and thus contribute to the final spectra. In this case, the combination of a detailed description of nicotine-water interactions, with directional effects such as hydrogen bonding, offered by the QM/FQ/PCM approach and the conformational sampling resulting from MD provides a great improvement with respect to other modeling strategies.

To retain a more accurate description of the directionality of the key O···H−O and N···H−O molecule−water interactions, in some works,[26,64,66,68] additional interaction sites (also known as virtual sites, VS) have been added to the molecular topology of the solutes. An improvement of the agreement between QM/FQ and experimental spectra was reported by Skoko et al.[64] in the UV−vis study of flavonoids in aqueous solution. For the test cases, the inclusion of VS mainly affects the relative intensity of the two bands and for Kaempferol, Quercetin, and Myricetin, an MD sampling done without VS underestimates the intensity of the first transition with respect to the experimental findings. The presence of explicit water molecules also modifies the conformational distribution and in turn the spreading of excitation energies for the different flavonoids. Individual results for quercetin are presented in Figure 3. Computed excitation energies of the first transition are plotted as a function of the δ dihedral angle in Figure 3b, whereas its corresponding absorption spectra are depicted in Figure 3c. Both these results suggest that a refined modeling of intermolecular interactions through the incorporation of VSs in the oxygen atoms seems to be fundamental to accomplish a rigorous reproduction of the overall spectral shape.

It is important to draw attention to the fact that MD offers some advantages in the long run as a result of the assorted distribution of solvent molecules around the different solute conformations. These merits are especially associated with the natural emerging of the inhomogeneous broadening of the bands in the absorption spectra, which stems from the spreading of the stick spectra of the individual snapshots as can be seen in Figure 4 for the anti-inflammatory drug, Ibuprofen. In addition, sticks can be color-coded to better understand the origin of some bands and choose excited states of interest for subsequent applications as in the case of Resonance Raman spectroscopy.[67]Figure 4 also illustrates that the main band of solvated anionic Ibuprofen (that at 222 nm) appears mainly as a consequence of electronic transitions to states S1, S2, and S3.

Note that unsatisfactory reproduction of reference data (e.g., experimental spectra, full QM calculations,...) may be due to the specific methods that are employed in one (or more) of the steps of the computational protocol (see Figure 1). For instance, a potential source of error might follow the so-called "geometry mismatch" problem, which emerges from the combination of classical MD simulations and quantum-chemical calculations.[91−97] In short, there is no consistency between conformations resulting from classical force fields and those coming from quantum descriptions. This may yield a poor description of spectral features. Some possible solutions to overcome this drawback have been proposed in the literature, among them, resorting to QM/MM MD simulations[93] or parametrizing the classical force fields so that they reproduce the QM description.[85]

4.2. Hydrogen-Bonding Description

The way of describing or representing hydrogen bonds (HBs) is crucial in the acquisition of accurate absorption spectra of aqueous systems. As was mentioned before, their correct directionality can be recovered by using VS in the MDs during the sampling stage, however, HBs appear to be also decisive for reproducing specific spectral details, with renowned consequences as shifts of a few nanometers in UV−vis spectra.[2] Admittedly, it always depends on the nature of solute and solvent, but for polar solvents like water and solutes with potential HB sites, it has been proved that the presence or absence of water molecules determines the molecular conformational distribution and thus, the spectral property at hand.

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

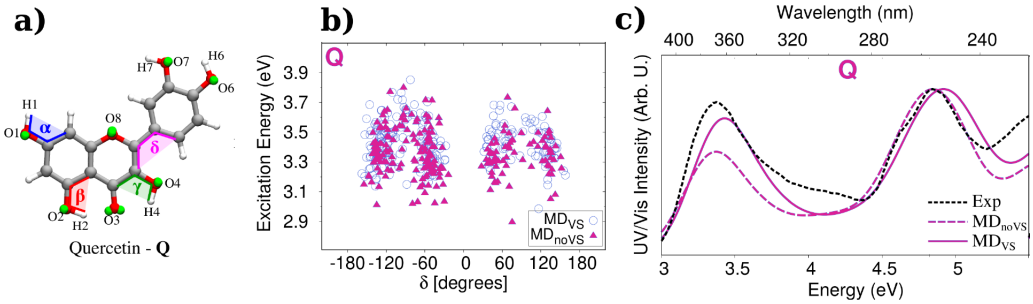


Figure 3. (a) Definition of dihedral angles and oxygen virtual sites of quercetin.[64] (b) QM/FQ excitation energies of the first electronic transition of quercetin in water as a function of its δ dihedral angle. (c) Experimental[90] and computed absorption spectra as calculated with QM/FQ when coupled to both MD_{VS} and MD_{noVS}. Quercetin was freely allowed to move during MD simulations. QM level: B3LYP/6-311+G(d, p). Images adapted with permission under a Creative Commons CC BY License from ref 64. Copyright 2020 MDPI.

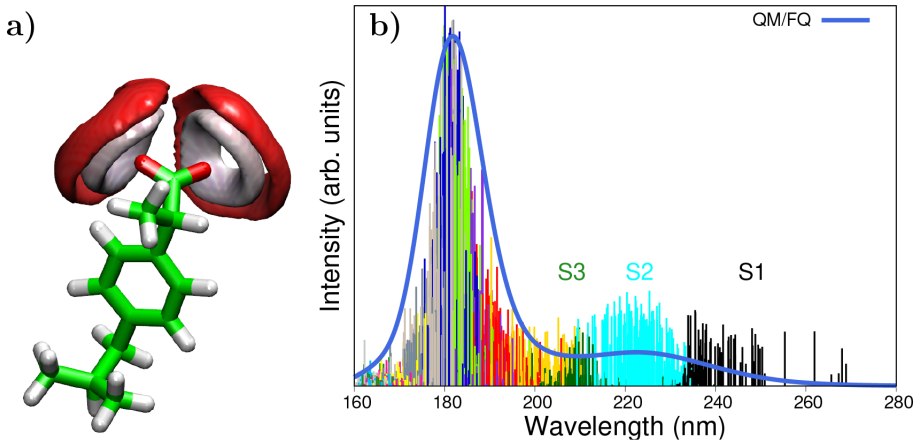


Figure 4. (a) Spatial distribution function of water oxygen (red) and hydrogen (white) atoms around ibuprofen.[67] (b) Sticks and convoluted QM/FQ UV−vis absorption spectrum of anionic Ibuprofen in aqueous solution. Colored sticks stand for different excited states, from which S1, S2, and S3 are labeled. Ibuprofen was freely allowed to move during MD simulations. QM level: CAM-B3LYP/6-311++G(d, p). Images adapted with permission under a Creative Commons CC BY License from ref 67. Copyright 2022 MDPI.

Consequences of using a static QM/continuum approach are known for the case of curcumin, a natural antioxidant, dissolved in aqueous solution.[63] The enol−keto (EK) form of curcumin (and other curcuminoids) has been reported[98] to be the lowest energy isomer and its UV−vis spectrum is experimentally dominated by a single vertical transition placed at about 429 nm.[99] Such a band is the result of a pure HOMO → LUMO transition, with $\pi - \pi^*$ character.[100] As can be seen in Figure 5, an implicit treatment of the solvent with the PCM approach makes the absorption maximum to lie at about 381 nm, whereas it is red-shifted toward the experimental value if QM/MM is exploited, lying at 435 nm for QM/FQ. This improvement offered by the atomistic approach coupled with a dynamical description of the solvation phenomenon is rationalized by analyzing the strong HB interactions detected in the radial distribution functions resulting from classical MD simulations. Despite the fact that MD predicts the EK structure to be predominantly planar (it means that just one conformer should be significant akin to the static QM/PCM picture), the arrangement of water molecules around the equilibrium position of curcumin is crucial to correctly account for the variability of the configurations that lead to the emergence of that main band.

In another case, the impact of the atomistic description of the solvent molecules has been reported by Giovannini et al.,[14] by investigating the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ vacuo-to-water solvatochromic shifts of acrolein. A hierarchy of solvation approaches for modeling the aqueous environment has been tested and some results are presented in Figure 6. It should be pointed out that since acrolein is free to move along the MD trajectories, solvatochromic shifts reported for it arise from both solute−solvent interactions and also from changes in its internal geometry when passing from vacuum to solution. The effect of hydrogen bonding solute−solvent interactions is immediately perceptible by comparing QM/PCM and QM/FQ (in all its variants) computed results with the experimental blueshift of 0.25 eV and redshift of −0.52 eV reported in ref 101 for $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions, respectively, when going from gas phase to a water solution (sign is just a convention). It is clear that the continuum QM/PCM totally fails at reproducing the experiments. From different papers in the literature,[2,4] it has been suggested that possible refinements can be achieved by adding water molecules near the solute in a

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

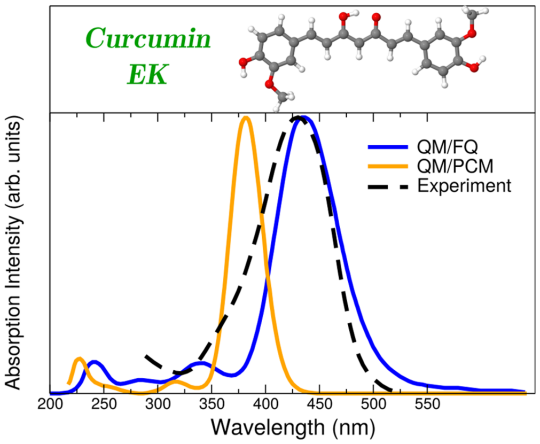


Figure 5. Simulated QM/FQ and QM/PCM absorption spectra for the enol−keto (EK) tautomer of curcumin in aqueous solution.[63] Curcumin was freely allowed to move during MD simulations. QM level: M06-2X/def2-TZVP. Experimental data taken from ref 99. Image reproduced and adapted with permission from ref 63. Copyright 2019 Royal Society of Chemistry.

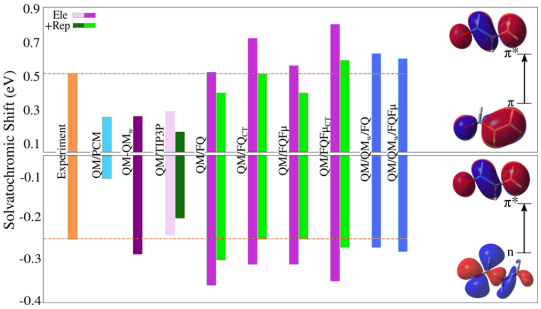


Figure 6. Computed and experimental $\pi \rightarrow \pi^*$ (top) and $n \rightarrow \pi^*$ (bottom) vacuo-to-water solvatochromic shifts of acrolein. Acrolein was freely allowed to move during MD simulations. QM level: CAM-B3LYP/aug-cc-pVDZ. To get a better picture, experimental values taken from ref 101 extend along the whole results. Orbitals involved in both transitions are also included. Image adapted from ref 14. Copyright 2019 American Chemical Society.

cluster-like QM/QM$_w$ model. Figure 6 shows that this approach does not alter the solvatochromic shift of the $\pi \rightarrow \pi^*$ transition of acrolein but it enhances the result for the $n \rightarrow \pi^*$, although it is still overestimated. That strategy can be extended to the QM/FQ methodology, where explicit water molecules are included in the QM portion in what is called QM/QM$_w$/FQ.[34] Sometimes results achieved in that way have been in better agreement with experiments[55,65,68] or have been taken as reference, but it also happens that the QM/FQ approach is more than enough in the description of the solvent since it outperforms other solvation models at a lower computational cost.[55]

Additional evidence of the reliability of the HB description given by QM/FQ coupled to MD sampling is given by Ibuprofen in water solution.[67] In particular, the strength of HBs around the solute (see the spatial distribution functions, sdfs, in Figure 4) has been quantified by using the stabilization energies provided by the Natural Bond Orbitals (NBO) framework.[102] After comparing NBO results from QM/FQ and from purely QM ibuprofen-water clusters, it appears that when 6 water molecules (plus the rest of the FQ layer) surround the solute, the dominant contributions are due to the same kind of charge transfer and the magnitudes of stabilization energies are similar to those found in micro-solvated aggregates.

In light of the above examples, the atomistic, albeit classic, treatment of the QM/FQ model seems to be a major ingredient to reproduce experimental findings.

4.3. Mutual Polarization Effects

The quality of the description of solute−solvent interactions rules the accuracy of computed spectral properties. For this reason, the physics lying behind the QM/MM model definitely matters and it should be able to represent how both layers "feel" each other. Although in the EE approach the MM layer does polarize the QM density, the opposite is not true, disregarding the fact that both solute and solvent are collectively affected by the presence of each other. Polarizable QM/MM approaches overcome this important shortcoming by endowing MM atoms with polarizable sources that vary as a result of the interaction with the QM density, and viceversa.

Different examples of how polarization effects exert influence on molecular response properties have been pointed out in the literature.[23,36,103] In this Perspective, we have picked a select set of them to comment on the relevance of mutual solute−solvent polarization effects. The first case involves acrolein whose solvatochromic shifts for the two lowest-lying electronic transitions are shown in Figure 6. For the $n \rightarrow \pi^*$ transition, the nonpolarizable QM/TIP3P (which places fixed charges at atomic sites[104]) well reproduces the experimental shift, underestimating it by just 4%, while the shift of the $\pi \rightarrow \pi^*$ transition is wrongly reproduced, with an error of 44% in absolute value.[14] In contrast, much better results are achieved when exploiting polarizable QM/MM approaches, with a striking 0% and 9% of error for the $\pi \rightarrow \pi^*$ transition in the contexts of QM/FQ and QM/FQFμ, respectively.

Polarization effects can be directly observed in the UV−vis spectra of the styrylpyridinium cyanine dye displayed in Figure 7a. For that chromophore, the absorption spectrum in water exhibits an absorption band at 446 nm and another band at a shorter wavelength of 278 nm (not shown in the plot).[44,105] Comparison between QM/FQ and QM/TIP3P spectra indicates that the mutual polarization is crucial to improve the description of experimental data, since QM/TIP3P places the absorption maximum in the furthest position with respect to the other solvation models, thus having the highest error.[55] This system is another peculiar example in which including explicit water molecules in the QM portion does not alter considerably the spectrum. This is due to the fact that there are no specific solute−solvent interactions according to the radial distribution functions from the MD sampling. Also, for NAGMA and NALMA dipeptides,[68] a shift toward the experimental results is recovered by using the QM/FQ model rather than QM/TIP3P. Concerning a very popular solute, caffeine, some repercussions are noted too. Caffeine is a xanthine, structurally similar to a purine. Its UV−vis absorption spectrum in water exhibits two bands centered at 205 and 273 nm.[106] Similar to the case of the Cyanine dye, Figure 7b shows that for solvated caffeine (it also applies for

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

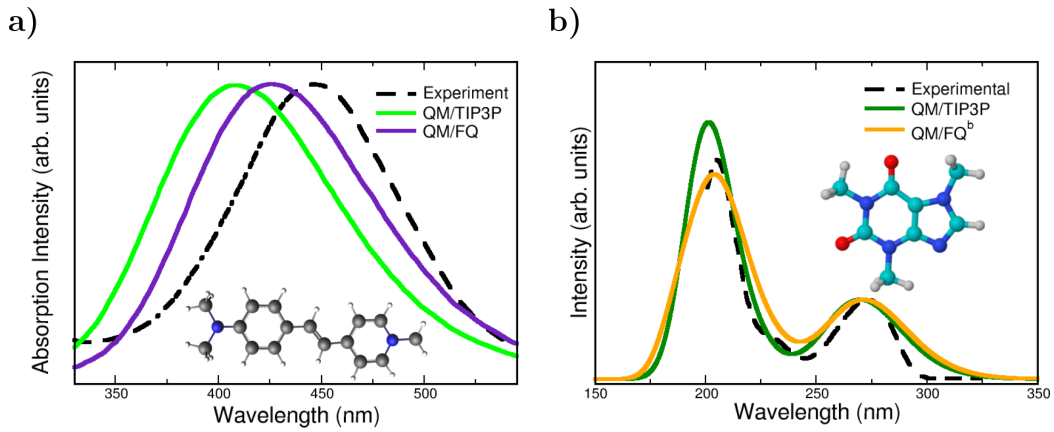


Figure 7. QM/TIP3P, QM/FQ, and experimental UV−vis spectra of (a) pyridinium dye[55] and (b) caffeine[56] in aqueous solution. Experimental spectra from refs 105 and 106. The pyridinium dye was rigid whereas Caffeine was freely allowed to move during MD simulations. QM level: CAM-B3LYP/6-311++G(*d*, *p*) and B3LYP/6-311++G(*d*, *p*) for pyridinium dye and caffeine, respectively. Image (a) reproduced and adapted from ref 55. Copyright 2019 John Wiley & Sons publications. Image (b) reproduced and adapted with permission from ref 56. Copyright 2020 Royal Society of Chemistry.

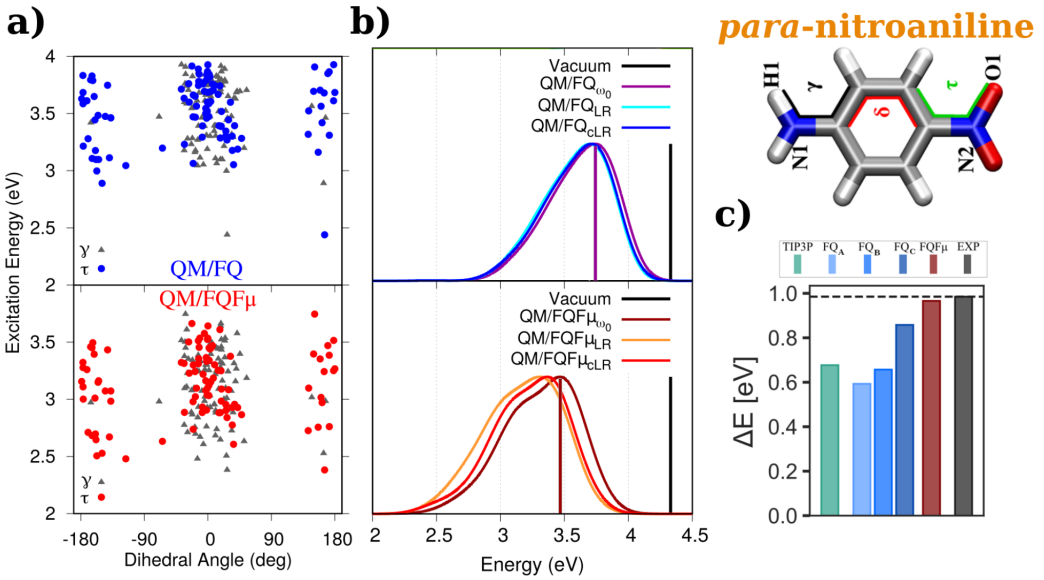


Figure 8. QM/FQ (top) and QM/FQFμ (bottom) (a) cLR excitation energies of PNA in aqueous solution as a function of selected dihedral angles and (b) ω_0, LR, and cLR spectra. Vertical bars mark the position of excitation energies in vacuum and those obtained with ω_0, i.e., the frozen density approximation.[13] c) PNA vacuo-to-water solvatochromic shifts, $\Delta E = E^{solv} - E^{vac}$, in water[54] computed with QM/EE and different QM/FQ parametrizations. PNA was freely allowed to move during MD simulations. QM level: CAM-B3LYP/aug-cc-pVDZ and CAMY-B3LYP/TZ2P. Images (a) and (b) reproduced and adapted from ref 13, with the permission of AIP Publishing, Copyright 2019 AIP Publishing. Image (c) reproduced and adapted with permission under a Creative Commons Attribution 4.0 International (CC BY 4.0) from ref 54. Copyright 2022 arXiv.

paraxanthine),[56] the accounting for polarization improves the relative intensities between the bands.

4.3.1. Different Treatments of Solvent Polarization: From QM/FQ to QM/FQFμ. At this point, it should be clear that a correct inclusion of mutual solute−solvent polarization and a detailed description of solute−solvent interactions (e.g., HBs) both need to be appropriately modeled to guarantee high-quality computed absorption spectra. Undoubtedly, polarizable QM/MM approaches are proper choices, but spectra can vary as a function of the chosen PE model. Basically, the alternatives differ in how the QM/MM coupling term is inserted in the QM Hamiltonian[1] and in the polarization sources under consideration. While charges rule polarization effects QM/FQ, charges and dipoles play that role in QM/FQFμ. Therefore, when QM/FQFμ is used, each MM atom is assigned a charge and a dipole which can vary according to the external electric potential and electric field. QM/FQFμ introduces a better description than QM/FQ of

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

short-range electrostatics and out-of-plane polarization effects.[13,42] Within the absorption spectroscopy scenario, the most notorious implications of using QM/FQFμ have been the shifting of the absorption bands and a more spread distribution of raw spectral data (sticks). Indeed, the application of QM/FQFμ has granted the calculation of accurate vacuo-to-water solvatochromic shifts not only in sign but also in value.[13,14]

Data for *para*-nitroaniline (PNA) are shown in Figure 8. Known for its solvatochromicity, PNA is typically used as a probe in the assessment of the quality of computational models for spectroscopy.[13,42,51,54] The effect of the conformational wealth on excitation energies of PNA is reported in Figure 8a. Small discrepancies between QM/FQ and QM/FQFμ vertical excitation energies are identified. The corresponding UV−vis spectra for PNA aqueous solutions are shown in Figure 8b. A few important points are immediately drawn: first, QM/FQFμ convoluted spectra show a larger in-homogeneous broadening with respect to QM/FQ, due to a broader distribution of the sticks. Second, the different description of electrostatic interactions given by QM/FQ and QM/FQFμ leads to a diverse value of solvatochromic shifts for the $\pi \rightarrow \pi^*$ transition of PNA (see the position of the vertical bars in Figure 8b). Experiments[107,108] suggest that the solvatochromic shift of PNA is 0.99 eV, which is a value better achieved in the case of the QM/FQFμ approach.[13,54]

Still commenting on PNA, it is well-known that a correct definition of the solvation regime, i.e., how to reliably describe the environment response following the electronic transition, is crucial to compute accurate excitation energies in solution. Specifically, here we focus on the LR and cLR approaches, which have already been extended to both QM/FQ and QM/FQFμ.[13] In brief, cLR, unlike LR, is able to catch the relaxation of the environment (the solvent in this case) as a response to the charge equilibration of the QM density to the specific excited state. Therefore, cLR is the most appropriate method when dealing with excitations that involve large density rearrangements.[32,109]Figure 8b shows computed excitation energies of solvated PNA[13] obtained either with LR or cLR regimes. In this specific case, LR and cLR excitation energies are similar, thus indicating that the LR scheme, where the response of the MM portion is adjusted to the QM transition density, is sufficient in the treatment of the excitation of this molecule. That is not always the case: in ref 42, it has been shown that relevant discrepancies between LR and cLR excitation energies might occur, for instance for 1-methyl-4-[(oxocyclohexadienylidene)ethylidene]-1,4-dihydropyridine (MER) and 2,6-diphenyl-4-(2,4,6-triphenylpyridin-1-ium-1-yl)phenolate (BET) and to a lesser extent for 1-methyl-8-oxy-quinolinium betaine (QB) when dissolved in different solvents (1,4-dioxane (DIO), tetrahydrofuran (THF), acetonitrile (ACN), ethanol (ETH), methanol (MET) and water (WAT)). Computed values are listed in Table 2.

4.3.2. How the Results Depend on the FQ Parametrization. In QM/MM approaches, the simulation of specific environments requires specific MM parametrization. As stated before, there is a list of predefined parameters that are used in the calculations to define the solvent's response. They are electronegativities and chemical hardnesses (and atomic polarizabilities) for QM/FQ (and QM/FQFμ). The hearth of the parametrization procedure is the fitting of calculated values of selected observables, such as interaction or total energies, with respect to reference data sets that can be obtained from full *ab initio* calculations on representative structures. Water or water−solute clusters of different molecularities have been often employed to find parameters for polarizable force fields. In this way, QM/FQ has been parametrized by Rick et al.,[29] Carnimeo et al.,[69] Giovannini et al.,[85] and lately by Ambrosetti et al.,[42] increasingly trying to refine computed spectral properties. QM/FQFμ parameters have been proposed once.[11] Many works have greatly benefited from one or another parametrization as can be read in Table 1. Overall, changing parameters improves the description of the position of the absorption bands[68] or relative intensities.[56] It is worth recalling here that what is changing when varying parameters is the degree of polarization that the solute can induce on the surrounding solvent molecules, thus directly affecting the solute's spectral response.[1]

Table 2. QM/FQ LR and cLR Excitation Energies (in eV) of QB, MER, and BET in Different Solvents[a]

	QB		MER		BET	
solvent	LR	cLR	LR	cLR	LR	cLR
DIO	2.62	2.55			2.43	2.22
THF	2.57	2.51	2.63	2.85		
ACN	2.78	2.74	2.83	2.98	2.94	2.79
MET	3.26	3.17			3.55	3.30
ETH			2.98	3.19		
WTR	3.41	3.35	3.18	3.33	3.88	3.71

[a]Data taken from Ref 42. QB, MER, and BET were kept frozen in their minimum energy structure during MD simulations.

Recently, the performance of different QM/MM embedding models and parametrizations to compute vacuo-to-water solvatochromic shifts has been investigated by Nicoli et al.[54] For PNA, those results are collected in Figure 8c. In that work, for QM/FQ calculations, three sets of parameters, namely, QM/FQ A,[29] QM/FQ B,[69] and QM/FQ C,[70] are reviewed. In all cases, the sign is correctly reproduced and after refinement, the solvatochromic shift moves toward the experimental value. Notwithstanding, QM/FQFμ outperforms the other models for this particular system and, although it is not the emphasis of this section, the shift given by the nonpolarizable QM/TIP3P method compares somewhat poorly to the experimental value.

To conclude this section, we refer interested readers to a recent assessment of the performance of different polarizable embedding approaches, ranging from QM/FQ to QM/Discrete Reaction Field (QM/DRF) and QM/FQFμ, in the reproduction of solvatochromic shifts of several dyes dissolved in aqueous solution.[54]

4.4. Inclusion of Nonelectrostatic Terms

As stated in section 2, most QM/MM approaches limit the QM/MM interaction term in eq 2 to electrostatic and polarization contributions. Nonelectrostatic interactions (repulsion and dispersion) are commonly included through parametric functions, such as the Lennard-Jones potential, which does not depend on the QM density. As a consequence, nonelectrostatic terms modify neither the GS nor the excited state QM density and only indirectly affect the final numerical values. The difficulty in treating QM/MM nonelectrostatic interactions stands in the fact that, since the MM part is classical, *ad hoc* models need to be constructed to properly model interactions of a purely quantum nature.

Here, we focus on how Pauli repulsion $E^{rep}_{QM/MM}$ can affect absorption properties of solvated systems, by exploiting a

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

method which has been developed in our group and is general enough to be applied to any kind of environment and QM/MM method.[14,15,57,70] The approach is rather simple: each MM molecule is endowed with a set of functions mimicking the QM density of the MM portion ρ_{MM}, and $E^{rep}_{QM/MM}$ is written as the opposite of an exchange integral, i.e.:[110,111]

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

The specific definition of the MM density ρ_{MM} for the case of water can be found elsewhere.[57] Here, it is worth remarking that the QM/MM repulsion energy is calculated as a two-electron integral, in which the MM density remains constant during the SCF cycles. As a consequence, $E^{rep}_{QM/MM}$ results in a one-electron contribution to the QM Hamiltonian, thus leading to a modification of the MO coefficients and energies. Therefore, Pauli repulsion gives an indirect contribution to excited state properties. Cases studied in the literature introducing these also-called quantum confinement effects include solvatochromic shifts evaluated with different levels of theory: nonpolarizable fixed-charges QM/TIP3P model, polarizable QM/FQ and QM/FQFμ models, and both polarizable models with the addition of repulsion forces.[14] To showcase how nonelectrostatics can affect absorption properties, we focus on the well-known case of vacuo-to-water solvatochromic shifts of acrolein.[14,20]

Some authors have become aware of the role of those effects when observing that solvatochromic shifts change by treating a few solvent molecules surrounding the solute at the QM level.[20] Of course the criticism/disadvantage of doing so lies in the expensiveness of the selection process of the solvent molecules (how many, where to put them) and the calculation itself. By using our approach, it is no longer necessary to take care of explicit water molecules to account for nonelectrostatic effects. By going back to Figure 6, we see that when the Pauli repulsion contribution between the QM solute and water molecules is considered (green bars) in addition to the electrostatic coupling, the solvatochromic shift for the $n \rightarrow \pi^*$ transition reduces and approaches the experiment. The same does not apply to the $\pi \rightarrow \pi^*$ transition, which requires the further consideration of charge exchange between water molecules to be perfectly simulated (see also section 2).[14] Similar conclusions can be drawn after including repulsion in the modeling of the UV−vis spectra of solvated nitrite.[65] These findings also reveal that nonelectrostatic effects in absorption spectroscopy are indeed not uniform, and nonobvious predictions can be made *a priori*.

Nonelectrostatic contributions may also be modeled by resorting to multiple-layer methods. In this scenario, excitation energies of solvated acrolein have been studied by using a multilayer polarizable embedding approach with frozen density embedding (FDE), QM/FDE/FQ,[61] or by coupling different quantum-embedding approaches with a third FQ layer.[51] In these approaches, the environment's density is retained, thus quantum forces are automatically included, without the need of resorting to specific parametrizations of the environment.

4.5. Brief Note on the Influence of Charge Transfer

Charge transfer (CT) effects can occur within the same molecule (intrasolute or intramolecular), between solute and solvent, and between solvent molecules. In the intrasolute case, excitations of CT character refer to transitions associated with significant movements of the electron density. The degree to which CT takes place can be assessed by means of different descriptors, such as the D_{CT} index based on the difference between the unrelaxed (or relaxed) excited-state density and ground-state density.[112] This spatial index indicates the charge displacement which is associated with the transition and has been exploited in connection to polarizable QM/MM approaches in various works.[13,55,56,59]

QM/FQ and QM/FQFμ can allow intermolecular solvent−solvent CT, i.e., the charge transfer between MM molecules (see section 2). Such effects have been evaluated for acrolein, pyridine, and pyrimidine in ref 14, and Figure 6 shows that, for acrolein, QM/FQ$_{CT}$ and QM/FQFμ_{CT} give solvatochromic shifts in almost perfect agreement with experiments when coupled to nonelectrostatic effects.

As for solute−solvent CT, neither QM/FQ nor QM/FQFμ (and none of the available QM/MM approaches) allows the exchange of charge between the two portions. When strong charge transfer occurs between solute and solvent, the orbitals involved in the transitions belong not only to the solute but also to the nearest solvent molecules. This happens in the case of nitrite[65] (see one of the molecular orbitals in Figure 9a),

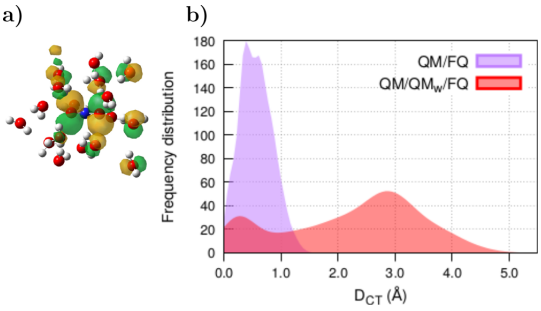


Figure 9. (a) Picture of one of the molecular orbitals involved in the UV−vis transitions of NO_2^- in water, obtained with the QM/QM$_w$/FQ approach. FQ water molecules are omitted. (b) Distribution of D_{CT} indices computed with QM/FQ and QM/QM$_w$/FQ on 200 snapshots extracted from MD trajectories of solvated NO_2^-. Twelve excited states are taken into account in each snapshot. NO_2^- was freely allowed to move during MD simulations. QM level: CAM-B3LYP/6-311++G(d, p). Numerical results are taken from ref 65.

where it is necessary to add the first, second, and third solvation shells in the QM portion—turning into QM/QM$_w$/FQ—in order to recover the experimental position of the main band in the UV−vis spectrum. Distributions of D_{CT} indexes for the 12 transitions of NO_2^- along the 200 snapshots (12 × 200 sticks) analyzed by Uribe et al.[65] are reported in Figure 9b. Clearly, the two distributions peak in different values, and the addition of water molecules to the QM portion makes the distribution flat, at the same time moving the main maximum to ≈2.9 Å, which confers to some excitations the feature of being of charge transfer nature. The need to include solute ↔ solvent charge transfer effects for a correct description of experimental spectra of small anions in aqueous solution[65,113] shows that even those apparently simple molecules hide a complex electronic structure that entails $n \rightarrow \pi^*$, $\pi \rightarrow \pi^*$ transitions and charge transfer states.

Some attempts to describe the charge transfer in multiscale simulations have been done by Lin et al.[114−116] who proposed a model with a flexible boundary characterized by *on-the-fly*

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

exchanges of partial charges and atoms between the QM and MM subsystems. However, to the best of our knowledge, such an approach has only been applied to the description of ground state properties.

4.6. Toward Nonaqueous Media

Until now, we have focused on the specific case of aqueous systems. QM/FQ has recently been extended to nonaqueous solvents of various polarity and hydrogen-bonding capability,[42] namely, 1,4-dioxane, tetrahydrofuran, acetonitrile, ethanol, and methanol. This implies the specification of the values of atomic electronegativities and chemical hardnesses, which are independent of the solute and/or the spectral observable but only depend on the solvent, which ultimately makes the parametrization transferable. Detailed results for PNA are reported in Figure 10, which shows that excitation energies

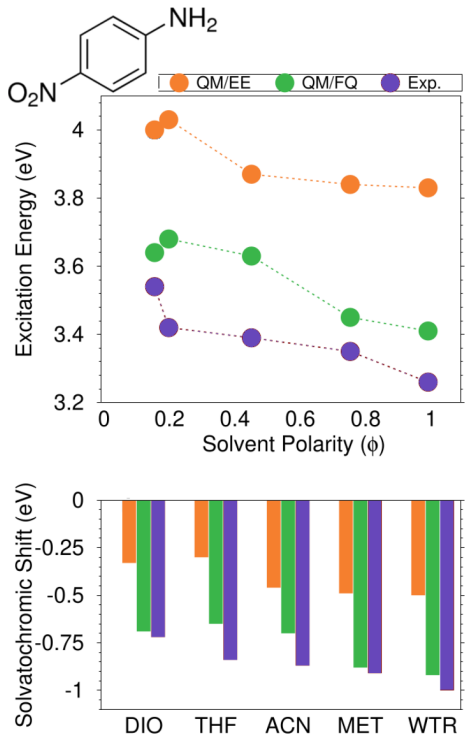


Figure 10. Top panel: QM/EE, QM/FQ, and experimental PNA excitation energies as a function of the solvent polarity. Bottom panel: solvatochromic shifts in diverse solvents, computed with respect to gas-phase.[42] PNA was kept frozen in its minimum energy structure during MD simulations. QM level: CAM-B3LYP/aug-cc-pVDZ. Image adapted from ref 42. Copyright 2021 American Chemical Society.

and, in turn, solvatochromic shifts increase with the polarity of the solvent. QM/FQ excitation energies are almost in line with the experimental trends, giving excellent values and in some cases exhibiting outstanding agreement with experimental data. Conversely, QM/EE strongly underestimates experimentally measured shifts, thus revealing once again the relevance of the solute−solvent polarization as highlighted in a previous section and in ref 42. Finally, it is worth noting that, for low-polar solvents, both QM/EE and QM/FQ fail at reproducing the experimental trends. This might be due to the lack of nonelectrostatic interactions in our modeling,[14,15,57,70] which may be particularly relevant in these cases.

5. CONCLUDING REMARKS

In this Perspective, we have shown, by means of clear-cut examples, the crucial role of diverse aspects in the simulation of electronic absorption spectra of solvated systems. Thus, a model intended to obtain a reliable description of experimental data and to have predictive power should be able to integrate essential elements, such as the phase-space sampling, solvation effects, and a physically sound calculation of the spectral signal. The dynamical aspects of the solvation phenomenon can be considered by resorting to classical MD simulations, whereas the computation of spectral properties can be performed by using multiscale QM/MM methodologies. As expected, spectral features depend on the reliability of the force fields used in MD and QM/MM calculations. For the latter, QM/FQ has proven to be particularly advantageous because of its flexibility to integrate the various effects in place in the solvated sample, thus treating all solute−solvent interactions that are meaningful for the generation of the spectroscopic response. Mutual polarization effects, nonelectrostatic contributions, relaxation effects in the solvent response, and even vibronic effects can be properly included. Broadly speaking, when QM/FQ in any of its flavors—QM/FQ or QM/FQFμ—has been challenged to reproduce absorption spectra or solvatochromic shifts of systems in solution, results have found better agreement with experimental data than continuum solvation and electrostatic embedding approaches.

Despite their success, further improvements for QM/FQ and QM/FQFμ can be foreseen. The first development line is their extensive parametrization for diverse environments, which will allow treating more and more complex systems, up to biological environments. The parametrization effort also extends to quantum repulsion, and eventually, quantum dispersion, which up to now have been only challenged for aqueous systems. In addition, dealing with large systems implies the immediate use of enhanced MD sampling techniques and a careful choice of the partitioning of the system. Also, in order to make the computational approach entirely coherent with the physics of the absorption phenomenon, charge transfer interactions (both between QM/MM or MM/MM moieties) would need to be included. However, universal protocols for treating such effects are not available yet, making it an interesting topic for future investigations. We finally note that, in the above discussion, effects arising from the electronic-nuclear coupling, which also affect the spectral shape, have not been emphasized. More extensive investigations, in line with recent studies,[117] would contribute to increase the quality of the description of experimental spectra.

AUTHOR INFORMATION

Corresponding Author

Chiara Cappelli − *Scuola Normale Superiore, Classe di Scienze, 56126 Pisa, Italy;* orcid.org/0000-0002-4872-4505; Email: chiara.cappelli@sns.it

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

Authors

Sara Gómez − *Scuola Normale Superiore, Classe di Scienze, 56126 Pisa, Italy;* orcid.org/0000-0002-5430-9228

Tommaso Giovannini − *Scuola Normale Superiore, Classe di Scienze, 56126 Pisa, Italy;* orcid.org/0000-0002-5637-2853

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphyschemau.2c00050

Author Contributions

CRediT: **Sara Gomez** data curation (lead), writing-original draft (lead); **Tommaso Giovannini** data curation (supporting), software (lead), writing-original draft (supporting); **Chiara Cappelli** conceptualization (lead), funding acquisition (lead), resources (lead), supervision (lead), writing-review & editing (lead).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 818064). We gratefully acknowledge the Center for High Performance Computing (CHPC) at SNS for providing the computational infrastructure.

■ REFERENCES

(1) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664−5677.

(2) Morzan, U. N.; Alonso de Armino, D. J.; Foglia, N. O.; Ramirez, F.; Gonzalez Lebrero, M. C.; Scherlis, D. A.; Estrin, D. A. Spectroscopy in complex environments from QM−MM simulations. *Chem. Rev.* **2018**, *118*, 4071−4113.

(3) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999−3094.

(4) Provorse, M. R.; Peev, T.; Xiong, C.; Isborn, C. M. Convergence of excitation energies in mixed quantum and classical solvent: Comparison of continuum and point charge models. *J. Phys. Chem. B* **2016**, *120*, 12148−12159.

(5) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Sorting out the relative contributions of electrostatic polarization, dispersion, and hydrogen bonding to solvatochromic shifts on vertical electronic excitation energies. *J. Chem. Theory Comput.* **2010**, *6*, 2829−2844.

(6) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(7) Dohn, A. O. Multiscale electrostatic embedding simulations for modeling structure and dynamics of molecules in solution: a tutorial review. *Int. J. Quantum Chem.* **2020**, *120*, No. e26343.

(8) Bondanza, M.; Nottoli, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. Polarizable embedding QM/MM: the future gold standard for complex (bio) systems? *Phys. Chem. Chem. Phys.* **2020**, *22*, 14433−14448.

(9) Giovannini, T.; Egidi, F.; Cappelli, C. Theory and algorithms for chiroptical properties and spectroscopies of aqueous systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864−22879.

(10) Cappelli, C. Integrated QM/Polarizable MM/Continuum Approaches to Model Chiroptical Properties of Strongly Interacting Solute-Solvent Systems. *Int. J. Quantum Chem.* **2016**, *116*, 1532−1542.

(11) Giovannini, T.; Puglisi, A.; Ambrosetti, M.; Cappelli, C. Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQFμ model. *J. Chem. Theory Comput.* **2019**, *15*, 2233−2245.

(12) Giovannini, T.; Grazioli, L.; Ambrosetti, M.; Cappelli, C. Calculation of ir spectra with a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles. *J. Chem. Theory Comput.* **2019**, *15*, 5495−5507.

(13) Giovannini, T.; Riso, R. R.; Ambrosetti, M.; Puglisi, A.; Cappelli, C. Electronic Transitions for a Fully Polarizable QM/MM Approach Based on Fluctuating Charges and Fluctuating Dipoles: Linear and Corrected Linear Response Regimes. *J. Chem. Phys.* **2019**, *151*, 174104.

(14) Giovannini, T.; Ambrosetti, M.; Cappelli, C. Quantum Confinement Effects on Solvatochromic Shifts of Molecular Solutes. *J. Phys. Chem. Lett.* **2019**, *10*, 5823−5829.

(15) Marrazzini, G.; Giovannini, T.; Egidi, F.; Cappelli, C. Calculation of linear and non-linear electric response properties of systems in aqueous solution: A polarizable quantum/classical approach with quantum repulsion effects. *J. Chem. Theory Comput.* **2020**, *16*, 6993−7004.

(16) Ma, H.; Ma, Y. Solvent effect on electronic absorption, fluorescence, and phosphorescence of acetone in water: Revisited by quantum mechanics/molecular mechanics (QM/MM) simulations. *J. Chem. Phys.* **2013**, *138*, 224505.

(17) Zuehlsdorff, T. J.; Isborn, C. M. Modeling absorption spectra of molecules in solution. *Int. J. Quantum Chem.* **2019**, *119*, No. e25719.

(18) Parac, M.; Doerr, M.; Marian, C. M.; Thiel, W. QM/MM calculation of solvent effects on absorption spectra of guanine. *J. Comput. Chem.* **2010**, *31*, 90−106.

(19) Reichardt, C. Solvatochromic dyes as solvent polarity indicators. *Chem. Rev.* **1994**, *94*, 2319−2358.

(20) Aidas, K.; Møgelhøj, A.; Nilsson, E. J.; Johnson, M. S.; Mikkelsen, K. V.; Christiansen, O.; Söderhjelm, P.; Kongsted, J. On the performance of quantum chemical methods to predict solvatochromic effects: The case of acrolein in aqueous solution. *J. Chem. Phys.* **2008**, *128*, 194503.

(21) Reichardt, C. Solvatochromism, thermochromism, piezochromism, halochromism, and chiro-solvatochromism of pyridinium N-phenoxide betaine dyes. *Chem. Soc. Rev.* **1992**, *21*, 147−153.

(22) Warshel, A. Calculations of chemical processes in solutions. *J. Phys. Chem.* **1979**, *83*, 1640−1652.

(23) Söderhjelm, P.; Husberg, C.; Strambi, A.; Olivucci, M.; Ryde, U. Protein influence on electronic spectra modeled by multipoles and polarizabilities. *J. Chem. Theory Comput.* **2009**, *5*, 649−658.

(24) Isborn, C. M.; Gotz, A. W.; Clark, M. A.; Walker, R. C.; Martínez, T. J. Electronic absorption spectra from MM and ab initio QM/MM molecular dynamics: Environmental effects on the absorption spectrum of photoactive yellow protein. *J. Chem. Theory Comput.* **2012**, *8*, 5092−5106.

(25) Rocha, F. S.; Gomes, A. J.; Lunardi, C. N.; Kaliaguine, S.; Patience, G. S. Experimental methods in chemical engineering: Ultraviolet visible spectroscopy−UV-Vis. *Can. J. Chem. Eng.* **2018**, *96*, 2512−2517.

(26) Macchiagodena, M.; Mancini, G.; Pagliai, M.; Del Frate, G.; Barone, V. Fine-tuning of atomic point charges: Classical simulations of pyridine in different environments. *Chem. Phys. Lett.* **2017**, *677*, 120−126.

(27) Stone, A.*The Theory of Intermolecular Forces*; Oxford University Press, 2013.

(28) Sanderson, R. An interpretation of bond lengths and a classification of bonds. *Science* **1951**, *114*, 670−672.

(29) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* **1994**, *101*, 6141−6156.

(30) Mayer, A. Formulation in terms of normalized propagators of a charge-dipole model enabling the calculation of the polarization properties of fullerenes and carbon nanotubes. *Phys. Rev. B* **2007**, *75*, 045407.

(31) Mennucci, B.; Cammi, R.; Tomasi, J. Excited states and solvatochromic shifts within a nonequilibrium solvation approach: A new formulation of the integral equation formalism method at the

self-consistent field, configuration interaction, and multiconfiguration self-consistent field level. *J. Chem. Phys.* **1998**, *109*, 2798−2807.
(32) Caricato, M.; Mennucci, B.; Tomasi, J.; Ingrosso, F.; Cammi, R.; Corni, S.; Scalmani, G. Formation and relaxation of excited states in solution: A new time dependent polarizable continuum model based on time dependent density functional theory. *J. Chem. Phys.* **2006**, *124*, 124520.
(33) Rinkevicius, Z.; Li, X.; Sandberg, J. A.; Mikkelsen, K. V.; Ågren, H. A hybrid density functional theory/molecular mechanics approach for linear response properties in heterogeneous environments. *J. Chem. Theory Comput.* **2014**, *10*, 989−1003.
(34) Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. Excitation energies in solution: the fully polarizable QM/MM/PCM method. *J. Phys. Chem. B* **2011**, *115*, 3027−3037.
(35) Loco, D.; Polack, É.; Caprasecca, S.; Lagardere, L.; Lipparini, F.; Piquemal, J.-P.; Mennucci, B. A QM/MM approach using the AMOEBA polarizable embedding: from ground state energies to electronic excitations. *J. Chem. Theory Comput.* **2016**, *12*, 3654−3661.
(36) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited states in solution through polarizable embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721−3734.
(37) Olsen, J. M. H.; Kongsted, J. Molecular properties through polarizable embedding. *Adv. Quantum Chem.* **2011**, *61*, 107−143.
(38) Schwabe, T. General theory for environmental effects on (vertical) electronic excitation energies. *J. Chem. Phys.* **2016**, *145*, 154105.
(39) Schröder, H.; Schwabe, T. Corrected Polarizable Embedding: Improving the Induction Contribution to Perichromism for Linear Response Theory. *J. Chem. Theory Comput.* **2018**, *14*, 833−842.
(40) Casida, M. E. In *Recent Advances in Density Functional Methods Part I*; Chong, D. P., Ed.; World Scientific: Singapore, 1995; pp 155−192.
(41) Guido, C. A.; Chrayteh, A.; Scalmani, G.; Mennucci, B.; Jacquemin, D. Simple Protocol for Capturing Both Linear-Response and State-Specific Effects in Excited-State Calculations with Continuum Solvation Models. *J. Chem. Theory Comput.* **2021**, *17*, 5155−5164.
(42) Ambrosetti, M.; Skoko, S.; Giovannini, T.; Cappelli, C. Quantum Mechanics/Fluctuating Charge Protocol to Compute Solvatochromic Shifts. *J. Chem. Theory Comput.* **2021**, *17*, 7146−7156.
(43) Jacquemin, D.; Mennucci, B.; Adamo, C. Excited-state calculations with TD-DFT: from benchmarks to simulations in complex environments. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16987−16998.
(44) Jacquemin, D.; Planchat, A.; Adamo, C.; Mennucci, B. TD-DFT assessment of functionals for optical 0−0 transitions in solvated dyes. *J. Chem. Theory Comput.* **2012**, *8*, 2359−2372.
(45) Fahim, Z. M. E.; Bouzzine, S. M.; Youssef, A. A.; Bouachrine, M.; Hamidi, M. Ground state geometries, UV/vis absorption spectra and charge transfer properties of triphenylamine-thiophenes based dyes for DSSCs: A TD-DFT benchmark study. *Comput. Theor. Chem.* **2018**, *1125*, 39−48.
(46) Suellen, C.; Freitas, R. G.; Loos, P.-F.; Jacquemin, D. Cross-comparisons between experiment, TD-DFT, CC, and ADC for transition energies. *J. Chem. Theory Comput.* **2019**, *15*, 4581−4590.
(47) Sarkar, R.; Boggio-Pasqua, M.; Loos, P.-F.; Jacquemin, D. Benchmarking TD-DFT and wave function methods for oscillator strengths and excited-state dipole moments. *J. Chem. Theory Comput.* **2021**, *17*, 1117−1132.
(48) Jacquemin, D.; Adamo, C. In *Density-Functional Methods for Excited States*; Ferré, N., Filatov, M., Huix-Rotllant, M., Eds.; Springer International Publishing: Cham, 2016; pp 347−375.
(49) Bursch, M.; Mewes, J.-M.; Hansen, A.; Grimme, S. Best-Practice DFT Protocols for Basic Molecular Computational Chemistry. *Angew. Chem.* **2022**, *134*, e202205735.
(50) Caricato, M.; Lipparini, F.; Scalmani, G.; Cappelli, C.; Barone, V. Vertical electronic excitations in solution with the EOM-CCSD method combined with a polarizable explicit/implicit solvent model. *J. Chem. Theory Comput.* **2013**, *9*, 3035−3042.
(51) Goletto, L.; Giovannini, T.; Folkestad, S. D.; Koch, H. Combining multilevel Hartree−Fock and multilevel coupled cluster approaches with molecular mechanics: a study of electronic excitations in solutions. *Phys. Chem. Chem. Phys.* **2021**, *23*, 4413−4425.
(52) Sneskov, K.; Schwabe, T.; Kongsted, J.; Christiansen, O. The polarizable embedding coupled cluster method. *J. Chem. Phys.* **2011**, *134*, 104108.
(53) Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. Linear response functions for coupled cluster/molecular mechanics including polarization interactions. *J. Chem. Phys.* **2003**, *118*, 1620−1633.
(54) Nicoli, L.; Giovannini, T.; Cappelli, C.Assessing the Quality of QM/MM Approaches to Describe Vacuo-to-water Solvatochromic Shifts. *arXiv (Chemical Physics)*, October 27, **2022**, 2210.15412, ver. 1. https://arxiv.org/abs/2210.15412 (accessed on 2022-11-02).
(55) Giovannini, T.; Macchiagodena, M.; Ambrosetti, M.; Puglisi, A.; Lafiosca, P.; Lo Gerfo, G.; Egidi, F.; Cappelli, C. Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling. *Int. J. Quantum Chem.* **2019**, *119*, No. e25684.
(56) Gómez, S.; Giovannini, T.; Cappelli, C. Absorption Spectra of Xanthines in Aqueous Solution: A Computational Study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5929−5941.
(57) Giovannini, T.; Lafiosca, P.; Cappelli, C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. *J. Chem. Theory Comput.* **2017**, *13*, 4854−4870.
(58) Egidi, F.; Russo, R.; Carnimeo, I.; D'Urso, A.; Mancini, G.; Cappelli, C. The Electronic Circular Dichroism of Nicotine in Aqueous Solution: A Test Case for Continuum and Mixed Explicit-Continuum Solvation Approaches. *J. Phys. Chem. A* **2015**, *119*, 5396−5404.
(59) Egidi, F.; Lo Gerfo, G.; Macchiagodena, M.; Cappelli, C. On the nature of charge-transfer excitations for molecules in aqueous solution: a polarizable QM/MM study. *Theor. Chem. Acc.* **2018**, *137*, 82.
(60) Lafiosca, P.; Gómez, S.; Giovannini, T.; Cappelli, C. Absorption Properties of Large Complex Molecular Systems: The DFTB/Fluctuating Charge Approach. *J. Chem. Theory Comput.* **2022**, *18*, 1765−1779.
(61) Egidi, F.; Angelico, S.; Lafiosca, P.; Giovannini, T.; Cappelli, C. A polarizable three-layer frozen density embedding/molecular mechanics approach. *J. Chem. Phys.* **2021**, *154*, 164107.
(62) Di Remigio, R.; Giovannini, T.; Ambrosetti, M.; Cappelli, C.; Frediani, L. Fully polarizable QM/fluctuating charge approach to two-photon absorption of aqueous solutions. *J. Chem. Theory Comput.* **2019**, *15*, 4056−4068.
(63) Puglisi, A.; Giovannini, T.; Antonov, L.; Cappelli, C. Interplay between conformational and solvent effects in UV-visible absorption spectra: Curcumin tautomers as a case study. *Phys. Chem. Chem. Phys.* **2019**, *21*, 15504−15514.
(64) Skoko, S.; Ambrosetti, M.; Giovannini, T.; Cappelli, C. Simulating Absorption Spectra of Flavonoids in Aqueous Solution: A Polarizable QM/MM Study. *Molecules* **2020**, *25*, 5853.
(65) Uribe, L.; Gómez, S.; Giovannini, T.; Egidi, F.; Restrepo, A. An efficient and robust procedure to calculate absorption spectra of aqueous charged species applied to NO 2-. *Phys. Chem. Chem. Phys.* **2021**, *23*, 14857−14872.
(66) Gómez, S.; Egidi, F.; Puglisi, A.; Giovannini, T.; Rossi, B.; Cappelli, C. Unlocking the Power of Resonance Raman Spectroscopy: The Case of Amides in Aqueous Solution. *J. Mol. Liq.* **2022**, *346*, 117841.
(67) Gómez, S.; Rojas-Valencia, N.; Giovannini, T.; Restrepo, A.; Cappelli, C. Ring Vibrations to Sense Anionic Ibuprofen in Aqueous Solution as Revealed by Resonance Raman. *Molecules* **2022**, *27*, 442.

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

(68) Gómez, S.; Bottari, C.; Egidi, F.; Giovannini, T.; Rossi, B.; Cappelli, C. Amide Spectral Fingerprints are Hydrogen Bonding-Mediated. *J. Phys. Chem. Lett.* **2022**, *13*, 6200−6207.
(69) Carnimeo, I.; Cappelli, C.; Barone, V. Analytical gradients for MP2, double hybrid functionals, and TD-DFT with polarizable embedding described by fluctuating charges. *J. Comput. Chem.* **2015**, *36*, 2271−2290.
(70) Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. Effective yet Reliable Computation of Hyperfine Coupling Constants in Solution by a QM/MM Approach: Interplay Between Electrostatics and Non-electrostatic Effects. *J. Chem. Phys.* **2019**, *150*, 124102.
(71) Frenkel, D.; Smit, B.*Understanding molecular simulation: from algorithms to applications*, 2nd ed.; Academic Press: San Diego, 2002.
(72) Pérez, J.; Restrepo, A.ASCEC V02:Annealing Simulado con Energía Cuántica. 2008; *Property, development, and implementation: Grupo de Química−Física Teórica*; Instituto de Química, Universidad de Antioquia, Medellín, Colombia.
(73) Pérez, J. F.; Hadad, C. Z.; Restrepo, A. Structural studies of the water tetramer. *Int. J. Quantum Chem.* **2008**, *108*, 1653−1659.
(74) Pérez, J. F.; Florez, E.; Hadad, C. Z.; Fuentealba, P.; Restrepo, A. Stochastic Search of the Quantum Conformational Space of Small Lithium and Bimetallic Lithium−Sodium Clusters. *J. Phys. Chem. A* **2008**, *112*, 5749−5755.
(75) Vainio, M. J.; Johnson, M. S. Generating conformer ensembles using a multiobjective genetic algorithm. *J. Chem. Inf. Model.* **2007**, *47*, 2462−2474.
(76) Ishikawa, Y. A script for automated 3-dimentional structure generation and conformer search from 2-dimentional chemical drawing. *Bioinformation* **2013**, *9*, 988.
(77) Poli, G.; Seidel, T.; Langer, T. Conformational sampling of small molecules with iCon: Performance assessment in comparison with OMEGA. *Front. Chem.* **2018**, *6*, 229.
(78) Gao, J. Monte Carlo Quantum Mechanical-Configuration Interaction and Molecular Mechanics Simulation of Solvent Effects on the n → pi* Blue Shift of Acetone. *J. Am. Chem. Soc.* **1994**, *116*, 9324−9328.
(79) Brooks, C. L.; Case, D. A.; Plimpton, S.; Roux, B.; van der Spoel, D.; Tajkhorshid, E. Classical molecular dynamics. *J. Chem. Phys.* **2021**, *154*, 100401.
(80) Iftimie, R.; Minary, P.; Tuckerman, M. E. Ab initio molecular dynamics: Concepts, recent developments, and future trends. *P. Natl. Acad. Sci.* **2005**, *102*, 6654−6659.
(81) Marx, D.; Hutter, J.*Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press, 2009.
(82) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **2017**, *45*, W331−W336.
(83) Cacelli, I.; Prampolini, G. Parametrization and validation of intramolecular force fields derived from DFT calculations. *J. Chem. Theory Comput.* **2007**, *3*, 1803−1817.
(84) Barone, V.; Cacelli, I.; De Mitri, N.; Licari, D.; Monti, S.; Prampolini, G. J oyce and U lysses: integrated and user-friendly tools for the parameterization of intramolecular force fields from quantum mechanical data. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3736−3751.
(85) Giovannini, T.; Del Frate, G.; Lafiosca, P.; Cappelli, C. Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution. *Phys. Chem. Chem. Phys.* **2018**, *20*, 9181−9197.
(86) Egidi, F.; Giovannini, T.; Del Frate, G.; Lemler, P. M.; Vaccaro, P. H.; Cappelli, C. A combined experimental and theoretical study of optical rotatory dispersion for (R)-glycidyl methyl ether in aqueous solution. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3644−3655.
(87) Brehm, M.; Thomas, M.; Gehrke, S.; Kirchner, B. TRAVIS−A free analyzer for trajectories from molecular simulation. *J. Chem. Phys.* **2020**, *152*, 164105.
(88) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J.; Melo, M. N.; Seyler, S. L.; Domanski, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M., et al.MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. *Proceedings of the 15th python in science conference*; 2016; p 105.
(89) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319−2327.
(90) Biler, M.; Biedermann, D.; Valentová, K.; Křen, V.; Kubala, M. Quercetin and its analogues: Optical and acido−basic properties. *Phys. Chem. Chem. Phys.* **2017**, *19*, 26870−26879.
(91) Renger, T.; Müh, F. Understanding photosynthetic light-harvesting: a bottom up theoretical approach. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3348−3371.
(92) Jurinovich, S.; Viani, L.; Curutchet, C.; Mennucci, B. Limits and potentials of quantum chemical methods in modelling photosynthetic antennae. *Phys. Chem. Chem. Phys.* **2015**, *17*, 30783−30792.
(93) Rosnik, A. M.; Curutchet, C. Theoretical characterization of the spectral density of the water-soluble chlorophyll-binding protein from combined quantum mechanics/molecular mechanics molecular dynamics simulations. *J. Chem. Theory Comput.* **2015**, *11*, 5826−5837.
(94) Curutchet, C.; Mennucci, B. Quantum chemical studies of light harvesting. *Chem. Rev.* **2017**, *117*, 294−343.
(95) Padula, D.; Lee, M. H.; Claridge, K.; Troisi, A. Chromophore-dependent intramolecular exciton−vibrational coupling in the FMO complex: Quantification and importance for exciton dynamics. *J. Phys. Chem. B* **2017**, *121*, 10026−10035.
(96) Maity, S.; Bold, B. M.; Prajapati, J. D.; Sokolov, M.; Kubar Kubař, T.; Elstner, M.; Kleinekathöfer, U. DFTB/MM molecular dynamics simulations of the FMO light-harvesting complex. *J. Phys. Chem. Lett.* **2020**, *11*, 8660−8667.
(97) Chaillet, M. L.; Lengauer, F.; Adolphs, J.; Müh, F.; Fokas, A. S.; Cole, D. J.; Chin, A. W.; Renger, T. Static disorder in excitation energies of the Fenna−Matthews−Olson protein: Structure-based theory meets experiment. *J. Phys. Chem. Lett.* **2020**, *11*, 10306−10314.
(98) Llano, S.; Gómez, S.; Londoño, J.; Restrepo, A. Antioxidant activity of curcuminoids. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3752−3760.
(99) Manolova, Y.; Deneva, V.; Antonov, L.; Drakalska, E.; Momekova, D.; Lambov, N. The effect of the water on the curcumin tautomerism: A quantitative approach. *Spectrochim. Acta, Part A* **2014**, *132*, 815−820.
(100) Galasso, V.; Kovac, B.; Modelli, A.; Ottaviani, M. F.; Pichierri, F. Spectroscopic and theoretical study of the electronic structure of curcumin and related fragment molecules. *J. Phys. Chem. A* **2008**, *112*, 2331−2338.
(101) Moskvin, A.; Yablonskii, O.; Bondar, L. An experimental investigation of the effect of alkyl substituents on the position of the K and R absorption bands in acrolein derivatives. *Theor. Exp. Chem.* **1968**, *2*, 469−472.
(102) Weinhold, F.; Landis, C.; Glendening, E. What is NBO analysis and how is it useful? *Int. Rev. Phys. Chem.* **2016**, *35*, 399−440.
(103) Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. Scrutinizing the effects of polarization in QM/MM excited state calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551−18560.
(104) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(105) Abdel-Halim, S. T.; Awad, M. K. Solvatochromism, molecular and electronic structures of trans and cis isomers of a typical styryl pyridinium cyanine dye. *J. Mol. Struct.* **2009**, *920*, 332−341.
(106) Chen, J.; Kohler, B. Ultrafast nonradiative decay by hypoxanthine and several methylxanthines in aqueous and acetonitrile solution. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10677−10682.
(107) Millefiori, S.; Favini, G.; Millefiori, A.; Grasso, D. Electronic spectra and structure of nitroanilines. *Spectrochim. Acta, Part A* **1977**, *33*, 21−27.
(108) Kovalenko, S.; Schanz, R.; Farztdinov, V.; Hennig, H.; Ernsting, N. Femtosecond relaxation of photoexcited para-nitroaniline: solvation, charge transfer, internal conversion and cooling. *Chem. Phys. Lett.* **2000**, *323*, 312−322.

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

(109) Guido, C. A.; Jacquemin, D.; Adamo, C.; Mennucci, B. Electronic excitations in solution: the interplay between state specific approaches and a time-dependent density functional theory description. *J. Chem. Theory Comput.* **2015**, *11*, 5782−5790.
(110) Mennucci, B.; Amovilli, C.; Tomasi, J. On the effect of Pauli repulsion and dispersion on static molecular polarizabilities and hyperpolarizabilities in solution. *Chem. Phys. Lett.* **1998**, *286*, 221−225.
(111) McWeeny, R.*Methods of Molecular Quantum Mechanics*; Academic Press Elsevier Science: London, 1992.
(112) Le Bahers, T.; Adamo, C.; Ciofini, I. A qualitative index of spatial extent in charge-transfer excitations. *J. Chem. Theory Comput.* **2011**, *7*, 2498−2506.
(113) Uribe, L.; Gómez, S.; Egidi, F.; Giovannini, T.; Restrepo, A. Computational hints for the simultaneous spectroscopic detection of common contaminants in water. *J. Mol. Liq.* **2022**, *355*, 118908.
(114) Zhang, Y.; Lin, H. Flexible-boundary quantum-mechanical/molecular-mechanical calculations: Partial charge transfer between the quantum-mechanical and molecular-mechanical subsystems. *J. Chem. Theory Comput.* **2008**, *4*, 414−425.
(115) Zhang, Y.; Lin, H. Flexible-boundary QM/MM calculations: II. Partial charge transfer across the QM/MM boundary that passes through a covalent bond. *Theor. Chem. Acc.* **2010**, *126*, 315−322.
(116) Pezeshki, S.; Lin, H. Molecular dynamics simulations of ion solvation by flexible-boundary QM/MM: on-the-fly partial charge transfer between QM and MM subsystems. *J. Comput. Chem.* **2014**, *35*, 1778−1788.
(117) Cerezo, J.; Aranda, D.; Avila Ferrer, F. J.; Prampolini, G.; Santoro, F. Adiabatic-molecular dynamics generalized vertical hessian approach: a mixed quantum classical method to compute electronic spectra of flexible molecules in the condensed phase. *J. Chem. Theory Comput.* **2020**, *16*, 1215−1231.

https://doi.org/10.1021/acsphyschemau.2c00050
ACS Phys. Chem Au XXXX, XXX, XXX−XXX

4.2. Drugs

4.2.1. Caffeine[2]

PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2020, **22**, 5929

Absorption spectra of xanthines in aqueous solution: a computational study†

Sara Gómez, [a] Tommaso Giovannini [b] and Chiara Cappelli *[a]

Received 3rd October 2019,
Accepted 11th February 2020

DOI: 10.1039/c9cp05420k

rsc.li/pccp

We present a detailed computational analysis of the UV/Vis spectra of caffeine, paraxanthine and theophylline in aqueous solution. A hierarchy of solvation approaches for modeling the aqueous environment have been tested, ranging from the continuum model to the non-polarizable and polarizable quantum mechanical (QM)/molecular mechanics (MM) models, with and without the explicit inclusion of water molecules in the QM portion. The computed results are directly compared with the experimental data, thus highlighting the role of electrostatic, polarization and hydrogen boding solute–solvent interactions.

1 Introduction

Xanthine is a chemical compound structurally comparable to purine.[1] It can be obtained from purine degradation and according to some metabolic pathways, it is thought to be converted to uric acid by means of oxidation processes.[2–4] Xanthine-derivatives, also called xanthines, and among them methylxanthines, have a wide range of pharmacological/clinical activities[5,6] including as mild stimulants[7] not only for the central nervous system but also for the respiratory center, depending on many factors such as dose and concentration.[8] Particular attention was paid to the fact that theophylline and other methylxanthines can block responses to adenosine, once it was discovered in 1970.[9–11]

Methylated xanthines include 8-chlorotheophylline (1,3-dimethyl-8-chloroxanthine), aminophylline (theophylline : ethylenediamine 2 : 1), IBMX (3-isobutyl-1-methylxanthine), pentoxifylline (1-(5-oxohexyl)-3,7-dimethylxanthine), paraxanthine (1,7-dimethylxanthine), theophylline (1,3-dimethylxanthine), theobromine (3,7-dimethylxanthine), and caffeine (1,3,7-trimethylxanthine), the latter being the most broadly known and consumed.[12–16] Some of these methylxanthines are displayed in Fig. 1, with

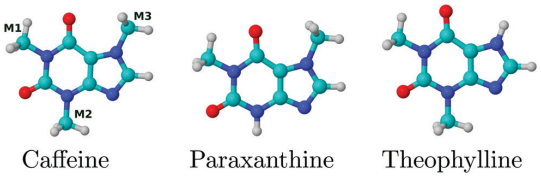


Fig. 1 Molecular structure of selected xanthines studied in this work. M1, M2 and M3 stand for the methyl groups, whose rotation is analyzed in Fig. 2. See also Fig. S1 in the ESI† for atom numbering.

water-solubility[17] being a common feature of all of them. As a matter of fact, a variety of tautomeric forms can exist for these xanthines, due to the influence of temperature, aggregation state, pH conditions, *etc.*[18–21]

Several analytical techniques, such as Fourier-transform infrared spectroscopy (FTIR), liquid chromatography-mass spectrometry (LC-MS), UV/Vis spectroscopy, nuclear magnetic resonance (NMR) and others can be used in the detection of xanthines.[22–28] For instance, UV absorption, which is the technique we have focused on in this paper, has been used in the identification and determination of caffeine in pharmaceuticals[29] as well as in coffee beans[30] and in brewed coffees.[31] The potential use of these methods has motivated many experimentalists and theoreticians to study the information provided by experimental techniques and their subsequent interpretation by using computational tools. Multiple works have been performed regarding electronic absorption spectra.[32,33] In particular, electronic absorption and fluorescence spectra of aqueous solutions of xanthine, caffeine, theophylline and theobromine have been studied at different pH values,[34] and the excited state lifetimes of hypoxanthine, paraxanthine, theophylline, theobromine, and caffeine were studied *via* femtosecond transient absorption spectroscopy performed in aqueous

[a] *Scuola Normale Superiore, Classe di Scienze, Piazza dei Cavalieri 7, 56126 Pisa, Italy. E-mail: chiara.cappelli@sns.it*
[b] *Department of Chemistry, Norwegian University of Science and Technology, 7491 Trondheim, Norway*

† Electronic supplementary information (ESI) available: Xanthines atom numbering; time evolution and dihedral distribution function for the methyl rotation of caffeine; the convergence test of computed UV spectra with respect to the number of extracted snapshots; linear combinations (on the NBO basis) for the HOMO and LUMO of caffeine; effect of the choice of DFT functional on vertical excitation energies and spectra; paraxanthine and theophylline tautomers; plots of the orbitals involved in the first transitions of paraxanthine and theophylline; stick spectra obtained by using the QM/FQ[b] solvation model; charge transfer index. See DOI: 10.1039/c9cp05420k

and acetonitrile solutions.[35] Other studies have investigated thermochemical processes, enthalpy of hydration, solubility and enthalpy of solution,[36] structure–activity relationships,[37] resonant two-photon ionization (R2PI) spectroscopy,[38,39] vibrational cooling time,[40] pharmacokinetics,[41] amongst others.

Xanthines have also been extensively explored from the computational point of view. Many techniques encompassing quantum mechanics (QM), molecular dynamics (MD),[42–44] combined quantum mechanics/molecular mechanics (QM/MM) and MD calculations,[45] free energy calculations[46] and QSAR investigations,[47] among others, have been used to study metabolic mechanisms,[48] formation of π-stacked complexes,[39,45,49] microsolvation,[50,51] proton transfer reactions,[52] recognition sites in molecular imprinting polymers,[53] vibrational[18,42] and electronic[54] spectra, spectroscopic signatures and structural motifs in isolated and hydrated xanthines, *etc.*[21,55] In the latter cases, solvent effects need to be considered if reliable simulated spectra are desired.[56,57]

A straightforward approach, which includes solvent effects in QM calculations of spectroscopic quantities, is to resort to the polarizable continuum model (PCM)[58–61] in which the solute is accommodated in a molecule-shaped cavity, whereas the solvent is modelled as a homogeneous continuum dielectric. However, for strongly interacting solute–solvent systems, different hydrogen bonding networks around the analyzed substrates could remarkably influence their photodynamical properties,[57,62–65] as pointed out by Guo *et al.* in the specific case of nucleobase analogues.[45] Strictly speaking, complexes formed between the compound of interest and solvent molecules, which are generally stabilized *via* hydrogen bonds[66–69] (HBs), could have a dramatic impact on the charge transfer taking place during electronic transitions. In this respect, continuum solvation is not able to properly account for HBs.

In order to accurately describe specific solute–solvent interactions, a combined QM/MM-MD approach,[70] in which MD simulations allow for sampling of the phase-space, can be exploited. Such a multiscale scheme implies the partition of the solvated system in two layers, in which the key ingredient is to accurately account for the interaction between the two portions. This is usually accomplished by including an electrostatic coupling,[71] which may or may not contain mutual polarization between the two layers, in polarizable or electrostatic embedding, respectively. Several polarizable QM/MM approaches have been developed so far, based on multipole expansions,[72,73] induced dipoles,[74–78] AMOEBA,[65,79] fluctuating charges (FQ)[64,70,80,81] and fluctuating charges and dipoles (FQFμ).[82–84] QM/MM non-electrostatic interactions may also be included.[85–89] Among the different polarizable QM/MM approaches, QM/FQ[57,70,80,90,91] is adopted in this work, because it has been successfully applied to the calculation of vertical electronic excitation energies (VEE) for several compounds in aqueous solution.[57,80,92–94]

In this work, we focus our attention on solvent effects on electronic transitions of a set of selected xanthines, for which experimental spectra exist.[35,95,96] In particular, we borrowed and analyzed the corresponding UV/Vis absorption spectra in aqueous solution of the three xanthine derivatives displayed in Fig. 1, which were reported by Chen and Kohler.[35] It should be noted that, theobromine and the parent compound xanthine are not studied in this work due to their poor water solubility. To fully characterize solvent effects on purine analogues in aqueous solution, different computational approaches, ranging from PCM to non-polarizable and polarizable QM/MM approaches, have been investigated. The present study also constitutes a good starting point to gain insight about the process related to solvation and UV radiation, and can potentially serve as a model for more complex cases like purine nucleobases in DNA.

The manuscript is organized as follows: in the next section the computational protocol adopted for the calculation of electronic absorption spectra is discussed. The subsequent section contains the results obtained by exploiting different solvation approaches for the simulation of the UV-Vis spectra of caffeine in aqueous solution. The procedure is then applied to other caffeine-related molecules, namely theophylline and paraxanthine. As an illustration of the performance of the various approaches, the computed results are compared with the experimental ones. Our summary and conclusions end the manuscript.

2 Computational protocol

Caffeine, theophylline and paraxanthine, (see Fig. 1 for their molecular structure), were initially optimized and fully characterized as minima on the potential energy surface (PES) at the B3LYP/6-311++G(d,p) level of theory, by including the water environment by means of PCM.[59] In the case of caffeine, the different conformers arising from the methyl group rotation were analyzed by means of relaxed scans. Additional geometry optimizations were performed for theophylline and paraxanthine because different tautomers can be found in aqueous solution. Once the lowest energy structures were identified, CM5 atomic charges[97] were calculated, and used to account for electrostatic interactions in the following MD simulations.

MD simulations were performed using GROMACS[98] with the GAFF force field.[99] The model was constructed by placing each xanthine in a cubic box of about 5800 TIP3P[100] water molecules. The steepest descent minimization algorithm was used in the energy minimization procedure. Two equilibration stages were carried out before the production stage. In the first one, with a total time of 0.2 ns, the system was heated to 298.15 K in a *NVT* ensemble using the velocity-rescaling method,[101] with an integration time step of 0.2 fs and a coupling constant of 0.1 ps. Then, 1 ns of *NPT* equilibration with a time step of 1 fs was carried out along with the constant pressure periodic boundary conditions, and using the Parrinello–Rahman barostat.[102] Once equilibrated, a further 50 ns MD simulation in the *NVT* ensemble for each xanthine–water system was conducted with an integration step of 2 fs. The LINCS algorithm was used to constrain the fastest internal degrees of freedom.[103] Electrostatic interactions were calculated using the particle mesh

Ewald (PME)[104] method with a grid spacing of 0.16 nm, a cubic interpolation and a distance for the Coulomb cut-off of 1.0 nm. Geometric combination rules were used to combine van der Waals (VdW)-parameters in the reciprocal part of Lennard Jones-PME. The classical equations were integrated using the leap-frog algorithm.[105] To the best of our knowledge, we report for the first time molecular dynamics simulations for paraxanthine and theophylline in aqueous solution.

250 uncorrelated snapshots were extracted from the MD simulations of each solvated xanthine. For each snapshot, a solute-centered sphere with a radius of 17 Å was cut (containing approximately 700 water molecules). On each droplet, non-polarizable QM/TIP3P,[100] QM/FQ and QM/QM$_w$/FQ (*vide infra*) vertical excitation energies were calculated. Computed spectra were convoluted with a Gaussian band shape with a full width at half maximum (FWHM) of 0.6 eV in all cases.

Vertical excitation energies and spectra were obtained following the computational protocol described in ref. 57. All time-dependent (TD) DFT calculations were run using CAM-B3LYP[106] and B3LYP,[107] both in combination with the 6-311++G(d,p) basis set. The first ten excited states were taken into account in each case. The details of each investigated solvation approach are summarized in the following points:

1. QM/PCM: the solvent is described as a continuum dielectric. As a consequence, HBs are not considered. Vertical excitation energy calculations were conducted over the lowest energy structures, found after both tautomeric and conformational analyses.

2. QM/QM$_w$/PCM: this model was considered in the case of caffeine only. Potential solute HB sites were saturated with water molecules treated at the QM level (QM$_w$). The resulting supermolecule was accommodated in a cavity surrounded by the implicit solvent treated at the PCM level. Geometry optimization and harmonic frequency calculations were performed at the B3LYP/6-311++G(d,p) level.

3. QM/TIP3P: the atomistic nature of the environment is taken into account. On each MM water atom a fixed charge is placed,[100] and the Coulomb interaction between such charges and the QM potential is included in the QM Hamiltonian. Final UV/Vis spectra are computed by averaging over the MD snapshots.

4. QM/FQ: similarly to QM/TIP3P, each MM water molecule is modeled by the FQ[70] force field, *i.e.* each atom is endowed with a charge which varies as a response to the difference in atomic electronegativites and QM potential. Such charges are found by solving a linear equation system. The FQs are then included in the QM Hamiltonian so that mutual polarization effects are modeled. The FQ parametrizations utilized in this work were taken from ref. 86 and 90. Final UV/Vis spectra are computed by averaging over the MD snapshots.

5. QM/QM$_w$/FQ: in order to investigate effects other than electrostatics and polarization, the closest water molecules were added to the solute in the QM portion for each MD snapshot, whereas the remaining water molecules were described by exploiting the FQ force field[70] with the parameters reported in ref. 90. The number of water molecules to be included in the QM portion was chosen by inspecting MD radial distribution functions (RDFs) for the atoms involved in the HBs (being 5 and 4 on average, for caffeine and theophylline/paraxanthine, respectively).

The character of each excitation was determined by resorting to natural bond orbital (NBO) analysis,[108,109] *i.e.* by analyzing canonical (delocalized) molecular orbitals (CMOs) in terms of their leading NBO contributions. In the CMO analysis, each MO, Φ_i, can be expressed in terms of the complete orthonormal set of NBOs, Ω_α as

$$\Phi_i = \sum_\alpha c_{\alpha i}\Omega_\alpha \quad (1)$$

where $c_{\alpha i}$ coefficients determine the percentage contribution of each NBO to the linear combination LCNBO-MO. Such a decomposition permits one to identify the specific localized regions involved in the transition, and to assign them to bonding orbitals (σ_{A-B} or π_{A-B}), lone pairs (n_X) or antibonding orbitals (σ^*_{A-B} or π^*_{A-B}), which are the most common ways to describe electronic transitions. In addition, we explored the charge transfer character of the first electronic transition in each xanthine. To do that, we evaluated the qualitative (sometimes even quantitative) D_{CT} index of spatial extent in charge-transfer excitations.[110] D_{CT} provides information about the spacial proximity between regions participating in a given transition and the index has proven to be extremely useful in the description of families of push–pull compounds and in the exploration of potential energy surfaces at the excited state.[57,94,110–113] We also exploited the well-established method of calculating the charge transfer of any chemical system in the ground state, both intramolecularly and intermolecularly, through the estimation of donor–acceptor interaction strength within the framework of the aforementioned NBO method. In short, the procedure consists of estimating a lowering of the orbital-donor energy, related to the transfer of charge onto the orbital which was initially fully un-populated in the unperturbed system. Thus, for each donor NBO (i) and acceptor NBO (j), the stabilization energy $E^{(2)}_{ij}$ is calculated from a second order perturbation theory approach.[109,114] This NBO methodology has been used by Agou *et al.*[115] to study phosphaborins.

All QM/PCM, QM/MM, QM/QM$_w$/PCM and QM/QM$_w$/FQ calculations were performed by using a locally modified version of the Gaussian16 suite of programs,[116] whereas NBO calculations were carried out with the NBO6.0 program.[108]

3 Results and discussion

This section includes the data computed for the three studied methylxanthines. We start by discussing conformational preferences for caffeine and the results obtained from MD simulation in aqueous solution. After that, we report on the absorption spectra and excitation energies by exploiting different solvation approaches, and finally assign the respective transitions. We then move on to simulate electronic absorption spectra of paraxanthine and theopylline. In all cases, a comparison with experimental data taken from the literature is presented.

3.1 Caffeine

3.1.1 Conformational analysis. The structure of caffeine is characterized by the presence of three methyl groups, all bonded to nitrogen atoms. In Fig. 2 we show the energy profiles for each possible rotation, calculated through relaxed scans at the B3LYP/6-311++(d,p) level by treating aqueous solutions by means of the implicit PCM. From these complementary plots we can extract some information about the structural preferences in this xanthine. Despite small barriers associated with the rotations, the caffeine structure only has one eclipsed C–H···C=O (just for the methyl group between two carbonyl groups) corresponding to the well defined minima on the PES. In this preferred conformation, the polar C=O bond linked to the two nitrogen atoms does not eclipse other bonds and the orientation of the methyl group in the five-membered ring is one in which the in-plane C–H bond eclipses another C–H bond. This is probably due to hyperconjugation, as has been reported in the case of thymine.[117]

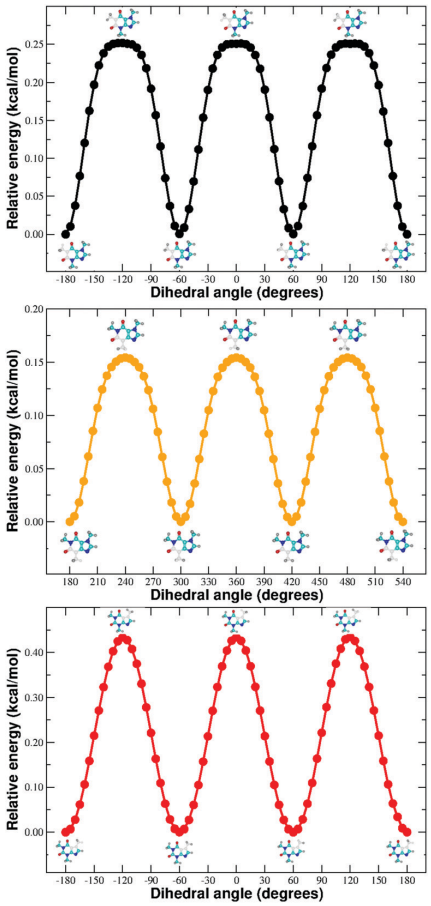


Fig. 2 B3LYP/6-311++G(d,p)/PCM caffeine PES scans of the methyl group rotation: M1 (top), M2 (middle) and M3 (bottom). See Fig. 1 for methyl group labelling.

In order to analyze the most important interactions involved in the stabilization of the caffeine structure, we performed NBO calculations to estimate the stabilization energies, $E_{ij}^{(2)}$, associated with delocalizations. A summary of the results is shown in Table 1 for those cases when the interaction energy exceeds a threshold of 10.0 kcal mol^{-1}. The overlapping between donor and acceptor NBOs is depicted in Fig. 3, for the NBOs taking part only in the frontier molecular orbitals.

As reported in Table 1, the $n_{N9} \rightarrow \pi^*_{N7-C8}$ interaction between a nitrogen lone pair and the N7–C8 antibond orbital gives the strongest stabilization (139.45 kcal mol^{-1}). Such an outcome can be rationalized in terms of the aromaticity of the compound, because the lone pair resides in a p-orbital (99.82% of p-character) and is a part of the electronic aromatic cloud. Other delocalizations involving lone pairs (with p-character) in the nitrogen atoms are also important for the molecule stabilization, but to a lower extent. Among the selected interactions, there are several types standing for the π–π^* charge transfer, with the donation of occupancy from the localized π orbitals (π_{N7-C8} and π_{C4-C5}, *etc.*) to the empty non-Lewis orbitals (for example π^*_{C4-C5}, π^*_{C6-O10}, π^*_{N7-C8}), moving the structure away from the idealized Lewis description, and also accounting for the so-called "delocalization" corrections. These interactions will be significant in the identification of the type of transition for the electronic spectra.

Through the NBO analysis we can also examine the nature of the second lone pair in the nitrogen (N9) of the five-membered cycle. This lone pair can be described as a $sp^{1.91}$ hybrid in the same plane as the σ bonds and projected outward from the ring, which leads to the potential formation of hydrogen bonds when caffeine is dissolved in polar solvents.

3.1.2 MD analysis. The results obtained by adopting an implicit treatment of solvent effects are confirmed by the atomistic MD simulations in aqueous solutions. In fact, the dihedral distribution functions for the methyl rotations support the small energy barriers already registered in the QM/PCM scans (see Fig. S2 in the ESI†). Essentially, the range

Table 1 Caffeine stabilization energies $E_{ij}^{(2)}$ (kcal mol^{-1}) between donor and acceptor intramolecular NBOs exceeding 10.0 kcal mol^{-1}

NBO donor	NBO acceptor	$E_{ij}^{(2)}$ (kcal mol^{-1})
n_{N9}	π^*_{N7-C8}	139.45
n_{N9}	π^*_{C4-C5}	75.31
n_{N3}	σ^*_{C2-O12}	63.08
n_{N1}	σ^*_{C2-O12}	59.11
n_{N1}	π^*_{C6-O10}	54.30
n_{N3}	π^*_{C4-C5}	48.18
π_{C4-C5}	π^*_{C6-O10}	32.98
n_{O10}	σ^*_{N1-C6}	28.66
n_{O12}	σ^*_{N1-C2}	26.35
n_{O12}	σ^*_{C2-N3}	25.12
n_{O10}	σ^*_{C5-C6}	17.30
π_{N7-C8}	π^*_{C4-C5}	15.30
π_{C4-C5}	π^*_{N7-C8}	13.69

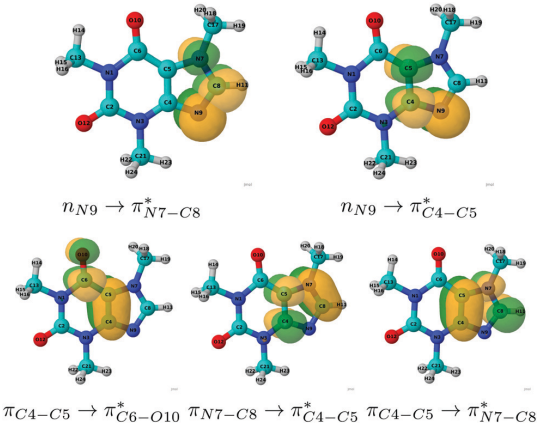


Fig. 3 Pictorial view of the overlap between donor and acceptor NBOs associated with intramolecular interactions in caffeine. See Table 1 for their stabilization energies.

of values from −180 to 180 degrees is entirely covered in the MD simulation time, indicating that the methyl group is a relatively free rotor at room temperature.

From the above NBO results and from an inspection of the caffeine molecule (see Fig. 1), it can be seen that there are three specific solvation sites, namely, the two oxygen atoms of the carbonyl groups (O10 and O12) and the nitrogen (N9) with its almost sp^2 hybridized lone pair, all highlighted in Fig. 4. To analyze HB interactions, $g(r)$ (RDFs) for the potentially charge-donor atoms in the HBs with the solvent were extracted from the MD trajectory, together with the corresponding running coordination numbers (RCNs), which represent the number of closest water molecules interacting with caffeine *via* primary HBs. The calculated RDFs obtained from the classical MD simulation are plotted in Fig. 4.

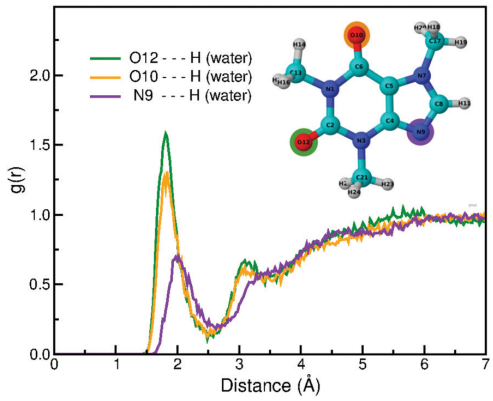


Fig. 4 Radial distribution function between selected sites of caffeine and water molecules. Selected sites are highlighted in the structures. Integration of RDFs leads to the following running coordination numbers (see text): $RCN_{O10} \approx 2$; $RCN_{O12} \approx 2$; $RCN_{N9} \approx 1$.

Caffeine RDFs exhibit a couple of peaks for each atom analyzed. The first peak for the two oxygen atoms is located at 1.9 Å, and at 2.1 Å for nitrogen. Similar distances have been reported for primary hydrogen bonds in water clusters,[118] aqueous solvation of ions and molecules, *etc.*[119] The second solvation peaks in the RDFs, associated with the second solvation sphere, are somewhat diffused and spread out in the region from 2.9 to 3.6 Å (Fig. 4). The total number of water molecules in the first solvation shell, RCN, was found to be 2 for the oxygen atoms, whereas the nitrogen atom has only one water molecule in its close vicinity. As a matter of fact, this number of water molecules coincides with the number of available lone pairs in caffeine. These solvent molecules were explicitly introduced in the QM portion, in the following QM/QM_w/PCM results. Recently, the same strategy was used to choose the number of the water molecules to be added either in the QM/QM_w/FQ or QM/QM_w/PCM calculations, in order to obtain reliable electronic spectra.[57] One of the many choices for hydrated caffeine with five water molecules is shown in Fig. 5. Interestingly, the same number of water molecules was used to analyze the acceleration of vibrational cooling due to hydrogen bond donors in some purine derivatives.[40] Moreover, in 2014, several caffeine–$(water)_n$ complexes with n = 1 and 2 were proposed to analyze the effect of the hydration on excited states.[21] Kim *et al.* also reported some caffeine : water 1 : 1 complexes to determine the binding site giving rise to a blue-shift in the resonant two-photon ionization (R2PI) spectrum when compared to the spectrum for bare caffeine.[38]

3.1.3 Excitation energies and absorption spectra. We now move on to discuss the comparison between the computed and experimental UV-Vis spectra of caffeine in aqueous solution. As stated before, the aqueous environment was modeled by exploiting the implicit (QM/PCM) approach and the non-polarizable QM/TIP3P and polarizable QM/FQ, performed by adopting two different parametrizations. Additional calculations were performed by including water molecules in the QM portion and by treating the bulk by either using PCM (QM/QM_w/PCM) or the FQ force field (QM/QM_w/FQ). QM/MM and QM/QM_w/FQ calculations were performed on 250 uncorrelated snapshots extracted from the MD trajectory.[62,63] Such a number of snapshots is more than enough to assure convergence (see Fig. S3 in the ESI†).

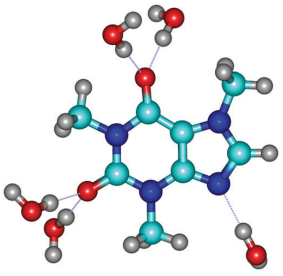


Fig. 5 One of the minima obtained for solvated caffeine, according to the RCNs extracted from MD simulations. Such a motif was used in the QM/QM_w/PCM calculations.

In Fig. 6, computed QM/PCM and QM/FQ[b] UV-Vis stick spectra are depicted. The two solvation approaches give totally different pictures, due to the fact that the explicit-dynamic QM/FQ predicts for each snapshot a different vertical transition energy and a different intensity. Such an outcome is primarily because the solute molecule feels a different environment by varying the snapshot, as a consequence of the dynamics of the water molecules during the MD runs. Therefore, this results in a large variability in both vertical excitation energies and intensities, and consequently, the inhomogeneous band-broadening is automatically included (see convoluted spectra in Fig. 6). The results obtained in the case of the implicit-static QM/PCM consist of a single stick for each transition of each conformer (Fig. 6).

The computed QM/PCM, QM/QM$_w$/PCM, QM/TIP3P, QM/FQ and QM/QM$_w$/FQ UV-Vis spectra are reported in Fig. 7 together with the experimental spectrum reproduced from ref. 35 and 96. All computed spectra are dominated by two main bands, the one at the lowest energy being characterized by a single electronic transition (see Fig. 6). The nature of the first transition was analyzed in terms of caffeine MOs (see Fig. S4 in the ESI†) by investigating the NBO results, in particular the CMO analysis which is based on the ground state density. The orbitals involved in the first excitation are the so-called frontier orbitals: HOMOs and LUMOs. These canonical delocalized molecular orbitals can be formulated as NBO-linear combinations (see Fig. S4 in the ESI† for the complete expressions). The nature of the first electronic transition for caffeine in an aqueous environment is of the $\pi \rightarrow \pi^*$ type, because the HOMO is derived primarily ($\sim$34%) from the π_{C4-C5} orbital, with weaker contributions from the n_{N3} lone pair (19%) and from π_{N7-C8}, while the LUMO has significant contributions from at least three mixtures of NBOs (35%, 31% and 16%, respectively), with π^* antibond characteristics in all cases. It is interesting to note that the NBOs with the greatest coefficients in each linear combination, namely, π_{C4-C5} and π^*_{C6-O10}, are remarkably involved in the charge transfer that confers a stabilization energy of 32.98 kcal mol^{-1} to caffeine, as previously shown in Table 1 and in Fig. 3. The same kind of $\pi \rightarrow \pi^*$ transitions have also been postulated in ref. 21 and 120.

The nature of the second visible band is difficult to assign, because it is due to more than one electronic transition, all characterized by mixed excitations (see Fig. 6).

We gained deeper insight into the nature of the first electronic transition by investigating the length of the charge transfer, measured by the D_{CT} index,[110] which is just 1.891 Å (QM/FQ[a] model, but similar to the other models), meaning that it has a low CT character and that there is not a strong density rearrangement going from the ground to the excited state. D_{CT} indexes for caffeine and the other studied molecules can be found in the ESI† (Table S2).

Let us now focus more deeply on the comparison between the computed and experimental spectra (see Fig. 7). The experimental spectrum of caffeine exhibits two main bands, placed at about 273 nm and 205 nm.[35,96] As explained above, all the considered approaches suggest that a single transition is responsible for the lowest energy band. The first experimental

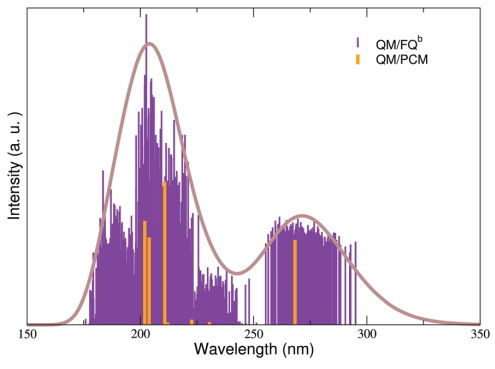


Fig. 6 B3LYP/6-311++G(d,p) computed QM/PCM (orange) and QM/FQ (purple) caffeine absorption stick spectra in aqueous solution. Convoluted QM/FQ spectra are also reported.

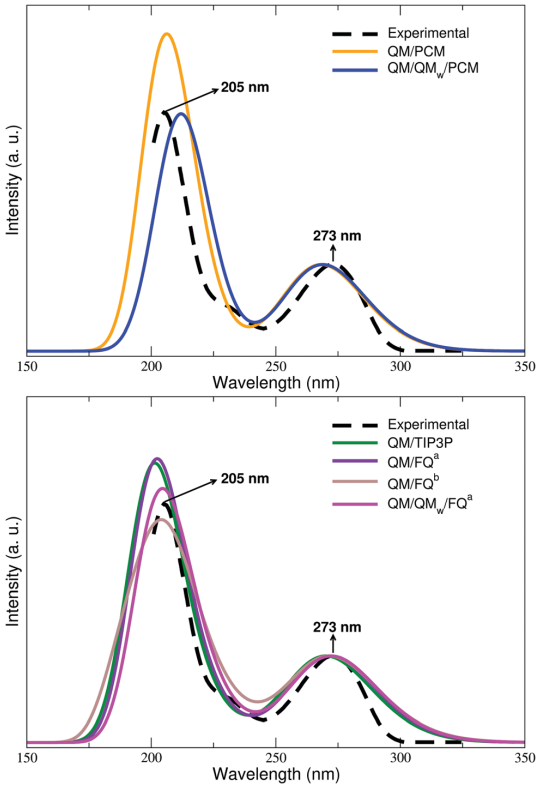


Fig. 7 Caffeine B3LYP/6-311++G(d,p) QM/PCM and QM/QM$_w$/PCM (top) and QM/TIP3P, QM/FQ[a], QM/FQ[b], QM/QM$_w$/FQ[a] (bottom) computed UV-Vis absorption spectra. The experimental spectrum has also been reported.[35,96] All computed spectra are normalized with respect to the first band intensity. [a] FQ parametrization from ref. 90. [b] FQ parametrization from ref. 86.

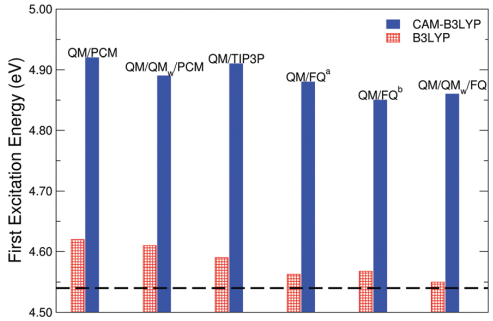


Fig. 8 Vertical excitation energies for the first electronic transition of caffeine in aqueous solution as computed by exploiting B3LYP or CAM-B3LYP density functionals. The experimental first excitation energy is 4.54 eV.[35]

band position is well-reproduced by all the considered solvation approaches, with a slight shifting toward the experimental results when explicit solvation models are considered. In particular, by also looking at Fig. 8, it can be seen that the first excitation energy is always overestimated (or blueshifted) with relative errors decreasing by moving from the continuum PCM to FQ. However, incorporating some explicit water molecules in the QM portion, either in the case of QM/PCM or QM/FQ, does not imply a significant improvement: the excitation energy obtained by using the QM/FQ model is quite similar to the QM/QM$_w$/FQ value. The latter gives the best agreement with the experimental results, however at much higher computational costs. Since all relative deviations with respect to the experimental value do not exceed 2% (see Table 3 below), even in the case of the fixed charges TIP3P, the inclusion of polarization effects into the solvation model seems not to be critical for the description of the first band (shape and λ_{max}). Note that our results referring to the first band are in line with those obtained by Chen and Kohler, who reported that vertical excitation energies for caffeine–water complexes are similar (4.62 eV *vs.* 4.60 eV) when the PCM/COSMO method is compared against caff$_1$–(H$_2$O)$_2$ clusters at the TD-DFT-B3LYP/6-311++G(d,p) level of theory.[35]

The various solvation approaches do differ in the reproduction of the second experimental band at about 205 nm (see Fig. 7 and Table 3). In particular, the best agreement is given by either QM/FQ^b or QM/QM$_w$/FQ, with QM/QM$_w$/PCM showing the largest discrepancies (about 0.2 eV). Such findings show that, although by adopting the cluster QM/QM$_w$/PCM approach explicit interactions are taken into account, the dynamical aspect of solvation is a crucial aspect that needs to be considered in the reproduction of absorption spectra. The various solvation approaches also differ in the reproduction of the relative intensities of the two main bands, with QM/PCM, QM/TIP3P and QM/FQ^a showing the largest deviations with respect to the experiments, and QM/FQ^b, QM/QM$_w$/PCM and QM/QM$_w$/FQ showing the best agreement. It is worth discussing here that differences between FQ^a and FQ^b are mainly owing to the fact that FQ^b allows for a larger polarization because of lower atomic chemical hardness parameters (see ref. 86 for further details). QM/FQ^b parametrization leads to a UV-Vis spectrum that resembles the experiment much more, especially for the band centered at 205 nm (Fig. 7). The discrepancy between excitation energies calculated by both sets is about 0.01 eV. Consequently, the set of FQ parameters does not influence the excitation energies very much, but it does affect the spectral shape and relative intensities of the peaks. By also considering the results obtained in the case of non-polarizable QM/TIP3P, we observe that polarization effects seem to be dominant in the reproduction of the spectral profile. In fact, it is worth noticing that all considered QM/classical (*i.e.* QM/MM and QM/PCM) approaches describe solute–solvent interactions by only accounting for electrostatics. Therefore, the reported differences between the computational approaches are primarily ascribed to a different description of such interactions, which is refined by moving from QM/PCM to QM/TIP3P to QM/FQ^a to QM/FQ^b (in which polarization effects are amplified with respect to QM/FQ^a [86]). In addition, the discrepancies between QM/QM$_w$/FQ and QM/FQ^b are probably related to non-electrostatic effects (especially repulsion) which are not included in QM/FQ^b.[85,121]

We note that the discussed results clearly show that in order to achieve a good agreement with the experimental counterparts, the dynamical and specific effects in the solvation phenomenon need to be adequately introduced in the theoretical model. In this context, it should be noted that the specific solute–solvent interactions as identified from the RDFs, depicted in Fig. 4, are not fundamental to the reproduction of excitation energies, but are important in the modeling of the peak at 205 nm and its intensity.

To end the discussion on caffeine, it is well known that the choice of DFT functional is crucial when studying electronic excitation energies.[56] CAM-B3LYP and B3LYP functionals are compared in Fig. 8 (see also Table 3 and Fig. S5 given in the ESI† for spectral shapes). From the depicted results, it is clear that B3LYP predicts excitation energies much closer to the experimental values, for all the selected solvation approaches. Additional functional tests were performed by using M062X in connection with PCM. For that functional, a relative error around 10% in the calculated excitation energy is obtained in comparison with the experimental values.

3.2 Paraxanthine and theophylline

Paraxanthine and theophylline depicted in Fig. 1 were also studied. Both molecules are the result of a single demethylation in the caffeine molecule, and constitute positional isomers.[122]

Paraxanthine and theophylline exhibit a N–H bond in their chemical structure. This feature leads up to the possibility of prototropy tautomerism, wherein the proton can change place and be linked to another nitrogen or to a carbonyl group.[39] Tautomeric options were carefully investigated in terms of Boltzmann populations. The main structures are depicted in Fig. S6 and S7 in the ESI.† In this work, the results obtained in the case of the lowest energy tautomers are reported.

Different from caffeine, the presence of the N–H bond gives rise to a new specific solvation site to be considered in

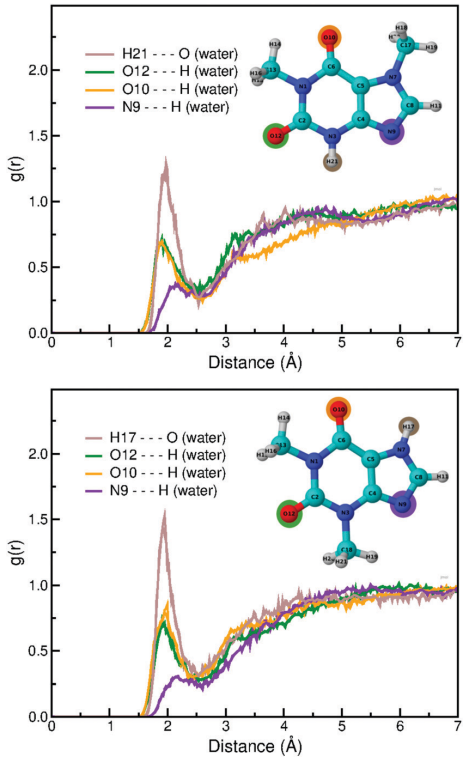


Fig. 9 RDFs of selected sites of paraxanthine (top panel), and theophylline (bottom panel) with water oxygen atoms or hydrogen atoms.

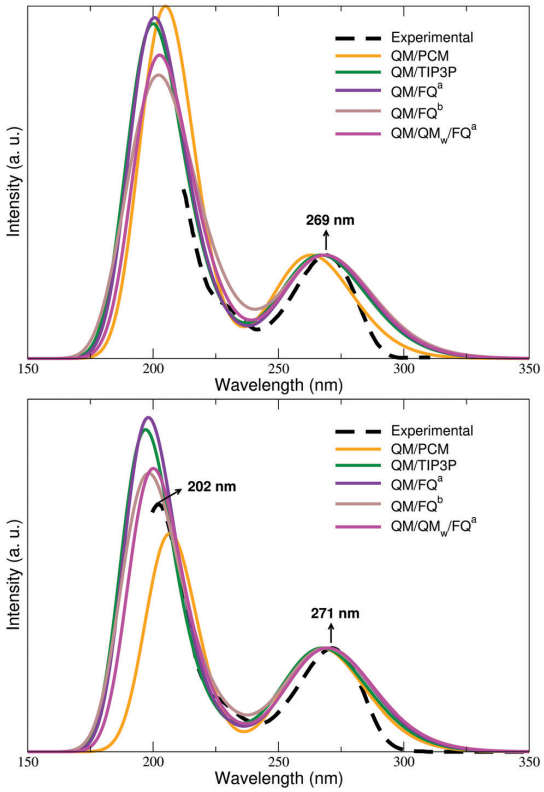


Fig. 10 B3LYP/6-311++G(d,p) QM/PCM, QM/TIP3P, QM/FQ[a], QM/FQ[b] and QM/QM$_w$/FQ computed UV-Vis absorption spectra for paraxanthine (top) and theophylline (bottom). Experimental spectra are reproduced from ref. 35 and 95 for paraxanthine and theophylline, respectively. All computed spectra are normalized with respect to the first band intensity. [a] FQ parametrization from ref. 90. [b] FQ parametrization from ref. 86.

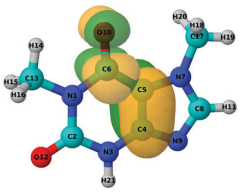
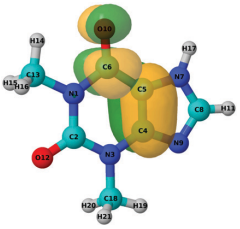
solute–solvent interactions. The radial distribution functions ($g(r)$) between such hydrogen atoms in paraxanthine (or theophylline) and water oxygen atoms are depicted in Fig. 9, along with other possible O$\cdots$H$_w$ and N$\cdots$H$_w$ pairs. RDFs for both methylxanthines have a similar behavior with reference to the approximated position of the peaks and the corresponding running integrals, which yielded a total of four water molecules broken down into a single water molecule interacting with each HB site. The analysis of Fig. 9 may also indicate that the location of the methyl group in the chemical structure does not exert any influence on the hydration pattern for these molecules.

Experimental and calculated spectra obtained with the B3LYP functional are displayed in Fig. 10. QM/MM and QM/QM$_w$/MM calculations were performed on 250 uncorrelated snapshots extracted from MD simulations. Stick spectra are reported in Fig. S8 in the ESI.† Similar to caffeine, all computed UV-Vis spectra for both molecules are characterized by two main bands.

To define the character of the transitions, we again resort to NBO analysis. HOMOs and LUMOs are found to be involved in the first excitation. Table 2 gives the NBO composition for the MOs of paraxanthine and theophylline showing only the dominant terms in the LCNBO-MO expansion. Entire frontier orbitals are plotted in Fig. S9 in the ESI.† We recall here that nitrogen lone pairs, taking part in the HOMO description, have p-character and they are in the same orientation as the rest of the ring π-cloud. This is a reason for considering them as a part of the π delocalized density in these molecules. Therefore, the nature of the first electronic transition is once again $\pi \rightarrow \pi^*$. This assignment is found to be consistent with that reported by Singh,[55] who suggested that the highest occupied molecular orbital (HOMO) of theophylline is a π-orbital, whose electron density is expected to be localized mainly on the C4–C5 fragment. As in the case of caffeine, the second transition is constituted by several electronic transitions, and therefore it is not easily characterized in terms of the involved excitations.

D_{CT} values of the first electronic transition for solvated paraxantine and theophylline are 1.774 Å and 1.970 Å, respectively, (see Table S2 in the ESI† for a complete list of indexes) and primarily differ in the position of the methyl groups. Such a charge transfer diagnostic, which uses the unrelaxed electronic density difference, shows that in both molecules the first excitation does not have a strong CT character.

Table 2 NBO analysis of canonical molecular orbitals (CMOs) involved in the electronic transitions of paraxanthine and theophylline as modelled by QM/PCM

Xanthine	CMO	LC-NBO	Plot of the main NBO interactions HOMO → LUMO
Paraxanthine	HOMO	$0.619\pi_{C4-C5}$ $0.428\pi_{N7-C8}$ $-0.410n_{N3}$	
	LUMO	$0.607\pi^*_{C6-O10}$ $0.551\pi^*_{N7-C8}$	
Theophylline	HOMO	$0.597\pi_{C4-C5}$ $-0.444n_{N3}$	
	LUMO	$0.594\pi^*_{C6-O10}$ $0.431\pi^*_{C8-N9}$ $-0.398\pi^*_{C4-C5}$	

We now move on to the comparison between the computed and experimental spectra. The experimental spectrum is characterized by a single transition at 269 nm (paraxanthine) and 271 nm (theophylline). A second intense peak is reported in the case of theophylline at about 202 nm.[95] It is worth noting that, although to the best of our knowledge there are no available experimental spectra in that region for paraxanthine, due to the similarity in the structure with respect to caffeine and theophylline, a second more intense peak should, in principle, be present.

The findings commented above for caffeine are also valid for both theophylline and paraxanthine. Certainly, the first transition energies are correctly reproduced by all the investigated methods, with the lowest error (0%) reported for QM/QM$_w$/FQ, which includes in the QM portion an average of four water molecules for both methylxanthines. The polarizable QM/FQ (both parametrizations) has a valuable performance with a minimal discrepancy (around 1%) for the first electronic transition of theophylline. From Fig. 10 it can be evidenced that for both systems, hydrogen bond interactions with solvent molecules (hydrogen bond acceptors and donors) seem not to play a pivotal role in the description of the first transition. In fact, it is not necessary to include explicit water molecules to reproduce, to a large extent, the excitation energy of this transition. Notwithstanding, the achievement of a well simulated second peak requires the use of a full polarizable model, as can be seen for theophylline's spectra in the bottom panel of Fig. 10. Even so, all the atomistic approaches shift the relative intensity towards the experimental results with respect to QM/PCM. Again, the increasing of polarization effects, which is achieved by adopting QM/FQ[b], results in a better reproduction of the experimental ratio between the first and second bands (see also Table 3). Such findings are perfectly in agreement with what we found for caffeine.

3.3 General performance of the FQ model

Vertical excitation energies and relative intensities computed for the three studied xanthines with the different solvation approaches are listed in Table 3. For the sake of comparison with the experiments, the percentage errors with respect to the values reported in ref. 35, 95 and 96 are also listed. In addition, results obtained by utilizing both B3LYP and CAM-B3LYP density functionals (DFs) are also given in Table 3.

First, it should be noted that the B3LYP results are in better agreement with the experiment as compared to CAM-B3LYP, thus justifying the choice adopted in this work. Such an improvement is particularly clear by looking at the excitation energies, but it also affects relative intensities between the two peaks.

It can be observed that all approaches at the B3LYP level yield reliable results with deviations of no more than 3% from

Table 3 Computed QM/PCM, QM/TIP3P, QM/FQ,[a] QM/FQ[b] and QM/QM$_w$/FQ vertical excitation energies (eV) and relative intensities (Int.) for the two main peaks of the studied xanthines. Results obtained by utilizing B3LYP and CAM-B3LYP density functionals (DFs) have also been reported. Experimental data are recovered from ref. 35, 95 and 96. All intensities are normalized with respect to the first vertical transition

			QM/PCM		QM/TIP3P		QM/FQ[a]		QM/FQ[b]		QM/QM$_w$/FQ[a]		Experimental	
DF	Xanthine	Peak	E (eV)	Int.	E (eV)	Int.	E (eV)	Int.	E (eV)	Int.	E (eV)	Int.	E (eV)	Int.
CAM-B3LYP	Caffeine	1	4.92 (8%)	1.00	4.91 (8%)	1.00	4.88 (7%)	1.00	4.85 (7%)	1.00	4.86 (7%)	1.00	4.54	1.00
		2	6.43 (6%)	3.45	6.61 (9%)	3.03	6.58 (9%)	3.03	6.56 (8%)	2.50	6.57 (9%)	2.86	6.05	2.75
	Paraxanthine	1	5.00 (8%)	1.00	4.95 (7%)	1.00	4.93 (7%)	1.00	4.91 (6%)	1.00	4.92 (7%)	1.00	4.61	—
		2	6.50 (—)	3.23	6.69 (—)	3.03	6.66 (—)	3.13	6.64 (—)	2.63	6.66 (—)	2.94	—	—
	Theophylline	1	4.93 (8%)	1.00	4.95 (8%)	1.00	4.93 (8%)	1.00	4.90 (7%)	1.00	4.91 (7%)	1.00	4.58	1.00
		2	6.53 (6%)	3.45	6.77 (10%)	3.23	6.73 (10%)	3.23	6.74 (10%)	2.78	6.72 (9%)	3.23	6.14	2.39
B3LYP	Caffeine	1	4.62 (2%)	1.00	4.59 (1%)	1.00	4.56 (0%)	1.00	4.57 (1%)	1.00	4.55 (0%)	1.00	4.54	1.00
		2	6.01 (1%)	3.70	6.16 (2%)	3.23	6.13 (1%)	3.33	6.08 (0%)	2.56	6.06 (0%)	2.94	6.05	2.75
	Paraxanthine	1	4.71 (2%)	1.00	4.64 (1%)	1.00	4.62 (0%)	1.00	4.62 (0%)	1.00	4.62 (0%)	1.00	4.61	—
		2	6.05 (—)	3.45	6.20 (—)	3.23	6.18 (—)	3.33	6.13 (—)	2.70	6.12 (—)	2.94	—	—
	Theophylline	1	4.64 (1%)	1.00	4.64 (1%)	1.00	4.61 (1%)	1.00	4.61 (1%)	1.00	4.60 (0%)	1.00	4.58	1.00
		2	6.00 (2%)	2.08	6.29 (3%)	3.13	6.26 (2%)	3.23	6.25 (2%)	2.70	6.20 (1%)	2.70	6.14	2.39

[a] FQ parametrization from ref. 90. [b] FQ parametrization from ref. 86.

experimental excitation energies. For the studied molecules, PCM excitation energies are always larger than the corresponding experimental values, whereas the atomistic-dynamical approaches given by QM/MM coupled with MD simulations always shift excitation energies towards the experiment. In particular, QM/FQ and QM/QM$_w$/FQ show an almost perfect matching with reference data and calculated excitation energies reporting zero relative errors for caffeine and paraxanthine.

In the case of theophylline and paraxanthine, QM/FQ and QM/QM$_w$/FQ methods show more or less the same relative deviations, which is an indication that the FQ approach is well suited to study these systems and does not require the explicit inclusion of water molecules (that is, of course, computationally more expensive) to provide excellent simulated spectra. A performance enhancement is achieved by using the parameters in ref. 86, particularly in the description of the second band of the spectra. Such a finding, together with the fact that both QM/PCM and QM/TIP3P wrongly estimate the relative intensities between the two main bands, shows that, for the studied xanthines, specific solute–solvent interactions, such as HBs, and a correct inclusion of mutual solute–solvent polarization both need to be appropriately modelled.

4 Summary and conclusions

In this work, we have presented a complete analysis of the role of solvent effects in the simulation of the absorption spectra of caffeine, paraxanthine and theophylline in aqueous solution. The investigated computational approaches range from the implicit QM/PCM to explicit QM/MM approaches, either polarizable or non-polarizable. In addition, supermolecular approaches, in which the closest water molecules are included in the QM region, were also explored. In particular, this has been done by both considering the remaining environment as described by the implicit PCM (QM/QM$_w$/PCM) and retaining its atomistic description (QM/QM$_w$/FQ). QM/MM and QM/QM$_w$/FQ calculations have been performed by averaging 250 snapshots extracted from MD simulations, thus accounting for the dynamical aspects of the solvation phenomenon, different from the static QM/PCM and QM/QM$_w$/PCM approaches. The utilized methods have allowed for a deep analysis of the differences between implicit-static and explicit-dynamic solvation methods.

By analyzing the radial distribution functions from MD simulations, we extracted information about solute–solvent interactions. The average number of water molecules directly involved in hydrogen bonds was found to be five for caffeine, and four for both paraxanthine and theophylline. This number of water molecules was included in the QM region in the supermolecular approaches (QM/QM$_w$/PCM and QM/QM$_w$/FQ). By averaging over the snapshots extracted from MD runs, UV-Vis absorption spectra were calculated. For the three studied systems, computed spectra show a two peak profile, which has also been experimentally reported.

The first band (*i.e.* the one at lower energy), which is due to a $\pi \rightarrow \pi^*$ transition as reported by NBO analysis, is correctly reproduced by all methods, in particular when B3LYP is used in TD-DFT calculations. The major differences are computed for the second band, which is the result of several electronic transitions. In the case of caffeine, we showed that first/second band relative excitation energies are wrongly reproduced by the supermolecular QM/QM$_w$/PCM approach, whereas a nice agreement with the experiment is registered in the case of all atomistic QM/(QM$_w$/)MM methods. Such findings show that dynamical effects, primarily due to the different arrangement of water molecules around the QM solute, play a relevant role. As for the other studied molecules, the experimental first/second band relative intensity ratio is recovered by exploiting the supermolecule QM/QM$_w$/FQ or QM/FQ^b approaches. QM/FQ^b parametrization allows for a bigger polarization with respect to QM/FQ^a. Therefore, our results, together with the fact that both QM/PCM and non-polarizable QM/TIP3P wrongly estimate the relative intensities between the two main bands, show that specific solute–solvent interactions, such as HBs, and a correct inclusion of mutual solute–solvent polarization need both to be appropriately modelled. To conclude, we remark that the small differences reported between QM/FQ^b and QM/QM$_w$/FQ can be ascribed to the fact that QM/FQ^b only accounts for electrostatic interactions, whereas QM/QM$_w$/FQ also accounts for non-electrostatic interactions, in particular Pauli repulsion.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We are thankful for the computer resources provided by the high performance computer facilities of the SMART Laboratory (http://smart.sns.it/). SG and CC gratefully acknowledge the support of H2020-MSCA-ITN-2017 European Training Network ''Computational Spectroscopy in Natural sciences and Engineering'' (COSINE), grant number 765739. TG acknowledges funding from the Research Council of Norway through the grant TheoLight (grant no. 275506).

Notes and references

1 V. R. Preedy, *Caffeine: chemistry, analysis, function and effects*, Royal Society of Chemistry, 2012.
2 X. B. Chen, F. D. D. Hovell, E. R. Ørskov and D. S. Brown, *Br. J. Nutr.*, 1990, **63**, 131–142.
3 A. Many, C. Hubel and J. Roberts, *Am. J. Obstet. Gynecol.*, 1996, **174**, 288–291.
4 J. Maiuolo, F. Oppedisano, S. Gratteri, C. Muscoli and V. Mollace, *Int. J. Cardiol.*, 2016, **213**, 8–14.
5 N. L. Benowitz, P. Jacob III, H. Mayan and C. Denaro, *Clin. Pharmacol. Ther.*, 1995, **58**, 684–691.

6 O. A. Gressner, B. Lahme, M. Siluschek and A. M. Gressner, *Liver Int.*, 2009, **29**, 886–897.
7 N. J. Wesensten, *Principles and Practice of Sleep Medicine*, W.B. Saunders, Philadelphia, 5th edn, 2011, pp. 809–813.
8 G. B. Kaplan, D. J. Greenblatt, B. L. Ehrenberg, J. E. Goddard, M. M. Cotreau, J. S. Harmatz and R. I. Shader, *J. Clin. Pharmacol.*, 1997, **37**, 693–703.
9 A. Sattin and T. W. Rall, *Mol. Pharmacol.*, 1970, **6**, 13–23.
10 S. H. Snyder, J. J. Katims, Z. Annau, R. F. Bruns and J. W. Daly, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 3260–3264.
11 R. F. Bruns, *Purines in Cellular Signaling*, New York, 1990, pp. 126–135.
12 J. Barone and H. Roberts, *Food Chem. Toxicol.*, 1996, **34**, 119–129.
13 M. Panagiotou, M. Meijer, J. H. Meijer and T. Deboer, *J. Psychopharmacol.*, 2019, **33**, 122–131.
14 M. C. Cornelis, *Nutrients*, 2019, **11**, 416.
15 J. de Paula Lima and A. Farah, *Coffee: Consumption and Health Implications*, The Royal Society of Chemistry, 2019, pp. 313–339.
16 J. de Paula Lima and A. Farah, *Coffee: Consumption and Health Implications*, The Royal Society of Chemistry, 2019, pp. 489–508.
17 J. Zhong, N. Tang, B. Asadzadeh and W. Yan, *J. Chem. Eng. Data*, 2017, **62**, 2570–2577.
18 M. E. Costas and R. Acevedo-Chávez, *J. Phys. Chem. A*, 1997, **101**, 8309–8318.
19 S. Gogia, A. Jain and M. Puranik, *J. Phys. Chem. B*, 2009, **113**, 15101–15118.
20 K. Rottger, R. Siewertsen and F. Temps, *Chem. Phys. Lett.*, 2012, **536**, 140–146.
21 V. B. Singh, *RSC Adv.*, 2014, **4**, 58116–58126.
22 M. Falk, M. Gil and N. Iza, *Can. J. Chem.*, 1990, **68**, 1293–1299.
23 H. G. M. Edwards, E. Lawson, M. de Matas, L. Shields and P. York, *J. Chem. Soc., Perkin Trans. 2*, 1997, 1985–1990.
24 B. R. Singh, M. A. Wechter, Y. Hu and C. Lafontaine, *Biochem. Educ.*, 1998, **26**, 243–247.
25 L. López-Martínez, P. L. L. de Alba, R. García-Campos and L. M. D. León-Rodríguez, *Anal. Chim. Acta*, 2003, **493**, 83–94.
26 A. Aresta, F. Palmisano and C. G. Zambonin, *Food Chem.*, 2005, **93**, 177–181.
27 N. D'Amelio, L. Fontanive, F. Uggeri, F. Suggi-Liverani and L. Navarini, *Food Biophys.*, 2009, **4**, 321–330.
28 A. F. M. Cláudio, A. M. Ferreira, M. G. Freire and J. A. P. Coutinho, *Green Chem.*, 2013, **15**, 2002–2010.
29 A. Abbaspour and R. Mirzajani, *J. Pharm. Biomed. Anal.*, 2005, **38**, 420–427.
30 A. Belay, K. Ture, M. Redi and A. Asfaw, *Food Chem.*, 2008, **108**, 310–315.
31 K. Fujioka and T. Shibamoto, *Food Chem.*, 2008, **106**, 217–221.
32 M. M. Stimson and M. A. Reuter, *J. Am. Chem. Soc.*, 1943, **65**, 153–155.
33 H. M. Kalckar and M. Shafran, *J. Biol. Chem.*, 1947, **167**, 429–443.
34 M. Shukla and P. Mishra, *J. Mol. Struct.*, 1994, **324**, 241–249.
35 J. Chen and B. Kohler, *Phys. Chem. Chem. Phys.*, 2012, **14**, 10677–10682.
36 S. Bruns, J. Reichelt and H. K. Cammenga, *Thermochim. Acta*, 1984, **72**, 31–40.
37 C. E. Mueller, D. Shi, M. Manning and J. W. Daly, *J. Med. Chem.*, 1993, **36**, 3341–3349.
38 D. Kim, H. M. Kim, K. Y. Yang, S. K. Kim and N. J. Kim, *J. Chem. Phys.*, 2008, **128**, 134310.
39 M. P. Callahan, Z. Gengeliczki, N. Svadlenak, H. Valdes, P. Hobza and M. S. de Vries, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2819–2826.
40 Y. Zhang, J. Chen and B. Kohler, *J. Phys. Chem. A*, 2013, **117**, 6771–6780.
41 A. Lelo, D. Birkett, R. Robson and J. Miners, *Br. J. Clin. Pharmacol.*, 1986, **22**, 177–182.
42 P. B. Balbuena, W. Blocker, R. M. Dudek, F. A. Cabrales-Navarro and P. Hirunsit, *J. Phys. Chem. A*, 2008, **112**, 10210–10219.
43 L. Tavagnacco, U. Schnupf, P. E. Mason, M.-L. Saboungi, A. Cesàro and J. W. Brady, *J. Phys. Chem. B*, 2011, **115**, 10957–10966.
44 M. Paloncýová, K. Berka and M. Otyepka, *J. Phys. Chem. B*, 2013, **117**, 2403–2410.
45 X. Guo, Y. Zhao and Z. Cao, *Phys. Chem. Chem. Phys.*, 2014, **16**, 15381–15388.
46 H. Gouda, I. D. Kuntz, D. A. Case and P. A. Kollman, *Biopolymers*, 2003, **68**, 16–34.
47 S. El-Taher, K. M. El-sawy and R. Hilal, *J. Chem. Inf. Comput. Sci.*, 2002, **42**, 386–392.
48 Y. Kang, J. Tao, Z. Xue, Y. Zhang, Z. Chen and Y. Xue, *Tetrahedron*, 2016, **72**, 2858–2867.
49 L. Redivo, R.-M. Anastasiadi, M. Pividori, F. Berti, M. Peressi, D. Di Tommaso and M. Resmini, *Phys. Chem. Chem. Phys.*, 2019, **21**, 4258–4267.
50 R. Senthilnithy, M. Weerasingha and D. Dissanayake, *Comput. Theor. Chem.*, 2014, **1028**, 60–64.
51 L. Tavagnacco, S. Di Fonzo, F. D'Amico, C. Masciovecchio, J. W. Brady and A. Cesàro, *Phys. Chem. Chem. Phys.*, 2016, **18**, 13478–13486.
52 M. K. Shukla and J. Leszczynski, *J. Phys. Chem. A*, 2000, **104**, 3021–3027.
53 S. Monti, C. Cappelli, S. Bronco, P. Giusti and G. Ciardelli, *Biosens. Bioelectron.*, 2006, **22**, 153–163.
54 M. K. Shukla and J. Leszczynski, *J. Phys. Chem. A*, 2003, **107**, 5538–5543.
55 V. B. Singh, *RSC Adv.*, 2015, **5**, 11433–11444.
56 D. Jacquemin, A. Planchat, C. Adamo and B. Mennucci, *J. Chem. Theory Comput.*, 2012, **8**, 2359–2372.
57 T. Giovannini, M. Macchiagodena, M. Ambrosetti, A. Puglisi, P. Lafiosca, G. Lo Gerfo, F. Egidi and C. Cappelli, *Int. J. Quantum Chem.*, 2019, **119**, e25684.
58 B. Mennucci and J. Tomasi, *J. Chem. Phys.*, 1997, **106**, 5151–5158.
59 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.

60 O. Cannelli, T. Giovannini, A. Baiardi, B. Carlotti, F. Elisei and C. Cappelli, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32544–32555.
61 C. Cappelli, B. Mennucci and S. Monti, *J. Phys. Chem. A*, 2005, **109**, 1933–1943.
62 T. Giovannini, M. Olszówka and C. Cappelli, *J. Chem. Theory Comput.*, 2016, **12**, 5483–5492.
63 T. Giovannini, M. Olszówka, F. Egidi, J. R. Cheeseman, G. Scalmani and C. Cappelli, *J. Chem. Theory Comput.*, 2017, **13**, 4421–4435.
64 T. Giovannini, M. Ambrosetti and C. Cappelli, *Theor. Chem. Acc.*, 2018, **137**, 74.
65 D. Loco, É. Polack, S. Caprasecca, L. Lagardere, F. Lipparini, J.-P. Piquemal and B. Mennucci, *J. Chem. Theory Comput.*, 2016, **12**, 3654–3661.
66 A. Zapata-Escobar, M. Manrique-Moreno, D. Guerra, C. Z. Hadad and A. Restrepo, *J. Chem. Phys.*, 2014, **140**, 184312.
67 S. Gómez, A. Restrepo and C. Z. Hadad, *Phys. Chem. Chem. Phys.*, 2015, **17**, 31917–31930.
68 N. Rojas-Valencia, C. Ibargüen and A. Restrepo, *Chem. Phys. Lett.*, 2015, **635**, 301–305.
69 E. Flórez, N. Acelas, C. Ibargüen, S. Mondal, J. L. Cabellos, G. Merino and A. Restrepo, *RSC Adv.*, 2016, **6**, 71913–71923.
70 C. Cappelli, *Int. J. Quantum Chem.*, 2016, **116**, 1532–1542.
71 H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.*, 2009, **48**, 1198–1229.
72 P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, H. Basch and D. Cohen, *J. Chem. Phys.*, 1996, **105**, 1968–1986.
73 V. Kairys and J. H. Jensen, *J. Phys. Chem. A*, 2000, **104**, 6656–6665.
74 B. T. Thole, *Chem. Phys.*, 1981, **59**, 341–350.
75 C. Curutchet, A. Muñoz-Losa, S. Monti, J. Kongsted, G. D. Scholes and B. Mennucci, *J. Chem. Theory Comput.*, 2009, **5**, 1838–1848.
76 J. M. H. Olsen and J. Kongsted, *Adv. Quantum Chem.*, Elsevier, 2011, vol. 61, pp. 107–143.
77 A. H. Steindal, K. Ruud, L. Frediani, K. Aidas and J. Kongsted, *J. Phys. Chem. B*, 2011, **115**, 3027–3037.
78 F. Lipparini, *J. Chem. Theory Comput.*, 2019, **15**, 4312–4317.
79 D. Loco and L. Cupellini, *Int. J. Quantum Chem.*, 2019, **119**, e25726.
80 F. Lipparini, C. Cappelli and V. Barone, *J. Chem. Theory Comput.*, 2012, **8**, 4153–4165.
81 F. Egidi, T. Giovannini, G. Del Frate, P. M. Lemler, P. H. Vaccaro and C. Cappelli, *Phys. Chem. Chem. Phys.*, 2019, **21**, 3644–3655.
82 T. Giovannini, A. Puglisi, M. Ambrosetti and C. Cappelli, *J. Chem. Theory Comput.*, 2019, **15**, 2233–2245.
83 T. Giovannini, L. Grazioli, M. Ambrosetti and C. Cappelli, *J. Chem. Theory Comput.*, 2019, **15**, 5495–5507.
84 T. Giovannini, R. R. Riso, M. Ambrosetti, A. Puglisi and C. Cappelli, *J. Chem. Phys.*, 2019, **151**, 174104.
85 T. Giovannini, P. Lafiosca and C. Cappelli, *J. Chem. Theory Comput.*, 2017, **13**, 4854–4870.
86 T. Giovannini, P. Lafiosca, B. Chandramouli, V. Barone and C. Cappelli, *J. Chem. Phys.*, 2019, **150**, 124102.
87 C. Curutchet, L. Cupellini, J. Kongsted, S. Corni, L. Frediani, A. H. Steindal, C. A. Guido, G. Scalmani and B. Mennucci, *J. Chem. Theory Comput.*, 2018, **14**, 1671–1681.
88 P. Reinholdt, J. Kongsted and J. M. H. Olsen, *J. Phys. Chem. Lett.*, 2017, **8**, 5949–5958.
89 H. Gokcan, E. G. Kratz, T. A. Darden, J.-P. Piquemal and G. A. Cisneros, *J. Phys. Chem. Lett.*, 2018, **9**, 3062–3067.
90 S. W. Rick, S. J. Stuart and B. J. Berne, *J. Chem. Phys.*, 1994, **101**, 6141–6156.
91 F. Lipparini and V. Barone, *J. Chem. Theory Comput.*, 2011, **7**, 3711–3724.
92 M. Caricato, F. Lipparini, G. Scalmani, C. Cappelli and V. Barone, *J. Chem. Theory Comput.*, 2013, **9**, 3035–3042.
93 F. Egidi, R. Russo, I. Carnimeo, A. D'Urso, G. Mancini and C. Cappelli, *J. Phys. Chem. A*, 2015, **119**, 5396–5404.
94 A. Puglisi, T. Giovannini, L. Antonov and C. Cappelli, *Phys. Chem. Chem. Phys.*, 2019, **21**, 15504–15514.
95 R. Singh, A. K. Shakya, R. Naik and N. Shalan, *Int. J. Anal. Chem.*, 2015, 1–12.
96 Agilent Technologies, Inc., https://www.agilent.com/cs/library/certificateofanalysis/8500-6917%20lot%20OC128158.pdf, accessed: 2019-05-30.
97 A. V. Marenich, S. V. Jerome, C. J. Cramer and D. G. Truhlar, *J. Chem. Theory Comput.*, 2012, **8**, 527–541.
98 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19–25.
99 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
100 W. L. Jorgensen, *J. Am. Chem. Soc.*, 1981, **103**, 335–340.
101 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
102 M. Parrinello and A. Rahman, *J. Chem. Phys.*, 1982, **76**, 2662–2666.
103 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
104 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
105 H. J. Berendsen and W. F. Van Gunsteren, *Molecular-dynamics simulation of statistical-mechanical systems*, 1986, pp. 43–65.
106 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.
107 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.
108 E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, C. R. Landis and F. Weinhold, *NBO 6.0*, Theoretical Chemistry Institute, University of Wisconsin, Madison, 2013.
109 E. D. Glendening, C. R. Landis and F. Weinhold, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 1–42.
110 T. Le Bahers, C. Adamo and I. Ciofini, *J. Chem. Theory Comput.*, 2011, **7**, 2498–2506.

111 C. Adamo, T. L. Bahers, M. Savarese, L. Wilbraham, G. García, R. Fukuda, M. Ehara, N. Rega and I. Ciofini, *Coord. Chem. Rev.*, 2015, **304–305**, 166–178.
112 F. Egidi, G. Lo Gerfo, M. Macchiagodena and C. Cappelli, *Theor. Chem. Acc.*, 2018, **137**, 82.
113 R. Di Remigio, T. Giovannini, M. Ambrosetti, C. Cappelli and L. Frediani, *J. Chem. Theory Comput.*, 2019, **15**, 4056–4068.
114 F. Weinhold, *Computational Methods in Photochemistry*, Taylor & Francis/CRC Press, 2005, pp. 393–476.
115 T. Agou, J. Kobayashi and T. Kawashima, *Org. Lett.*, 2005, **7**, 4373–4376.
116 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian-16 Revision B.01*, Gaussian Inc., Wallingford CT, 2016.
117 E. Rengifo, S. Gómez, J. C. Arce, F. Weinhold and A. Restrepo, *Comput. Theor. Chem.*, 2018, **1130**, 58–62.
118 S. Gómez, J. Nafziger, A. Restrepo and A. Wasserman, *J. Chem. Phys.*, 2017, **146**, 074106.
119 C. Hadad, E. Florez, N. Acelas, G. Merino and A. Restrepo, *Int. J. Quantum Chem.*, 2019, **119**, e25766.
120 H. Farrokhpour and F. Fathi, *J. Comput. Chem.*, 2011, **32**, 2479–2491.
121 T. Giovannini, M. Ambrosetti and C. Cappelli, *J. Phys. Chem. Lett.*, 2019, **10**, 5823–5829.
122 A. Roberts, *Nutraceuticals*, Academic Press, Boston, 2016, pp. 417–434.

4.2.2. Anionic Ibuprofen[3]

molecules

MDPI

Article

Ring Vibrations to Sense Anionic Ibuprofen in Aqueous Solution as Revealed by Resonance Raman

Sara Gómez [1,*], Natalia Rojas-Valencia [2], Tommaso Giovannini [1], Albeiro Restrepo [2] and Chiara Cappelli [1,*]

[1] Scuola Normale Superiore, Classe di Scienze, Piazza dei Cavalieri 7, 56126 Pisa, Italy; tommaso.giovannini@sns.it
[2] Instituto de Química, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellin 050010, Colombia; nandrea.rojas@udea.edu.co (N.R.-V.); albeiro.restrepo@udea.edu.co (A.R.)
* Correspondence: sara.gomezmaya@sns.it (S.G.); chiara.cappelli@sns.it (C.C.)

Abstract: We unravel the potentialities of resonance Raman spectroscopy to detect ibuprofen in diluted aqueous solutions. In particular, we exploit a fully polarizable quantum mechanics/molecular mechanics (QM/MM) methodology based on fluctuating charges coupled to molecular dynamics (MD) in order to take into account the dynamical aspects of the solvation phenomenon. Our findings, which are discussed in light of a natural bond orbital (NBO) analysis, reveal that a selective enhancement of the Raman signal due to the normal mode associated with the C–C stretching in the ring, $\nu_{C=C}$, can be achieved by properly tuning the incident wavelength, thus facilitating the recognition of ibuprofen in water samples.

Keywords: resonance Raman; UV-vis; spectroscopy; NBO; ibuprofen

Citation: Gómez, S.; Rojas-Valencia, N.; Giovannini, T.; Restrepo, A.; Cappelli, C. Ring Vibrations to Sense Anionic Ibuprofen in Aqueous Solution as Revealed by Resonance Raman. *Molecules* **2022**, *27*, 442. https://doi.org/10.3390/molecules27020442

Academic Editor: Eric Glendening

Received: 10 December 2021
Accepted: 4 January 2022
Published: 10 January 2022

1. Introduction

Ibuprofen, a non-steroidal anti-inflammatory agent, has become and remains one of the most consumed medications around the world [1]. Indeed, the worldwide ibuprofen market is projected to increase over the period 2020–2024 [2]. The wide use of ibuprofen is related to its capability to provide therapeutic action for a variety of diseases,[1] via the inhibition of the COX enzyme [3] during the production of prostaglandins [4]. According to its chemical structure, ibuprofen is classified as an amphiphilic molecule as it presents an aromatic ring doubly substituted with a methyl propyl group and with a propionic acid, the latter constituting the polar part of the drug. The connection between the non-polar and polar groups not only allows the molecule to be soluble in both polar and apolar environments but also gives rise to an enantiomeric carbon from which the *S*-configuration has been regarded as the active species (see Figure 1) [5].

The study of ibuprofen in aqueous solutions is a hot research topic from biological and environmental points of view [6–8]. From the biological perspective, the human body is around 70% water, therefore it is expected that water$\cdots$ibuprofen interactions will play an important role in the drug behavior and hence in its therapeutic action [9,10]. On the other hand, ibuprofen is of environmental importance because it is consumed in large quantities and it has been reported as a pollutant that must be removed from residual waters (especially from healthcare and wastewater treatment facilities) [8,11–13]. In both cases, the identification of ibuprofen in the aquatic bodies is quite relevant taking into account its protonation state since ibuprofen undergoes structural transitions at acidic or basic pH, and its properties highly depend on it [14,15]. In this regard, ibuprofen as well as, to a greater extent, the anionic ibuprofen (a-Ibu, resulting from deprotonation) have been described as strong chelating agents, acting as monodentate or bidentate ligands through the carboxylate oxygens [16].

Over the years, the removal of ibuprofen from aqueous solutions has drawn great interest. Researchers have suggested several promising strategies as chemical adsorption on different materials,[17–21] advanced oxidation processes [13,22], coagulation–flocculation [23], photocatalytic degradation [24], among others, focusing on the increase of the removal rate. As for the detection of ibuprofen in water and wastewater, electrochemical techniques [25–28] seem to be the most investigated, as well as microextraction methodologies [29]. In addition, spectroscopic techniques are well established fundamental tools to identify several analytes in a given aqueous matrix. Among these techniques, UV-vis is one of the most commonly used (by itself or coupled to a separation technique such as HPLC), and calculated UV-vis spectra are also available in the literature to better understand the experimental data [30].

According to the experimental studies carried out by Olaru and Patras [31] and Du et al. [32], it is possible to identify ibuprofen in water by means of UV-vis spectroscopy. However, the spectra from both the protonated and deprotonated drug perfectly match each other, thus suggesting that the excited electronic states involving the apolar part of the drug, which is the only difference between the structures, play a major role than those of its polar component. In order to get better detection and distinction pictures, other spectroscopic techniques have been used. For example, neutral and anionic ibuprofen have been widely studied and characterized by Raman, Raman optical activity (ROA), and IR spectroscopies [33–38]. However, due to the very strong IR absorption of water in the 1700–1400 cm^{-1} region, the use of Raman is recommended [16]. It has been shown that the vibrational modes, especially those of the aromatic ring, change depending on the chemical environment surrounding the molecule, and due to that, they have been regarded as sensors [35]. Notwithstanding, very diluted ibuprofen-containing solutions continue to be a challenge in the detection processes [39,40].

Resonance Raman (RR) has been proposed as a technique that can significantly increase band intensity and resolution compared to conventional Raman, because the incident wavelength is at resonance with molecular electronic transitions [41–44]. Therefore, RR allows the detection of analytes even at very low concentrations [45,46]. Despite the obvious importance of the detection of ibuprofen in aqueous solution, to the best of our knowledge, there are neither computational nor experimental studies of RR applied to that system, albeit ultraviolet resonance Raman (UVRR) spectra of human serum albumin and its complexes with three types of ligands, including ibuprofen, have been measured at 240 nm [47]. In this work, based on computational simulations of UV-vis, Raman, and resonance Raman spectra of a-Ibu in aqueous environment, we investigate the potentialities of RR to identify this drug in water. The Natural Bond Orbitals (NBO)-aided detailed description of the different electronic transitions that are involved in UV-vis spectra and could be finely tuned in RR, allows us to gather complementary information to understand the underlying causes of intensity enhancements in Raman spectra, which have not been previously explored.

The paper is organized as follows: after describing the details of the computations in the next section, a discussion focused on hydration patterns under the MD and NBO contexts is presented. Then, the results for diverse spectroscopies, UV-vis, Raman and resonance Raman applied to solvated a-Ibu are reported and analyzed. Finally, conclusions are drawn.

2. Methods

Ibuprofen can occur in protonated and deprotonated forms depending on the pH of the solution. An initial analysis of protonation with respect to pH was carried out using Marvin Beans version 19.20 [48] and according to the speciation plots, at neutral pH (pH = 7.0) the dominant form of ibuprofen bears a deprotonated carboxylic group. Besides, physicochemical analyses of the quality of wastewater from different effluents have revealed that regardless of the facility, season, level of treatment, sampling point, etc., pH ranges vary from 6 to 10 [49] with an average pH value found to be around 7.5 [50].

With the ibuprofen pKa being 4.91, this compound will primarily exist in the dissociated form in wastewater too. Therefore, we used the a-Ibu in all the following calculations.

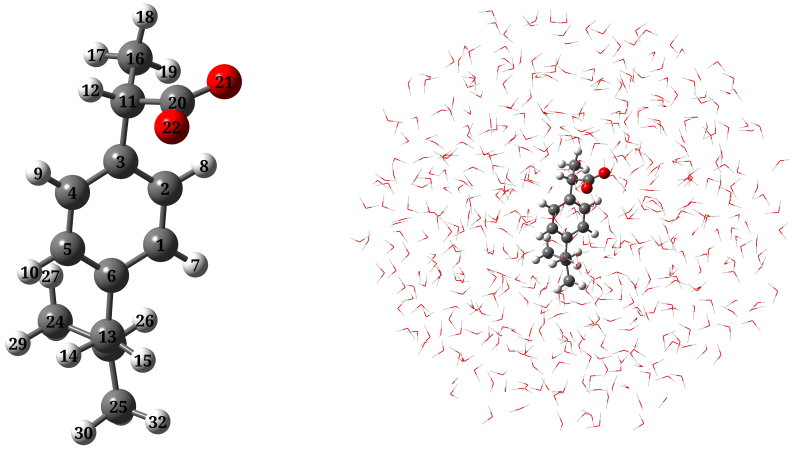


Figure 1. Molecular structure of a-Ibu and atom labeling (**left**); a pictorial view of a-Ibu dissolved in aqueous solution as treated in QM/MM calculations (**right**).

The molecular geometry of a-Ibu (see Figure 1, left panel) was optimized by employing the CAM-B3LYP density functional [51] combined with the 6-311++G(d,p) basis set [52,53]. To ensure that the geometries and calculated properties are actually representative of the overall solvation environment, we carried out a Molecular Dynamics (MD) simulation to sample the solute-solvent phase space [54]. To do that, the optimized structure of a-Ibu was placed in a cubic box with 7390 water molecules, setting the smallest atom$\cdots$wall distance to 1 nm. The MD run was conducted using the GAFF force field [55] with the TIP3P water model [56]. Parameters for a-Ibu were generated from the electrostatic potential and the CM5 charges [57]. Simulations were run in GROMACS 2019.3, [58] for a total production length of 30 ns, and a step-size of 2 fs. Temperature and pressure were maintained at 298.15 K and 1 bar, respectively, using a modified Berendsen thermostat [59] and Parrinello-Rahman barostat [60], with a coupling constant, τ, of 0.1 ps for each. The system was also equilibrated in two ensembles, NVT and NPT, of 1 ns runs, with solvent and solute in separate temperature coupling groups. At the end of the production stage, we sampled 200 uncorrelated configurations at intervals of 10 ps, discarding the first 10 ns. In the chosen frames, the closest water molecules within a radius of 14 Å of a-Ibu were included in a sphere-shaped cut, as depicted in the right panel of Figure 1. Geometric clustering was conducted to identify similar conformations sampled during the MD run. To this end, the *gromos* clustering method from Ref. [61] with a cutoff of 0.13 nm (the average RMSD) was used. This cluster analysis identified 7 representative structures of a-Ibu, illustrated in Figure S1 in the Supplementary Material (SM). Their corresponding occurrences along the MD trajectory and the size of the cluster are plotted in Figure S2 in the SM.

To get insight into solute–solvent hydrogen bonding interactions, we analyzed the MD trajectory by means of radial distribution functions (RDFs), spatial distribution functions (SDFs), and the average number of hydrogen bonds (HBs) between water molecules and the carboxylate group, by using the TRAVIS package [62,63]. Orbital interactions associated to hydrogen bonding were analyzed at the NBO level [64–66] and the interaction energies were obtained via second order perturbation corrections to the Fock matrix with the NBO7.0 program [67]. We also borrowed and reoptimized the structures reported in Ref. [7] as being the lowest energy motifs in the quantum mechanics (QM) potential energy surface (PES) for the microsolvation of a-Ibu: W_1S_1 ($\%x_i = 94.3$), W_1S_2 ($\%x_i = 63.4$) and W_1S_3

(%$x_i = 44.9$), and applied NBO methodologies to quantify and compare the strength of the interactions.

In spectral calculations, solvent effects were described by means of the quantum mechanics/fluctuating charges (QM/FQ) model [54], applied to each of the extracted snapshots. FQ parameters reported in Ref. [68] were exploited. TD-DFT calculations (15 excited states) for each one of the configurations extracted from the MD trajectory were then carried out at the CAM-B3LYP/6-311++G(d,p) level of theory. It is worth noticing that with this DFT functional and large number of excited states not only intensities and band shapes are accurately modeled, but also the experimental spectra are better reproduced. Reported averaged absorption spectrum was obtained by convoluting peak intensities with Gaussian functions, with a full width at half maximum (FWHM) of 0.5 eV. The orbitals involved in the main transitions were identified and by means of a canonical molecular orbitals (CMO) analysis, we tabulated the leading NBO contributions (bonding, nonbonding, or antibonding) to each canonical Molecular Orbital (MO). Afterwards, the set of geometries extracted from MD simulations was partially optimized [69] and the vibrational calculations were performed on each converged minima. We also recomputed electronic absorption spectra with the optimized snapshots and we obtained essentially the same results as for the non-optimized ones (see Figure S3 in the SM). The Raman scattering spectrum of a-Ibu in aqueous solution was calculated at 532 nm with the QM/FQ protocol [69]. Finally, the QM/FQ methodology described in Ref. [70] was applied to calculate resonance Raman spectra, choosing several incident wavelengths to build the resonance Raman excitation profile (RREP). The Franck Condon Vertical Gradient approximation, successfully used in other works [70–73], was exploited. For a better visualization, the Raman and RR stick bands were convoluted with a Lorentzian line shape with FWHM of 20 cm^{-1}. Convergence tests indicated that the inclusion of extra (more than 200) uncorrelated snapshots yield unaltered UV-vis, Raman, and RR spectra (see Figures S6–S8 in the SM). All QM calculations were performed using a locally modified version of the Gaussian 16 package [74]. It should be noted that all calculations refer to the *S*-enantiomer.

3. Results and Discussion

3.1. Hydration Patterns

Hydration patterns sampled by MD simulations were analyzed in terms of two descriptors: SDFs and RDFs. Figure 2 shows SDF of solvent atoms (O_w and H_w) near oxygen atoms of a-Ibu. Such plots clearly identify the space region occupied by water molecules; the almost symmetrically distributed surfaces indicate that the two oxygen atoms in the CO_2^- motif behave in a very similar way. In order to refine the analysis, the RDF between the carboxylate group and water hydrogen atoms was also calculated (Figure 2, right panel). A sharp peak in the *g(r)* can be recognized at 1.7 Å, which integrates for around 3 water molecules located close to each oxygen in the first hydration sphere, whereas a less pronounced peak is present at 3.0 Å, thus denoting a second solvation shell formed by 7 ($\times 2$) water molecules. These $O\cdots H_w$ distances and arrangements of the solvent molecules surrounding the solute are in line with the findings by Zapata-Escobar et al. [7] who reported a detailed study of the microsolvation of a-Ibu and found that water molecules in direct contact with the solute prefer to aggregate around the carboxylate oxygen atoms via cyclic or bridged charge assisted HBs.

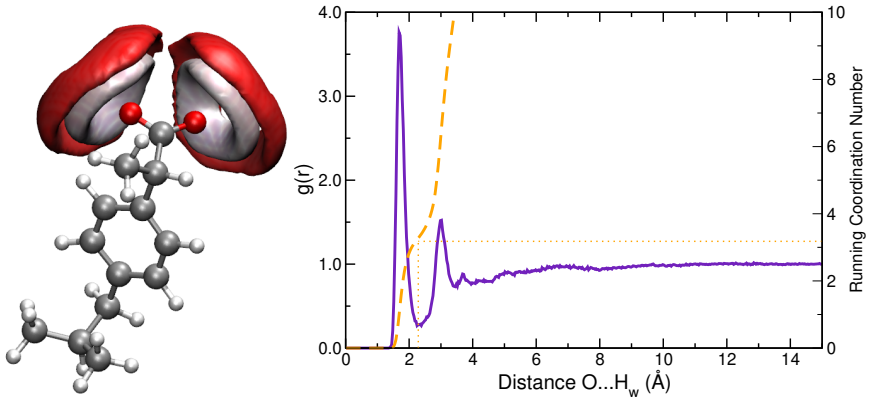


Figure 2. Left panel: Spatial distribution function of water oxygen (red) and hydrogen (white) atoms around a-Ibu. Calculated SDF isodensity values are equal to 70 and 100 nm^{-3} for hydrogen and oxygen atoms, respectively. **Right panel**: RDF between ibuprofen oxygens and water hydrogens. The curve is similar for both oxygen atoms in the carboxylate group. Running Coordination Numbers are also included (dashed orange line).

A more comprehensive examination of hydrogen-bonding motifs for the solvated a-Ibu can be done by means of the stabilization energies given by the second order perturbation theory analysis of the Fock matrix on the NBO basis. Such energies allow determining the strength of the donor–acceptor interactions between the solute and the water molecules in its vicinities. Figure 3 compares the intermolecular NBO interactions for anionic ibuprofen and water molecules for the most stable structures reported in the [a-Ibu(H_2O)$_n$] n = 1, 2, 3 PES [7], with those obtained for a single snapshot (Ibu1 in the SM) extracted from the MD when several water molecules were included in the QM portion. To quantify and compare the relative strengths of the interactions in these molecular clusters, we list the corresponding energies, $E_{d\rightarrow a^{(2)}}$, in Table 1.

Table 1. $n_O \rightarrow \sigma^*_{H-O}$ orbital interaction energies in solvated a-Ibu. The largest stabilization energies in kcal/mol are reported in each case. See Figure 3 for atom labeling.

Configuration	Donor (d)	Acceptor (a)	$-E^{(2)}_{d\rightarrow a}$	Type
W_1S_1	O21	O33-H35	6.77	$CO_2^-\cdots$H–O–H
	O22	O33-H34	6.81	$CO_2^-\cdots$H–O–H
W_2S_1	O22	O33-H35	9.37	$CO_2^-\cdots$H–O–H
	O21	O36-H37	26.16	$CO_2^-\cdots$H–O–H
	O36	O33-H34	7.57	$H_2O\cdots$H–O–H
W_3S_1	O21	O33-H34	16.38	$CO_2^-\cdots$H–O–H
	O22	O36-H38	7.84	$CO_2^-\cdots$H–O–H
	O22	O39-H41	4.61	$CO_2^-\cdots$H–O–H
	O33	O39-H40	9.30	$H_2O\cdots$H–O–H
	O36	O33-H35	5.14	$H_2O\cdots$H–O–H
	O39	O36-H37	3.21	$H_2O\cdots$H–O–H
[a-Ibu(H_2O)$_6$]/FQ	O21	O33-H35	18.82	$CO_2^-\cdots$H–O–H
	O21	O36-H37	3.19	$CO_2^-\cdots$H–O–H
	O21	O39-H41	5.13	$CO_2^-\cdots$H–O–H
	O22	O42-H43	10.63	$CO_2^-\cdots$H–O–H
	O22	O45-H46	9.60	$CO_2^-\cdots$H–O–H
	O22	O48-H49	9.65	$CO_2^-\cdots$H–O–H
	O42	O33-H34	3.84	$H_2O\cdots$H–O–H

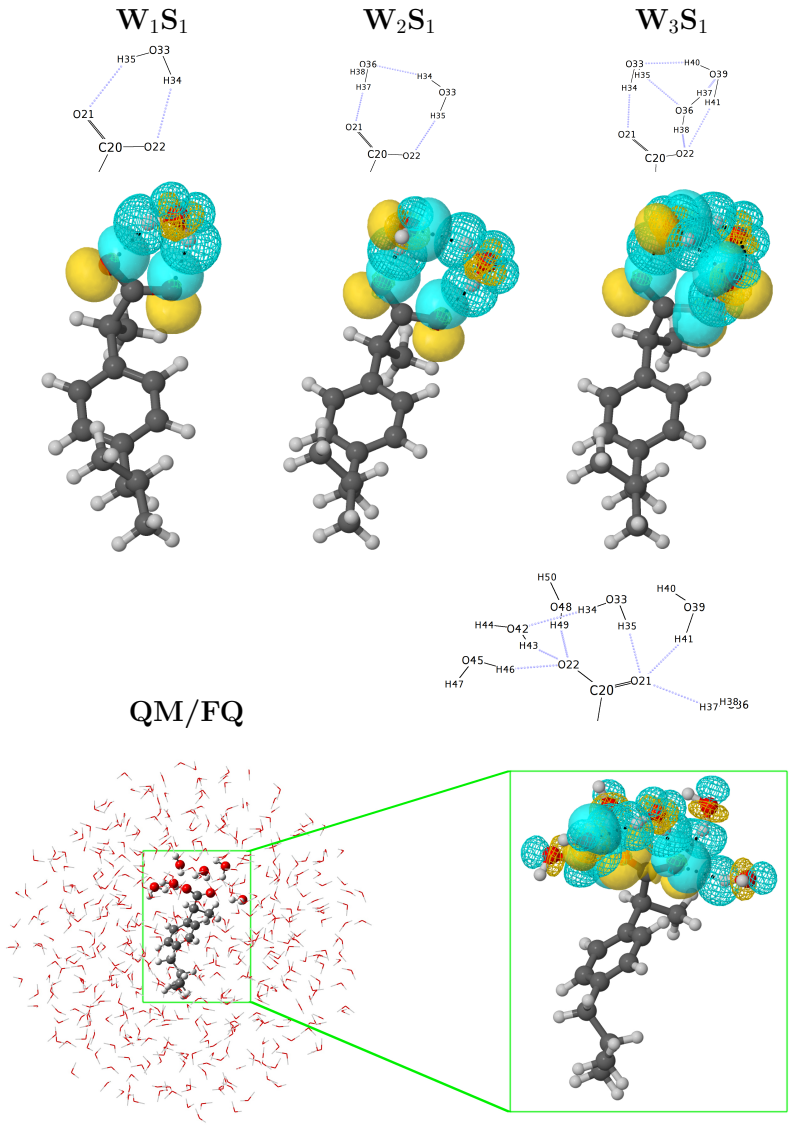


Figure 3. Orbital representation within the NBO picture for the intermolecular interactions in solvated a-Ibu. Structures for W_1S_1, W_2S_1, and W_3S_1 were taken from Ref. [7]. The QM/FQ structure with six explicit waters exhibits the most representative conformation of a-Ibu during the MD run. See also Figure S1 in the SM.

Even if in the NBO framework there are distinct types of interactions which stabilize the configurations/clusters of a-Ibu in water, the dominant contribution is given by the lone pair $\rightarrow$ antibonding ($n_O \rightarrow \sigma^*_{H-O}$) form. In all cases, the strongest interactions arise from the charge transfer between a lone pair in the carboxylate group and an antibonding σ^*_{H-O} orbital of a neighboring water molecule, affording stabilization energies up to 6.81, 26.16, 16.38, and 18.82 kcal/mol when 1, 2, 3 and 6 water molecules (plus the rest of the FQ layer) surround the solute. For clusters with more than one water molecule, the two oxygen atoms are not exactly equivalent and water$\cdots$water interactions are also possible but in general, the $n_O \rightarrow \sigma^*_{H-O}$ overlap leads to small orbital interaction energies, though they are comparable or stronger than the HB in the reference water dimer [75,76] due to the formal charge in a-Ibu. Note that, as predicted by the RDFs, for the configuration considered for

this analysis, there are three water molecules located around each oxygen in the a-Ibu. In addition, the $E^{(2)}_{d\to a}$ energy values, calculated for the interactions in that snapshot, are in the same ranges as those found in a-Ibu/water clusters coming from exhaustive explorations of the PESs, reinforcing the fact that randomly choosing configurations from classical mechanics simulations are reliable sources to gain deep insight about inter-fragment bonding, a strategy that has been recently used in several works [10,77–80].

3.2. UV-Vis Spectrum

We now move to discuss the comparison between computed and experimental UV-vis spectra. Electronic absorption spectra of a-Ibu are usually measured for the ibuprofen sodium salt in solution. When recorded in the 200–300 nm range, the spectrum is characterized by an intense peak with a maximum absorption wavelength (λ_{max}) of 222 nm, associated with a $\pi \to \pi^*$ transition [24,32,81,82]. Other authors claim the main absorption bands to be located at 222 nm and 190 nm, with the latter value being a rough estimation because below 200 nm the spectral data is not conclusive [24]. The QM/FQ calculated absorption spectrum for a-Ibu in aqueous solution is shown in Figure 4, left panel, and the position of absorption maxima (222 nm and 182 nm) in the analyzed interval give a first indication of the good agreement between calculated and experimental data, thus also validating the level of theory here employed. Interestingly, the destruction of the drug by photocatalytic degradation is followed by the disappearance of the band corresponding to ibuprofen at 222 nm [24].

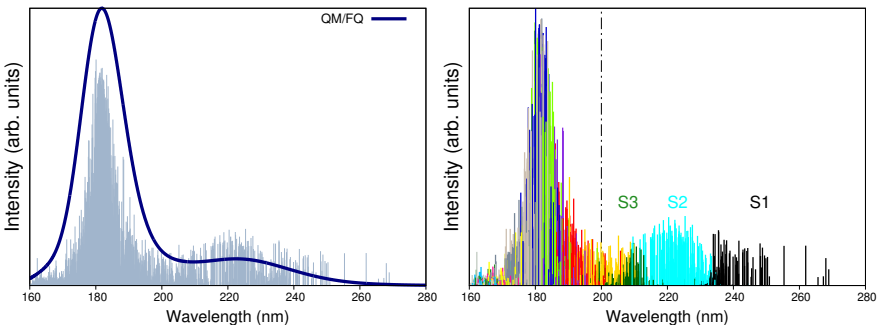


Figure 4. Left panel: Convoluted QM/FQ UV-vis absorption spectrum of ibuprofen in aqueous solution. Convolution was done with Gaussian functions and using a FWHM of 0.5 eV. **Right panel**: Distinction of the fifteen excited states converged in the TD-DFT calculations, where each one of them is associated to a different stick color. Labels S1 to S3 indicate the first three excited states. Experimental reports collected from Refs. [24,32,81,82] determine the first maximum to appear at 222 nm. Dashed vertical line indicates that there is no experimental information below 200 nm.

As can be seen in Figure 4, left panel, the fact of using a robust sampling methodology that includes several snapshots from the MD brings in a natural broadening in the UV-vis spectra, otherwise lost or arisen from a mere convolution if just a few configurations/conformations are taken into account as it is done in cluster-like approaches or implicit solvation [54]. On the right panel of Figure 4, the stick-like spectrum of a-Ibu was separated by colors according to the excited state the sticks described, in such a way that it is possible to distinguish the excited states mainly contributing to each band. It is clear from such a qualitative separation that S_1, S_2 and S_3 lead to the appearance of the band at higher wavelengths, whereas the second band is due to a combination of all the remaining excitations. In Table 2, the MOs contributions to S_1–S_3 transitions are listed. For instance, S_1 (black sticks in Figure 4) is predominantly a HOMO $\to$ LUMO + 1 excitation, whereas S_2 includes the HOMO as well but the receptor of the charge transfer is the LUMO. The involved MOs are depicted in Figure 5 for the most representative snapshot in light of the clustering method [61].

Table 2. Decomposition of the excited states mainly contributing to the onset of the band centered at 222 nm in the absorption spectrum of a-Ibu in solution. The decomposition is done based on the canonical molecular orbitals (CMO) and the natural bonding orbitals (NBO).

Excited State	Orbitals Involved		Contribution (%)
	CMO	NBO	
S_1	HOMO → LUMO + 1	$\pi_{ring} \rightarrow \pi^*_{ring}$	49.72
	HOMO → LUMO	$\pi_{ring} \rightarrow \pi^*_{C=O}$	29.36
	HOMO - 1 → LUMO	$\pi_{ring} \rightarrow \pi^*_{C=O}$	20.92
S_2	HOMO → LUMO	$\pi_{ring} \rightarrow \pi^*_{C=O}$	79.80
	HOMO → LUMO + 1	$\pi_{ring} \rightarrow \pi^*_{ring}$	20.20
S_3	HOMO - 1 → LUMO + 1	$\pi_{ring} \rightarrow \pi^*_{ring}$	50.50
	HOMO → LUMO + 3	$\pi_{ring} \rightarrow \pi^*_{ring}, \pi^*_{C=O}$	49.50

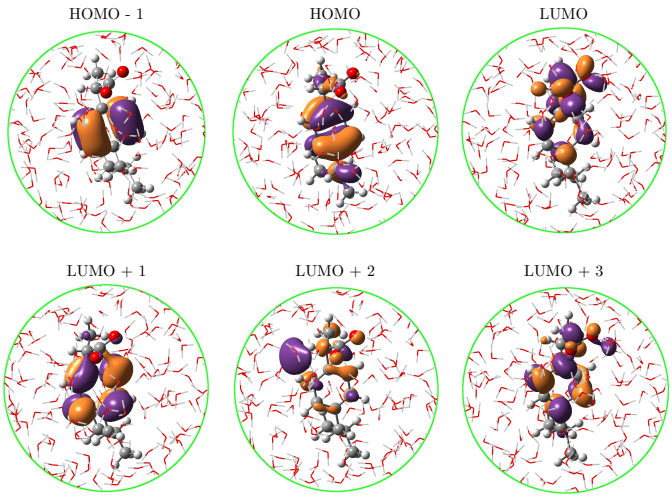


Figure 5. Molecular orbitals for the lowest three excited states of a-Ibu in aqueous solution. See Table 2 for their contributions to the transitions.

The NBO analysis of CMO reported in Table 2 allows to further decompose the MOs in NBOs and assign the nature of the transitions. By looking at the values in the linear combinations for the MO already mentioned, it turns out that though the occupied orbitals involved in the transitions might look somewhat different, the NBOs forming them belong to the π aromatic ring in the molecule. Visual inspection of the LUMO orbital in Figure 5 reveals that the carboxylate group plays an important role in that orbital and the CMO decomposition confirms that hypothesis by attributing it primarily to the antibonding $\pi^*_{C=O}$ orbital. Gathering all this information, the band at 222 nm in the resonance Raman spectrum of the solvated a-Ibu is a superposition of three excited states, which can be briefly summarized as a redistribution of the electron density from the π-cloud on the ring to either the ring itself or the COO^- groups, in a $\pi_{ring} \rightarrow \pi^*_{ring,C=O}$ charge transfer according to the NBO description.

3.3. Raman Spectrum

The experimental far-from-resonance Raman spectrum of solid IbuNa salt and of its solution in water were reported by Bonora et al. [16] along with the corresponding assignments of the normal modes. Therefore, in what follows, just a summary of the main findings will be discussed. Figure 6 reports the QM/FQ simulated Raman spectrum of

a-Ibu in aqueous solution and its experimental counterpart. Overall, there is an outstanding agreement of almost all peak positions and relative intensities.

It has been reported that the most relevant difference between the Raman spectra of the solid undissociated acid form of ibuprofen and a-Ibu in aqueous solution is due to the $\nu_{C=O}$ stretching mode that for the carboxylic group appears at 1647 cm^{-1} (computed at 1709 cm^{-1} in Ref. [16]), which is of course not present neither in the diluted IbuNa water solution nor in the solid IbuNa, because in those systems, the asymmetrical ν_{as,COO^-} and symmetrical ν_{s,COO^-} stretching vibrations take place [16].

Concerning the assignments of the bands appearing in Raman spectrum, the most important signals in the 600–1800 cm^{-1} interval account for an intense peak at 1613 cm^{-1} (theoretically located at ≈1690 cm^{-1}), attributed to stretching vibrations of the benzene ring ($\nu_{C=C}$). The second most intense peak appears as a doublet with maxima at 1450 and 1455 cm^{-1} (simulated at 1500 and 1510 cm^{-1}) and can be assigned to the antisymmetric deformations of the methyl groups, the CH_2 scissors mode and the ν_{s,COO^-} vibration. Vargek [83] claimed that the splitting could originate from the close proximity of the phenyl ring and the CH_3 group in the chiral carbon atom. Another couple of peaks popping up in the spectrum is that at 1288 and 1342 cm^{-1}, which originates from the *out-of-plane* CCCH motion, and the aliphatic β_{HCH} and β_{HCC} bendings. The two Raman peaks at around 800 cm^{-1} are attributable to torsions τ_{HCCC} and ν_{CC} stretchings, whereas the peak at 640 cm^{-1} can be assigned to the phenyl CH *out-of-plane* deformation, though the bending β_{CCC} is also involved. For the upcoming Resonance Raman spectra, it is worth describing the two peaks appearing at 1188 and 1210 cm^{-1}, which are the result of a complex combined τ_{HCCC}, β_{HCC}, out_{CCCC}, and aliphatic ν_{CC} vibrations including both the propyl and the isobutyl group, located at the *para*- position from the propanoate (propyl group + carboxylate). CH-bending modes of the benzene ring also contribute to these peaks [47]. The normal modes with frequencies around 1200 and 1690 cm^{-1} are depicted in Figure 7 for the most representative snapshot selected *via* the *gromos* clustering method. These vibrational modes display large displacements of the C–C bonds.

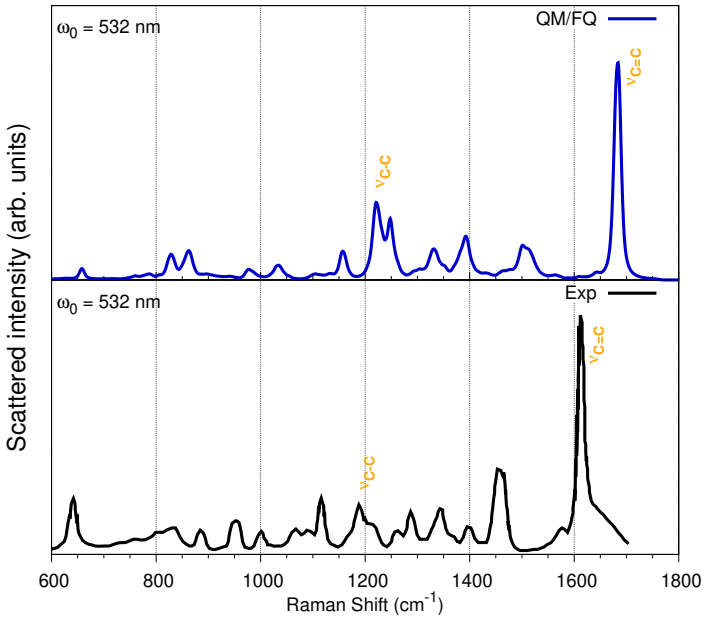


Figure 6. Convoluted QM/FQ (top) and experimental (bottom) Raman spectrum of a-Ibu in aqueous solution. Experimental data taken from Refs. [83–85].

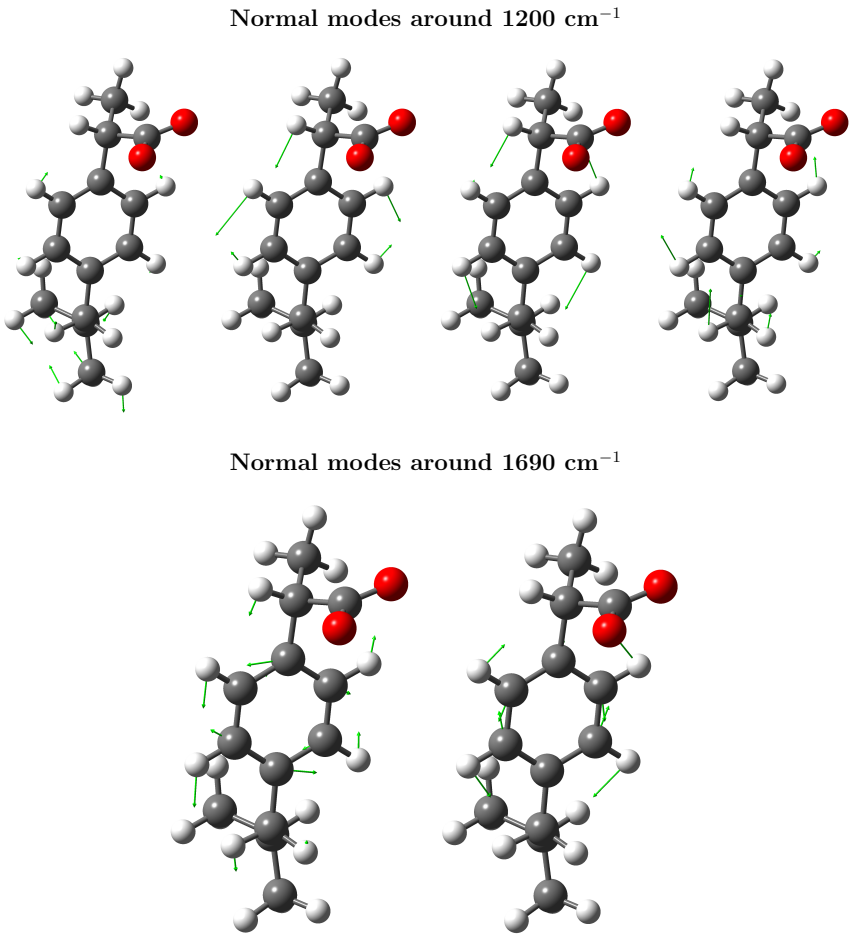


Figure 7. Important vibrational modes giving rise to the most enhanced peaks in resonance Raman. Top panel: normal modes especially enhanced at incident frequencies, $\omega_0 \approx 55{,}000$ cm^{-1} (182 nm). Bottom panel: normal modes always enhanced according to the RREP, but selectively more enhanced at $\omega_0 \approx 45{,}000$ cm^{-1} (222 nm).

3.4. Resonance Raman

From the simulated UV-vis spectrum of a-Ibu (Figure 4), the lowest electronic absorption with significant oscillator strength was predicted to arise from three excitations, consistent with the appearance of a single, broad band in the aqueous phase absorption spectrum having a maximum (at 222 nm) quite close to what is experimentally observed. Nevertheless, there is another more intense band (at 182 nm) as a result of several transitions with the highest calculated oscillator strengths. The computed RR spectra of a-Ibu when these two incident wavelengths are used to irradiate the sample, are depicted in Figure 8. Spectral profiles for ω_0 ranging from 165 to 285 nm are displayed in Figure S4 in the SM. For full clarity, we also computed RR spectra taking into account just the first three excited states, the results are found in the SM (Figure S5).

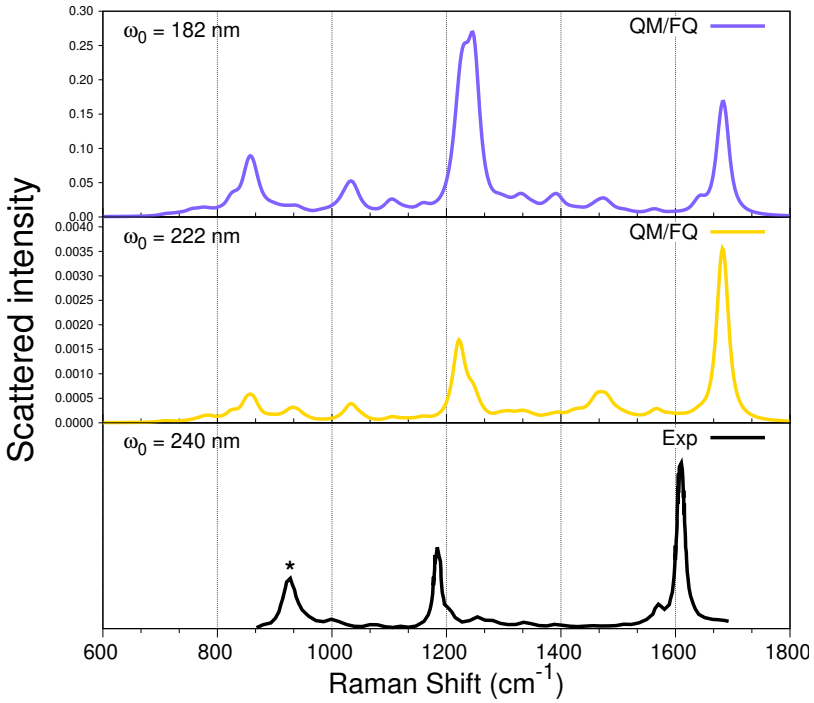


Figure 8. QM/FQ UV-resonance Raman spectra of a-Ibu in aqueous solution, calculated when the two absorption maxima wavelengths are used to irradiate the system (top, 182 nm and middle, 222 nm). RR intensities (in $cm^2mol^{-1}sr^{-1}$) were calculated with a damping factor of 200 cm^{-1} and broadened using Lorentzian functions with FWHM = 20 cm^{-1}. For visualization purposes, the intensity of the highest peak was normalized to 1. In the experimental spectrum [47] (bottom panel), the intensity was calibrated with the ClO_4^- band marked with *.

When compared against the off-resonance Raman spectra in Figure 6, the RR spectra in Figure 8 immediately show evidence of selective resonance enhancement, as is characteristic of RR spectroscopy [86–88]. That is particularly evidenced by the slight changes in the position and strong changes in the intensity of some of the peaks described above. Indeed, the major features of the RR spectra of solvated a-Ibu are the enhancements of either the peak around 1200 cm^{-1} or the peak located at ≈1690 cm^{-1}, both associated mostly with carbon–carbon vibrations in the aliphatic regions and ring, respectively, and to movements of the methyl groups of ibuprofen.

Two important findings are noticeable when comparing the two simulated spectra in Figure 8. On the one hand, the RR intensities at 182 nm are about 100 times larger than those at 222 nm (see also Figure 9 below), thus reflecting the relative intensities in the absorption spectra; and on the other hand, there is a selective enhancement of the peak at ≈1200 cm^{-1} at $\omega_0 = 182$ nm. In fact, at that incident wavelength, this peak exhibits the largest RR intensity, while the peaks located at 1400 cm^{-1} almost disappeared. Furthermore, notice that at $\omega_0 = 182$ nm, the intensity of the peak at about 800 cm^{-1} is enhanced.

The only work we found concerning UVRR studies of a-Ibu in aqueous solution, reported measurements with a laser excitation set up to 240 nm,[47] a near-resonance incident wavelength. For the sake of comparison, that spectrum is included in the bottom panel of Figure 8. The authors pointed out that the most prominent peaks are located at 1613 and 1188 cm^{-1}, assigned to the ring stretching and CH-bending modes of the benzene ring, respectively. Those peaks are characteristic of p-disubstituted benzene derivatives. Our simulations are perfectly consistent with their findings, except perhaps for a small shift in the frequencies.

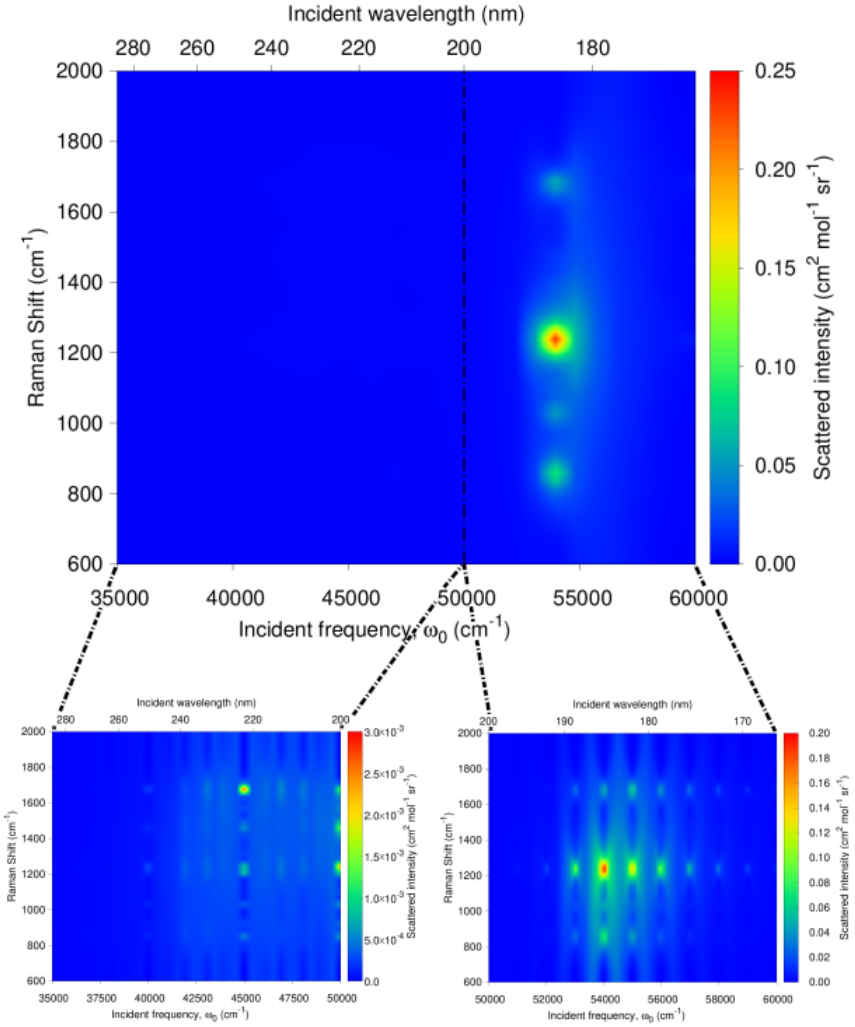


Figure 9. Calculated QM/FQ resonance Raman excitation profile (RREP) of a-Ibu in aqueous solution. Dashed vertical line indicates that there is no experimental information below 200 nm.

Going beyond these results at particular excitation wavelengths, the RREP of a-Ibu is shown in Figure 9 as well as a zoom to see in more detail the excitation profile in two specific incident frequency regions: (285–200 nm) and (200–165 nm) under the available experimental information for the absorption spectra. At a first glance, there are at least four peaks (≈800, 1000, 1200 and 1700 cm^{-1}) in the RR spectra expected to increase their intensity when an incident wavelength ω_0 = 182 nm is used. Nonetheless, the separation of the regions allows us to recognize that if the threshold of the intensities is lowered, it would be possible to find more enhanced peaks, as seen in the green patches of the RREP maps. While one might argue that this sacrifice in intensity jeopardizes the aim of the detection, for finely tunable devices (like a synchrotron source) used to carry out UVRR experiments, the exploitable range of incident wavelengths is 127–280 nm and moreover, the spectral resolution can be set up for synchrotron radiation-based UVRR experiments in order to have a sufficiently high signal to noise ratio [70,89–91].

The most enhanced peaks in the RR spectrum of solvated a-Ibu were monitored individually by sectioning the RREP maps. The excitation wavelength dependence of the

intensity of the enhanced peaks is shown in Figure 10 along with its Raman cross-section ratios. As expected, the calculated excitation profiles for $\nu_{C=C}$ and ν_{C-C} reach their maxima at the excitation wavelengths found in the simulated absorption spectra, though the ν_{C-C} peak shows essentially the same low Raman cross section from 200 to 260 nm.

The selective enhancement of the two signals is also clear from the sinusoidal behavior of the intensity ratios (Figure 10, bottom). Notice that the $\nu_{C=C}$-peak becomes up to two-fold more intense than the ν_{C-C} peak, for incident wavelengths approaching the second absorption maximum, whereas the ratio equals 0.5 in the vicinities of 182 nm, thus favoring the ν_{C-C} peak. This fact could be explained by resorting to the information provided by NBO, the analysis of the orbitals involved in the transitions leading to that band, and of course to the normal modes whose signals generate those peaks (see Figure 7). As stated above, the $\pi \rightarrow \pi^*$ transition predominantly involves orbitals confined to the aromatic ring or the carboxylate group of ibuprofen. What is more, the peak at 1690 comprises normal modes essentially concentrated in the ring, i.e., $\nu_{C=C}$; therefore it is understandable that even at longer wavelengths (200–260 nm) and most likely at 222 nm, an increase in the intensity of this peak is to be observed.

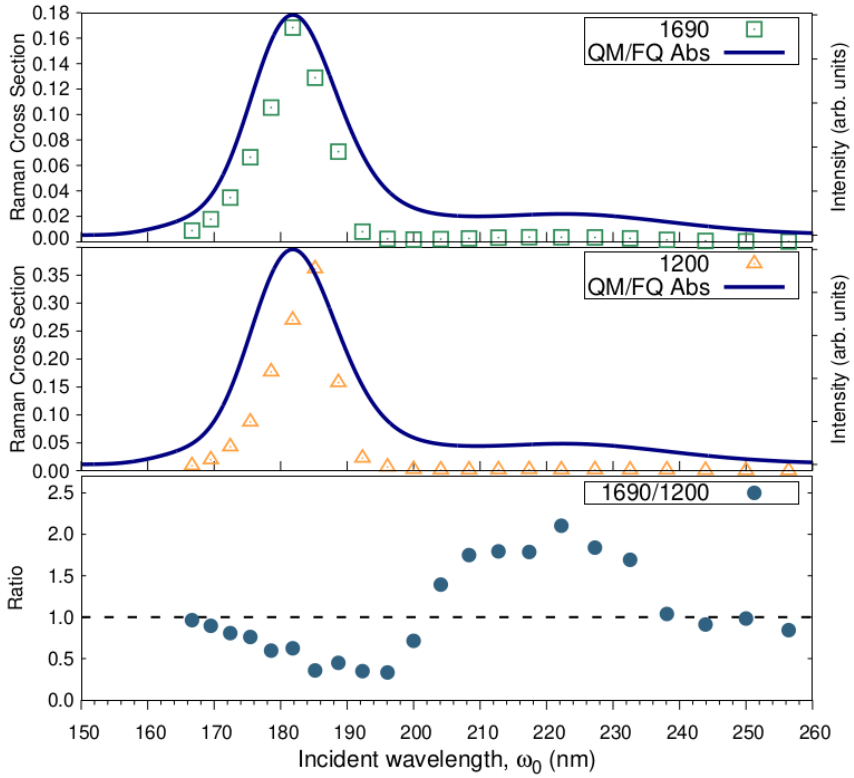


Figure 10. Excitation wavelength dependence of the $\nu_{C=C}$ and ν_{C-C} bands intensity, and of $\nu_{C=C}/\nu_{C-C}$ Raman cross-section ratios. RR intensities (in $cm^2mol^{-1}sr^{-1}$) were calculated with a damping factor of 200 cm^{-1} and broadened using Lorentzian functions with FWHM = 20 cm^{-1}.

Finally, these results suggest that there are at least two signals to unequivocally trace ibuprofen in aqueous solution through UVRR techniques: (*i*) the aromatic ring $\nu_{C=C}$ stretching mode occurring at 1690 cm^{-1} and (*ii*) the aliphatic ν_{C-C} vibrations emerging at 1200 cm^{-1}, which can be selectively enhanced by choosing the proper incident wavelength. We emphasize here that even if devices with tunable excitation sources are not available for shorter wavelengths, the ring vibrations will appear with an increased intensity when

typical 226 nm lasers are used, thus becoming the preferable signals to detect ibuprofen in water.

4. Conclusions

In the present study, the potentialities of resonance Raman to detect ibuprofen in water have been investigated by exploiting the mixed quantum–classical polarizable QM/MM scheme based on fluctuating charges to treat in a refined way the solute–solvent interactions, thus obtaining reliable spectral profiles. We additionally have ensured that the anionic ibuprofen$\cdots$water geometries taken for the study are representative of the overall solvated environment by sampling configurations from classical molecular dynamics simulations. As a general rule, hydration patterns indicated that solvent molecules prefer to bind the carboxylate group in a-Ibu, leading to the formation of charge assisted networks of strong hydrogen bonds as confirmed by the NBO analysis.

To validate the application of the QM/FQ + MD methodology for the target system, we have shown that computed vertical absorption energies display a minimal deviation from experimental results, and on top of this, our calculations also capture most of the features of the experimental Raman spectra of a-Ibu in aqueous solution. The validated approach was then used to simulate resonance Raman spectra. The variation of the intensities as a function of ω_0 revealed that a couple of peaks in the RR spectrum might be selectively enhanced by properly choosing the incident wavelength. With an NBO-guided assignment of the nature of the electronic transitions, we have related the location of the MOs involved in the excited states with the characteristics of the normal modes whose intensities are enhanced in the RR spectrum. For example, we have outlined that as the $\nu_{C=C}$ normal mode has a large contribution of the C–C stretching in the ring and the ν_{C-C} does not, therefore the use of $\omega_0 = 222$ nm (the maximum absorption for the lowest energy band, which in turn is due to a $\pi_{\text{ring}} \rightarrow \pi^*_{\text{ring}}$ transition) would enhance the RR intensity of the $\nu_{C=C}$ normal mode or in general, those normal modes having ring components.

Although the findings in this work provide solid evidence that Resonance Raman can be suggested as a powerful technique to detect ibuprofen in very diluted aqueous solutions, our results would need further experimental verification. Indeed, we are currently in the process of collecting data on that matter, and this will be the subject of a future paper.

Supplementary Materials: Most representative conformers of a-Ibu under clustering method (Figures S1, S2), UV-vis spectra from non-optimized structures (Figure S3), RR spectra at different incident wavelengths (Figure S4), and taking into account just the first three excited states (Figure S5), Convergence tests for UV-vis, Raman and RR spectra (Figures S6, S7, S8)

Author Contributions: Conceptualization, S.G., N.R.-V., T.G. and A.R.; methodology, S.G. and T.G.; validation, S.G. and N.R.-V.; formal analysis, S.G., N.R.-V., T.G., A.R. and C.C.; investigation, S.G., N.R.-V., T.G., C.C. and A.R.; resources, S.G., T.G. and C.C.; writing—original draft preparation, S.G., N.R.-V., T.G., A.R. and C.C.; writing—review and editing, S.G., N.R.-V., T.G., C.C. and A.R.; project administration, C.C. and A.R.; funding acquisition, C.C. and A.R. All authors have read and agreed to the published version of the manuscript.

Funding: Partial funding for this project from H2020-MSCA-ITN-2017 European Training Network "Computational Spectroscopy In Natural sciences and Engineering" (COSINE), grant number 765739 is acknowledged. Internal support from Universidad de Antioquia via "Estrategia para la sostenibilidad" is also acknowledged. We gratefully acknowledge the Center for High Performance Computing (CHPC) at SNS for providing the computational infrastructure.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Most representative conformers of a-Ibu under clustering method, UV-vis spectra from non-optimized structures, RR spectra at different incident wavelengths, and taking into account just the first three excited states, Convergence tests for UV-vis, Raman and RR spectra.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Sample Availability: Not applicable.

Abbreviations

The following abbreviations are used in this manuscript:

(UV)RR	(Ultra–Violet) Resonance Raman
NBO	Natural Bond Orbitals
FQ	Fluctuating Charges
PES	Potential Energy Surface
MO	Molecular Orbital
CMO	Canonical Molecular Orbital
RREP	Resonance Raman Excitation Profile
QM	Quantum Mechanics

References

1. Conaghan, P.G. A turbulent decade for NSAIDs: Update on current concepts of classification, epidemiology, comparative efficacy, and toxicity. *Rheumatol. Int.* **2012**, *32*, 1491–1502. doi:10.1007/s00296-011-2263-6.
2. Research and Market. The World's Largest Market Research Store. Outlook on the Ibuprofen Global Market to 2024—Insights & Forecast with Potential Impact of COVID-19. 2021. Available online: https://www.globenewswire.com/ (accessed on 20 October 2021).
3. Vane, J.; Botting, R. Anti-inflammatory drugs and their mechanism of action. *Inflamm. Res.* **1998**, *47*, 78–87.
4. Goodwin, G. *Prostaglandins: Biochemisty, Functions, Types and Roles*; Nova Science Publisher: New York, NY, USA, 2010.
5. Evans, A.M. Comparative pharmacology of S (+)-ibuprofen and (RS)-ibuprofen. *Clin. Rheumatol.* **2001**, *20*, 9–14.
6. Wood, M.; Morales, M.; Miller, E.; Braziel, S.; Giancaspro, J.; Scollan, P.; Rosario, J.; Gayapa, A.; Krmic, M.; Lee, S. Ibuprofen and the Phosphatidylcholine Bilayer: Membrane Water Permeability in the Presence and Absence of Cholesterol. *Langmuir* **2021**, *37*, 4468–4480.
7. Zapata-Escobar, A.; Manrique-Moreno, M.; Guerra, D.; Hadad, C.Z.; Restrepo, A. A combined experimental and computational study of the molecular interactions between anionic ibuprofen and water. *J. Chem. Phys.* **2014**, *140*, 184312. doi:10.1063/1.4874258.
8. Chopra, S.; Kumar, D. Ibuprofen as an emerging organic contaminant in environment, distribution and remediation. *Heliyon* **2020**, *6*, e04087.
9. Rojas-Valencia, N.; Lans, I.; Manrique-Moreno, M.; Hadad, C.Z.; Restrepo, A. Entropy drives the insertion of ibuprofen into model membranes. *Phys. Chem. Chem. Phys.* **2018**, *20*, 24869–24876. doi:10.1039/C8CP04674C.
10. Rojas-Valencia, N.; Gómez, S.; Montillo, S.; Manrique-Moreno, M.; Cappelli, C.; Hadad, C.; Restrepo, A. Evolution of Bonding during the Insertion of Anionic Ibuprofen into Model Cell Membranes. *J. Phys. Chem. B* **2020**, *124*, 79–90. doi:10.1021/acs.jpcb.9b09705.
11. Baccar, R.; Sarrà, M.; Bouzid, J.; Feki, M.; Blánquez, P. Removal of pharmaceutical compounds by activated carbon prepared from agricultural by-product. *Chem. Eng. J.* **2012**, *211*, 310–317.
12. Petrie, B.; Barden, R.; Kasprzyk-Hordern, B. A review on emerging contaminants in wastewaters and the environment: current knowledge, understudied areas and recommendations for future monitoring. *Water Res.* **2015**, *72*, 3–27.
13. Klavarioti, M.; Mantzavinos, D.; Kassinos, D. Removal of residual pharmaceuticals from aqueous systems by advanced oxidation processes. *Environ. Int.* **2009**, *35*, 402–417.
14. Liu, W.; Zhang, S.; Meng, F.; Tang, L. Molecular simulation of ibuprofen passing across POPC membrane. *J. Theor. Comput. Chem.* **2014**, *13*, 1450033.
15. Hiller, E.; Šebesta, M. Effect of temperature and soil pH on the sorption of ibuprofen in agricultural soil. *Soil Water Res.* **2017**, *12*, 78–85.
16. Bonora, S.; Pisi, A.; Ottani, S.; Cesini, D.; Maris, A.; Foggia, M.D. Raman and SERS study on ibuprofen metal complexes with biomedical interest. *Vib. Spectrosc.* **2014**, *73*, 45–55. doi:10.1016/j.vibspec.2014.04.009.
17. Xiong, P.; Zhang, H.; Li, G.; Liao, C.; Jiang, G. Adsorption removal of ibuprofen and naproxen from aqueous solution with Cu-doped Mil-101(Fe). *Sci. Total. Environ.* **2021**, *797*, 149179. doi:10.1016/j.scitotenv.2021.149179.
18. Xiao, R.; He, Z.; Diaz-Rivera, D.; Pee, G.Y.; Weavers, L.K. Sonochemical degradation of ciprofloxacin and ibuprofen in the presence of matrix organic compounds. *Ultrason. Sonochem.* **2014**, *21*, 428–435. doi:10.1016/j.ultsonch.2013.06.012.
19. Khazri, H.; Ghorbel-Abid, I.; Kalfat, R.; Trabelsi-Ayadi, M. Removal of ibuprofen, naproxen and carbamazepine in aqueous solution onto natural clay: Equilibrium, kinetics, and thermodynamic study. *Appl. Water Sci.* **2017**, *7*, 3031–3040.
20. Rafati, L.; Ehrampoush, M.; Rafati, A.; Mokhtari, M.; Mahvi, A. Removal of ibuprofen from aqueous solution by functionalized strong nano-clay composite adsorbent: Kinetic and equilibrium isotherm studies. *Int. J. Environ. Sci. Technol.* **2018**, *15*, 513–524.

21. Rafati, L.; Ehrampoush, M.H.; Rafati, A.A.; Mokhtari, M.; Mahvi, A.H. Fixed bed adsorption column studies and models for removal of ibuprofen from aqueous solution by strong adsorbent Nano-clay composite. *J. Environ. Health Sci. Eng.* **2019**, *17*, 753–765.
22. Feng, L.; van Hullebusch, E.D.; Rodrigo, M.A.; Esposito, G.; Oturan, M.A. Removal of residual anti-inflammatory and analgesic pharmaceuticals from aqueous systems by electrochemical advanced oxidation processes. A review. *Chem. Eng. J.* **2013**, *228*, 944–964.
23. Boyd, G.R.; Reemtsma, H.; Grimm, D.A.; Mitra, S. Pharmaceuticals and personal care products (PPCPs) in surface and treated waters of Louisiana, USA and Ontario, Canada. *Sci. Total. Environ.* **2003**, *311*, 135–149.
24. Choina, J.; Kosslick, H.; Fischer, C.; Flechsig, G.U.; Frunza, L.; Schulz, A. Photocatalytic decomposition of pharmaceutical ibuprofen pollutions in water over titania catalyst. *Appl. Catal. B Environ.* **2013**, *129*, 589 –598. doi:10.1016/j.apcatb.2012.09.053.
25. Motoc, S.; Manea, F.; Iacob, A.; Martinez-Joaristi, A.; Gascon, J.; Pop, A.; Schoonman, J. Electrochemical Selective and Simultaneous Detection of Diclofenac and Ibuprofen in Aqueous Solution Using HKUST-1 Metal-Organic Framework-Carbon Nanofiber Composite Electrode. *Sensors* **2016**, *16*, 1719. doi:10.3390/s16101719.
26. Manea, F.; Motoc, S.; Pop, A.; Remes, A.; Schoonman, J. Silver-functionalized carbon nanofiber composite electrodes for ibuprofen detection. *Nanoscale Res. Lett.* **2012**, *7*, 331.
27. Amin, S.; Soomro, M.T.; Memon, N.; Solangi, A.R.; Qureshi, T.; Behzad, A.R.; Behzad, A.R. Disposable screen printed graphite electrode for the direct electrochemical determination of ibuprofen in surface water. *Environ. Nanotechnol. Monit. Manag.* **2014**, *1*, 8–13.
28. Motoc, S.; Manea, F.; Pop, A.; Pode, R.; Burtica, G. Determination of ibuprofen in water using Ag-doped zeolite-expanded graphite composite electrode. *Adv. Sci. Eng. Med.* **2011**, *3*, 7–12.
29. Wang, C.H.; Su, H.; Chou, J.H.; Lin, J.Y.; Huang, M.Z.; Lee, C.W.; Shiea, J. Multiple solid phase microextraction combined with ambient mass spectrometry for rapid and sensitive detection of trace chemical compounds in aqueous solution. *Anal. Chim. Acta* **2020**, *1107*, 101–106.
30. Rocha, F.S.; Gomes, A.J.; Lunardi, C.N.; Kaliaguine, S.; Patience, G.S. Experimental methods in chemical engineering: Ultraviolet visible spectroscopy—UV-Vis. *Can. J. Chem. Eng.* **2018**, *96*, 2512–2517.
31. Olaru, M.; Cotofana, C.; Patras, X. Design of novel PVA-based hybrid composites for drug delivery. *Optoelectron. Adv. Mater. Rapid Commun.* **2009**, *3*, 1273–1275.
32. Du, L.; Liu, X.; Huang, W.; Wang, E. A study on the interaction between ibuprofen and bilayer lipid membrane. *Electrochim. Acta* **2006**, *51*, 5754–5760. doi:10.1016/j.electacta.2006.03.009.
33. Ertürk, A.S.; Gürbüz, M.U.; Metin, T. New-generation Jeffamine® D230 core amine, TRIS and carboxylterminated PAMAM dendrimers: Synthesis, characterization and the solubility application for a model NSAID drug Ibuprofen. *Marmara Pharm. J.* **2017**, *21*, 385–399.
34. Qandil, A.M.; Obaidat, A.A.; Ali, M.A.M.; Al-Taani, B.M.; Tashtoush, B.M.; Al-Jbour, N.D.; Al Remawi, M.M.; Al-Sou'od, K.A.; Badwan, A.A. Investigation of the Interactions in Complexes of Low Molecular Weight Chitosan with Ibuprofen. *J. Solut. Chem.* **2009**, *38*, 695–712. doi:10.1007/s10953-009-9405-4.
35. Levin, C.S.; Kundu, J.; Janesko, B.G.; Scuseria, G.E.; Raphael, R.M.; Halas, N.J. Interactions of Ibuprofen with Hybrid Lipid Bilayers Probed by Complementary Surface-Enhanced Vibrational Spectroscopies. *J. Phys. Chem. B* **2008**, *112*, 14168–14175. doi:10.1021/jp804374e.
36. Rossi, B.; Verrocchio, P.; Viliani, G.; Mancini, I.; Guella, G.; Rigo, E.; Scarduelli, G.; Mariotto, G. Vibrational properties of ibuprofen–cyclodextrin inclusion complexes investigated by Raman scattering and numerical simulation. *J. Raman Spectrosc.* **2009**, *40*, 453–458. doi:10.1002/jrs.2150.
37. Ewing, A.V.; Clarke, G.S.; Kazarian, S.G. Attenuated total reflection-Fourier transform infrared spectroscopic imaging of pharmaceuticals in microfluidic devices. *Biomicrofluidics* **2016**, *10*, 024125. doi:10.1063/1.4946867.
38. Matkovic, S.; Valle, G.; Briand, L. Quantitative analysis of ibuprofen in pharmaceutical formulations through FTIR spectroscopy. *Lat. Am. Appl. Res.* **2005**, *35*, 189–195.
39. Jampilek, J.; Dolowy, M.; Pyka-Pajak, A. Estimating limits of detection and quantification of ibuprofen by TLC-densitometry at different chromatographic conditions. *Processes* **2020**, *8*, 919.
40. Khan, N.A.; Khan, S.U.; Ahmed, S.; Farooqi, I.H.; Yousefi, M.; Mohammadi, A.A.; Changani, F. Recent trends in disposal and treatment technologies of emerging-pollutants—A critical review. *TrAC Trends Anal. Chem.* **2020**, *122*, 115744.
41. Long, D.A. Vibrational Resonance Raman Scattering. In *The Raman Effect*; John Wiley & Sons: Hoboken, NJ, USA, 2002; Chapter 7, pp. 221–270. doi:10.1002/0470845767.ch7.
42. Myers Kelley, A. Resonance Raman and Resonance Hyper-Raman Intensities: Structure and Dynamics of Molecular Excited States in Solution. *J. Phys. Chem. A* **2008**, *112*, 11975–11991. doi:10.1021/jp805530y.
43. Myers, A.B. Resonance Raman Intensities and Charge-Transfer Reorganization Energies. *Chem. Rev.* **1996**, *96*, 911–926. doi:10.1021/cr950249c.
44. Efremov, E.V.; Ariese, F.; Gooijer, C. Achievements in resonance Raman spectroscopy: Review of a technique with a distinct analytical chemistry potential. *Anal. Chim. Acta* **2008**, *606*, 119–134. doi:10.1016/j.aca.2007.11.006.

45. Rossi, B.; Bottari, C.; Catalini, S.; D'Amico, F.; Gessini, A.; Masciovecchio, C. Chapter 13—Synchrotron-based ultraviolet resonance Raman scattering for material science. In *Molecular and Laser Spectroscopy*; Gupta, V., Ozaki, Y., Eds.; Elsevier: Amsterdam, The Netherlands, 2020; pp. 447–482. doi:10.1016/B978-0-12-818870-5.00013-7.
46. Catalini, S.; Rossi, B.; Foggi, P.; Masciovecchio, C.; Bruni, F. Aqueous solvation of glutathione probed by UV resonance Raman spectroscopy. *J. Mol. Liq.* **2019**, *283*, 537–547. doi:10.1016/j.molliq.2019.03.113.
47. Hashimoto, S.; Yabusaki, T.; Takeuchi, H.; Harada, I. Structure and ligand-binding modes of human serum albumin studied by UV resonance raman spectroscopy. *Biospectroscopy* **1995**, *1*, 375–385. doi:10.1002/bspy.350010603.
48. Marvin. *Calculator Plugins Were Used for Structure Property Prediction and Calculation, Marvin 19.20*; ChemAxon: Budapest, Hungary, 2019. Available online: http://www.chemaxon.com (accessed on 31 October 2021).
49. Odjadjare, E.E.; Okoh, A.I. Physicochemical quality of an urban municipal wastewater effluent and its impact on the receiving environment. *Environ. Monit. Assess.* **2010**, *170*, 383–394.
50. Majumder, A.; Gupta, A.K.; Ghosal, P.S.; Varma, M. A review on hospital wastewater treatment: A special emphasis on occurrence and removal of pharmaceutically active compounds, resistant microorganisms, and SARS-CoV-2. *J. Environ. Chem. Eng.* **2020**, *9*, 104812.
51. Yanai, T.; Tew, D.P.; Handy, N.C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57. doi:10.1016/j.cplett.2004.06.011.
52. Krishnan, R.; Binkley, J.S.; Seeger, R.; Pople, J.A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650. doi:10.1063/1.438955.
53. Clark, T.; Chandrasekhar, J.; Spitznagel, G.W.; Schleyer, P.V.R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li–F. *J. Comput. Chem.* **1983**, *4*, 294–301. doi:10.1002/jcc.540040303.
54. Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: A theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664–5677. doi:10.1039/C9CS00464E.
55. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. doi:10.1002/jcc.20035.
56. Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954–9960.
57. Marenich, A.V.; Jerome, S.V.; Cramer, C.J.; Truhlar, D.G. Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8*, 527–541. doi:10.1021/ct200866d.
58. Abrahama, M.J.; Murtola, T.; Schulz, R.; Pálla, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
59. Berendsen, H.J.; Van Gunsteren, W.F. Practical algorithms for dynamic simulations. *Mol.-Dyn. Simul. Stat.-Mech. Syst.* **1986**, Enrico Fermi Summer School, Varenna, Italy, 43–65.
60. Parrinello, M.; Rahman, A. Strain fluctuations and elastic constants. *J. Chem. Phys.* **1982**, *76*, 2662–2666. doi:10.1063/1.443248.
61. Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W.F.; Mark, A.E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem. Int. Ed.* **1999**, *38*, 236–240. doi:10.1002/(SICI)1521-3773(19990115)38:1/2<236::AID-ANIE236>3.0.CO;2-M.
62. Brehm, M.; Kirchner, B. TRAVIS - A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories. *J. Chem. Inf. Model.* **2011**, *51*, 2007–2023. doi:10.1021/ci200217w.
63. Brehm, M.; Thomas, M.; Gehrke, S.; Kirchner, B. TRAVIS—A free analyzer for trajectories from molecular simulation. *J. Chem. Phys.* **2020**, *152*, 164105. doi:10.1063/5.0005078.
64. Weinhold, F. Natural bond critical point analysis: Quantitative relationships between natural bond orbital-based and QTAIM-based topological descriptors of chemical bonding. *J. Comput. Chem.* **2012**, *33*, 2440–2449.
65. Weinhold, F.; Landis, C.R. *Discovering Chemistry with Natural Bond Orbitals*; Wiley-VCH: Hoboken, NJ, USA, 2012; 319p.
66. Weinhold, F.; Landis, C.; Glendening, E. What is NBO analysis and how is it useful? *Int. Rev. Phys. Chem.* **2016**, *35*, 399–440. doi:10.1080/0144235X.2016.1192262.
67. Glendening, E.D.; Badenhoop, J.K.; Reed, A.E.; Carpenter, J.E.; Bohmann, J.A.; Morales, C.M.; Karafiloglou, P.; Landis, C.R.; Weinhold, F. *NBO 7.0*; Theoretical Chemistry Institute, University of Wisconsin: Madison, WI, USA, 2018.
68. Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. Effective yet reliable computation of hyperfine coupling constants in solution by a QM/MM approach: Interplay between electrostatics and non-electrostatic effects. *J. Chem. Phys.* **2019**, *150*, 124102. doi:10.1063/1.5080810.
69. Giovannini, T.; Del Frate, G.; Lafiosca, P.; Cappelli, C. Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution. *Phys. Chem. Chem. Phys.* **2018**, *20*, 9181–9197. doi:10.1039/C8CP00487K.
70. Gómez, S.; Egidi, F.; Puglisi, A.; Giovannini, T.; Rossi, B.; Cappelli, C. Unlocking the Power of Resonance Raman Spectroscopy: The Case of Amides in Aqueous Solution. *J. Mol. Liq.* **2021**, 117841. doi:10.1016/j.molliq.2021.117841.
71. Santoro, F.; Cappelli, C.; Barone, V. Effective Time-Independent Calculations of Vibrational Resonance Raman Spectra of Isolated and Solvated Molecules Including Duschinsky and Herzberg—Teller Effects. *J. Chem. Theory Comput.* **2011**, *7*, 1824–1839. doi:10.1021/ct200054w.
72. Olszówka, M.; Russo, R.; Mancini, G.; Cappelli, C. A computational approach to the resonance Raman spectrum of doxorubicin in aqueous solution. *Theor. Chem. Accounts* **2016**, *135*, 27.

73. Egidi, F.; Bloino, J.; Cappelli, C.; Barone, V. A Robust and Effective Time-Independent Route to the Calculation of Resonance Raman Spectra of Large Molecules in Condensed Phases with the Inclusion of Duschinsky, Herzberg–Teller, Anharmonic, and Environmental Effects. *J. Chem. Theory Comput.* **2014**, *10*, 346–363. doi:10.1021/ct400932e.
74. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16 Revision B.01*; Gaussian Inc.: Wallingford, CT, USA, 2016.
75. Reed, A.E.; Curtiss, L.A.; Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **1988**, *88*, 899–926.
76. Gómez, S.; Nafziger, J.; Restrepo, A.; Wasserman, A. Partition-DFT on the water dimer. *J. Chem. Phys.* **2017**, *146*, 074106.
77. Gómez, S.; Rojas-Valencia, N.; Gómez, S.A.; Cappelli, C.; Merino, G.; Restrepo, A. A molecular twist on hydrophobicity. *Chem. Sci.* **2021**, *12*, 9233–9245. doi:10.1039/D1SC02673A.
78. Rojas-Valencia, N.; Gómez, S.; Núñez-Zarur, F.; Cappelli, C.; Hadad, C.; Restrepo, A. Thermodynamics and Intermolecular Interactions during the Insertion of Anionic Naproxen into Model Cell Membranes. *J. Phys. Chem. B* **2021**, *125*, 10383–10391. doi:10.1021/acs.jpcb.1c06766.
79. Gomez, S.A.; Rojas-Valencia, N.; Gomez, S.; Egidi, F.; Cappelli, C.; Restrepo, A. Binding of SARS-CoV-2 to cell receptors: A tale of molecular evolution. *ChemBioChem* **2021**, *22*, 724–732. doi:10.1002/cbic.202000618.
80. Gomez, S.A.; Rojas-Valencia, N.; Gomez, S.; Cappelli, C.; Restrepo, A. The role of spike protein mutations in the infectious power of SARS-COV-2 variants: A molecular interaction perspective. *ChemBioChem* **2021**, *22*. doi:10.1002/cbic.202100393.
81. Frade, T.; Gomes, A. *ZnO Nanostructured Thin Films Applied on Ibuprofen Photoelectrodegradation*; Solid Compounds of Transition Elements II; Trans Tech Publications: Zurich, Switzerland, 2013; Volume 194, pp. 258–261. doi:10.4028/www.scientific.net/SSP.194.258.
82. Tian, H.; Fan, Y.; Zhao, Y.; Liu, L. Elimination of ibuprofen and its relative photo-induced toxicity by mesoporous BiOBr under simulated solar light irradiation. *RSC Adv.* **2014**, *4*, 13061–13070. doi:10.1039/C3RA47304J.
83. Vargek, M. *Instrumentation Advances and New Applications in Raman Optical Activity*; ProQuest LLC., Ann Arbor, MI, USA, 1997.
84. Vargek, M.; Freedman, T.B.; Nafie, L.A. Resonance and Non-Resonance Raman Optical Activity in Pharmaceutical Analgesic Molecules. In *Spectroscopy of Biological Molecules: Modern Trends*; Carmona, P., Navarro, R., Hernanz, A., Eds.; Springer: Dordrecht, The Netherlands, 1997; pp. 459–460. doi:10.1007/978-94-011-5622-6_206.
85. Vargek, M.; Freedman, T.B.; Lee, E.; Nafie, L.A. Experimental observation of resonance Raman optical activity. *Chem. Phys. Lett.* **1998**, *287*, 359–364. doi:10.1016/S0009-2614(98)00017-7.
86. Guthmuller, J.; Champagne, B. Time dependent density functional theory investigation of the resonance Raman properties of the julolidinemalononitrile push-pull chromophore in various solvents. *J. Chem. Phys.* **2007**, *127*, 164507. doi:10.1063/1.2790907.
87. Xiong, K. UV Resonance Raman Spectroscopy: A Highly Sensitive, Selective and Fast Technique for Environmental Analysis. *J. Environ. Anal. Chem.* **2014**, *2*, e107. doi:10.4172/2380-2391.1000e107.
88. Guthmuller, J. Calculation of Vibrational Resonance Raman Spectra of Molecules Using Quantum Chemistry Methods. In *Molecular Spectroscopy: A Quantum Chemistry Approach*; Ozaki, Y., Wójcik, M.J., Popp, J., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2019. doi:10.1002/9783527814596.ch17.
89. Bottari, C.; Catalini, S.; Foggi, P.; Mancini, I.; Mele, A.; Perinelli, D.R.; Paciaroni, A.; Gessini, A.; Masciovecchio, C.; Rossi, B. Base-specific pre-melting and melting transitions of DNA in presence of ionic liquids probed by synchrotron-based UV resonance Raman scattering. *J. Mol. Liq.* **2021**, *330*, 115433. doi:10.1016/j.molliq.2021.115433.
90. Bottari, C.; Mancini, I.; Mele, A.; Gessini, A.; Masciovecchio, C.; Rossi, B. Conformational stability of DNA in hydrated ionic liquid by synchrotron-based UV resonance raman. In *UV and Higher Energy Photonics: From Materials to Applications 2019*; Lérondel, G., Cho, Y.H., Taguchi, A., Kawata, S., Eds.; International Society for Optics and Photonics, SPIE: Bellingham, WA, USA, 2019; Volume 11086, pp. 33–40. doi:10.1117/12.2529077.
91. Rossi, B.; Catalini, S.; Bottari, C.; Gessini, A.; Masciovecchio, C. Frontiers of UV resonant raman spectroscopy by using synchrotron radiation: The case of aqueous solvation of model peptides. In *UV and Higher Energy Photonics: From Materials to Applications 2019*; Lérondel, G., Cho, Y.H., Taguchi, A., Kawata, S., Eds.; International Society for Optics and Photonics, SPIE: Bellingham, WA, USA, 2019; Volume 11086, pp. 23–32. doi:10.1117/12.2529172.

4.3. Prototypical systems for the peptide bond

4.3.1. Small Amides: Acetamide, NMA, and DMA[4]

Journal of Molecular Liquids 346 (2022) 117841

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Unlocking the power of resonance Raman spectroscopy: The case of amides in aqueous solution

Sara Gómez [a], Franco Egidi [a], Alessandra Puglisi [a], Tommaso Giovannini [a], Barbara Rossi [b,c], Chiara Cappelli [a,*]

[a] Scuola Normale Superiore, Classe di Scienze, Piazza dei Cavalieri 7, 56126 Pisa, Italy
[b] Elettra Sincrotrone Trieste S.C.p.A., S. S. 14 Km 163.5 in Area Science Park, I-34149 Trieste, Italy
[c] Department of Physics, University of Trento, via Sommarive 14, I-38123 Povo, Trento, Italy

ARTICLE INFO

Article history:
Received 6 August 2021
Accepted 12 October 2021
Available online 29 October 2021

Keywords:
UV Resonance Raman
Fluctuating charges
Amides
NMA
QM/MM

ABSTRACT

We report a joined experimental and computational study of Raman and Resonance Raman spectra of amides in aqueous solution. By employing state-of-the-art QM/MM methods combined with synchrotron–based UV Resonance Raman spectroscopy, we propose a protocol to interpret and reliably predict Resonance Raman spectra for amide systems in water, which are prototypical system for the peptide bond. We demonstrate that the main experimental spectral features can be correctly reproduced by simultaneously taking into account the dynamical aspects of the solvation phenomenon, specific solute–solvent hydrogen bond interactions and mutual solute–solvent polarization effects.

1. Introduction

Resonance Raman (RR) spectroscopy is a powerful tool to unveil structural and electronic properties of systems under different conditions. Much molecular information can be obtained from RR measurements, therefore numerous applications encompassing fields like analytical, physical and biophysical chemistry have been proposed [1–4]. Among the most experimentally studied systems using RR spectroscopy, are proteins [5–8], peptides [9–11], model peptides [12–19], aminoacids [20–22], (DNA) nucleobases [23–27] and a variety of inorganic compounds [28,29].

RR is a mixed electronic and vibrational spectroscopy, because it probes a system's vibrational degrees of freedom by employing an electromagnetic impulse which is in resonance with an electronic transition. Therefore, the approach combines the advantages of both types of techniques, i.e. the spectra can be directly connected to the vibrational degrees of freedom and thus structural aspects, while the ability to tune the spectra to a specific electronic transition can focus the signal on a single chromophore/portion of the system. Also, the ability to vary the wavelength of the probing laser permits to investigate the effect of the electronic transition upon the spectrum, enriching the description. Therefore, RR and especially UVRR, that provides selectivity and sensitivity through enhancement of particular vibrations associated to specific chromophores, has a very large potential [30–33]. However its full exploitation is hindered by the need for special experimental setups which are not as readily available as in the case of more common techniques. At the molecular level, through the resonance enhancement some selectivity is reached because the vibrational modes observed are only those whose motions couple to the electronic density change taking place in the electronic transition [34,35].

Evidently, the full exploitation of UVRR spectroscopy to selectively analyze vibrations in the system is contingent upon the experimental need of sources with appropriate characteristics of intensity, tunability and wavelength range extension. This is particularly critical in the case of deep ultraviolet (DUV) range of excitation (between about 150–300 nm), where lasers, which are widely used for visible Raman spectroscopy [8,36], suffer of some important limitations. The use of synchrotron radiation (SR) as source for UVRR experiments [37] appears to be an excellent choice because it offers advantages with respect to conventional laser sources. For instance, the possibility of extending the UV domain of excitation above 7 eV that would provide the chance to cover the whole range of outer electronic transitions in matter. In most laser-based UVRR studies, only a few discrete excitation wavelengths are used. The continuous tunability of SR enables a fine mapping of the whole resonance landscape of the sample in order to achieve a fine matching between the exciting radiation energy and the resonance conditions of specific chromophores. This allows to perform, for example, accurate UVRR measurements

* Corresponding author.
E-mail address: chiara.cappelli@sns.it (C. Cappelli).

https://doi.org/10.1016/j.molliq.2021.117841
0167-7322/

not biased by self-absorption effects and/or to also detect the pre-resonance Raman scattered signal. Recently, the frontiers of UVRR spectroscopy were investigated by using synchrotron radiation, especially for the case study of investigation of peptides dissolved in their natural environment, i.e. aqueous solutions [18].

Due to the complexity of information that is hidden behind RR spectral patterns (especially for aqueous systems), their interpretation benefits from the coupling with reliable theoretical simulations. Such calculations are generally doable for isolated systems, or systems in solution described by means of the Polarizable Continuum Model (PCM) [38–41]. However, in general, substantial variations of the positions and relative intensities of the bands are observed after comparison with experimental spectra [38,40,42], due to lack of any description of hydrogen bonding (HB) interactions. To overcome this limitation, researchers commonly resort to the so-called "cluster" methodologies, where explicit solvent molecules are included surrounding the potential hydrogen bond sites on the solute; and the entire system is then treated with Quantum Mechanical (QM) descriptions [43,44]. Nevertheless, this procedure has its own weakness and not in all cases captures missing features.

As an alternative, a widespread practice is to split the entire system in layers to be treated at different model chemistries, with the part responsible of the property examined at the QM level and the rest computed at a much less expensive level of approximation, such as Molecular Mechanics (MM) [45,46].

For the most part, QM/MM calculations of Raman spectra are oriented to the off-resonance (spontaneous) regime [46], whereas a few QM/MM applications to Resonance Raman spectra are available in literature [47,48]. It is important to note that an atomistic description of the solvation shell only affects the spectroscopic properties of the solute directly, only if it is polarizable, and is therefore able to dynamically respond to the probing electromagnetic field [49]. Therefore, in order to properly describe the physico-chemical nature of the solute–solvent interaction and how it affects the final spectrum, we resort to a polarizable QM/MM method based on fluctuating charges (FQ) known as QM/FQ [50,51], which has become a powerful tool to model a large variety of spectral signals of aqueous solutions [49].

The aim of this work is to offer a detailed view of the effect of the solvation description on the simulation of Resonance Raman spectra of model peptides, by combining state-of-the-art theoretical techniques with an experimental setup based on SR radiation that allows us to fully exploit the potential of this technique. To this end, we extend QM/FQ for the first time to model RR, and we challenge it to reproduce experimental UVRR spectra of aqueous Acetamide (ACA), N-methyl acetamide (NMA), N,N-dimethyl acetamide (DMA) (see Fig. 1). Such systems were not chosen by chance, rather they serve as models for the extensively documented peptide bond [30].

2. Methodology

2.1. QM/FQ approach to RR spectroscopy

The spontaneous Raman scattering cross-section is usually calculated at the DFT level using response theory by differentiating the dynamic electric polarizability with respect to the normal mode displacements, calculated for a perturbation with angular frequency ω corresponding to the one of the light source (e.g. a laser or synchrotron light). Given the vibrational transition polarizability α_i corresponding to an excitation of the i-th normal mode, then the cross-section σ_i can be expressed in terms of the Raman rotational invariants:

$$a_i^2 = \frac{1}{9}\sum_{ab}\alpha^*_{aa,i}\alpha_{bb,i} = \frac{1}{9}\mid \alpha_{xx,i} + \alpha_{yy,i} + \alpha_{zz,i}\mid^2 \tag{1}$$

$$g_i^2 = \frac{1}{2}\sum_{ab}(3\alpha^*_{ab,i}\alpha_{ab,i} - \alpha^*_{aa,i}\alpha_{bb,i}) \tag{2}$$

$$\sigma_i = \left(\frac{\omega-\omega_i}{c}\right)^4 \frac{45a_i^2 + 7g_i^2}{45} \tag{3}$$

The harmonic approximation is usually invoked to describe the molecule's potential energy surface, and in the non-resonant regime the vibrational transition polarizability is expanded in a Taylor series to first order, and only the first derivative term is usually retained which is calculated either numerically or analytically. This method assumes that the imaginary part of the polarizability is negligible, which is true when the incident radiation is far-from-resonance, however in the resonant case the full sum-over-state expression must be considered

$$\alpha_{ab,i} = \frac{1}{\hbar}\sum_{m'}\frac{\langle i \mid \mu_a \mid m'\rangle\langle m' \mid \mu_b \mid 0\rangle}{\omega_{m'} - \omega_i - \omega - i\gamma} \tag{4}$$

where the summation runs over all vibronic states belonging to the potential energy surface of the resonant electronic state, while γ is the excited state's phenomenological damping constant. This is the method used in this work to calculate the RR cross section, as detailed in Ref. [38]. In particular, we employ the Vertical-Gradient, Franck–Condon approximation, where the vibrational frequencies and normal modes of the excited state are assumed to be the same as the ground state, and the transition dipole moments are considered to be independent of the molecular geometry [38]. The method employed to calculate the RR spectrum is considerably more involved compared to the one for spontaneous Raman, however it should be emphasized that attempting to simulate the Raman spectrum using the same methodology employed in the non-resonant case by simply altering the incident frequency so that it is close to that of the electronic transition would lead to completely erroneous results, unless a method that explicitly includes the imaginary part of polarizability is used [52].

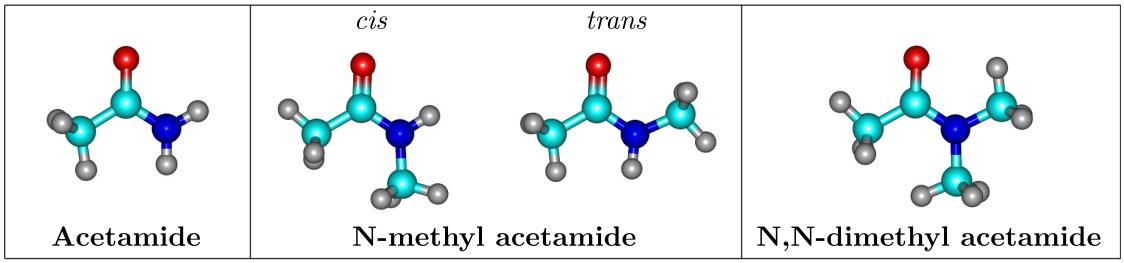


Fig. 1. Structure of amides studied in this work. For NMA both *cis* and *trans* conformers are depicted.

The method employed in this work is also known as time-independent [38,40] (TI) method, though an equivalent time-dependent [42,53,54] (TD) formulation has also been described. Calculations of vibrational Resonance Raman spectra of isolated and solvated molecules have been performed by resorting to these frameworks, combining different strategies for the description of the excited states [55–57,24,25,58,31,48], though none have used a polarizable QM/MM model to include solvent effects in all terms within Eq. 4, to the best of our knowledge. The presence of the solvent must be carefully considered and included in all steps of the simulation.

In this paper, solvent effects are described by means of the FQ force field, in which each atom of the classical layer is endowed with a charge, whose value is not fixed, but is allowed to vary as a response to the electric potential produced by the QM density [49]. In recent years, the QM/FQ method has been extended to the calculation of analytical energy third derivatives, which allow for the calculation of spontaneous Raman spectra [59], excitation energies [60], and excited state gradients [61]. In Eq. 4 all terms include solvation effects evaluated at the QM/FQ level of theory: the ground state geometry of the molecule is first optimized in the presence of the solvent shell, then the electronic density and harmonic potential energy surface are modelled by taking the reaction field due to the water molecules into account. The contribution due to the water molecules also enters the response equations that are solved to calculate excitation energies and to model the excited state potential energy surface. Therefore, in order to properly simulate the RR spectrum of a system in solution, one requires a solvation method with the flexibility to model a wide array of molecular properties, involving both electronic and vibrational degrees of freedom, as well as excited states. In the last few years, the QM/FQ method has indeed been extended to the treatment of a vast set of spectroscopies, allowing us to finally tackle RR, which can be regarded as one of the most complex due to the interplay of all these features [49–51].

3. Experimental procedure

Acetamide, N-methylacetamide and N,N-dimethyl acetamide were purchased by Sigma Aldrich and D.B.A. Italia and used without further purification. The aqueous solutions of the three molecules were prepared by dissolving ACA, NMA and DMA in high-purity water, deionized through a MilliQTM water system (>18 M cm resistivity), in order to obtain the desired molar fractions x (where x is defined as mole of solute/total number of moles of the solution) ranging from 0.2 to 0.01. It has been accurately checked that, at these values of concentration, the solutes are totally dissolved and the solutions appear limpid. All the samples were freshly prepared and placed into optical quartz cells for the Raman scattering measurements.

Out of Resonance Raman spectra were collected on the solutions of ACA, NMA and DMA by means of a micro-Raman setup (Horiba-JobinYvon, LabRam Aramis) in backscattering geometry and using the exciting radiation at 632.8 nm provided by a He-Ne laser. The resolution was set at about 1.2 cm^{-1}/pixel.

UV Resonance Raman (UVRR) measurements were collected at the BL10.2-IUVS beamline of Elettra-Sincrotrone Trieste (Italy) [37] using 210, 226 and 266 nm as excitation wavelengths. The exciting wavelength was set by adjusting the gap parameters of the undulator and by using a Czerny-Turner monochromator (Acton SP2750, Princeton Instruments) equipped with 1800 and 3600 grooves/mm gratings to monochromatize the incoming SR. The final radiation power on the samples was kept about 10–15 μW. The Raman scattered radiation was collected in back-scattered geometry and analyzed by using a single pass of a Czerny-Turner spectrometer (Trivista 557, Princeton Instruments). Depending on the excitation wavelength, the resolution was set between 1.8 and 2.8 cm^{-1}/pixel, in order to ensure enough resolving power and count-rate of the spectra. The calibration of the spectrometer was standardized using cyclohexane (spectroscopic grade, Sigma Aldrich). Any possible photo-damage effect due to a prolonged exposure of the samples to UV radiation was avoided by continuously spinning the sample cell during the measurements. The comparison between the individual spectra acquired for each sample evidences that no gradual changes to the spectra with respect to accumulation number were observed, confirming that any sample photodegradation due to UV exposure is not occurred in the experiments.

4. Computational details

All QM calculations were carried out at the B3LYP/aug-cc-pVDZ level of theory, using a locally modified version of the GAUSSIAN16 suite of programs [62]. Equilibrium geometries and ten vertical excitation energies of the three amides in water were obtained by using PCM to treat environmental effects [41]. For the three amides, CM5 point charges [63] were also calculated with the aim of using them in the Molecular Dynamics (MD) simulations.

Classical MDs were performed by using GROMACS 5.0.5 [64]. A single amide molecule (the most populated PCM conformer in the case of NMA, see Fig. S1 in the Supplementary Material - SM) was inserted in a cubic box with an edge of 4.52 nm and solvated with around 3000 water molecules. In order to recover the correct directionality of solute–solvent hydrogen bonds, dummy atoms [65,66] (virtual sites) were placed on the oxygen atom of the carbonyl group in each case [65–69], specifically at the centroid positions, determined by the Boys localization procedure [70]. Bonding and non-bonding interactions were modeled according to the General Amber Force Field (GAFF) [71]. After minimizing the energy of the solvated systems, a short (500 ps) simulation was performed at 298.15 K for thermalization purposes by adopting the canonical ensemble (NVT) with a velocity-rescale thermostat, and periodic boundary conditions applied in all directions. Production runs were performed in the isothermal–isobaric ensemble (NPT) by using the velocity-rescale method [72] with a coupling constant of 0.1 ps and a Berendsen barostat with time constant of 1.0 ps, saving coordinates every 10 ps. The total time of the simulation was set to 30 ns, with a time step of 2 fs. The MD trajectories were analyzed with the TRAVIS package [73].

From the last 10 ns of the MD runs, 200 uncorrelated snapshots were extracted to be used in QM/FQ calculations, and for each of them, a sphere-shape of radius 15 Å centered on the solute molecule was cut. For each snapshot, the solute geometry was optimized at the QM/FQ level by keeping fixed the solvent molecules, by using the Berny algorithm .[74]. On the optimized geometries, frequencies, excitation energies and excited state gradients were calculated as well. In all QM/FQ calculations, the FQ parametrization proposed in Ref. [75] was exploited. Note that the selected number of snapshots is sufficient to reach the convergence of the calculated spectrum (see Fig. S2 in the SM).

QM/PCM and QM/FQ spontaneous Raman spectra were computed in the dynamic regime by setting the incident frequency (ω_0) to match the experimental value of 633 nm, using analytical response theory as implemented for QM/FQ [59]. UVRR spectra were computed using a time-independent sum-over-state methodology [38]. In particular, we employed the Vertical Gradient Franck–Condon (VG|FC) approximation for the modeling of the excited-state Potential Energy Surface (PES) and transition dipole moments, by considering ten excited states (see also Fig. S3 in the SM). RR spectra were computed using an array of inci-

dent frequencies, producing Raman excitation profiles for all bands. Final QM/FQ UV/Vis and Raman, both spontaneous and RR, spectra were obtained by averaging the spectra from all the snapshots. For absorption spectra, Gaussian functions and a full width at half maximum (FWHM) of 0.66 eV were chosen, while spontaneous and RR sticks were convoluted with Lorentzian profiles and FWHM values of 8 and 20 cm^{-1}, respectively.

5. Results

Spectral features are often rationalized in terms of chromophores and functional groups. From this point of view the three amides considered in this work may appear very similar, however they present important structural and chemical differences that affect their intrinsic properties as well as their interaction with the aqueous environment. All three share the amide functional group characterizing their chemistry and can act as hydrogen bond acceptor through the non-bonding electron pairs of the carbonyl group (see Fig. 1). However they differ in the number of hydrogen atoms bonded to the amide nitrogen (those to be potentially donated to the water molecules), which changes the way they interact with the solvent, and this may strongly affect the resulting spectrum. In addition, NMA differs from the other two amides by virtue of *cis/trans* conformational freedom connected to the rotation of the NC bond, combined with the presence of two different moieties bonded to the nitrogen (see Figs. 1 and S1 in the SM). In this section, we first discuss the results obtained for trans NMA (i.e. the most stable conformer), followed by the other two systems. All spectral features are commented in terms of hydration patterns, as obtained from MD simulations. Notice that, because NMA is the simplest molecular model of the peptide linkage in proteins, it has motivated many experimental and theoretical studies [16,76] with particular emphasis on solvent effects on RR spectra [14,44,43,77].

5.1. NMA

5.1.1. Spontaneous Raman spectra

Computed QM/PCM and QM/FQ spontaneous Raman spectra are shown in Fig. 2(a) along with our own experimental Raman measurements. In Fig. 2(a), top panel, almost all QM/PCM bands for aqueous NMA are slightly narrower than QM/FQ (middle panel). This is due to the dispersion in the vibrational energies within the set of extracted snapshots (see Fig. S4 in the SM for QM/FQ raw data), and is a feature that is missing in a static model such as QM/PCM, which is based on calculations performed on a single minimum-energy structure.

The most relevant vibrational modes in the experimental NMA Raman spectra are amide I (1626–1646 cm^{-1}), amide II (1566–1584 cm^{-1}) and amide III (1313 cm^{-1}) [16,17], which are all correctly reproduced by QM/FQ, which improves the description provided by QM/PCM, particularly in the regions where the strong amide bands are located (see Section S1.4 in the SM for a graphical depiction of the normal modes).

5.1.2. UV Resonance Raman spectra

UVRR spectra of aqueous NMA were measured by using 210, 226, and 266 nm as excitation wavelengths (ω_0), though the 266 nm-excited spectrum should be considered as pre-resonant because the probing wavelength is far from the maximum of the absorption spectrum (see Sec. S1.5 in the SM). The experimental UVRR spectrum collected by setting ω_0 = 226 nm is graphically depicted in Fig. 2(b) (see Fig. S7 in the SM for UVRR spectra measured at different excitation wavelengths).

The value for the incident frequency to be exploited in both QM/PCM and QM/FQ calculations must perfectly reproduce the experimental conditions in order to preserve the resonance enhancement. Selecting the same ω_0, used for experimental measurements, would introduce systematic errors into the calcula-

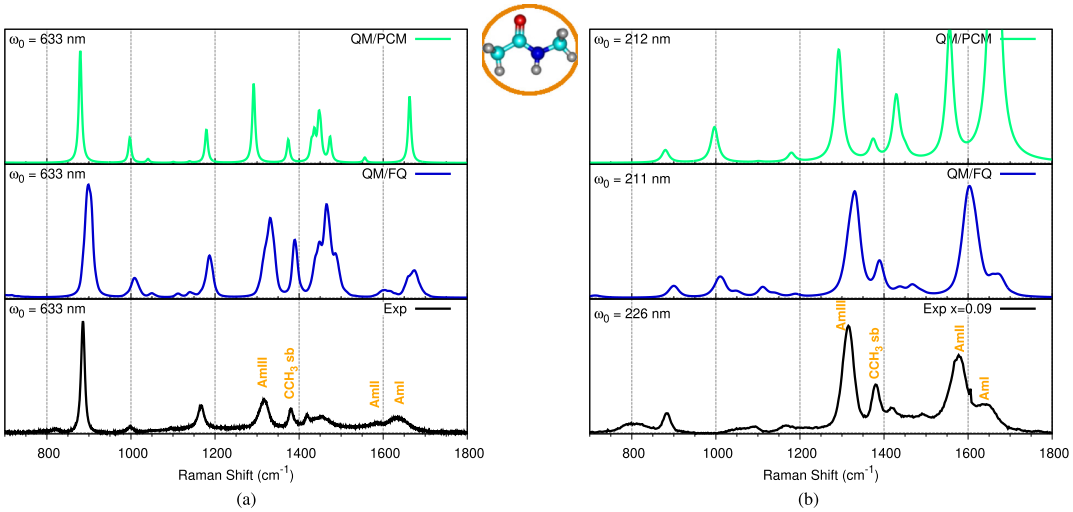


Fig. 2. Raman spectra of **NMA** calculated with different solvent descriptions and compared to experimental spectra measured in aqueous solution at room temperature. (a) Spontaneous (Far From Resonance) Raman spectra simulated and measured using 633 nm as excitation wavelength. FWHM: 8 cm^{-1}. (b) UV Resonance Raman spectra. In the UVRR experiments, the external excitation wavelengths have been set to 226 nm. x = 0.09 is the NMA molar fraction in water. RR intensities were calculated with a damping factor of 200 cm^{-1} and broadened using Lorentzian functions with FWHM = 20 cm^{-1}. The RR band at around 800 cm^{-1} is a spurious signal arising from the quartz cuvette.

tion because any electronic structure method carries some error in the description of electronic excited states which, within the frame of response theory, are obtained from the poles of the system's response function. Therefore, we follow the method described in [38], suitably adapted to QM/FQ calculations. In essence, we select a frequency that is at the same energy gap to the simulated absorption maximum as that observed experimentally (see Tab. S1 in the SM). QM/PCM and QM/FQ UVRR spectra are reported in Fig. 2(b).

Comparing experimental UVRR and spontaneous Raman spectra in Fig. 2, different patterns can be highlighted: (*i*) a weaker or absent Amide I (C═O) stretching vibration (1626–1646 cm^{-1}), (*ii*) an enhancement of the signal involving C—N stretching, which is characteristic of Amide II (1566–1584 cm^{-1}), as well as the CH_3 umbrella bending (1380 cm^{-1}), (*iii*) an enhancement of the Amide III band (1313 cm^{-1}), gaining intensity relative to Amide I [16,76,14,78,44]. Such spectral features are perfectly reproduced by QM/FQ, which is for instance able to correctly predict smaller Amide I:Amide II intensity ratio when moving from off-resonant to resonant regime. Moreover, the atomistic approach outperforms QM/PCM in the description of most experimental bands. In fact, Amide I band is predicted as the strongest band in QM/PCM UVRR spectrum, and also, an overestimated enhancement for the methyl modes (peak erroneously emerging at 1440 cm^{-1}) is observed. The discrepancies between the QM/PCM and the experiment, together with the almost perfect agreement provided by QM/FQ, clearly demonstrate the importance of specific solute–solvent interactions, appropriately coupled to a physically consistent treatment of polarization effects.

An aforementioned feature of the synchrotron-based UVRR setup is the ability to tune the excitation wavelength to obtain RR excitation profiles (RREP) for every band. In Fig. 3(b), computed QM/FQ RREP for NMA in aqueous solution are reported as 3D graphs where the Raman intensity is plotted as a function of both the incident wavelength (150 nm$< \omega_0 <$ 250 nm) and the Raman shifts. Sections of the 3D plots are shown in Fig. 3(b), where we keep track of the enhanced Amide I, II and III peaks as well as the peak at 1380 cm^{-1} (CCH_3 symmetric bend - sb), which have been previously described. As expected, Fig. 3 reveals that all the analyzed band excitation profiles present maxima around 178 nm, which corresponds to the vertical absorption of the $\pi \rightarrow \pi^*$ transition. We also note that Amide I/Amide II relative Raman cross-section ratios vary without any apparent trend in the studied wavelength range. However, amide I band is always weaker in intensity, in total contrast with QM/PCM results in Fig. 2(b). Such findings also reveal that QM/FQ is notably able to account for the almost threefold enhancement of Amide II with respect to Amide I intensity that is experimentally measured[76].

5.2. Acetamide

5.2.1. Spontaneous Raman spectra

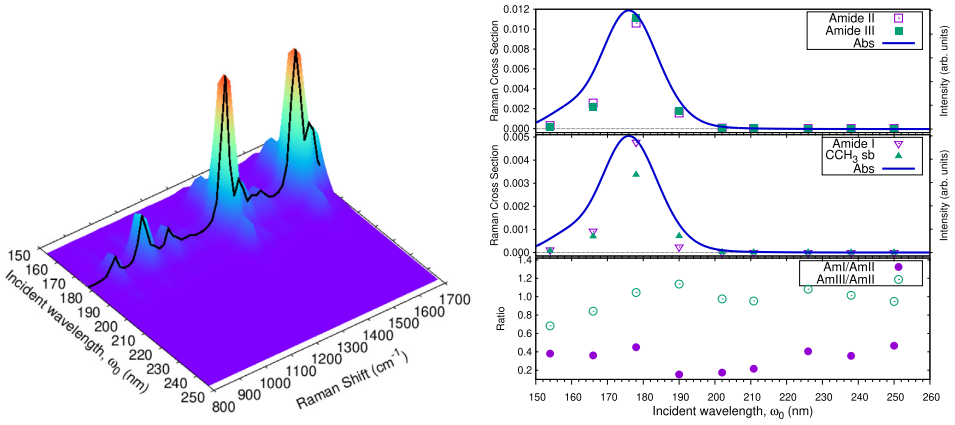
QM/PCM and QM/FQ computed spontaneous Raman spectra of acetamide (ACA) in aqueous solution are reported in Fig. 4(a) together with their experimental counterpart measured by using ω_0 = 633 nm.

We first notice that almost all peak relative intensities are correctly reproduced by both the solvation approaches, however the inhomogeneous broadening, and also the position of most bands, are better described by QM/FQ with respect to the implicit description. As in the case of NMA, in solvated acetamide, the amide I band located at 1662 cm^{-1} originates primarily from C═O stretching, whereas an amide II-like band occurs at 1616 cm^{-1} and is associated with NH_2 bending, even if contains a small contribution from C═O stretching (see Sec. S2.1 in the SM). The amide III-like band of ACA, located at 1404 cm^{-1} derives from the C—N stretching, but also has minor contributions of C—C stretching, NH_2 rocking, and symmetric deformations of the NCO group and CH_3 groups [13].

Similarly to NMA, Amide I and Amide II bands are located in the region around 1600 cm^{-1}, where the broad H—O—H bending vibration of the water molecules contributes to the Raman intensity [13,79,80].

5.2.2. UV Resonance Raman spectra

QM/PCM and QM/FQ UVRR spectra of ACA in aqueous solution are reported in Fig. 4(b), together with our spectra measured by using ω_0 = 226 nm (see Fig. S10 in the SM for experimental UVRR measured with ω_0 = 210, 226 and 266 nm). We first note that all experimental Raman peaks increase in intensity as the excitation wavelength decreases from 633 to 226 nm, with Amide I, II, and III bands showing the largest resonance enhancement, that in agreement with previous findings [13]. The same is valid for QM/FQ, which is able to provide an almost perfect agreement with the experiment in terms of relative intensities, band positions, and band broadening. Conversely, QM/PCM UVRR spectrum dra-



(a) (b)

Fig. 3. (a) Calculated QM/FQ Resonance Raman Excitation Profiles (RREP) of **NMA** in aqueous solution. (b) Excitation wavelength dependence of the Amide I, II, III and the 1380 cm^{-1} CCH_3 symmetric bend (sb) band intensity, and of Amide I/Amide II and the Amide III/Amide II band Raman cross-section ratios. RR intensities (in $cm^2mol^{-1}sr^{-1}$) were calculated with a damping factor of 200 cm^{-1} and broadened using Lorentzian functions with FWHM = 20 cm^{-1}.

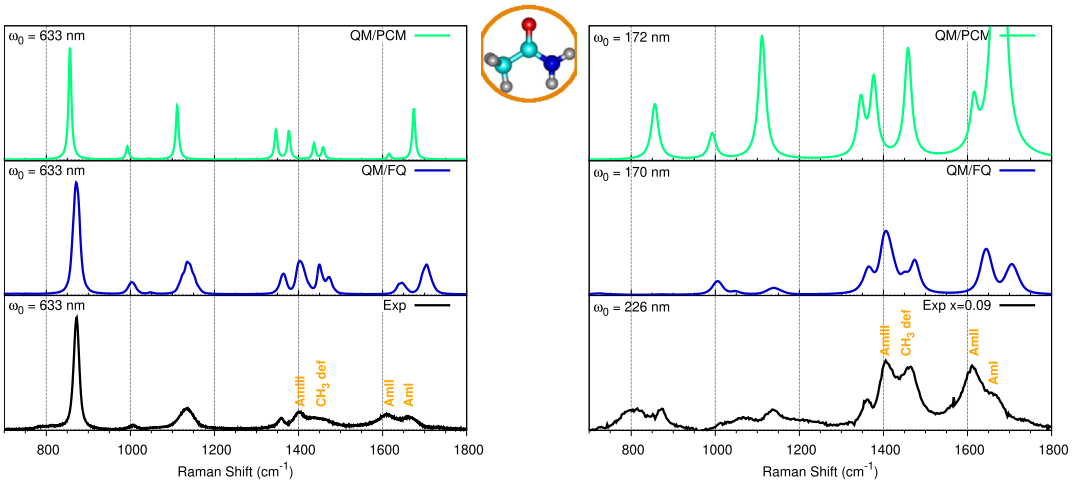


Fig. 4. Raman spectra of **Acetamide** calculated with different solvent descriptions (top, QM/PCM and middle, QM/FQ) and compared to the experimental spectra (bottom) measured in aqueous solution at room temperature. (a) Spontaneous (Far From Resonance) Raman spectra simulated and measured using 633 nm as excitation wavelength. FWHM: 8 cm^{-1}. (b) UV Resonance Raman spectra. In the UVRR experiments, the external excitation wavelengths have been set to 226 nm. x = 0.09 stands for the Acetamide molar fraction in water. RR intensities were calculated with a damping factor of 200 cm^{-1} and broadened using Lorentzian functions with FWHM = 20 cm^{-1}. The RR band at around 800 cm^{-1} is a spurious signal arising from the quartz cuvette.

matically differs from the experimental counterpart. In particular, the amide I mode ($\sim$ 1700 cm^{-1}) is much more intense than in the experiments, similarly to what has been already commented for NMA. The nature of the normal mode, which is mainly characterized by a C=O stretching, suggests that an explicit description of the water molecules is needed.

We now move to comment on the RREP, i.e. the Raman spectrum as a function of the external wavelength, of ACA in aqueous solution as calculated by means of QM/FQ (see Fig. 5. Note that an experimental RREP has been measured in Ref. [22]. The excitation profiles of Amide I (C=O stretching), Amide II (NH_2 bending), Amide III (C-N stretching) as well as the CH_3 rocking vibration (1457 cm^{-1}) are reported in Fig. 5(b). Two clear-cut trends can be seen in Fig. 5(b): on one hand, similar to NMA, the amide I, II and III bands are resonance enhanced in ACA with Amides II and III increasing in intensity more than than Amide I. As expected, the enhancement is maximum facing 172 nm (see Sec. S2.3 in the SM), which is the computed QM/FQ excitation energy of the $\pi \rightarrow \pi^*$ transition (see Tab. S1 in the SM). As observed in Fig. 5 (b), this transition of the peptide group starts to be significant at

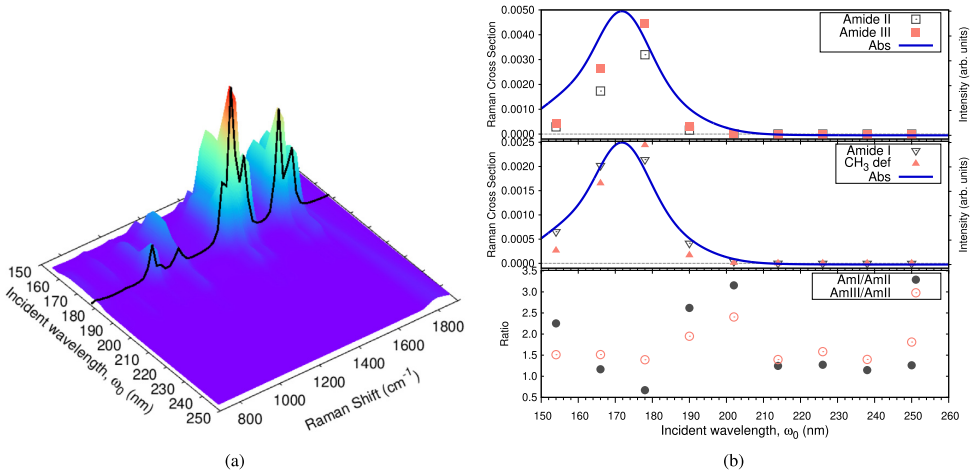


Fig. 5. (a) Calculated QM/FQ Resonance Raman Excitation Profiles (RREP) of **Acetamide** in aqueous solution. (b) Excitation wavelength dependence of the Amide I, II, III and and 1457 cm^{-1} CCH_3 deformation (def) bands intensity, and of Amide I/Amide II and the Amide III/Amide II band Raman cross-section ratios. RR intensities (in $cm^2 mol^{-1} sr^{-1}$) were calculated with a damping factor of 200 cm^{-1} and broadened using Lorentzian functions with a FWHM of 20 cm^{-1}.

excitation wavelengths near the maximum of the absorption curve, as reported in Ref. [22], where the authors found a relationship between the excitation wavelength and the Raman band positions for the same signals, associating it to the HB strength at the NH_2 and C=O sites.

5.3. DMA

5.3.1. Spontaneous Raman spectra

In Fig. 6(a), the QM/PCM and QM/FQ spontaneous Raman spectra of DMA in aqueous solution are reported together with the measured experimental spectrum (ω_0 = 633 nm). Overall, the Raman spectrum of DMA is dominated by the same amide bands previously mentioned for NMA and ACA (see also Sec. S3.1 in the SM). No significant differences between QM/PCM and QM/FQ solvation models can be noticed (Fig. 6(a)). Both environmental treatments lead to peak positions close to the experimental ones, with the only exception being the Amide I band ($\sim$ 1650 cm^{-1}), for which QM/FQ also provide a more accurate prediction of its relative intensity. This is not unexpected because, as for the other two amides, Amide I band is largely affected by specific solute–solvent interactions.

5.3.2. UV Resonance Raman spectra

Simulated UVRR spectra of DMA in aqueous solution and experimental spectra, recorded using 226 nm as excitation wavelength, are depicted in Fig. 6(b) (see also Fig. S14 in the SM). The overall picture that emerges from Fig. 6(b) is of rather good agreement between theory and experiment, at least in terms of the major features. This overview is similar for both QM/PCM and QM/FQ models, even though a significant improvement is reached by QM/FQ, as judged by the description of the Amide I and Amide II, which occur at 1650 and 1490 cm^{-1}, respectively. This shows that a better description of the solute–solvent interactions, provided when the discrete water molecules are employed, is important also in this system but not as crucial as in the case of ACA and NMA, due to the presence of the two methyl groups. Small discrepancies between computed and experimental UVRR results can be explained considering that for DMA it has been reported that the amide bands are sensitive to changes in concentration [78]. Remarkably, in both solvation approaches, we are assuming an infinite dilute solution, which is not the case for experimental conditions.

The dependence of QM/FQ UVRR as a function of the excitation wavelength is reported in Fig. 7(a), whereas the enhancements of selected normal modes (Amide I, II, III and CH_3 antisymmetric bend) are illustrated in Fig. 7(b). An enhancement of the Amide II band relative to Amide I and Amide III is immediately perceived for excitation wavelengths near the absorption maximum (184 nm, see also Sec. S3.3 in the SM). Indeed, for this spectral region, the Amide II band becomes fivefold more intense than both the Amide I and Amide II bands. It is also clear from Fig. 7(b) that the amide III band cross section shows essentially no resonance enhancement, differently from NMA and ACA, for which its intensity does change (see Figs. 2(b) and 4(b)). Furthermore, by comparing our non–resonant (Fig. 6)) and pre-resonant (Fig. 6(b)) Raman results, we notice that the overcrowded region between 1400 and 1500 cm^{-1} turns into a couple of bands located at 1437 and 1487 cm^{-1}. For the first one, there is also an enhancement as incident light approaches the maximum of the absorption spectra as shown in Fig. 7(b), middle panel.

6. Discussion

The results discussed in the previous sections clearly show that QM/FQ outperforms continuum solvation, and gives an almost perfect description of experimental UVRR spectra of all three amides. In this section, we discuss such findings on the basis of physicochemical properties of the systems when dissolved in aqueous solution. In particular, the performance of the two solvation models in the description of both spontaneous and RR of ACA, NMA and DMA can be investigated by analyzing the resutls of MD simulations, which can provide a qualitative and quantitative overview on specific, directional solute–solvent interactions. To this end, in

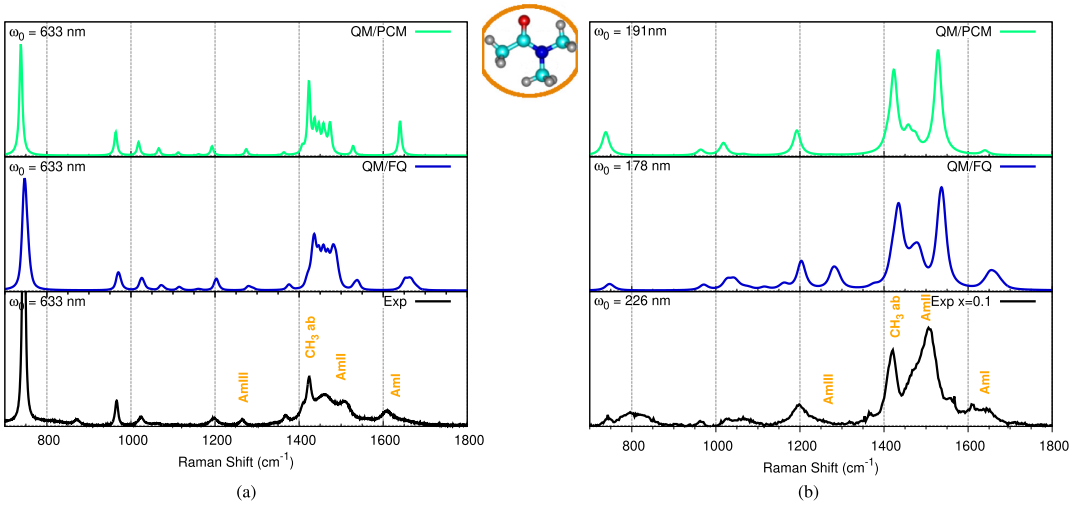


Fig. 6. Raman spectra of **DMA** calculated with different solvent descriptions (top, QM/PCM and middle, QM/FQ) and compared to the experimental spectra (bottom) measured in aqueous solution at room temperature. (a) Spontaneous (Far From Resonance) Raman spectra simulated and measured using 633 nm as excitation wavelength. FWHM: 4 cm^{-1}. (b) UV Resonance Raman spectra. In the UVRR experiments, the external excitation wavelengths have been set to 226 nm. x = 0.09 stands for the DMA molar fraction in water. RR intensities were calculated with a damping factor of 200 cm^{-1} and broadened using Lorentzian functions with a FWHM of 20 cm^{-1}. The RR band at around 800 cm^{-1} is a spurious signal arising from the quartz cuvette.

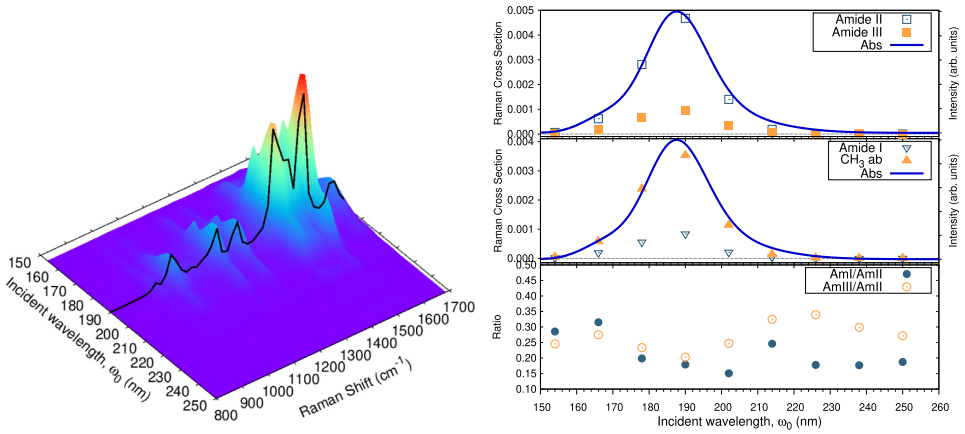


Fig. 7. (a) Calculated QM/FQ Resonance Raman Excitation Profile (RREP) of **DMA** in aqueous solution. (b) Excitation wavelength dependence of the Amide I, II, III and the 1430 cm^{-1} CCH_3 antisymmetric bend (ab) bands intensity, and of Amide I/Amide II and the Amide III/Amide II band Raman cross-section ratios. RR intensities (in $cm^2mol^{-1}sr^{-1}$) were broadened by a Lorentzian function with a FWHM of 20 cm^{-1} and a damping factor of 200 cm^{-1}.

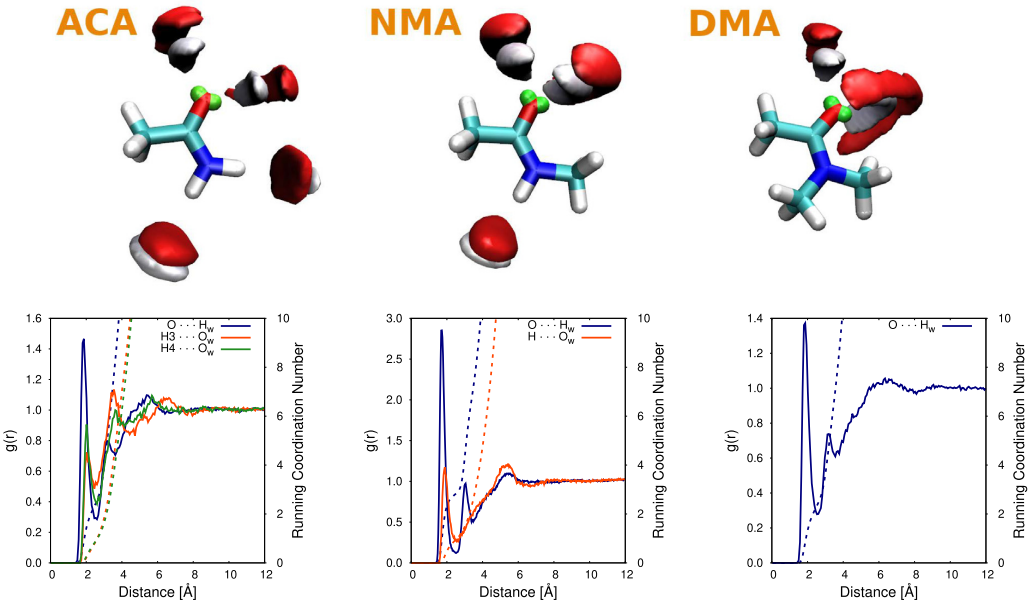


Fig. 8. Top panel: SDFs for the three amides studied in this work. Calculated SDF isodensity values are equal to 70 and 80 nm^{-3} for water hydrogen (white surfaces) and oxygen (red surfaces) atoms, respectively. Bottom panel: Radial distribution function, *g(r)*, (solid lines) and running coordination numbers (dashed lines) for the hydrogen bonds CO···H_w and NH···O_w in the solvated amides.

Fig. 8 we report the Spatial and Radial Distribution Functions (SDF and RDF, respectively), which were extracted from the last 10 ns of MD simulations. The two plots are complementary: RDFs are used to quantify the strength of HB interactions in terms of the average solute–solvent distance, whereas SDFs provide a graphical 3D depiction of the most favorable positions of water Hydrogen and Oxygen atoms.

From the inspection of Fig. 8 it is clear that the three amides can form strong HBs with the surrounding water molecules. Remarkably, SDFs show the importance of including virtual sites in the description of the solute's Oxygen atom, perfectly describing the directionality which characterizes HB interactions. For this reason, QM/FQ outperforms the implicit approximation (QM/PCM) in the description of Amide I and II bands, both involving the C═O bond, because the implicit approach lacks of any specific/directional interaction. In case of ACA and NMA, the importance of an atomistic description of the environment is also demonstrated by HB interactions which are established between the amide Hydrogen (s) and water molecules (see both RDFs and SDFs). The absence of the amide Hydrogen in DMA can explain the better agreement

between QM/PCM and QM/FQ results. In fact, the NH group is involved in the most relevant vibrations in the studied region (700–1800 cm^{-1}). Nevertheless, also in this case, we highlight that a collection of configurations representing the dynamical fluctuations of the solvent molecules around the solute guarantees a much better agreement with experimental results.

To conclude, we note that the differences between QM/PCM and QM/FQ results are more pronounced for RR than for spontaneous Raman, thus confirming the necessity of an accurate and reliable description of solvation, as it is provided in this work by the coupling of the polarizable QM/FQ with classical MD simulations.

7. Summary and conclusions

We have presented a joined computational and experimental work analyzing Raman and Resonance Raman (RR) spectra of three simple amides, acetamide, *trans-N*-methyl acetamide and *N,N* -dimethyl acetamide, in aqueous solution. In particular, we focused on the effect that the aqueous environment induces on the spectra by comparing computational results obtained with the commonly employed continuum model and the polarizable QM/FQ approach, that permits to account for both polarization and specific hydrogen bonding effects. The study was also possible thanks to availability of SR-based UVRR experimental setup that allowed us to record accurate experimental spectra at low concentration of amides in water and with a variety of incident excitation wavelengths, an indispensable requirement for analyzing the effect of the resonant enhancement.

Our results demonstrate once again the severe shortcomings of the popular polarizable continuum model (PCM) in simulating the aqueous environment. In fact, Raman and RR spectra differ dramatically from their experimental counterpart at the PCM level, and the discrepancy is much more pronounced for systems having more potential sites for specific solute–solvent interactions. Being a mixed electronic and vibrational spectroscopy, RR requires an unbiased solvation model that is capable of capturing all solute–solvent interactions, and we have shown that our QM/FQ model gives highly accurate results. In fact, the agreement between theory and experiment is so high that the main features of the spectra are easily identified, in particular the assignment to the Amide I, Amide II and Amide III bands, that in primary amides (e.g. acetamide) are essentially due to C═O stretching, N—H bending and C—N stretching vibrations, respectively, while in secondary amides (e.g. NMA) the Amide II and III bands arise from combined N—H bending and C—N stretching vibrations. On the other hand, in tertiary amides (e.g. DMA), the lack of the amide hydrogen causes different contributions for Amide II and Amide III signals. Computed QM/FQ Resonance Raman Excitation Profiles (RREP) indicate that amide bands are particularly sensitive to variations in the incident wavelength, with the strongest changes occurring by approaching absorption spectra maxima, as earlier reported in experimental works [19].

Given the success of the combination of theory and experiment in reproducing RR spectra of simple amides, the next step is to study more complex amino acids and peptides, which are essential building blocks of biological systems, and whose natural environment is water. The application of our method to these systems is in progress in our group, and will undoubtedly shed light into the complex mechanisms that cooperate to produce the final spectral response, thus building an understanding that is essential to and rationalize the spectroscopic features of complex protein systems.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

SG and CC gratefully acknowledge the support of H2020-MSCA-ITN-2017 European Training Network “Computational Spectroscopy In Natural sciences and Engineering” (COSINE), Grant No. 765739. We gratefully acknowledge computational resources available at SNS Center for High Performance Computing (CHPC). We thank Dr. A. Gessini of the IUVS beamline at Elettra for the technical support.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.molliq.2021.117841.

References

[1] S.A. Asher, Uv resonance raman studies of molecular structure and dynamics: Applications in physical and biophysical chemistry, Annu. Rev. Phys. Chem. 39 (1) (1988) 537–588.
[2] S.A. Asher, Uv resonance raman spectroscopy for analytical, physical, and biophysical chemistry. part 1, Anal. Chem. 65 (2) (1993) 59A–66A.
[3] S.A. Asher, Uv resonance raman spectroscopy for analytical, physical, and biophysical chemistry. part 2, Anal. Chem. 65 (4) (1993) 201A–210A.
[4] K. Xiong, Uv resonance raman spectroscopy: A highly sensitive, selective and fast technique for environmental analysis, J. Environ. Anal. Chem 2 (1) (2014) e107.
[5] Z. Chi, S.A. Asher, Uv resonance raman determination of protein acid denaturation: Selective unfolding of helical segments of horse myoglobin, Biochem. 37 (9) (1998) 2865–2872.
[6] Z. Ahmed, I.A. Beta, A.V. Mikhonin, S.A. Asher, Uv–resonance raman thermal unfolding study of trp-cage shows that it is not a simple two-state miniprotein, J. Am. Chem. Soc. 127 (31) (2005) 10943–10950.
[7] C.-Y. Huang, G. Balakrishnan, T.G. Spiro, Early events in apomyoglobin unfolding probed by laser t-jump/uv resonance raman spectroscopy, Biochem. 44 (48) (2005) 15734–15742.
[8] X. Zhao, R. Chen, C. Tengroth, T.G. Spiro, Solid-state tunable khz ultraviolet laser for raman applications, Appl. Spectrosc. 53 (10) (1999) 1200–1205.
[9] A.V. Mikhonin, S.A. Asher, Uncoupled peptide bond vibrations in α-helical and polyproline ii conformations of polyalanine peptides, J. Phys. Chem. B 109 (7) (2005) 3047–3052.
[10] G. Balakrishnan, Y. Hu, G.M. Bender, Z. Getahun, W.F. DeGrado, T.G. Spiro, Enthalpic and entropic stages in α-helical peptide unfolding, from laser t-jump/uv raman spectroscopy, J. Am. Chem. Soc. 129 (42) (2007) 12801–12808.
[11] S. Catalini, B. Rossi, P. Foggi, C. Masciovecchio, F. Bruni, Aqueous solvation of glutathione probed by uv resonance raman spectroscopy, J. Mol. Liq. 283 (2019) 537–547.
[12] K. Kaya, S. Nagakura, Electronic absorption spectra of hydrogen bonded amides, Theor. Chim. Acta 7 (Apr 1967) 124–132.
[13] J.M. Dudik, C.R. Johnson, S.A. Asher, Uv resonance raman studies of acetone, acetamide, and n-methylacetamide: models for the peptide bond, J. Phys. Chem. 89 (18) (1985) 3805–3814.
[14] L.C. Mayne, B. Hudson, Resonance raman spectroscopy of n-methylacetamide: overtones and combinations of the carbon-nitrogen stretch (amide ii') and effect of solvation on the carbon-oxygen double-bond stretch (amide i) intensity, J. Phys. Chem. 95 (8) (1991) 2962–2967.
[15] X.G. Chen, R. Schweitzer-Stenner, S. Krimm, N.G. Mirkin, S.A. Asher, N-methylacetamide and its hydrogen-bonded water molecules are vibrationally coupled, J. Am. Chem. Soc. 116 (24) (1994) 11141–11142.
[16] X.G. Chen, R. Schweitzer-Stenner, S.A. Asher, N.G. Mirkin, S. Krimm, Vibrational assignments of trans-n-methylacetamide and some of its deuterated isotopomers from band decomposition of ir, visible, and resonance raman spectra, J. Phys. Chem. 99 (10) (1995) 3074–3083.
[17] X.G. Chen, S.A. Asher, R. Schweitzer-Stenner, N.G. Mirkin, S. Krimm, Uv raman determination of the.pi.pi.* excited state geometry of n-methylacetamide: Vibrational enhancement pattern, J. Am. Chem. Soc. 117 (10) (1995) 2884–2895.
[18] B. Rossi, S. Catalini, C. Bottari, A. Gessini, C. Masciovecchio, Frontiers of UV resonant raman spectroscopy by using synchrotron radiation: the case of aqueous solvation of model peptides, in: G. Lérondel, Y.-H. Cho, A. Taguchi, S. Kawata (Eds.), UV and Higher Energy Photonics: From Materials to

Applications 2019, vol. 11086, International Society for Optics and Photonics, SPIE, 2019, pp. 23–32.
[19] F. D'Amico, B. Rossi, G. Camisasca, F. Bencivenga, A. Gessini, E. Principi, R. Cucini, C. Masciovecchio, Slow-to-fast transition of hydrogen bond dynamics in acetamide hydration shell formation, Phys. Chem. Chem. Phys. 17 (2015) 10987–10992.
[20] Q. Wu, G. Balakrishnan, A. Pevsner, T.G. Spiro, Histidine photodegradation during uv resonance raman spectroscopy, J. Phys. Chem. A 107 (40) (2003) 8047–8051.
[21] J. Kneipp, G. Balakrishnan, R. Chen, T.-J. Shen, S.C. Sahu, N.T. Ho, J.L. Giovannelli, V. Simplaceanu, C. Ho, T.G. Spiro, Dynamics of allostery in hemoglobin: Roles of the penultimate tyrosine h bonds, J. Mol. Biol. 356 (2) (2006) 335–353.
[22] M. Saito, F. D'Amico, G. Camisasca, F. Bencivenga, R. Cucini, A. Gessini, E. Principi, T. Ogura, C. Masciovecchio, Resonance raman spectroscopy with chemical state selectivity on histidine and acetamide using synchrotron radiation, Bull. Chem. Soc. Jpn. 88 (4) (2015) 591–596.
[23] S.P. Fodor, T.G. Spiro, Ultraviolet resonance raman spectroscopy of dna with 200–266-nm laser excitation, J. Am. Chem. Soc. 108 (12) (1986) 3198–3205.
[24] S. Ng, S. Yarasi, P. Brost, G.R. Loppnow, Initial excited-state structural dynamics of thymine are coincident with the expected photochemical dynamics, J. Phys. Chem. A 112 (41) (2008) 10436–10437.
[25] S. Yarasi, S. Ng, G.R. Loppnow, Initial excited-state structural dynamics of uracil from resonance raman spectroscopy are different from those of thymine (5-methyluracil), J. Phys. Chem. B 113 (43) (2009) 14336–14342.
[26] F. Bianchi, L. Comez, R. Biehl, F. D'Amico, A. Gessini, M. Longo, C. Masciovecchio, C. Petrillo, A. Radulescu, B. Rossi, F. Sacchetti, F. Sebastiani, N. Violini, A. Paciaroni, Structure of human telomere G-quadruplex in the presence of a model drug along the thermal unfolding pathway, Nucleic Acids Res. 46 (2018) 11927–11938.
[27] C. Bottari, S. Catalini, P. Foggi, I. Mancini, A. Mele, D.R. Perinelli, A. Paciaroni, A. Gessini, C. Masciovecchio, B. Rossi, Base-specific pre-melting and melting transitions of dna in presence of ionic liquids probed by synchrotron-based uv resonance raman scattering, J. Mol. Liq. 330 (2021) 115433.
[28] A. Ianoul, T. Coleman, S.A. Asher, Uv resonance raman spectroscopic detection of nitrate and nitrite in wastewater treatment processes, Anal. Chem. 74 (6) (2002) 1458–1461.
[29] D.D. Tuschel, A.V. Mikhonin, B.E. Lemoff, S.A. Asher, Deep ultraviolet resonance raman excitation enables explosives detection, Appl. Spectrosc. 64 (4) (2010) 425–432.
[30] S.A. Oladepo, K. Xiong, Z. Hong, S.A. Asher, Elucidating peptide and protein structure and dynamics: Uv resonance raman spectroscopy, J. Phys. Chem. Lett. 2 (4) (2011) 334–344.
[31] M.S. Barclay, T.J. Quincy, D.B. Williams-Young, M. Caricato, C.G. Elles, Accurate assignments of excited-state resonance raman spectra: A benchmark study combining experiment and theory, J. Phys. Chem. A 121 (41) (2017) 7937–7946.
[32] C.R. Johnson, S.A. Asher, A new selective technique for characterization of polycyclic aromatic hydrocarbons in complex samples: Uv resonance raman spectrometry of coal liquids, Anal. Chem. 56 (12) (1984) 2258–2261.
[33] S.A. Asher, Ultraviolet resonance raman spectrometry for detection and speciation of trace polycyclic aromatic hydrocarbons, Anal. Chem. 56 (4) (1984) 720–724.
[34] D.A. Long, Vibrational Resonance Raman Scattering, John Wiley & Sons Ltd, 2002, pp. 221–270 (ch. 7).
[35] Y. Ozaki et al., Molecular Spectroscopy: A Quantum Chemistry Approach, John Wiley & Sons, 2019.
[36] S. Bykov, I. Lednev, A. Ianoul, A. Mikhonin, C. Munro, S.A. Asher, Steady-state and transient ultraviolet resonance raman spectrometer for the 193–270 nm spectral region, Appl. Spectrosc. 59 (Dec 2005) 1541–1552.
[37] B. Rossi, C. Bottari, S. Catalini, F. D'Amico, A. Gessini, C. Masciovecchio, Chapter 13 - synchrotron-based ultraviolet resonance raman scattering for material science, in: V. Gupta, Y. Ozaki (Eds.), Molecular and Laser Spectroscopy, Elsevier, 2020, pp. 447–482.
[38] F. Egidi, J. Bloino, C. Cappelli, V. Barone, A robust and effective time-independent route to the calculation of resonance raman spectra of large molecules in condensed phases with the inclusion of duschinsky, herzberg-teller, anharmonic, and environmental effects, J. Chem. Theory Comput. 10 (1) (2014) 346–363.
[39] F.J. Avila Ferrer, V. Barone, C. Cappelli, F. Santoro, Duschinsky, herzberg-teller, and multiple electronic resonance interferential effects in resonance raman spectra and excitation profiles. the case of pyrene, J. Chem. Theory Comput. 9 (8) (2013) 3597–3611.
[40] F. Santoro, C. Cappelli, V. Barone, Effective time-independent calculations of vibrational resonance raman spectra of isolated and solvated molecules including duschinsky and herzberg-teller effects, J. Chem. Theory Comput. 7 (6) (2011) 1824–1839.
[41] B. Mennucci, Polarizable continuum model, WIREs Comput. Mol. Sci. 2 (3) (2012) 386–404.
[42] A. Baiardi, J. Bloino, V. Barone, A general time-dependent route to resonance-raman spectroscopy including franck-condon, herzberg-teller and duschinsky effects, J. Chem. Phys. 141 (11) (2014) 114108.
[43] L.M. Markham, B.S. Hudson, Ab initio analysis of the effects of aqueous solvation on the resonance raman intensities of n-methylacetamide, J. Phys. Chem. 100 (7) (1996) 2731–2737.
[44] N.S. Myshakina, Z. Ahmed, S.A. Asher, Dependence of amide vibrations on hydrogen bonding, J. Phys. Chem. B 112 (38) (2008) 11873–11877.
[45] H.M. Senn, W. Thiel, QM/MM methods for biomolecular systems, Angew. Chem. Int. Ed. 48 (7) (2009) 1198–1229.
[46] U.N. Morzan, D.J. Alonso de Armiño, N.O. Foglia, F. Ramírez, M.C. González Lebrero, D.A. Scherlis, D.A. Estrin, Spectroscopy in complex environments from qm-mm simulations, Chem. Rev. 118 (7) (2018) 4071–4113.
[47] H. Ren, J. Jiang, S. Mukamel, Deep uv resonance raman spectroscopy of β-sheet amyloid fibrils: A qm/mm simulation, J. Phys. Chem. B 115 (47) (2011) 13955–13962.
[48] V. Macaluso, L. Cupellini, G. Salvadori, F. Lipparini, B. Mennucci, Elucidating the role of structural fluctuations, and intermolecular and vibronic interactions in the spectroscopic response of a bacteriophytochrome, Phys. Chem. Chem. Phys. 22 (2020) 8585–8594.
[49] T. Giovannini, F. Egidi, C. Cappelli, Molecular spectroscopy of aqueous solutions: a theoretical perspective, Chem. Soc. Rev. 49 (2020) 5664–5677.
[50] C. Cappelli, Integrated qm/polarizable mm/continuum approaches to model chiroptical properties of strongly interacting solute-solvent systems, Int. J. Quant. Chem. 116 (21) (2016) 1532–1542.
[51] T. Giovannini, F. Egidi, C. Cappelli, Theory and algorithms for chiroptical properties and spectroscopies of aqueous systems, Phys. Chem. Chem. Phys. 22 (2020) 22864–22879.
[52] L. Jensen, L.L. Zhao, J. Autschbach, G.C. Schatz, Theory and method for calculating resonance raman scattering from resonance polarizability derivatives, J. Chem. Phys. 123 (17) (2005) 174110.
[53] E.J. Heller, R. Sundberg, D. Tannor, Simple aspects of raman scattering, J. Phys. Chem. 86 (10) (1982) 1822–1833.
[54] D.J. Tannor, E.J. Heller, Polyatomic raman scattering for general harmonic potentials, J. Chem. Phys. 77 (1) (1982) 202–218.
[55] A.B. Myers, Resonance raman intensity analysis of excited-state dynamics, Acc. Chem. Res. 30 (12) (1997) 519–527.
[56] G.N.R. Tripathi, Electronic structure of para aminophenoxyl radical in water, J. Chem. Phys. 118 (3) (2003) 1378–1391.
[57] J. Guthmuller, B. Champagne, Time dependent density functional theory investigation of the resonance raman properties of the julolidinemalononitrile push-pull chromophore in various solvents, J. Chem. Phys. 127 (16) (2007) 164507.
[58] J. Guthmuller, Comparison of simplified sum-over-state expressions to calculate resonance raman intensities including franck-condon and herzberg-teller effects, J. Chem. Phys. 144 (6) (2016) 064106.
[59] T. Giovannini, M. Olszówka, F. Egidi, J.R. Cheeseman, G. Scalmani, C. Cappelli, Polarizable embedding approach for the analytical calculation of raman and raman optical activity spectra of solvated systems, J. Chem. Theory Comput. 13 (9) (2017) 4421–4435.
[60] F. Lipparini, C. Cappelli, V. Barone, Linear response theory and electronic transition energies for a fully polarizable QM/classical hamiltonian, J. Chem. Theory Comput. 8 (11) (2012) 4153–4165.
[61] I. Carnimeo, C. Cappelli, V. Barone, Analytical gradients for mp2, double hybrid functionals, and td-dft with polarizable embedding described by fluctuating charges, J. Comput. Chem. 36 (31) (2015) 2271–2290.
[62] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, et al, Gaussian 16 Revision A.03, 2016. Gaussian Inc., Wallingford CT.
[63] A.V. Marenich, S.V. Jerome, C.J. Cramer, D.G. Truhlar, Charge model 5: An extension of hirshfeld population analysis for the accurate description of molecular interactions in gaseous and condensed phases, J. Chem. Theory Comput. 8 (2) (2012) 527–541.
[64] M.J. Abrahama, T. Murtola, R. Schulz, S. Pálla, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25.
[65] M. Macchiagodena, G. Mancini, M. Pagliai, V. Barone, Accurate prediction of bulk properties in hydrogen bonded liquids: Amides as case studies, Phys. Chem. Chem. Phys. 18 (2016) 25342–25354.
[66] M. Macchiagodena, G. Mancini, M. Pagliai, G. Del Frate, V. Barone, Fine-tuning of atomic point charges: Classical simulations of pyridine in different environments, Chem. Phys. Lett. 677 (2017) 120–126.
[67] T. Giovannini, M. Macchiagodena, M. Ambrosetti, A. Puglisi, P. Lafiosca, G. Lo Gerfo, F. Egidi, C. Cappelli, Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling, Int. J. Quant. Chem. 119 (1) (2019) e25684.
[68] A. Puglisi, T. Giovannini, L. Antonov, C. Cappelli, Interplay between conformational and solvent effects in uv-visible absorption spectra: curcumin tautomers as a case study, Phys. Chem. Chem. Phys. 21 (2019) 15504–15514.
[69] T. Giovannini, M. Ambrosetti, C. Cappelli, A polarizable embedding approach to second harmonic generation (shg) of molecular systems in aqueous solutions, Theor. Chem. Acc. 137 (May 2018) 74.
[70] S.F. Boys, Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another, Rev. Mod. Phys. 32 (1960) 296–299.
[71] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, J. Comput. Chem. 25 (9) (2004) 1157–1174.
[72] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (1) (2007) 014101.

[73] M. Brehm, B. Kirchner, Travis - a free analyzer and visualizer for monte carlo and molecular dynamics trajectories, J. Chem. Inf. Model. 51 (8) (2011) 2007–2023.
[74] X. Li, M.J. Frisch, Energy-represented direct inversion in the iterative subspace within a hybrid geometry optimization method, J. Chem. Theory Comput. 2 (3) (2006) 835–839.
[75] S.W. Rick, S.J. Stuart, B.J. Berne, Dynamical fluctuating charge force fields: Application to liquid water, J. Chem. Phys. 101 (7) (1994) 6141–6156.
[76] S.A. Asher, Z. Chi, P. Li, Resonance Raman examination of the two lowest amide $\pi\pi^*$ excited states, J. Raman Spectrosc. 29 (10–11) (1998) 927–931.
[77] D. Punihaole, R.S. Jakubek, E.M. Dahlburg, Z. Hong, N.S. Myshakina, S. Geib, S.A. Asher, Uv resonance raman investigation of the aqueous solvation dependence of primary amide vibrations, J. Phys. Chem. B 119 (10) (2015) 3931–3939.
[78] N.E. Triggs, J.J. Valentini, An investigation of hydrogen bonding in amides using raman spectroscopy, J. Phys. Chem. 96 (17) (1992) 6922–6931.
[79] T. Giovannini, M. Olszówka, C. Cappelli, Effective fully polarizable qm/mm approach to model vibrational circular dichroism spectra of systems in aqueous solution, J. Chem. Theory Comput. 12 (11) (2016) 5483–5492.
[80] T. Giovannini, G. Del Frate, P. Lafiosca, C. Cappelli, Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution, Phys. Chem. Chem. Phys. 20 (2018) 9181–9197.

4.3.2. NAGMA and NALMA dipeptides[5]

THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL

Letter

Amide Spectral Fingerprints are Hydrogen Bonding-Mediated

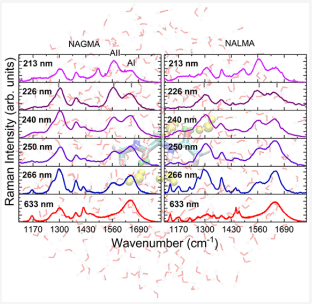
Sara Gómez,* Cettina Bottari, Franco Egidi, Tommaso Giovannini, Barbara Rossi, and Chiara Cappelli*

Cite This: *J. Phys. Chem. Lett.* 2022, 13, 6200−6207

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The origin of the peculiar amide spectral features of proteins in aqueous solution is investigated, by exploiting a combined theoretical and experimental approach to study UV Resonance Raman (RR) spectra of peptide molecular models, namely *N*-acetylglycine-*N*-methylamide (NAGMA) and *N*-acetylalanine-*N*-methylamide (NALMA). UVRR spectra are recorded by tuning Synchrotron Radiation at several excitation wavelengths and modeled by using a recently developed multiscale protocol based on a polarizable QM/MM approach. Thanks to the unparalleled agreement between theory and experiment, we demonstrate that specific hydrogen bond interactions, which dominate hydration dynamics around these solutes, play a crucial role in the selective enhancement of amide signals. These results further argue the capability of vibrational spectroscopy methods as valuable tools for refined structural analysis of peptides and proteins in aqueous solution.



Amide bands are considered sensitive probes of the secondary structures of proteins and peptides enabling prediction that provides a significant advance in the knowledge of protein activity and function. For this reason, vibrational spectroscopy experiments such as Infrared and Raman are well-established methods to identify and quantify distinct secondary structure motifs of proteins and polypeptides through exploration of the Amide fingerprint region.

The proper interpretation of ever more accurate experimental measurements makes the availability of reference studies analyzing the deep nature and physical origin of the spectroscopic response down to the atomistic detail highly desirable. *N*-Acetyl-glycine-methylamide (NAGMA) and *N*-acetyl-leucine-methylamide (NALMA) (see Figure 1) can be employed as minimal prototypes to model certain protein properties and behaviors.[1−12] Compared to simple unmodified amino acids, both their C and N termini are modified to model the peptide bonding, while maintaining a small size and conformational flexibility; therefore, they are more suitable as "peptide models" than single amino acids, which do not exhibit the chemical heterogeneity and interactions that characterize a protein backbone. Being minimal models for larger molecular structures, they allow for extensive and highly detailed investigations into their physicochemical properties, both intrinsic and relating to their molecular environment, as well as the details of their spectroscopic properties,[1,13−22] which is particularly crucial as advanced spectroscopic techniques are then applied to more complex biological polymers.

The simplicity of the molecular models, however, hides the complexity that exists, as these systems are dissolved in aqueous solution, i.e. their physiological environment. Any attempt at modeling these peptides must therefore focus on the nature of solvation and its effects upon the properties of the systems. Indeed, the investigation of the hydration dynamics at hydrophilic and hydrophobic biomolecular sites of simple peptides has paved the way to appropriate models for understanding the dynamics of the first hydration shell of proteins.[1,14,23]

The unique combination of the capabilities of UV Resonance Raman (UVRR) spectroscopy (enhanced detection limit, high selectivity of specific chemical groups, no interfering signal of water solvent in the amide fingerprint spectral region)[24−26] and the features of these small peptides result in local molecular probes for focusing on hydrogen bond (HB) rearrangements at specific sites of peptides, even under high diluted conditions. In this regard, it has already been shown[22] that, due to the advantages of Synchrotron Radiation (SR)-based UVRR, the fine-tuning of the excitation wavelength allows the experimenter to select the best working conditions that ensure one can reliably detect the spectral changes of amide signals, as a function of peptide concentration and temperature.[27,28] However, there are still some experimental features that are not fully understood. First, as also demonstrated in this work, a selective enhancement of the Amide II (AII) mode is experimentally observed at the shortest excitation wavelengths for both NAGMA and NALMA in aqueous solution and not for their microcrystalline form. *Selectivity* in this context means that excitation of $\pi \rightarrow \pi^*$ transitions of the amide peptide bonds, lead to strong UVRR

Received: April 28, 2022
Accepted: June 22, 2022

https://doi.org/10.1021/acs.jpclett.2c01277
J. Phys. Chem. Lett. 2022, 13, 6200−6207

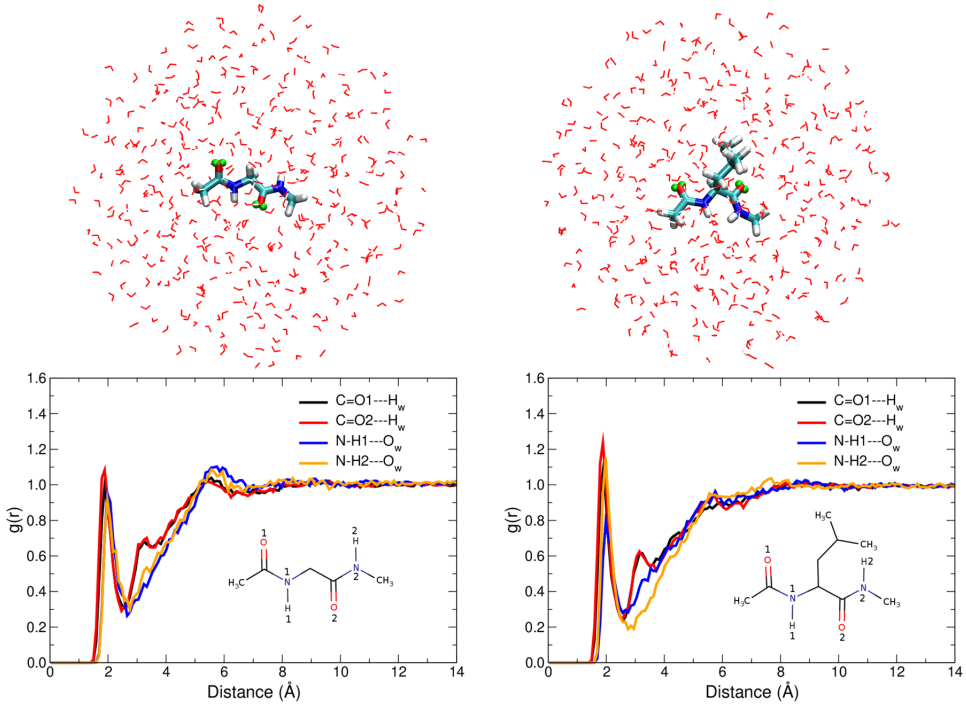


Figure 1. Top: QM/FQ representation of NAGMA (left) and NALMA (right) in aqueous solution. Virtual sites (VS) in the C═O groups are depicted in green. VS are interaction sites constructed to improve the description of the hydrogen bonding. Bottom: Radial Distribution Functions (RDFs) for the intermolecular O···H interactions in solvated dipeptides.

enhancement of Raman signals associated with the vibrations that have large components of C−N stretching, while smaller enhancement occurs for vibrations with strong C═O stretching. Second, Amide I (AI) and AII shapes and positions change going from microcrystalline/hydrated powders to solutions of both peptides. Third, in the measured spectra a dependence of wavenumbers of amide bands on concentration is reported. In this work, we explain the physical origin of the three aforementioned experimental findings, by interplaying experiments and simulations.

Often, simulations are unable to reproduce experimental findings of (Resonance) Raman spectra due to the intrinsic limitations in standard solvation models.[29] Solvation affects molecular response properties both directly and indirectly by altering a molecule's conformational landscape. The effects of solvation are commonly estimated either by implicit, continuum approaches or by simulation studies (MonteCarlo or Molecular Dynamics (MD)) in which the solvent molecules are explicitly considered in both the sampling and the calculation of the properties.[30−34] Recently, some of the present authors have designed a multiscale computational protocol, combining MD and the Quantum Mechanics/Fluctuating Charges (QM/FQ) approach, a polarizable embedding model that keeps a fully atomistic representation of the solvent and provides accurate excitation energies and RR spectra of various systems, including small amides.[35,36]

Disentangling the role of each type of interaction in the generation of the UVRR spectra of these systems is a necessary first step if one hopes to fully rationalize and explain the spectroscopic behavior of peptides in aqueous solution, and due to the aforementioned complexities, this result can only be achieved by combining the highest possible level of techniques, both theoretical and experimental. To tackle these problems we therefore used the QM/FQ model in combination with SR-UVRR experiments at different excitation wavelengths and interpret, at the molecular level, the selective enhancement of the Amide II band experimentally detected at the shortest wavelengths. To this end, we performed extensive simulations using a hierarchy of solvation approaches on NAGMA and NALMA (see Computational Methods in the Supporting Information (SI)) and investigated the role of hydrogen bonding in the spectroscopic behavior of these systems, by comparing simulated results[37,38] with the multiwavelengths experimental measurements made possible by the fine energy tunability of a SR source, which affords a much greater degree of flexibility compared to standard experimental setups, where the excitation radiation is provided by energy-fixed laser sources. Experimental procedure and terminology are part of the SI.

We began by examining the performance of QM/FQ, coupled to MD sampling, to describe UV−vis, Raman, and RR spectra of the two peptide systems (see Figure 1 for a picture of the representative systems used in the simulations). Furthermore, we evaluated gas phase, Polarizable Continuum Model (PCM), and cluster calculations as summarized in Table 1. All studied systems are depicted in Table S1 in the SI; their structural parameters are listed in Tables S2 and S3.

Due to their conformational flexibility, NAGMA and NALMA monomers may feature intramolecular HBs. Consistent with previous reports,[8,10,12] we found that in the gas

https://doi.org/10.1021/acs.jpclett.2c01277
J. Phys. Chem. Lett. 2022, 13, 6200−6207

Table 1. Computed Vertical Excitation Wavelengths (in nm) for NAGMA and NALMA in Different Environments, Calculated at the B3LYP/6-311++G(*d*, *p*) Level of Theory[c]

Motif	Environment	Absorption maxima (nm)	
		NAGMA	NALMA
Monomer *C5*	Gas phase	200.4, 180.5	190.1
Monomer *C7*	Gas phase	196.6	196.0
Monomer β_2	Gas phase	194.6	200.2
Monomer *C5*	PCM	178.9	184.1
Monomer *C7*	PCM	182.6	187.9
Monomer β_2	PCM	184.5	189.0
Monomer + 4W	PCM	181.7	183.8
Solution (366:1)	QM/MM NP	174.0	177.3
Solution (366:1)	QM/FQ[a]	175.2	178.5
Solution (366:1)	QM/FQ[b]	175.3	177.5
Solution (366:1)	QM/QM_w/FQ	178.3	181.5
Dimer	Gas phase	197.0	199.5
Dimer solvated	Gas phase	190.9	193.9

[a]FQ parametrization from ref 42. [b]FQ parametrization from ref 43. [c]NP stands for the Non-polarizable TIP3P.[39] Numbers in parentheses indicate the ratio water molecules: peptide molecules. In amides/small peptides, the first allowed electronic transition is experimentally reported to occur at ca. 190 nm[40,41]

phase the stable conformers are located mainly in *C5* (β_L) and *C7* (γ) regions, while solvent effects increase the prevalence of minima in the α and β_2 (δ_L) regions. Ramachandran maps in Figure S2 in the SI indicate that β_2 is the most representative conformer sampled by MD runs.

In dilute aqueous solution intramolecular hydrogen bonds compete with intermolecular interactions between solute and solvent, with the latter becoming dominant, as supported by the intra- and intermolecular radial distribution functions (RDFs) for the O···H contacts and by the HB strength estimated using Natural Bond Orbitals (NBO)[44] (Figures S3 and S4 in the SI). Intermolecular RDFs are shown in Figure 1. For both molecules, the C═O and N−H groups of the backbone are involved in the same number of HBs (2, 1 each). Building upon this observation, we have saturated all potential HB sites using 4 water molecules to build what is generally called a "supermolecule" (see Table S1 in the SI).

Despite sharing some features, the maxima in the RDFs are always located at slightly shorter distances for NAGMA, confirming that the strength of the solute−solvent interactions is higher for NAGMA than for NALMA, due to the hydrophobic residue in the latter. This finding is consistent with the detected wavenumber shift between AI and AII signals in the UVRR spectra of NAGMA and NALMA that reflects a different strength of peptide−solvent interactions for the two peptides.[22]

Considering that strong UVRR intensities are to be expected when properly tuning excitation wavelengths, we calculated the UV−vis spectra of both molecules in the different environments considered in this study. Table 1 lists vertical excitation energies in each case. The bands in the electronic absorption

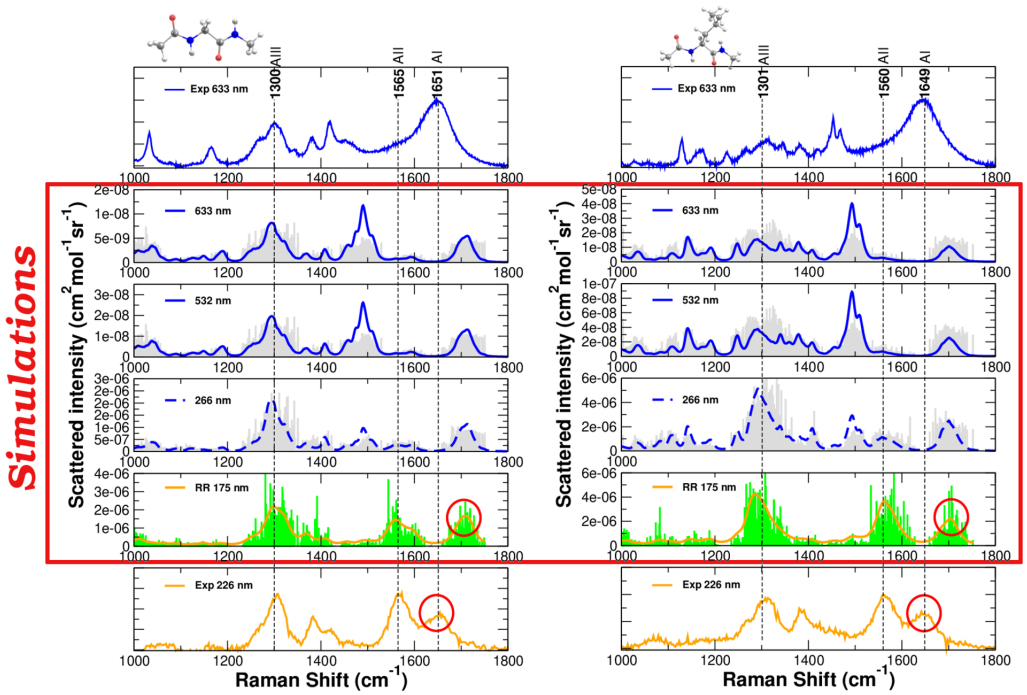


Figure 2. Raman (blue) and Resonance Raman (orange) spectra of NAGMA and NALMA, left and right panels, respectively. Experimental spectra were measured at room temperature in aqueous solution at a concentration which corresponds to 366 molecules of water for each molecule of peptide. QM/FQ results for spontaneous (Far From Resonance) and Resonance Raman spectra were broadened using Lorentzian functions with an fwhm of 8 and 20 cm^{-1} respectively. RR intensities were calculated with a damping factor of 200 cm^{-1}. Sticks in the simulated spectra are also included. The dashed blue curve indicates a preresonance condition.

https://doi.org/10.1021/acs.jpclett.2c01277
J. Phys. Chem. Lett. 2022, 13, 6200−6207

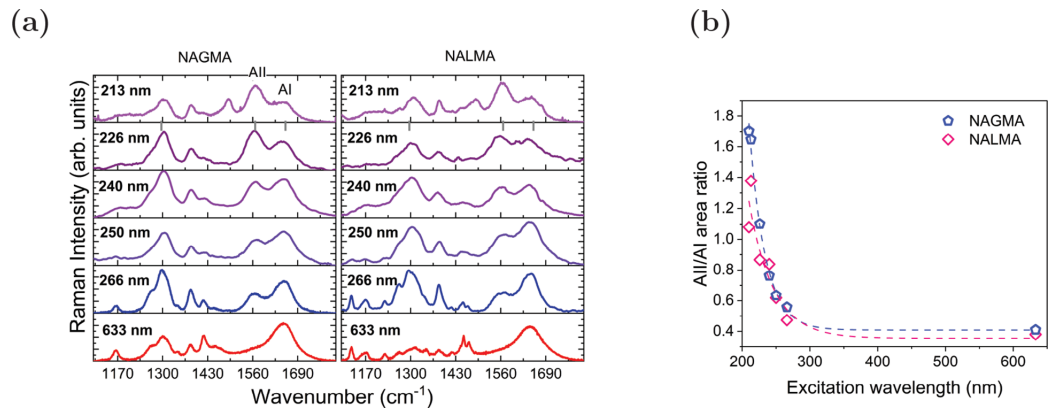


Figure 3. Experimental results for NAGMA and NALMA dissolved in water at a concentration corresponding to 366 molecules of water for each molecule of peptide (a) FFR and UVRR spectra collected using different excitation wavelengths ranging from visible to deep UV energies. (b) Estimated ratio of the areas of amide modes AII/AI as a function of the excitation wavelength.

spectra are very similar in appearance (Figure S5 in the SI), with the aqueous solution bands being blue-shifted by up to 20 nm relative to *C7*, the most stable conformer in the gas phase. In contrast, and regardless of the solvent representation, slight differences (5 nm at most) are noted in the case of the aqueous solution. We assign the strong electronic absorption band computed at ∼180 nm (190 nm is the experimental report) for both systems as the $\pi \rightarrow \pi^*$ transition based on prior studies[40] and the orbitals involved (Figure S6 in the SI). Thereupon, excitation within this absorption band will give rise to enhancements of the peptide bond vibrations.

Raman (Far From Resonance, FFR) and UVRR spectra of solvated NAGMA and NALMA are displayed in Figure 2. At first glance, all the typical vibrational features can be recognized in both the measured and computed spectra, namely, (*i*) the presence of the Amide I (AI) band with a Raman signal around 1650 cm^{-1} which is mainly associated with the stretching vibration of the C═O of the amide linkage in the peptide backbone; (*ii*) the Raman signal at ∼1560 cm^{-1}, assigned to the Amide II (AII) and resulting by the out-of-phase combination of N−H bending and C−N stretching movements of the groups in the amide linkage; and (*iii*) the Amide III (AIII) band, whose vibrational mode is mostly due to the in-phase combination of N−H bending and C−N stretching and which appears around 1260 cm^{-1}. Notice that in the FFR visible spectra of peptides the Amide II mode is completely absent (as expected), while in the UVRR spectra this band clearly appears even in preresonance conditions (excitation wavelength at 266 nm). Normal modes are drawn in Figures S7 and S8 in the SI.

The above results show that the mutually polarizable QM/FQ approach describes well Raman and RR spectra of the two systems, while fairly preserving positions and relative intensities, unlike results in PCM or using cluster/supermolecule approaches (see Figure S9 in the SI). However, it can be noted that when the FQ parameters taken from ref 42 are utilized, the accuracy in predicting AI band position is not entirely satisfactory (see red circles in Figure 2). One direction of improvement could be the usage of a different parametrization, where electrostatics and polarization effects are more accurately accounted for in the QM Hamiltonian.[43] In fact, when FQ^b parameters are used, AI and AII bands approach each other, thus moving computed results toward experimental data (Figure S10 in the SI). Detailed analysis of the normal modes obtained with the parameters of ref 43 suggests that AI and AII modes are coupled.

From the UVRR spectra in Figure 2, AI, AII, and AIII are the signals found to be particularly affected by the resonance enhancement, making their simultaneous measurement the ideal experimental target to be used to directly determine protein secondary structure.[45] Nevertheless, the experimental scanning along the excitation wavelength reveals a selective enhancement of the AII mode at the shortest wavelengths, as shown in Figure 3a. This feature has been already pointed out in literature for peptides and proteins[22,46,47] and is also relevant here, even though it is slightly more intense for NAGMA than for NALMA (compare for example spectra at 226 and 213 nm in Figure 3a) as can be seen in the ratio of the areas in Figure 3b. Interestingly, such an effect is observed only for solutions of peptides and not for the microcrystalline form of the molecules (see Figure S11 in the SI). This latter experimental evidence suggests a crucial role played by the hydration shell around the peptides in determining the spectral features of Amide Raman signals.

An explanation for the selective AII intensity increase has been proposed[48] in terms of the stabilization of the ground state dipolar resonance structure $-O-C{=}NH_2^+$ that becomes more favored in aqueous solution with respect to the O═C−NH_2 resonance form that lacks charge separation. Thus, the formation of HBs at the amide and carbonyl sites leads to a contraction of the C−N distance (see Table S2 in the SI) and the vibrations having significant contributions from the C−N stretching mode, namely, amide I, II, and III, will have appreciable intensity enhancements in the Raman lines. Our calculations of resonance structures using Natural Resonance Theory (NRT)[49] indicate that 3 out of 4 main NAGMA resonance hybrids favor the dipolar form, namely, negative charges in the O atoms and positive charges on the N atoms (Figure S12 in the SI), and more importantly, the percentage values of all the hybrids bearing formal charges increase in solution. Indeed, when explicit water molecules are considered, dipolar structures become the leading resonance hybrids (Table S4 in the SI). The special preference for enhancing AII when excitation wavelengths approach the maxima in the

https://doi.org/10.1021/acs.jpclett.2c01277
J. Phys. Chem. Lett. 2022, 13, 6200−6207

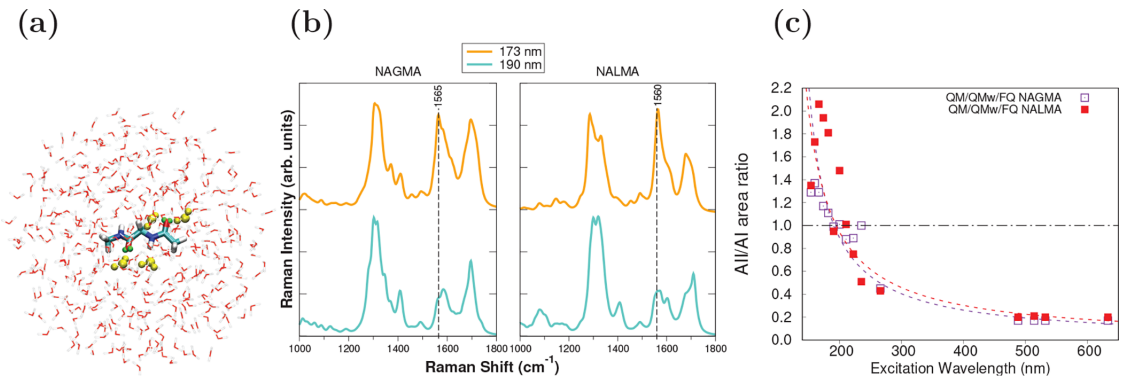


Figure 4. Computational results obtained with the QM/QM$_w$/FQ approach applied to NAGMA and NALMA in aqueous solution. (a) Representative structure, where selected water molecules (in yellow) are part of the QM portion. (b) Comparison between UVRR spectra simulated at 173 and 190 nm. (c) Computed ratio of the areas of the amide modes AII/AI as a function of excitation wavelength.

absorption spectra might be further analyzed based on the orbitals involved in the in-plane $\pi \rightarrow \pi^*$ transition. Computing 15 TD-DFT excited states on dipeptides structures revealed that the excited states with the highest oscillator strengths have an important charge transfer contribution, predominantly from HOMO and HOMO−1 to a wide assortment of virtual orbitals. For glycine and leucine dipeptides in solution, such states have larger contributions from the C−N regions of the molecules (Figure S6 in the SI). In addition, it is evident from inspection of Figure S6 (in the SI) that the orbitals involved not only belong to the solute but also involve the nearest water molecules, thus implying that charge transfer between peptides and solvent is an active component of molecular orbitals contributing to the 190 nm band. In view of the above findings, modeling the selective enhancement would require the inclusion of a few solvent molecules in the QM portion of the system.

Computed Resonance Raman Excitation Profiles (RREPs) are shown in the SI, Figures S13−S17 and S18−S22, for NAGMA and NALMA, respectively. The ratio between the areas of AI and AII in the simulated spectra using excitation wavelengths ranging from 633 to 166 nm is plotted in Figure S23 in the SI. Although QM/FQ with the two sets of parameters gives reasonable descriptions of the spectra, both parametrizations produce always an AI band with higher intensity and area than AII, leading to discrepancies when compared to the experiments. Conversely, it seems to be an important trend for the resonance enhancement of AII in the modeled RREPs of solvated NAGMA in the supermolecule case.

According to our molecular dynamics simulations and those reported by Boopathi and Kolandaive,[6] an average of two and three water molecules form persistent interactions with NALMA and NAGMA dipeptides, respectively. Also, it is well-known that the pattern of the relative intensities of the amide bands is determined by the properties in the electronic excited states for the various conformations.[45,47] Hence, to gain a deeper insight into the effect of these solvent molecules on the selective intensity alteration, we have investigated the intensity dependence on the exciting frequency of the AII by quantum-mechanically treating all water molecules that fulfill at least one of the following geometric criteria: $d_{H_w \cdots O=C} < 2.5$ Å or $d_{O_w \cdots H-N} < 2.5$ Å, hereafter noted as the QM/QM$_w$/FQ approach (details in Table S5 in the SI). This is graphically depicted in Figure 4b for NAGMA. A better reproduction of the experimental trend shown in Figure 3a was found when explicit water molecules are part of the QM layer, as seen in Figure 4b for two selected wavelengths. Furthermore, there is excellent agreement between the ratio of the areas, dashed curves in Figure 4c, and its experimental counterpart, Figure 3b. This constitutes an important piece of evidence of the role of hydrogen bonding and, specifically, its quantum-mechanical covalent component, in the intensity enhancement because just with solvent molecules linked to the C═O and N−H groups, the distances within the peptide are properly modified. Moreover, the orbitals of the transition are more concentrated in the C−N regions of the molecules, which ultimately triggers the enhancement. Regarding the NALMA case, the selective enhancement is seen in the QM/QM$_w$/FQ solvation model only. So, we deduce that there is a similar solvation behavior for the two peptides and that the hydrophilic part of the two molecules dominates the spectral properties, in addition to the solvation dynamics explained in ref 22.

Next, the different shapes and positions for AI and AII are studied by focusing on the UVRR spectra of the microcrystalline forms (Figure S11 in the SI), hydrated powders (Figure S24 in the SI), and solvated (Figure 2) samples. Explicitly in the amides region, some clear distinctions are evident: on one hand, a significant difference in the shape of AI for the two peptides, with probably two subcomponents with slight changes in frequency and relative intensity. Such an experimental splitting of the AI band (NAGMA case) is also recovered by the calculated spectra and explained by having a look at the normal modes, where two sets of C═O oscillators (see modes 21 and 22) with slightly different strengths can be identified. In contrast, they are more compactly gathered in the NALMA case. This observation is in line with the two types of C═O···H interactions found in the hydration patterns (Figure S3 in the SI), due to the arrangements of the molecules' backbone. On the other hand, it is also observed from Figure S24 in the SI that the position of Amide II (AII) has a red shift of about 5 cm^{-1} in the NALMA case. We recall here that the effect of the hydrophobic portion of NALMA is the lengthening of the N−H···O$_w$ distances with respect to

https://doi.org/10.1021/acs.jpclett.2c01277
J. Phys. Chem. Lett. 2022, 13, 6200−6207

NAGMA, thus affecting the N−H bending associated with that band.

Finally, in order to rationalize the concentration-dependence of wavenumber positions, it is worth mentioning that at high dilution conditions, like those in our simulations, solute−solute interactions are weak (see NBO interaction energies, Figure S4 in the SI) and then the C═O is expected to be more involved in HBs with solvent molecules. In particular, it has been seen[22] that for concentrations ≤50 mg/mL of both molecules in water (i.e., 144 molecules of water for each molecule of peptide), the position of the AI band becomes strongly dependent on concentration (severe red shift upon the increment of water content). Based upon the NRT results (Figure S12 and Table S4 in the SI), we argue that in aqueous solution, with the stabilization of the dipolar resonance structures, the double bond is more localized between C and N, leading to a shortening of the C−N bond and in turn to a lengthening of the C═O bond. Therefore, the force constant of a bond having a stronger single character compared to the resonance form that does not have a charge separation induces a decrease in the oscillation frequency for the C═O stretching and so a red shift of AI. The same explanation would apply for the concentration-dependent trend of AII in concentrations >50 mg/mL (144:1 water/peptide) where probably the solute−solute interactions become important. Aggregation propensities of these peptides in solution have been examined before in the literature.[4] Here, we explore solute···solute interactions through an approximate treatment, namely, the dimeric and solvated dimeric forms of NAGMA and NALMA. Their modeled RR can be found in Figure S25 in the SI. While in the UVRR spectra of concentrated dipeptide solutions it is reported[22] that there are no significant changes in the AI frequency position, AII, arising from the combination of C−N stretching and N−H bending motions, experiences a slightly monotonic blue shift, likely due to the strengthening of the C−N bond, which causes an increase in the force constant and, in turn, in the oscillation frequency. To verify the reliability of these arguments, we carefully checked the distances of the C═O and C−N in Tables S2 and S3 in the SI and concluded that, in all our representations of the solvent, the C═O and C−N bond lengths are actually longer and shorter, respectively, compared to the isolated cases. Consequently, it is again verified that the solvent does play a role in determining the positions, shapes, and intensities of the RR bands, as also pointed out by some authors.[28,50−52]

In summary, the interplay of computational and experimental data highlights two important observations: (*i*) The selective enhancement of the amides signals is hydrogen bonding-induced because it is intimately linked to the effect that water molecules exert on the C═O and N−H, C−N vibrations. We demonstrated that the inclusion of explicit water molecules concentrates the orbitals involved in the charge transfer in the C−N zones, which ultimately leads to the strong UVRR enhancement of vibrations that have large components of C−N stretching, particularly the AII signal. Thus, quantum effects must be present in any modeling of the solute−solvent interactions of RR spectroscopy for such systems. The unprecedented accordance found between the theoretical calculation and experimentally collected Raman and Resonance Raman spectra further testifies the reliability of our model. (*ii*) Due to the constant movement of the solute and its surrounding water molecules, a single snapshot (or cluster composed of the solute and some surrounding water molecules) is not representative of the dynamical nature of the system and can lead to heavily biased results if taken to be representative for the ensemble, which is more correctly modeled through an explicit average over a large set of structures. Through our investigation we put forward an explanation at the molecular level of the origin of the selective enhancements, emphasizing the crucial role of the backbone conformations and the dynamics of their surrounding waters. These results provide an important starting point to calibrate wavenumbers and intensities of the experimental Amide signals for the quantitative determination of structural parameters of protein and peptide in solutions. A key requirement of computational approaches to be truly useful and complementary for experimental measurements is the ability to properly describe and reproduce spectral features, and not just the energetics of a system. We have shown that the QM/QM_w/FQ fulfills this by promisingly going beyond standard methodologies based on more crude approximations and is, therefore, expected to open up new possibilities for novel applications in a truly synergistic partnership with advanced experimental techniques applied to biologically relevant samples, as well as shed new light regarding the details of physicochemical phenomena that characterize the functioning of life. In this respect, our study could be also extended to Raman Optical Activity (ROA) spectroscopy, which, due to its sensitivity to chirality, constitutes an alternative to Raman/Resonance Raman when examining the structure and behavior of peptides, proteins, and biomolecules.[53−56]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.2c01277.

Description of the experimental and computational methods, structural analysis, hydration patterns, measured and calculated spectra, Resonance Raman Excitation profiles, normal modes. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Sara Gómez − *Scuola Normale Superiore, 56126 Pisa, Italy;* orcid.org/0000-0002-5430-9228; Email: sara.gomezmaya@sns.it

Chiara Cappelli − *Scuola Normale Superiore, 56126 Pisa, Italy;* orcid.org/0000-0002-4872-4505; Email: chiara.cappelli@sns.it

Authors

Cettina Bottari − *Elettra Sincrotrone Trieste S.C.p.A., I-34149 Trieste, Italy*

Franco Egidi − *Scuola Normale Superiore, 56126 Pisa, Italy;* Present Address: Software for Chemistry & Materials BV, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands; orcid.org/0000-0003-3259-8863

Tommaso Giovannini − *Scuola Normale Superiore, 56126 Pisa, Italy;* orcid.org/0000-0002-5637-2853

Barbara Rossi − *Elettra Sincrotrone Trieste S.C.p.A., I-34149 Trieste, Italy; Department of Physics, University of Trento, I-38123 Povo, Trento, Italy;* orcid.org/0000-0003-1357-8074

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.2c01277

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge Elettra Sincrotrone Trieste for providing access to its synchrotron radiation facilities and for financial support (Proposal Number 20200432). We thank A. Gessini and M. Tortora for assistance in using beamline IUVS. This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 818064). We gratefully acknowledge the Center for High Performance Computing (CHPC) at SNS for providing the computational infrastructure.

■ REFERENCES

(1) Qvist, J.; Halle, B. Thermal Signature of Hydrophobic Hydration Dynamics. *J. Am. Chem. Soc.* **2008**, *130*, 10345−10353.
(2) Murarka, R. K.; Head-Gordon, T. Single Particle and Collective Hydration Dynamics for Hydrophobic and Hydrophilic Peptides. *J. Chem. Phys.* **2007**, *126*, 215101.
(3) Johnson, M. E.; Malardier-Jugroot, C.; Murarka, R. K.; Head-Gordon, T. Hydration Water Dynamics Near Biological Interfaces. *J. Phys. Chem. B* **2009**, *113*, 4082−4092.
(4) Nerenberg, P. S.; Jo, B.; So, C.; Tripathy, A.; Head-Gordon, T. Optimizing Solute−Water van der Waals Interactions To Reproduce Solvation Free Energies. *J. Phys. Chem. B* **2012**, *116*, 4524−4534.
(5) Panuszko, A.; Pieloszczyk, M.; Kuffel, A.; Jacek, K.; Biernacki, K. A.; Demkowicz, S.; Stangret, J.; Bruździak, P. Hydration of Simple Model Peptides in Aqueous Osmolyte Solutions. *Int. J. Mol. Sci.* **2021**, *22*, 9350.
(6) Boopathi, S.; Kolandaivel, P. Molecular Dynamics Simulations and Density Functional Theory Studies of NALMA and NAGMA Dipeptides. *J. Biomol. Struct. Dyn.* **2013**, *31*, 158−173.
(7) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J. L.; Marcoccia, J. F.; Csizmadia, I. G. Peptide Models. 1. Topology of Selected Peptide Conformational Potential Energy Surfaces (Glycine and Alanine Derivatives). *J. Am. Chem. Soc.* **1991**, *113*, 6256−6265.
(8) Gould, I. R.; Cornell, W. D.; Hillier, I. H. A Quantum Mechanical Investigation of the Conformational Energetics of the Alanine and Glycine Dipeptides in the Gas Phase and in Aqueous Solution. *J. Am. Chem. Soc.* **1994**, *116*, 9250−9256.
(9) Bisetty, K.; Catalan, J. G.; Kruger, H. G.; Perez, J. J. Conformational Analysis of Small Peptides of the Type Ac−X−NHMe, where X= Gly, Ala, Aib and Cage. *J. Mol. Struct.: THEOCHEM* **2005**, *731*, 127−137.
(10) Masman, M. F.; Lovas, S.; Murphy, R. F.; Enriz, R. D.; Rodríguez, A. M. Conformational Preferences of N-Acetyl-l-leucine-N'-methylamide. Gas-Phase and Solution Calculations on the Model Dipeptide. *J. Phys. Chem. A* **2007**, *111*, 10682−10691.
(11) Leavitt, C. M.; Moore, K. B., III; Raston, P. L.; Agarwal, J.; Moody, G. H.; Shirley, C. C.; Schaefer, H. F., III; Douberly, G. E. Liquid Hot NAGMA Cooled to 0.4 K: Benchmark Thermochemistry of a Gas-Phase Peptide. *J. Phys. Chem. A* **2014**, *118*, 9692−9700.
(12) Cormanich, R. A.; Rittner, R.; Buehl, M. Conformational Preferences of Ac-Gly-NHMe in Solution. *RSC Adv.* **2015**, *5*, 13052−13060.
(13) Russo, D.; Murarka, R. K.; Copley, J. R.; Head-Gordon, T. Molecular View of Water Dynamics near Model Peptides. *J. Phys. Chem. B* **2005**, *109*, 12966−12975.
(14) Russo, D.; Ollivier, J.; Teixeira, J. Water Hydrogen Bond Analysis on Hydrophilic and Hydrophobic Biomolecule Sites. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4968−4974.
(15) Born, B.; Weingärtner, H.; Bründermann, E.; Havenith, M. Solvation Dynamics of Model Peptides Probed by Terahertz Spectroscopy. Observation of the Onset of Collective Network Motions. *J. Am. Chem. Soc.* **2009**, *131*, 3752−3755.
(16) Murarka, R. K.; Head-Gordon, T. Dielectric Relaxation of Aqueous Solutions of Hydrophilic versus Amphiphilic Peptides. *J. Phys. Chem. B* **2008**, *112*, 179−186.
(17) Mazur, K.; Heisler, I. A.; Meech, S. R. Ultrafast Dynamics and Hydrogen-Bond Structure in Aqueous Solutions of Model Peptides. *J. Phys. Chem. B* **2010**, *114*, 10684−10691.
(18) Perticaroli, S.; Comez, L.; Paolantoni, M.; Sassi, P.; Morresi, A.; Fioretto, D. Extended Frequency Range Depolarized Light Scattering Study of N-Acetyl-leucine-methylamide−Water Solutions. *J. Am. Chem. Soc.* **2011**, *133*, 12063−12068.
(19) Perticaroli, S.; Nakanishi, M.; Pashkovski, E.; Sokolov, A. P. Dynamics of Hydration Water in Sugars and Peptides Solutions. *J. Phys. Chem. B* **2013**, *117*, 7729−7736.
(20) Comez, L.; Lupi, L.; Morresi, A.; Paolantoni, M.; Sassi, P.; Fioretto, D. More Is Different: Experimental Results on the Effect of Biomolecules on the Dynamics of Hydration Water. *J. Phys. Chem. Lett.* **2013**, *4*, 1188−1192.
(21) Comez, L.; Perticaroli, S.; Paolantoni, M.; Sassi, P.; Corezzi, S.; Morresi, A.; Fioretto, D. Concentration Dependence of Hydration Water in a Model Peptide. *Phys. Chem. Chem. Phys.* **2014**, *16*, 12433−12440.
(22) Rossi, B.; Catalini, S.; Bottari, C.; Gessini, A.; Masciovecchio, C. In *UV and Higher Energy Photonics: From Materials to Applications 2019, Frontiers of UV Resonant Raman Spectroscopy by Using Synchrotron Radiation: The Case of Aqueous Solvation of Model Peptides*; Lérondel, G., Cho, Y.-H., Taguchi, A., Kawata, S., Eds.; SPIE: 2019; Vol. *11086*; pp 23−32.
(23) Russo, D.; Teixeira, J.; Kneller, L.; Copley, J. R. D.; Ollivier, J.; Perticaroli, S.; Pellegrini, E.; Gonzalez, M. A. Vibrational Density of States of Hydration Water at Biomolecular Sites: Hydrophobicity Promotes Low Density Amorphous Ice Behavior. *J. Am. Chem. Soc.* **2011**, *133*, 4882−4888.
(24) Efremov, E. V.; Ariese, F.; Gooijer, C. Achievements in Resonance Raman Spectroscopy: Review of a Technique with a Distinct Analytical Chemistry Potential. *Anal. Chim. acta* **2008**, *606*, 119−134.
(25) Chowdhury, J. In *Molecular and Laser Spectroscopy, Chapter 7 - Resonance Raman Spectroscopy: Principles and Applications*; Gupta, V., Ed.; Elsevier: 2018; pp 147−164.
(26) Rossi, B.; Bottari, C.; Catalini, S.; D'Amico, F.; Gessini, A.; Masciovecchio, C. In *Molecular and Laser Spectroscopy, Chapter 13 - Synchrotron-based Ultraviolet Resonance Raman Scattering for Material Science*; Gupta, V., Ozaki, Y., Eds.; Elsevier: 2020; pp 447−482.
(27) Catalini, S.; Rossi, B.; Tortora, M.; Foggi, P.; Gessini, A.; Masciovecchio, C.; Bruni, F. Hydrogen Bonding and Solvation of a Proline-Based Peptide Model in Salt Solutions. *Life* **2021**, *11*, 824.
(28) Catalini, S.; Rossi, B.; Foggi, P.; Masciovecchio, C.; Bruni, F. Aqueous Solvation of Glutathione Probed by UV Resonance Raman Spectroscopy. *J. Mol. Liq.* **2019**, *283*, 537−547.
(29) Egidi, F.; Bloino, J.; Cappelli, C.; Barone, V. A Robust and Effective Time-Independent Route to the Calculation of Resonance Raman Spectra of Large Molecules in Condensed Phases with the Inclusion of Duschinsky, Herzberg−Teller, Anharmonic, and Environmental Effects. *J. Chem. Theory Comput.* **2014**, *10*, 346−363.
(30) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular Spectroscopy of Aqueous Solutions: A Theoretical Perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664−5677.
(31) Giovannini, T.; Lafiosca, P.; Cappelli, C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. *J. Chem. Theory Comput.* **2017**, *13*, 4854−4870.
(32) Giovannini, T.; Riso, R. R.; Ambrosetti, M.; Puglisi, A.; Cappelli, C. Electronic Transitions for a Fully Polarizable QM/MM Approach Based on Fluctuating Charges and Fluctuating Dipoles: Linear and Corrected Linear Response Regimes. *J. Chem. Phys.* **2019**, *151*, 174104.
(33) Egidi, F.; Giovannini, T.; Del Frate, G.; Lemler, P. M.; Vaccaro, P. H.; Cappelli, C. A Combined Experimental and Theoretical Study of Optical Rotatory Dispersion for (R)-glycidyl methyl ether in Aqueous Solution. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3644−3655.

https://doi.org/10.1021/acs.jpclett.2c01277
J. Phys. Chem. Lett. 2022, 13, 6200−6207

(34) Gómez, S.; Giovannini, T.; Cappelli, C. Absorption Spectra of Xanthines in Aqueous Solution: A Computational Study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5929−5941.
(35) Gómez, S.; Rojas-Valencia, N.; Giovannini, T.; Restrepo, A.; Cappelli, C. Ring Vibrations to Sense Anionic Ibuprofen in Aqueous Solution as Revealed by Resonance Raman. *Molecules* **2022**, *27*, 442.
(36) Gómez, S.; Egidi, F.; Puglisi, A.; Giovannini, T.; Rossi, B.; Cappelli, C. Unlocking the Power of Resonance Raman Spectroscopy: The Case of Amides in Aqueous Solution. *J. Mol. Liq.* **2021**, 117841.
(37) Guthmuller, J.; Champagne, B. Resonance Raman Scattering of Rhodamine 6G as Calculated by Time-Dependent Density Functional Theory: Vibronic and Solvent Effects. *J. Phys. Chem. A* **2008**, *112*, 3215−3223.
(38) Santoro, F.; Cappelli, C.; Barone, V. Effective Time-Independent Calculations of Vibrational Resonance Raman Spectra of Isolated and Solvated Molecules Including Duschinsky and Herzberg−Teller Effects. *J. Chem. Theory Comput.* **2011**, *7*, 1824−1839.
(39) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. B* **2001**, *105*, 9954−9960.
(40) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A. Elucidating Peptide and Protein Structure and Dynamics: UV Resonance Raman Spectroscopy. *J. Phys. Chem. Lett.* **2011**, *2*, 334−344.
(41) Asher, S. A.; Chi, Z.; Li, P. Resonance Raman Examination of the Two Lowest Amide $\pi\pi^*$ Excited States. *J. Raman Spectrosc.* **1998**, *29*, 927−931.
(42) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical Fluctuating Charge Force Fields: Application to Liquid Water. *J. Chem. Phys.* **1994**, *101*, 6141−6156.
(43) Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. Effective yet Reliable Computation of Hyperfine Coupling Constants in Solution by a QM/MM Approach: Interplay between Electrostatics and Non-Electrostatic Effects. *J. Chem. Phys.* **2019**, *150*, 124102.
(44) Weinhold, F.; Landis, C.; Glendening, E. What Is NBO Analysis and How Is It Useful? *Int. Rev. Phys. Chem.* **2016**, *35*, 399−440.
(45) Chi, Z.; Chen, X.; Holtz, J. S.; Asher, S. A. UV Resonance Raman-Selective Amide Vibrational Enhancement: Quantitative Methodology for Determining Protein Secondary Structure. *Biochemistry* **1998**, *37*, 2854−2864.
(46) Harada, I.; Sugawara, Y.; Matsuura, H.; Shimanouch, T. Preresonance Raman Spectra of Simple Amides using Ultraviolet Lasers. *J. Raman Spectrosc.* **1975**, *4*, 91−98.
(47) Sugawara, Y.; Harada, I.; Matsuura, H.; Shimanouchi, T. Preresonance Raman Studies of poly(L-lysine), poly(L-glutamic acid), and Deuterated N-methylacetamides. *Biopolymers* **1978**, *17*, 1405−1421.
(48) Punihaole, D.; Jakubek, R. S.; Dahlburg, E. M.; Hong, Z.; Myshakina, N. S.; Geib, S.; Asher, S. A. UV Resonance Raman Investigation of the Aqueous Solvation Dependence of Primary Amide Vibrations. *J. Phys. Chem. B* **2015**, *119*, 3931−3939.
(49) Glendening, E. D.; Badenhoop, J.; Weinhold, F. Natural Resonance Theory: III. Chemical Applications. *J. Comput. Chem.* **1998**, *19*, 628−646.
(50) Manas, E. S.; Getahun, Z.; Wright, W. W.; DeGrado, W. F.; Vanderkooi, J. M. Infrared Spectra of Amide Groups in α-Helical Proteins: Evidence for Hydrogen Bonding between Helices and Water. *J. Am. Chem. Soc.* **2000**, *122*, 9883−9890.
(51) Mikhonin, A. V.; Bykov, S. V.; Myshakina, N. S.; Asher, S. A. Peptide Secondary Structure Folding Reaction Coordinate: Correlation between UV Raman Amide III Frequency, Ψ Ramachandran Angle, and Hydrogen Bonding. *J. Phys. Chem. B* **2006**, *110*, 1928−1943.
(52) Myshakina, N. S.; Ahmed, Z.; Asher, S. A. Dependence of Amide Vibrations on Hydrogen Bonding. *J. Phys. Chem. B* **2008**, *112*, 11873−11877.
(53) Nafie, L. A. *Vibrational Optical Activity: Principles and Applications*; John Wiley & Sons: 2011.
(54) Nafie, L. A. Vibrational Optical Activity: From Discovery and Development to Future Challenges. *Chirality* **2020**, *32*, 667−692.
(55) Barron, L. D. *Molecular Light Scattering and Optical Activity*; Cambridge University Press: 2009.
(56) Berova, N.; Polavarapu, P. L.; Nakanishi, K.; Woody, R. W. *Comprehensive Chiroptical Spectroscopy: Applications in Stereochemical Analysis of Synthetic Compounds, Natural Products, and Biomolecules*; John Wiley & Sons: 2012; Vol. 2.

https://doi.org/10.1021/acs.jpclett.2c01277
J. Phys. Chem. Lett. 2022, 13, 6200−6207

Chapter 5

Drugs and Cell Membranes

The strategy of extracting snapshots from MD runs is also translated to the detailed analysis of the evolution of the intermolecular interactions during the insertion of drugs, namely anionic Ibuprofen[6] and Naproxen[7], into model cell membranes. In the first place, MD simulations at different temperatures and coupled to umbrella sampling methods, are run and the potential of mean force is obtained along the entire insertion path. The contributions from enthalpy and entropy are quantified as well to define the thermodynamic aspects of the drug's travel. Later on, different configurations are taken from the last part of the MD, and the fundamental interactions, at the molecular level, that are partly responsible for the drug insertion process, are explored using rigorous theoretical tools rooted in the formalism of quantum mechanics, including the NBO, QTAIM, and NCI indices methods. Interestingly, both drugs face the same energy barrier ($\approx$ 5 kJ/mol) and are located at the same equilibrium position, at the polar/non-polar interphase. Thus, the only appreciable quantitative difference between the insertion of anionic Naproxen and anionic Ibuprofen is the depth of the energy well: $\approx$ 5 and 16.5 kJ/mol, respectively, suggesting a larger structural affinity for the model membrane in favor of anionic Ibuprofen, which is more similar to the phospholipids forming the bilayer. Additionally, all the quantum descriptors of chemical bonding indicate that there is a collective action of many non-covalent weak interactions stabilizing the tertiary membrane/water/drug system as the drug goes from the aqueous phase to the interior of the membrane. The UV-Vis absorption spectra of both drugs as a function of their position from the purely aqueous phase to the approximate center of the membrane are also modeled and the negligible changes experimentally reported are explained by considering the constant presence of water molecules in the immediate vicinity of the ring/aromatic portion of the drugs, a feature that maintains similar local surroundings regardless of the environment and does not affect to a large extent the electronic properties of these two compounds.[10]

5.1. Insertion of NSAIDS into Lipid Bilayers

5.1.1. Thermodynamics and Intermolecular interactions[6,7]

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

Cite This: J. Phys. Chem. B 2020, 124, 79−90

pubs.acs.org/JPCB

Evolution of Bonding during the Insertion of Anionic Ibuprofen into Model Cell Membranes

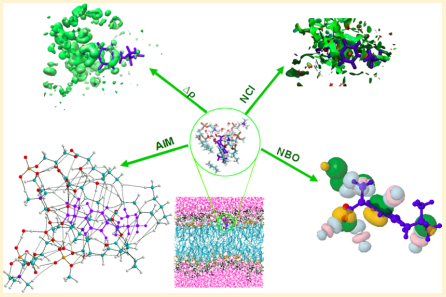
Natalia Rojas-Valencia,[†] Sara Gómez,[‡] Sebastian Montillo,[†] Marcela Manrique-Moreno,[†] Chiara Cappelli,[‡] Cacier Hadad,[†] and Albeiro Restrepo*[,†]

[†]Instituto de Química, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia

[‡]Scuola Normale Superiore, Classe di Scienze, Piazza dei Cavalieri 7, 56126, Pisa, Italy

Supporting Information

ABSTRACT: Descriptors of chemical bonding derived from five different analysis tools based on quantum mechanics (natural charges, electron density differences, atoms in molecules (AIM), natural bond orbitals (NBO), and non-covalent interactions (NCI) index) consistently afford a picture of a wall of weak, non-covalent intermolecular interactions separating anionic Ibuprofen from the environment. This wall, arising from the cumulative effect of a multitude of individual weak charge transfer interactions to the interstitial region between fragments, stabilizes the drug at all equilibrium positions in the free energy profile for its insertion into model cell membranes. The formal charge in anionic Ibuprofen strengthens all intermolecular interactions, having a particularly strong effect in the network of water to water hydrogen bonds in the solvent. Electron redistribution during the insertion process leads to a sensible reduction of electron delocalization in both the $-CO_2^-$ group and the aromatic ring of Ibuprofen. Here, we conclusively show that, despite their purely classical origin, randomly chosen configurations from molecular dynamics simulations provide deep insight into the purely quantum nature of bonding interactions.



INTRODUCTION

Non-steroidal antiinflamatory drugs (NSAIDs), Ibuprofen among them, are heavily used worldwide to relieve pain, fever, inflammation, and other maladies.[1] Ibuprofen consumption is unregulated, therefore, it is available over the counter, sold as a sodium salt, which once ingested dissociates into the corresponding ions in physiological aqueous environments. A sensible problem, which has been tied to unwanted side effects such as bleeding of the intestinal tract, is the interaction of anionic Ibuprofen (Ibu^- in what follows) with the lipids in the gastric mucous that acts as the first protective barrier in the stomach and intestines.[1−4]

It has recently been shown via classical molecular dynamics (MD) simulations[5] that insertion of anionic Ibuprofen into model cell membranes is a spontaneous, entropy driven process that faces a small energy barrier (ca. 5 kJ/mol). Once this barrier is overcome, anionic Ibuprofen reaches an energy minimum in the hydrophobic/polar interphase of the membrane, without reaching (as opposed to neutral Ibuprofen[6]) the midpoint of the lipid bilayer. This detailed classical molecular dynamics view of the insertion process indicates that severe changes in the chemical environment are experienced by anionic Ibuprofen in its path from the aqueous media, external to the membrane, to its equilibrium position. These changes are brought by loss of solvation water molecules, structural rearrangements, and increasing number of available microstates.

Regrettably, nothing is known about the fundamental interactions, at the molecular level, that are partly responsible for the Ibuprofen insertion process. In this work, we explore this issue, using rigorous theoretical tools rooted in the formalism of quantum mechanics, including the natural bond orbitals (NBO), quantum theory of atoms in molecules (QTAIM), and non-covalent interaction (NCI) indices methods, which have been thoroughly discussed elsewhere.[7−15] For this purpose, we adopt the common practice[16−22] of subjecting individual configurations afforded by MD simulations to quantum mechanical analysis. This procedure rests in the hypothesis that, because of the rigorous statistical nature of classical MD simulations and because of the accurate parametrization of force fields to reproduce experimental and/or *ab initio* results, randomly chosen configurations from equilibrium conditions would afford an insightful picture of the quantum molecular interactions. As we will show, the excellent agreement between the results in this work and the most important known facts of the thermodynamics of the insertion process on one hand and the atomistic

Received: October 15, 2019
Revised: November 28, 2019
Published: December 2, 2019

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

insight gathered in the description of bonding interactions on the other, provide strong support for the validity of this hypothesis for the title system.

Let us emphasize that the MD study from which individual conformations are borrowed exquisitely matches the experimental observations (for example, calorimetry measurements[23] rule out net heat transfer between anionic Ibuprofen and the lipids) and uncovers the reasons behind the entropy driven insertion process, thus adding support to their use to study bonding interactions in spite of their classical mechanics origin. We must also keep in mind at all times that equilibrium configurations extracted from a molecular dynamics simulation correspond to cases consistent with the Boltzmann distribution, which nonetheless are structures involved in a dynamic situation of change, so if the sample is sufficiently representative, the interactions involved (of interest in this study) are also representative interactions of the dynamic process despite not necessarily being configurations of minimal energy in the classical surface of the MD simulation, and even probably further from local minima in the quantum PES at the level of theory used for the description of the interactions.

METHODS

We build upon the classical dynamics simulations at 317 K under umbrella sampling reported by Rojas-Valencia and coworkers for the insertion of anionic Ibuprofen into model cell membranes. In short, the membranes consisted of twin layers of 64 units of dimyristoylphosphatidylcholine (DMPC) per layer, in aqueous environments provided by a total of 6200 explicit water molecules. See the original paper for detailed descriptions of the simulation conditions.[5] An equilibrium geometry for the tertiary membrane/water/Ibuprofen system after insertion of the drug is shown in Figure 1, along with a density of functional groups distribution in the membrane, and the free energy profile for the insertion process. The free energy profile at the bottom of Figure 1 highlights three relevant chemical environments surrounding Ibu^- as a function of the distance from the center of the lipid bilayer: (i) the aqueous phase, at ∼3.00 nm, where Ibu^- only interacts with water molecules, (ii) the top of the barrier, at ∼2.25 nm, where Ibu^- interacts with the phosphate and choline groups of the polar head and with water molecules, and (iii) the energy minimum, at ∼1.30 nm, where Ibu^- is mainly in contact with the two ester groups linking the polar and non-polar regions of the phospholipids, with the alkyl chains, with a reduced number of water molecules, and marginally with a few phosphate and choline groups.

MD simulations provide large scale, statistically sound pictures of dynamical processes undergone by systems composed of multitudes of individual units. Nonetheless, however accurate MD simulations may be, because of its very foundations, classical MD cannot offer insight into the fundamental nature of interactions at the molecular level. Therefore, according to the working hypothesis, exposed in the Introduction, in order to dissect the intermolecular interactions along the insertion path, we took 20 snapshots from the last 7 ns on the MD trajectories exploring the energy landscape in the plane perpendicular to the insertion axis[5] at each of the 3.00, 2.25, and 1.30 nm distances from the center of the lipid bilayer. Since the total simulation time on each case was 15 ns, taking snapshots after 8 ns ensures near or good equilibrium conditions. This approach puts us in a position to accurately study intermolecular interactions in the aqueous phase, at the top of the barrier, and at the energy minimum, the three relevant chemical environments. Radial distribution functions describing the density of molecules around Ibu^- were calculated using all possible configurations afforded by the MD simulation for each case. The corresponding plots are shown in Figure 2. The first maximum for each case (solid vertical lines) was used as a cutoff radius to determine, in addition to Ibuprofen, how many molecules (or lipid fragments) to include during the very expensive and delicate electronic structure calculations and in the subsequent analysis of the obtained molecular wave function. If the non-polar region of a given lipid extends beyond the cutoff radius, it is

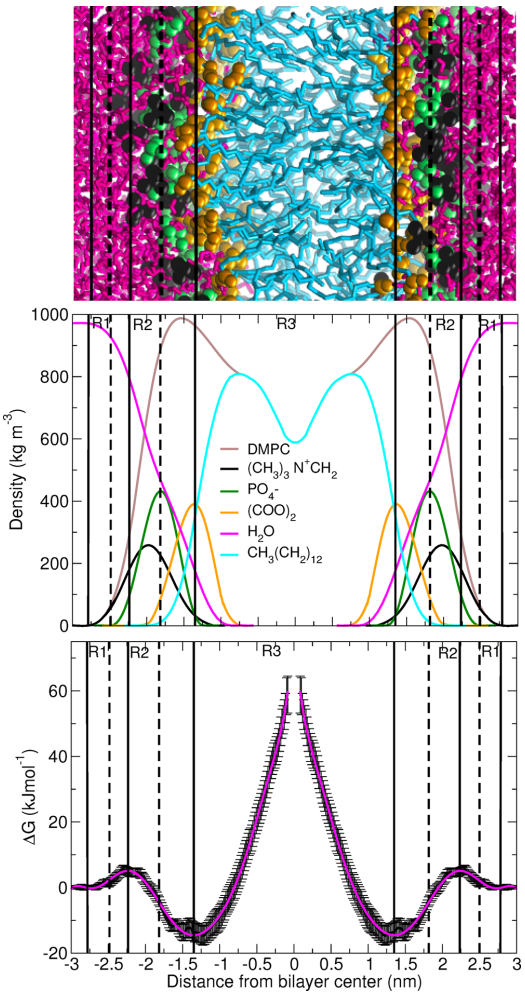


Figure 1. Top: A snapshot of the geometry for the model dimyristoylphosphatidylcholine (DMPC) membrane when the drug is at the equilibrium position (∼1.35 nm from the center of the lipid bilayer) at 317 K. Middle: Density of functional groups in the tertiary membrane/water/Ibuprofen system. Bottom: Free energy profile for the insertion process. Adapted with permission from ref 5. Copyright 2018 Royal Society of Chemistry.

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

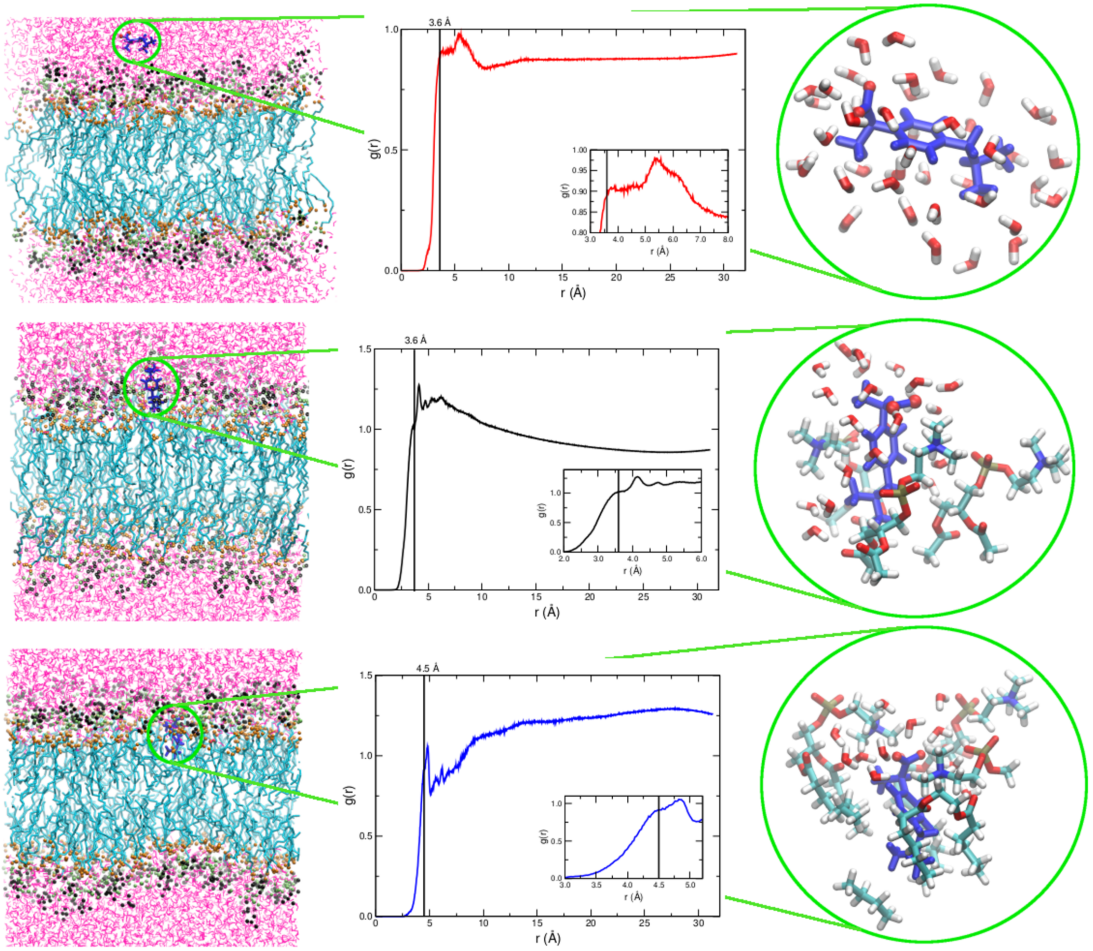


Figure 2. Selection of the systems of study: 20 snapshots for each equilibrium position (aqueous phase, top of the barrier, energy minimum at the top, middle, and bottom panels, respectively) in the free energy profile (Figure 1) of anionic Ibuprofen insertion into model DMPC membranes at 317 K where chosen. A particular snapshot is shown at the left, the corresponding radial distribution (for the last 8 ns of each simulation) for the distances of molecules falling within the cutoff radius for interaction with Ibuprofen (vertical solid lines, see text) at the middle, and a magnification of the system to be computed at the right (see Table 1).

truncated and the local valence of the terminal carbon restored with hydrogen atoms.

Analysis of the bonding interactions was carried out using the tools provided by the quantum theory of atoms in molecules (QTAIM),[10,11] natural bond orbitals (NBO),[7,24] and non-covalent interaction index (NCI)[14,15,25] methods. All of these methods have been extensively discussed elsewhere, so we make no effort to describe the corresponding formalisms, as they are outside the scope of this work. We used the Gaussian 09 suite[26] for all electronic structure calculations (60 structures in total) at the CAM-B3LYP functional proven to afford accurate results in this line of problems[18,20,27] in conjunction with the 6-31+G* basis set. The NBO 6.0 program[28] as implemented in Gaussian was used to study orbital interactions (60 structures in total). The AIMALL suite[29] was used to derive all QTAIM quantities (one representative structure per chemical environment). We used the NCIPLOT program[14,25] to calculate NCI surfaces (one representative structure per chemical environment).

We provide in Figure 3 the equilibrium geometry for isolated anionic Ibuprofen with atom numbering to serve as a reference for the discussion that follows. We find it useful to separate two regions in Ibu^-; thus, hereafter we take $—CO_2^-$ as the polar region and everything from the ring down to the isobutyl group as the non-polar region.

RESULTS AND DISCUSSION

According to the above discussion, 20 snapshots from each equilibrium position of the free energy profile (Figure 1) for the insertion of anionic Ibuprofen into model DMPC membranes were chosen for further analysis from the late stages of the corresponding MD simulations at 317 K. Application of our criterion for the cutoff radius for the

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

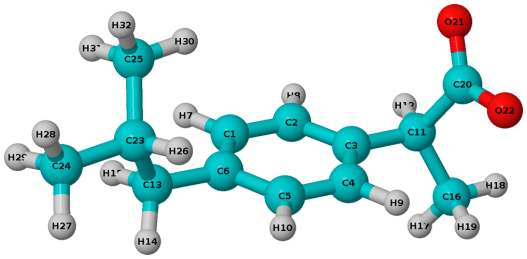


Figure 3. Equilibrium structure for isolated anionic (*S*)-(+)-Ibuprofen (Ibu^-) at the CAM-B3LYP/6-31+G(d) level. Atom numbers are included for reference.

Ibuprofen ↔ media interactions leads to the average number of water molecules and lipid residues or fragments listed in Table 1.

Table 1. Inventory of the Average Number of Water Molecules (n_W) and Lipid Residues (or Fragments) (n_L) within the Cutoff Radius in the 20 Snapshots Chosen from the MD Simulations of Each Equilibrium Position in the Free Energy Profile (Figure 1) for the Insertion of Ibu^- into Model DMPC Membranes

location	cutoff radius (Å)	n_W	n_L
aqueous phase	3.6	40	0
top of the barrier	3.6	27	3
energy minimum	4.5	14	6

Charge Redistribution. We study how the electron density on Ibu^- changes when switching chemical environments as the drug inserts into the model membrane from the aqueous phase to its equilibrium position, in two ways:

(1) We calculate how the CAM-B3LYP natural charges change relative to isolated anionic Ibuprofen (left panel in Figure 4). As a general rule, individual atom charges only change modestly, with carbon atoms becoming more negative and hydrogen atoms becoming more positive. Notwithstanding the small magnitudes for the change in individual atom charges, the cumulative effect is significant. Indeed, for the aqueous phase, accounting for all atoms in Ibu^-, a total of 0.24 and 0.32 |*e*| are gained and lost, respectively. At the top of the barrier, the corresponding values are 0.20 and 0.30 |*e*|, and for the position at the energy minimum, 0.16 and 0.29 |*e*| are collectively gained and lost at Ibu^-, respectively. These numbers indicate a subtle but important point: the interactions with the environment produce a larger polarization of the electron distribution in Ibu^- at the aqueous phase, conversely, a sensibly smaller polarization is seen at the equilibrium position, at the bottom of the energy profile (Figure 1), a region where mostly interactions between the alkyl chains of the phospholipids and the non-polar region of Ibu^- are at play. Nicely, at the top of the barrier, an intermediate situation is observed.

It is quite reassuring that the observations just exposed seem to be independent of two crucial factors: on one hand, the choice of functional leads to very small changes in the natural atom charges (left panel in Figure 5), and on the other hand, computing natural atom charges keeping the basis set at the positions of the removed atoms leads to negligible differences (right panel in Figure 5).

(2) In order to have an overall picture of the interactions responsible for the preferred location of the drug in the model membrane, we subtracted the electron density of the participating molecules from the total density of the tertiary Ibu^-/membrane/water system for one of the snapshots at the minimum of the free energy profile. In the right panel of Figure 4, we provide a picture of the regions where electron density is accumulated around Ibu^- as a consequence of the interaction. Remarkably, small gains of electron density in the empty space (in the absence of interaction) between molecules are revealed, thus creating a sort of attractive wall of weak interactions.

Non-Covalent Interaction Index (NCI). Deep insight into the nature of intermolecular bonding interactions as a function of the position of the drug in the insertion process can be obtained by analyzing the non-covalent interaction index.[14,15,25] Plots for individual snapshots at each equilibrium position are included in Figure 6. By convention, blue surfaces

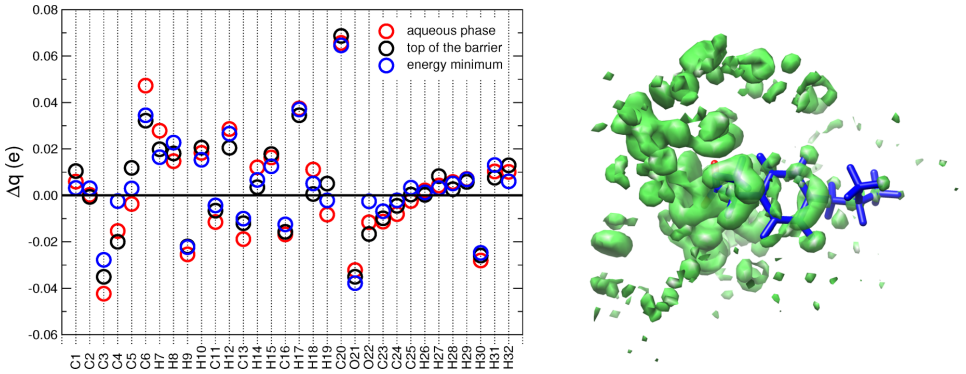


Figure 4. Left: Average natural relative charges on anionic Ibuprofen atoms (relative to isolated Ibu^-, solid line at $\Delta q = 0$) for the three equilibrium positions in the free energy profile (Figure 1) for the insertion of anionic Ibuprofen into model DMPC membranes. All charges were derived from CAM-B3LYP/6-31+G(d) computations on all available snapshots. See Figure 3 for atom numbers. Right: Difference in the electron density (only the surface corresponding to electron density gain is shown) from the tertiary Ibu^-/membrane/water system and the isolated components taking one particular snapshot at the energy minimum.

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

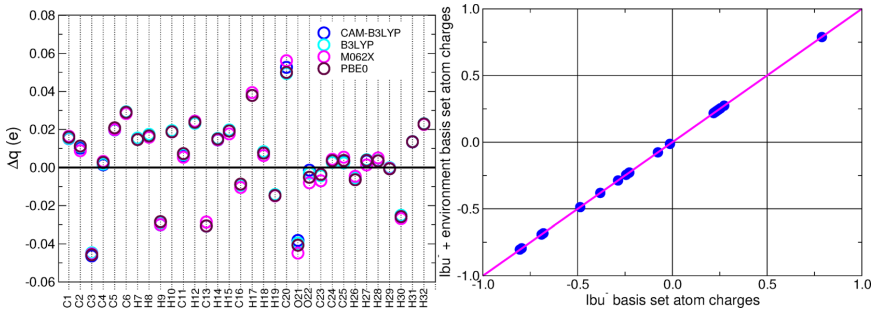


Figure 5. Effect of the functional and of the size of the basis set on the natural atom charges at Ibu^-. Left: Four different functionals are considered. Right: Natural atom charges calculated using the basis set centered at the positions of the removed atoms vs isolated Ibu^- at the CAM-B3LYP/6-31+G* level. A snapshot corresponding to the energy minimum in Figure 1 was randomly chosen.

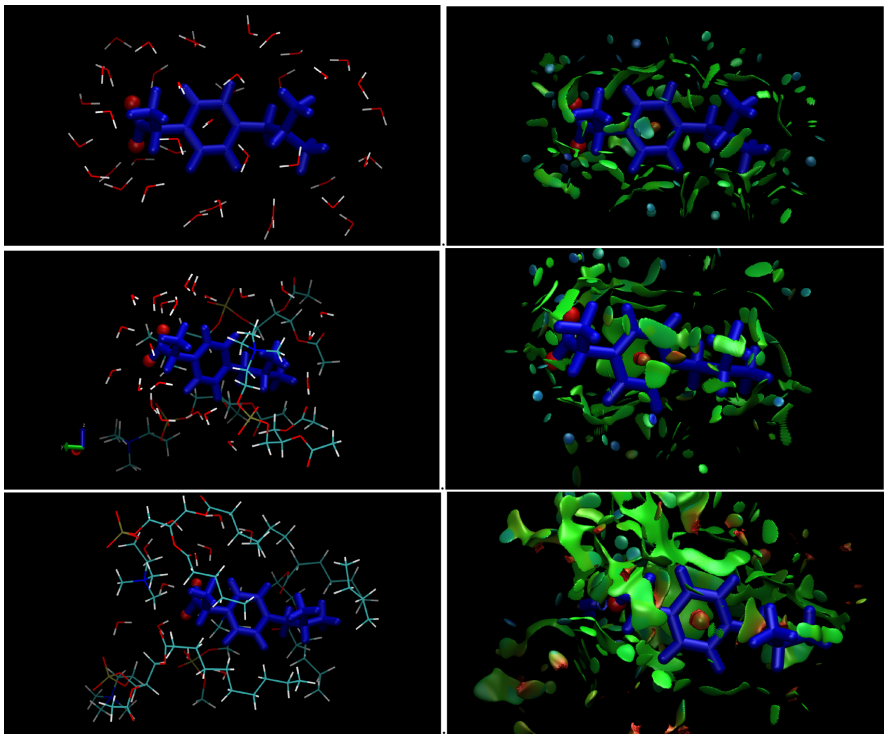


Figure 6. Non-covalent interactions (NCIs) surrounding anionic Ibuprofen in the aqueous phase (top), top of the barrier (middle), and energy minimum (bottom) of the free energy profile in Figure 1. The interacting molecules are shown in the left panels. All calculations were performed at the CAM-B3LYP/6-31+G(d) level.

are reserved for strong attractive interactions and green surfaces for weak dispersive interactions. A picture consistent with the above-mentioned attractive wall of weak interactions emerges: at the energy minimum of the free energy profile (Figure 1), Ibu^- is surrounded by a thicker dispersive attractive wall than at the top of the barrier, whose wall is in turn thicker than that in the aqueous phase. Again, despite the weakness of individual interactions, it is the cumulative effect which produces large stabilization energies at the bottom of the well. Interestingly, a wide attractive region arising from the interaction between the aromatic ring in Ibu^- and the methyl substituents from a choline group in a neighboring lipid is clearly discerned. Attractive non-covalent surfaces have been reported before, for example, for the interactions of methane with water cages.[30]

Natural Bond Orbital (NBO) Analysis. For the insertion of anionic Ibuprofen into model DMPC membranes, an inventory of orbital interactions may be established at this point. There are intramolecular orbital interactions within and among several groups of the same molecule, as well as

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

Table 2. Largest NBO Energies, $-E_{ij}^{(2)}$, in kcal/mol and Electron Densities, $\rho(r_c)$, in a.u. at the Bond Critical Points for All Interaction Types Found in a Randomly Chosen Frame of the MD Simulations at the Free Energy Minimum in Figure 1[a]

			$-E_{ij}^{(2)}$			$\rho(r_c)$			
label	fragments	orbitals involved	energy minimum	top of barrier	aqueous phase	BCP	energy minimum	top of barrier	aqueous phase
			Intramolecular						
1	$-CO_2^-$ in Ibu^-	$n_O \rightarrow \pi^*_{C=O}$	124.56	135.18	124.55	C═O	0.37	0.39	0.36
2	—COOR in DMPC	$n_O \rightarrow \pi^*_{C=O}$	76.92	72.72	N/A	C═O	0.34	0.34	N/A
3	—non-polar Ibu^-	$\pi_{C=C} \rightarrow \pi^*_{C=C}$	38.89	33.48	37.81	C═C	0.34	0.31	0.33
4	$-PO_2^-$ in DMPC	$n_O \rightarrow \sigma^*_{P=O}$	31.81	32.24	N/A	P═O	0.25	0.24	N/A
			Intermolecular						
5	$-CO_2^-$ in Ibu^- → water	$n_O \rightarrow \sigma^*_{O-H}$	29.19	21.37	21.36	O···H	0.07	0.05	0.05
6	water ↔ water	$n_O \rightarrow \sigma^*_{O-H}$	21.25	23.37	26.09	O···H	0.04	0.04	0.04
7	$-PO_2^-$ in DMPC → water	$n_O \rightarrow \sigma^*_{O-H}$	17.18	17.84	N/A	O···H	0.06	0.06	N/A
8	—COOR in DMPC → water	$n_O \rightarrow \sigma^*_{O-H}$	6.95	13.35	N/A	O···H	0.02	0.03	N/A
9	—choline in DMPC → water	$n_O \rightarrow \sigma^*_{C-H}$	5.08	3.97	N/A	O···H	0.02	0.02	N/A
10	$-CO_2^-$ in Ibu^- → choline in DMPC	$n_O \rightarrow \sigma^*_{C-H}$	5.02	0.38	N/A	O···H	0.03	0.01	N/A
11	$-CO_2^-$ in Ibu^- → choline in DMPC	$\pi_{C=O} \rightarrow \sigma^*_{C-H}$	1.23	1.72	N/A	O···H	0.01	0.01	N/A
12	non-polar Ibu^- → non-polar DMPC	$\sigma_{C-H} \rightarrow \sigma^*_{C-H}$	0.95	0.45	N/A	H···H	0.01	0.01	N/A
13	non-polar Ibu^- → non-polar DMPC	$\pi_{C=C} \rightarrow \sigma^*_{C-H}$	0.17	N/A	N/A	C···H	<0.01	N/A	N/A
14	choline in DMPC → non-polar Ibu^-	$\sigma_{C-H} \rightarrow \pi^*_{C=C}$	0.16	0.12	N/A	C···H	0.01	0.01	N/A
15	non-polar Ibu^- → choline in DMPC	$\pi_{C=C} \rightarrow \sigma^*_{C-H}$	0.27	0.58	N/A	C···H	0.01	0.01	N/A
16	non-polar Ibu^- → water	$\pi_{C=C} \rightarrow \sigma^*_{O-H}$	N/A	0.05	7.18	C···H	N/A	0.01	0.02
17	non-polar Ibu^- → water	$\sigma_{C-H} \rightarrow \sigma^*_{O-H}$	0.09	0.33	0.17	H···H	<0.01	0.01	0.01
18	water → non-polar Ibu^-	$n_O \rightarrow \sigma^*_{C-H}$	0.06	1.12	1.60	O···H	<0.01	0.01	0.01
19	water → non-polar Ibu^-	$n_O \rightarrow \sigma^*_{C-H}$	N/A	0.34	1.28	O···H	N/A	<0.01	0.01

[a]Figure 7 shows the specific involved orbitals in each case. Dotted lines represent intermolecular interactions. All calculations were done at the CAM-B3LYP/6-31+G* level. See Table S1 in the Supporting Information for calculations with other functionals.

intermolecular interactions, which include water ↔ water, water ↔ Ibu^-, water ↔ lipid, Ibu^- ↔ lipid, and lipid ↔ lipid. Intermolecular interactions may be further divided into polar ↔ polar, polar ↔ non-polar, and non-polar ↔ non-polar types. To illustrate the relative strengths of all of these interactions from the point of view of donor → acceptor orbital interactions according to NBO, we chose one random frame from each equilibrium position of the free energy profile in Figure 1. In Table 2, the largest $E_{ij}^{(2)}$ values found in each case are listed.

For the three equilibrium positions along the free energy profile in Figure 1, according to NBO, the largest donor → acceptor orbital interaction energies are of intramolecular type. This is a very insightful result because, on one hand, it shows severe local changes in the electronic configurations of both Ibu^- and the lipids during the insertion process and, on the other hand, it shows that those changes arise from the interaction with water molecules! (The five largest intermolecular interactions involve solvent molecules.) In this context, electron activity in the lone pairs of the $-CO_2^-$ group appears as the dominant orbital factor during the insertion process. Notice that, in all cases, large intramolecular interaction energies are associated with local delocalization of electrons within specific functional groups of the system (see Table 2 and Figure 7). For example, the non-polar Ibu^- $\pi_{C=C} \rightarrow \pi^*_{C=C}$ interactions (interaction 3 in Table 2) just describe delocalization of the π cloud within the aromatic ring.

It is interesting to notice the effect of the formal charge on the interactions.[31−38] Take, for example, the $n_O \rightarrow \sigma^*_{O-H}$ orbital interaction responsible for hydrogen bonding in the isolated water dimer: the $E_{ij}^{(2)}$ interaction energy has been reported to be −5.40 kcal/mol.[39] Our results indicate that, when this hydrogen bond is assisted by the formal charges in the environment (interaction 6, Table 2), $n_O \rightarrow \sigma^*_{O-H}$ interaction energies among water molecules in the solvent are considerably larger (−26.09, −23.37, and −21.25 for the aqueous phase, top of the barrier, and energy minimum, respectively). This charge induced strengthening translates to other intermolecular interactions as well: for the $-PO_2^-$··· water contacts, donor → acceptor orbital interaction energies do not exceed 14 kcal/mol in the microsolvation of isolated dimethylphosphate.[34] Here, those values rise up to ∼18 kcal/mol at the top of the barrier (interaction 7 in Table 2). The charge assisted strengthening of intermolecular interactions has already been reported for the microsolvation of charged species.[31−38]

A very interesting scenario of large stabilization effects resulting from the collective action of a large number of small individual interactions emerges. Comparatively, the $\sigma_{C-H} \rightarrow \sigma^*_{C-H}$ charge transfer from the non-polar region of Ibu^- (mostly from the isobutyl group) to the aliphatic chain in DMPC is quite small (0.95 and 0.45 kcal/mol for the energy minimum and at the top of the barrier, respectively, as seen in interaction 12, Table 2). $\pi_{C=C} \rightarrow \sigma^*_{C-H}$ charge transfer from the aromatic ring in Ibu^- to the lipid chains is even smaller (0.27 and 0.17 kcal/mol, interactions 13 and 15 in Table 2). Thus, despite the small magnitudes of individual interactions originating in the non-polar region of Ibu^- with the environment, the preferred location for the drug within the membrane after insertion leads to a situation where this large number of orbital interactions is the dominant stabilizing factor. This observation is directly tied to the non-covalent wall (see above) stabilizing the tertiary system shown in Figure 6.

We have already argued that electron activity within the $-CO_2^-$ group in Ibu^- appears as the dominant orbital factor during the insertion process. Accordingly, we analyze the role

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

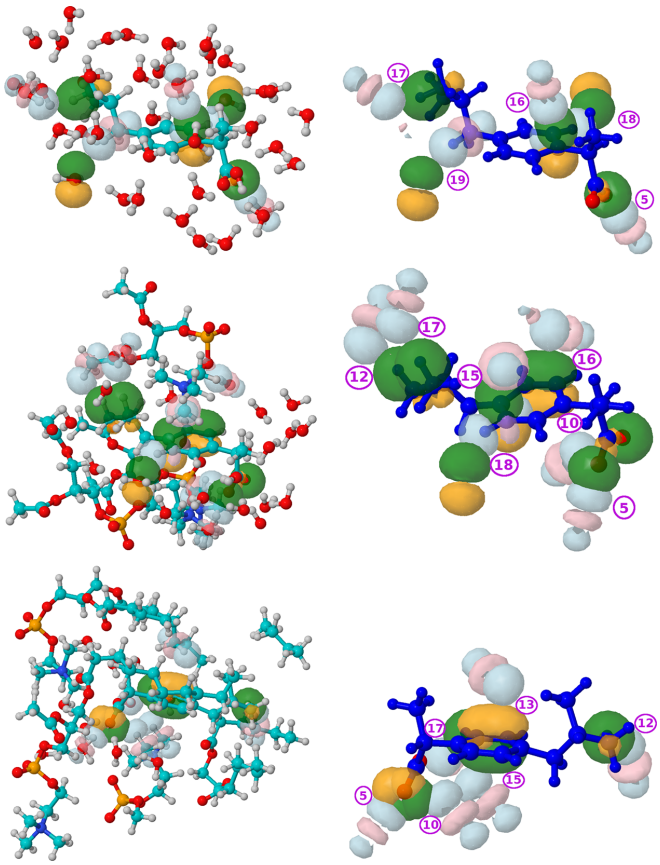


Figure 7. Assorted orbital interactions leading to intermolecular contacts in randomly chosen frames of the MD simulations for the aqueous phase (top), top of the barrier (middle), and energy minimum (bottom) in the free energy profile in Figure 1 for the insertion of anionic Ibuprofen into model DMPC membranes. The entire system is shown in the left, while only Ibu^- is shown in the right. See Table 2 for interaction labels (circled numbers). Green/orange surfaces correspond to donor, bonding orbitals. Light blue/pink surfaces correspond to acceptor, antibonding orbitals.

of the electronic structure of $—CO_2^-$ along the insertion path using the collection of plots shown in Figure 8 as follows:

(1) In all cases, the orange line describing charge transfer associated with the $—CO_2^-$ group in isolated Ibuprofen sits above the corresponding interactions for all environments. Therefore, although it remains quite important, charge delocalization in $—CO_2^-$ is a diminishing factor along the insertion path. This observation is fully consistent with the conclusions drawn by Zapata-Escobar and co-workers[35] in the sense that, since $—CO_2^-$ leads to two charge assisted hydrogen bonds of unequal strengths, explicit interactions between Ibu^- and individual solvent molecules have a localizing effect on the formal charge along one of the C═O bonds.

(2) The fact that the formal charge is slightly more localized in one of the two C═O bonds of the $—CO_2^-$ group is neatly supported by the data in Figure 4, where an asymmetry in the charge distribution of the two oxygens is observed, thus, on average, O21 becomes more negative than O22 along the insertion path.

(3) By far, the largest intramolecular charge transfer (and the largest deviations from the isolated Ibu^- case) occurs from the lone pairs in the oxygen atoms (O21, O22) of the $—CO_2^-$ group to the π^* orbitals involved in C20—O21 and C20—O22 (top left panel in Figure 8). Second in magnitude are the collective charge transfers from the same lone pairs to the σ^* orbitals also involved in C20—O21 and C20—O22 (top middle and right panels in Figure 8). Therefore, despite the localizing effect just mentioned, a high degree of delocalization of the formal charge within the $—CO_2^-$ group is still observed. This holds regardless of the equilibrium position for all frames chosen.

(4) The C11—C20 bond connecting the $—CO_2^-$ group to the aromatic ring (see Figure 3) also plays a significant role in the redistribution of the electron density during the insertion process. Indeed, charge transfer to the $\sigma^*_{C11—C20}$ from the lone pairs in $—CO_2^-$ is significant, although comparatively smaller (bottom left and middle panels in Figure 8) than charge transfer to the two σ^* orbitals involved in C20—O21 and C20—O22.

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

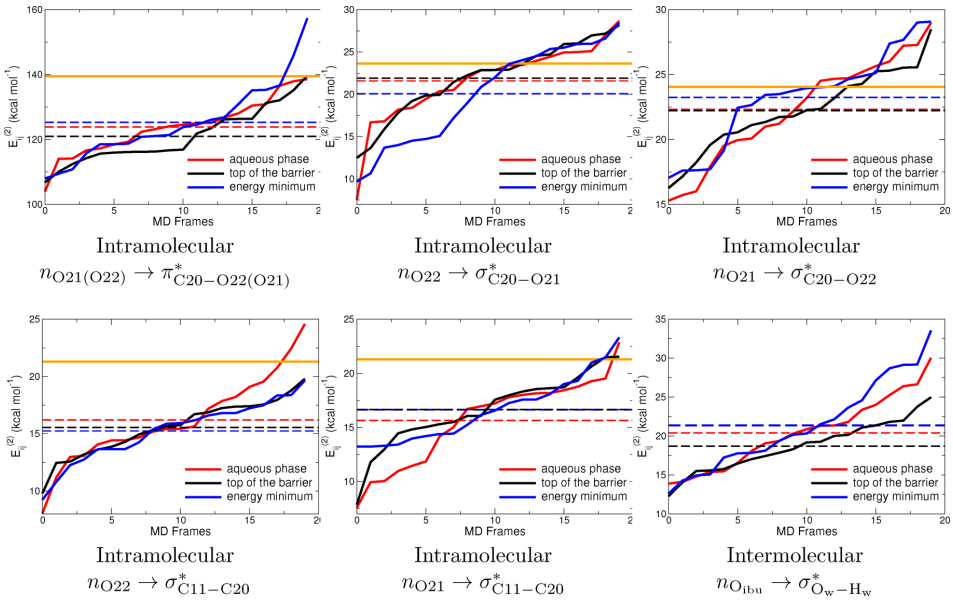


Figure 8. Energies for the donor → acceptor orbital interactions associated with the —CO_2^- group during the insertion process of Ibu^- into model DMPC membranes. Dashed horizontal lines indicate the averages over the 20 MD frames chosen for each equilibrium position in the free energy profile in Figure 1. For the intramolecular cases, the reference value for the corresponding orbital interaction at the isolated Ibu^- is represented by a solid horizontal line in each case. See Figure 3 for atom numbers.

(5) From the perspective of donor → acceptor orbital charge transfer, the interactions just analyzed, leading to electron delocalization in the —CO_2^- group, are comparatively as strong as the charge assisted hydrogen bonds between Ibu^- and water molecules (bottom right panel, Figure 4).

(6) Ibu^-···water interactions appear stronger at the energy minimum (bottom right panel, Figure 4); thus, it may be argued that, despite the somewhat smaller number of water molecules surrounding the drug as compared against the aqueous phase and against the top of the barrier, a not negligible portion of the stabilization energy at the minimum arises from these contacts.

Calorimetric titration experiments showed that there are no appreciable changes in enthalpy for increasing amounts (within the experimental limits) of Ibu^- added to model cell membranes after equilibration. Thus, the insertion process was postulated to be governed by entropy.[23] Recently, Rojas-Valencia and co-workers[5] decomposed the calculated free energy profiles for the insertion process into their entropy and enthalpy components, providing solid evidence to back up and rationalize that idea. Here, Table 2 shows that reorganization of the electron density in the —CO_2^- group of Ibu^- (largest amounts of charge transferred from and to its orbitals) is the major player in the interactions with the environment. Accordingly, we calculate the total intermolecular interaction energy between the NBO orbitals in —CO_2^- and the NBO orbitals in the environment (water + membrane) and list them in Table 3 along the relevant energy quantities.

Since charge transfer (via orbital interaction energies) constitutes a sizable chunk of the internal energies, the numbers in Table 3 are quite revealing: the fact that the averaged sums running over all frames for the orbital interaction energies of the —CO_2^- group with the environment behave in the exact same qualitative fashion as the changes in enthalpy along the insertion path leads us to argue that orbital interaction energies effectively contribute to a large degree to enthalpy changes during the insertion process. The fact that a mismatch occurs only at the top of the barrier indicates that at the transition state other interactions are not negligible.

Table 3. Differences in Gibbs Free Energies (ΔG), Entropies (ΔS), Enthalpies (ΔH), and Total Intermolecular (—CO_2^- with the Environment) Donor → Acceptor Orbital Interaction Energies (the Aqueous Phase Is Taken as a Reference)[a]

location	ΔG	$-T\Delta S$	ΔH	$\Delta(\sum E_{ij}^{(2)})$
aqueous phase	0	0	0	0
top of the barrier	5	−109	115	30
energy minimum	−15	−48	34	24

[a] $\sum E_{ij}^{(2)}$ is actually averaged over all 20 MD snapshots taken at each of the three equilibrium positions in the free energy profile for insertion of Ibu^- into a model DMPC membrane (Figure 1). All energies are in kJ mol^{-1} taking the aqueous phase as a reference.

Topology of the Electron Densities. A number of criteria to derive insight into the nature of bonding interactions from the properties of the electron density at bond critical points have been developed and are thoroughly described in the literature.[8,11,40−42] Following the procedure described above, we analyze intra- and intermolecular interactions in the randomly selected frames (same frames as those for the NBO and NCI analyses) of the MD simulations for each equilibrium position in the free energy profile for the insertion of Ibu^- into model cell membranes (Figure 1).

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

Table 2 lists the accumulation of electron densities around the bond critical points for the selected interactions that maximize $E_{ij}^{(2)}$ (see above). It is clear that, for formal bonds involved in all intramolecular interactions, electron densities around the corresponding BCPs are one entire order of magnitude larger than electron densities surrounding BCPs for the comparatively weaker intermolecular interactions. In the same line, primary $n_O \rightarrow \sigma^*_{O-H}$ hydrogen bonds accumulate 2−3 times as much electron density around the O···H—O BCPs than secondary hydrogen bonds where the recipient of electron density is a σ^*_{C-H} orbital. This hierarchy of interaction energies is fully consistent with the strengths of the interactions. For microsolvation of charged species,[31−38] it has been shown that hydrogen bonding among water molecules is strengthened because of the action of the formal charge when compared against pure water clusters.[43−47] Here, water molecules in contact with the highly locally charged molecules suffer the same effect; that is, electron densities around BCPs for water to water hydrogen bonding are larger than those for pure water clusters. Indeed, H_2O···H—O—H HBs (interaction 6, Table 2) almost double the 2.3×10^{-2} a.u. reported for the isolated water dimer.[33] Nicely, $E_{ij}^{(2)}$ orbital interaction energies (Table 2) provide a fully compatible picture with the QTAIM results,[48] not only regarding the ordering of interaction energies but also describing the strengthening effect of the formal charge in the surrounding network of hydrogen bonds.

Figure 9 shows the intricate molecular graphs that include all bond paths and the associated bond critical points for each case. For the chosen frames, a systematic increase in the number of intermolecular contacts between the non-polar groups of Ibu^- and the environment (more specifically, with the aliphatic tails of the lipids) is clearly seen as the drug traverses the membrane from the aqueous phase to its equilibrium position at the energy minimum. As pointed out in the NCI and NBO sections, the collective action of these interactions provides the large (largest along the insertion path) stabilization energy at the bottom of the energy well and creates the attractive wall of weak interactions surrounding Ibu^-.

Without exception, for all bond critical points found in this work associated with intermolecular interactions, $\nabla^2\rho(\mathbf{r}_c) > 0$. Thus, they represent local minima in the corresponding electron density. Accordingly, there is local depletion of charge in the vicinities of intermolecular BCPs and concentration of electron charge toward the nuclei originating the bonding path. This situation describes non-covalent, long-range interactions. An assortment of other properties calculated at bond critical points that yield useful information about the nature of bonding interactions are collected in the plots of Figure 10.

We found exponential decays of the electron densities at BCPs as a function of the distance between the two atoms bridged by the corresponding bond path (plots at the left of Figure 10).[45,49−52] Interestingly, this exponential decay (with different trend lines for different interactions) is obeyed regardless of the position of Ibu^- along the insertion path and regardless of the type of interaction. Exponential decays of electron densities at BCPs as a function of separation between the interacting atoms make good physical sense because they correctly describe that electron densities vanish asymptotically for larger distances. It is quite insightful that intermolecular interactions between the non-polar groups in Ibu^- and non-polar groups in the lipids (bottom row, Figure 10) are connected by the same trend line. These non-polar···non-polar contacts are well characterized by bonding paths with the corresponding BCPs, which do not belong to either one of the fragments. Rather, the collective accumulation of these small electron densities in the region separating the fragments are the physical source of what we call attractive walls, already characterized by the gain in electron density due to the

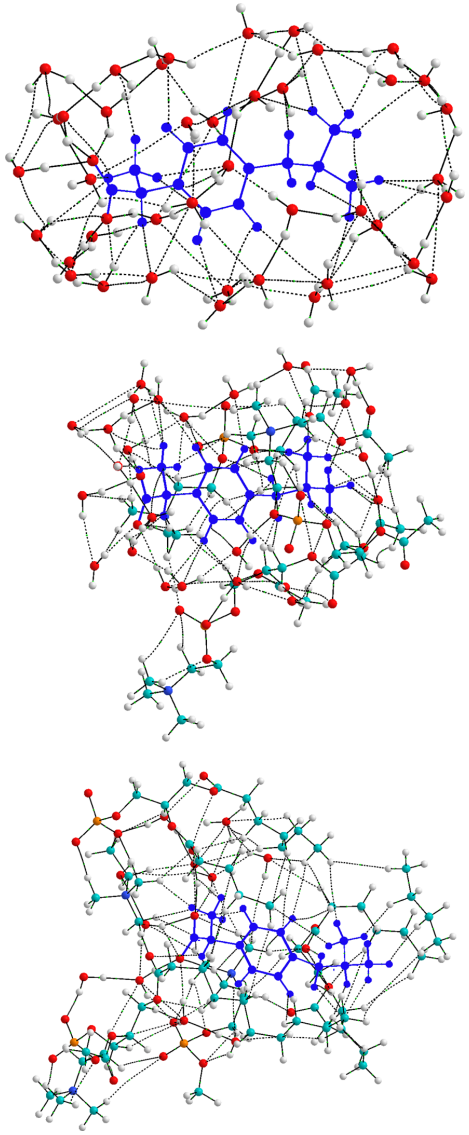


Figure 9. Bonding paths and bond critical points for randomly selected frames (same frames used in Figures 5, 6, and 7 and in Table 2) of the MD simulations at the equilibrium positions in the free energy profile of Figure 1. Top: Aqueous phase. Middle: Top of the barrier. Bottom: Energy minimum. In all cases, Ibu^- is highlighted in dark blue, except for the oxygen atoms in $—CO_2^-$, which are highlighted in red.

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

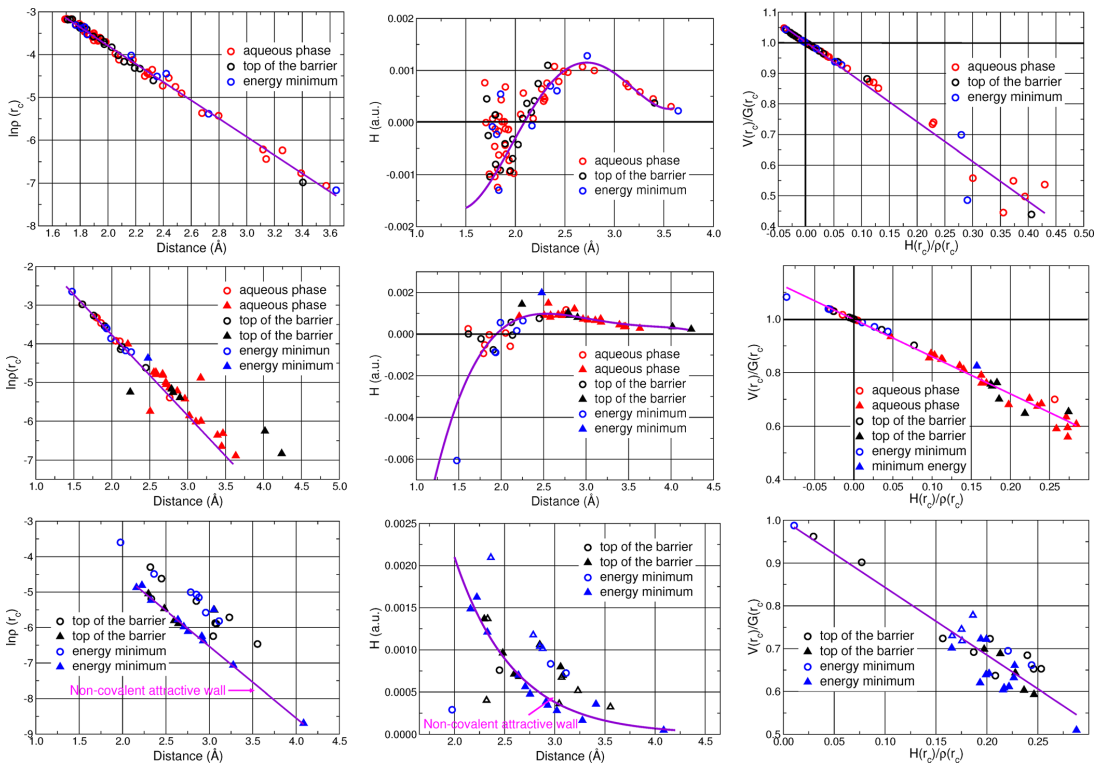


Figure 10. Properties of the bond critical points for intermolecular interactions in randomly selected frames (same frames as in the NCI (Figure 6), NBO (Figure 7), and QTAIM (Figure 9) analyses). Top: Water to water interactions. Middle: Ibu^- to water interactions. Bottom: Ibu^- to lipid interactions. In all cases, open symbols refer to interactions between the polar region of Ibu^- and the environment, while filled symbols refer to interactions between the non-polar region of Ibu^- and the environment.

interaction between fragments (Figure 4) and in the non-covalent surfaces (Figure 6).

Energy densities are dimensionally equivalent to pressure.[11] Thus, local negative pressures correspond to regions that suck electrons in. Accordingly, negative total energy densities at BCPs are equated to covalent bonding. Conversely, local positive pressures force electrons away from the BCPs and are tied to long-range interactions. In molecular systems, electronic kinetic energies are always positive and repulsive, while electronic potential energies are always negative and attractive. Thus, at BCPs, the sign of the total energy density, $\mathcal{H} = \mathcal{G} + \mathcal{V}$, which is the result of a local tug of war for dominance between the two terms, provides useful information about the nature of bonding interactions. A more quantitative criterion for the strength of bonding interactions, derived from local application of the virial theorem at BCPs, is found in the work by Espinosa and co-workers.[41] There, the $|\mathcal{V}|/\mathcal{G}$ ratio describes bonding interactions according to the following scheme

$$\frac{|\mathcal{V}(\mathbf{r}_c)|}{\mathcal{G}(\mathbf{r}_c)} = \begin{cases} <1 & \text{closed shell (ionic, long range, etc.)} \\ \in[1,2] & \text{intermediate} \\ >2 & \text{covalent} \end{cases} \quad (1)$$

In this context, we calculated total energy densities and $|\mathcal{V}|/\mathcal{G}$ ratios for all BCPs in the chosen frames and plotted them in the middle and right panels of Figure 10, respectively. The first general conclusion drawn from these plots and holding for the three equilibrium positions is that there is not a single covalent intermolecular interaction. Moreover, because of the positive energy densities and because of the $|\mathcal{V}|/\mathcal{G} < 1$ ratios, most of the intermolecular interactions are classified as long-range, with just a few exceptions, mostly found in the aqueous phase. Recall that from the NBO analysis we found an unusually strong contact (interaction 5 in Table 2) for the chosen frame at the energy minimum, which is nicely reproduced in having both the most negative energy density and the largest $|\mathcal{V}|/\mathcal{G}$ ratio. These strong Ibu^-···H—O—H contacts are consistent with the detailed characterization of bonding interactions in the binary system in the absence of the lipids found in the work by Zapata-Escobar and co-workers.[35] It is not surprising that, at the energy minimum, all Ibu^-··· lipid interactions, including those arising from the $—CO_2^-$ group, are characterized as long-range, weak contacts. Finally, notice that all interactions leading to the formation of non-covalent attractive walls are consistently weaker that the rest.

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

SUMMARY AND CONCLUSIONS

Insertion of anionic Ibuprofen into cell membranes is a complex process. Here, we use five different analysis tools based on the formalism of quantum mechanics (natural charges, electron density differences, atoms in molecules (AIM), natural bond orbitals (NBO), and non-covalent interactions (NCI) index) to gain insight into the nature and into the evolution of chemical bonding as the drug traverses model DMPC membranes from the aqueous phase to its preferred location at the interface between the polar/non-polar regions of the lipid bilayer. All descriptors provide a consistent picture, where minute amounts of electron density are transferred, because of individual intermolecular contacts, to the interstitial region between fragments. The collective action of a large number of these individual intermolecular contacts originates a wall of weak, non-covalent interactions surrounding Ibu^-, providing stabilization for the entire tertiary water/Ibu^-/membrane system. The formal charge in anionic Ibuprofen seems to be a dominant factor during the insertion process because, on one hand, it strengthens the hydrogen bond network among the solvent molecules and because of a sensible charge redistribution within the $—CO_2^-$ group, on the other. In this context, electron delocalization in $—CO_2^-$ and in the aromatic ring is reduced in going from isolated Ibu^- to the solvated phases. Our calculations show that orbital interaction energies of the $—CO_2^-$ group with the environment behave in the exact same qualitative fashion as the changes in enthalpy along the insertion path. Thus, since charge transfer (via orbital interaction energies) constitutes a sizable chunk of the internal energy, we argue that orbital interaction energies effectively contribute to a large degree to enthalpy changes during the insertion process.

We think that a fundamental contribution of this work is that we conclusively show that for the insertion of anionic Ibuprofen into model cell membranes, despite their purely classical origin, randomly chosen configurations from MD simulations provide deep insight into the purely quantum nature of bonding interactions. Furthermore, this is true despite those random configurations being representative of highly dynamic situations, not necessarily corresponding to well-defined minima within a given potential energy surface (classical or quantum). We see no reason why this approach of randomly choosing configurations from classical mechanics simulations to study bonding interactions should not translate to other systems; thus, we anticipate that this may become a general practice.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.9b09705.

Cartesian coordinates for the randomly chosen snapshots for which bonding descriptors were calculated at each equilibrium position in the free energy profile for the insertion of Ibu^- into model DMPC membranes (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: albeiro.restrepo@udea.edu.co.

ORCID

Sara Gómez: 0000-0002-5430-9228
Chiara Cappelli: 0000-0002-4872-4505
Cacier Hadad: 0000-0002-2799-301X
Albeiro Restrepo: 0000-0002-7866-7791

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Partial funding for this project from H2020-MSCA-ITN-2017 European Training Network "Computational Spectroscopy In Natural sciences and Engineering" (COSINE), grant number 765739, is acknowledged. We also thank Colciencias for partial financial support for this work, grant 111571249844, contract 378−2016. Internal support from Universidad de Antioquia via "Estrategia para la sostenibilidad" is also acknowledged. N.R.-V. thanks Colciencias for her doctoral scholarship.

REFERENCES

(1) Conaghan, P. G. A turbulent decade for NSAIDs: update on current concepts of classification, epidemiology, comparative efficacy, and toxicity. *Rheumatol. Int.* **2012**, *32*, 1491−1502.
(2) Lichtenberger, L. M.; Zhou, Y.; Dial, E. J.; Raphael, R. M. NSAID injury to the gastrointestinal tract: evidence that NSAIDs interact with phospholipids to weaken the hydrophobic surface barrier and induce the formation of unstable pores in membranes. *J. Pharm. Pharmacol.* **2006**, *58*, 1421−1428.
(3) Lichtenberger, L. M.; Zhou, Y.; Jayaraman, V.; Doyen, J. R.; O'Neil, R. G.; Dial, E. J.; Volk, D. E.; Gorenstein, D. G.; Boggara, M. B.; Krishnamoorti, R. Insight into NSAID-induced membrane alterations, pathogenesis and therapeutics: Characterization of interaction of NSAIDs with phosphatidylcholine. *Biochim. Biophys. Acta, Mol. Cell Biol. Lipids* **2012**, *1821*, 994−1002.
(4) Wallace, J. L. Prostaglandins, NSAIDs, and Gastric Mucosal Protection: Why Doesn't the Stomach Digest Itself? *Physiol. Rev.* **2008**, *88*, 1547−1565.
(5) Rojas-Valencia, N.; Lans, I.; Manrique-Moreno, M.; Hadad, C. Z.; Restrepo, A. Entropy drives the insertion of ibuprofen into model membranes. *Phys. Chem. Chem. Phys.* **2018**, *20*, 24869−24876.
(6) Boggara, M. B.; Mihailescu, M.; Krishnamoorti, R. Structural Association of Nonsteroidal Anti-Inflammatory Drugs with Lipid Membranes. *J. Am. Chem. Soc.* **2012**, *134*, 19669−19676.
(7) Glendening, E. D.; Landis, C. R.; Weinhold, F. Natural bond orbital methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 1−42.
(8) Bader, R. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, U.K., 1990.
(9) Bader, R. F. W. A quantum theory of molecular structure and its applications. *Chem. Rev.* **1991**, *91*, 893−928.
(10) Bader, R. F. W. The Quantum Mechanical Basis of Conceptual Chemistry. *Monatsh. Chem.* **2005**, *136*, 819−854.
(11) Popelier, P. L. *Atoms in Molecules: An Introduction*; Prentice Hall: London, 2000.
(12) Popelier, P. On the full topology of the Laplacian of the electron density. *Coord. Chem. Rev.* **2000**, *197*, 169−189.
(13) Becke, A. *The quantum theory of atoms in molecules: from solid state to DNA and drug design*; John Wiley & Sons: 2007.
(14) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498−6506.
(15) DiLabio, G. A.; Otero-de-la Roza, A. *Reviews in Computational Chemistry*; John Wiley & Sons, Ltd: 2016; Chapter 1, pp 1−97.
(16) Cappelli, C. Integrated QM/polarizable MM/continuum approaches to model chiroptical properties of strongly interacting solute−solvent systems. *Int. J. Quantum Chem.* **2016**, *116*, 1532−1542.

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

(17) Etienne, T.; Very, T.; Perpète, E. A.; Monari, A.; Assfeld, X. A QM/MM Study of the Absorption Spectrum of Harmane in Water Solution and Interacting with DNA: The Crucial Role of Dynamic Effects. *J. Phys. Chem. B* **2013**, *117*, 4973−4980.
(18) Egidi, F.; Lo Gerfo, G.; Macchiagodena, M.; Cappelli, C. On the nature of charge-transfer excitations for molecules in aqueous solution: a polarizable QM/MM study. *Theor. Chem. Acc.* **2018**, *137*, 82.
(19) Giovannini, T.; Del Frate, G.; Lafiosca, P.; Cappelli, C. Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution. *Phys. Chem. Chem. Phys.* **2018**, *20*, 9181−9197.
(20) Giovannini, T.; Macchiagodena, M.; Ambrosetti, M.; Puglisi, A.; Lafiosca, P.; Lo Gerfo, G.; Egidi, F.; Cappelli, C. Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling. *Int. J. Quantum Chem.* **2019**, *119*, No. e25684.
(21) Puglisi, A.; Giovannini, T.; Antonov, L.; Cappelli, C. Interplay between conformational and solvent effects in UV-visible absorption spectra: curcumin tautomers as a case study. *Phys. Chem. Chem. Phys.* **2019**, *21*, 15504−15514.
(22) Cwiklik, L.; Aquino, A. J. A.; Vazdar, M.; Jurkiewicz, P.; Pittner, J.; Hof, M.; Lischka, H. Absorption and Fluorescence of PRODAN in Phospholipid Bilayers: A Combined Quantum Mechanics and Classical Molecular Dynamics Study. *J. Phys. Chem. A* **2011**, *115*, 11428−11437.
(23) Manrique-Moreno, M.; Garidel, P.; Suwalsky, M.; Howe, J.; Brandenburg, K. The membrane-activity of Ibuprofen, Diclofenac, and Naproxen: A physico-chemical study with lecithin phospholipids. *Biochim. Biophys. Acta, Biomembr.* **2009**, *1788*, 1296−1303.
(24) Weinhold, F.; Landis, C. R. *Discovering Chemistry with Natural Bond Orbitals*; Wiley-VCH: Hoboken NJ, 2012.
(25) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. NCIPLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theory Comput.* **2011**, *7*, 625−632.
(26) Frisch, M. J.; et al. *Gaussian 09*, revision E.01; Gaussian, Inc.: Wallingford, CT, 2009.
(27) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange−correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51−57.
(28) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Landis, C. R.; Weinhold, F. *NBO 6.0*; University of Wisconsin: Madison, WI, 2013.
(29) Keith, T. *AIMALL*, version 13.05.06); 2013; aim.tkgristmill.com.
(30) Salazar-Cano, J.-R.; Guevara-García, A.; Vargas, R.; Restrepo, A.; Garza, J. Hydrogen bonds in methane−water clusters. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23508−23515.
(31) Hadad, C.; Florez, E.; Acelas, N.; Merino, G.; Restrepo, A. Microsolvation of small cations and anions. *Int. J. Quantum Chem.* **2019**, *119*, No. e25766.
(32) Flórez, E.; Acelas, N.; Ramírez, F.; Hadad, C.; Restrepo, A. Microsolvation of F^-. *Phys. Chem. Chem. Phys.* **2018**, *20*, 8909−8916.
(33) Flórez, E.; Acelas, N.; Ibargüen, C.; Mondal, S.; Cabellos, J. L.; Merino, G.; Restrepo, A. Microsolvation of NO_3^-: structural exploration and bonding analysis. *RSC Adv.* **2016**, *6*, 71913−71923.
(34) Rojas-Valencia, N.; Ibargüen, C.; Restrepo, A. Molecular interactions in the microsolvation of dimethylphosphate. *Chem. Phys. Lett.* **2015**, *635*, 301−305.
(35) Zapata-Escobar, A.; Manrique-Moreno, M.; Guerra, D.; Hadad, C. Z.; Restrepo, A. A combined experimental and computational study of the molecular interactions between anionic ibuprofen and water. *J. Chem. Phys.* **2014**, *140*, 184312.
(36) Romero, J.; Reyes, A.; David, J.; Restrepo, A. Understanding microsolvation of Li^+: structural and energetical analyses. *Phys. Chem. Chem. Phys.* **2011**, *13*, 15264−15271.
(37) Gonzalez, J. D.; Florez, E.; Romero, J.; Reyes, A.; Restrepo, A. Microsolvation of Mg^{2+}, Ca^{2+}: strong influence of formal charges in hydrogen bond networks. *J. Mol. Model.* **2013**, *19*, 1763−1777.
(38) Acelas, N.; Flórez, E.; Hadad, C.; Merino, G.; Restrepo, A. A Comprehensive Picture of the Structures, Energies, and Bonding in $[SO_4(H_2O)_n]^{2-}$, n = 1−6. *J. Phys. Chem. A* **2019**, *123*, 8650−8656.
(39) Reed, A. E.; Weinhold, F. Natural bond orbital analysis of near-Hartree−Fock water dimer. *J. Chem. Phys.* **1983**, *78*, 4066−4073.
(40) Grabowski, S. J. What Is the Covalency of Hydrogen Bonding? *Chem. Rev.* **2011**, *111*, 2597−2625.
(41) Espinosa, E.; Alkorta, I.; Elguero, J.; Molins, E. From weak to strong interactions: A comprehensive analysis of the topological and energetic properties of the electron density distribution involving X−H···F−Y systems. *J. Chem. Phys.* **2002**, *117*, 5529−5542.
(42) Cremer, D.; Kraka, E. A description of the chemical bond in terms of local properties of electron density and energy. *Croat. Chem. Acta* **1984**, *57*, 1259−1281.
(43) Gómez, S.; Nafziger, J.; Restrepo, A.; Wasserman, A. Partition-DFT on the water dimer. *J. Chem. Phys.* **2017**, *146*, 074106.
(44) Pérez, J. F.; Hadad, C. Z.; Restrepo, A. Structural studies of the water tetramer. *Int. J. Quantum Chem.* **2008**, *108*, 1653−1659.
(45) Ramírez, F.; Hadad, C.; Guerra, D.; David, J.; Restrepo, A. Structural studies of the water pentamer. *Chem. Phys. Lett.* **2011**, *507*, 229−233.
(46) Hincapié, G.; Acelas, N.; Castaño, M.; David, J.; Restrepo, A. Structural Studies of the Water Hexamer. *J. Phys. Chem. A* **2010**, *114*, 7809−7814.
(47) Acelas, N.; Hincapié, G.; Guerra, D.; David, J.; Restrepo, A. Structures, energies, and bonding in the water heptamer. *J. Chem. Phys.* **2013**, *139*, 044310.
(48) Weinhold, F. Natural bond critical point analysis: Quantitative relationships between natural bond orbital-based and QTAIM-based topological descriptors of chemical bonding. *J. Comput. Chem.* **2012**, *33*, 2440−2449.
(49) Farfán, P.; Echeverri, A.; Diaz, E.; Tapia, J. D.; Gómez, S.; Restrepo, A. Dimers of formic acid: Structures, stability, and double proton transfer. *J. Chem. Phys.* **2017**, *147*, 044312.
(50) Alkorta, I.; Rozas, I.; Elguero, J. Bond Length−Electron Density Relationships: From Covalent Bonds to Hydrogen Bond Interactions. *Struct. Chem.* **1998**, *9*, 243−247.
(51) Knop, O.; Rankin, K. N.; Boyd, R. J. Coming to Grips with N−H···N Bonds. 1. Distance Relationships and Electron Density at the Bond Critical Point. *J. Phys. Chem. A* **2001**, *105*, 6552−6566.
(52) Knop, O.; Rankin, K. N.; Boyd, R. J. Coming to Grips with N−H···N Bonds. 2. Homocorrelations between Parameters Deriving from the Electron Density at the Bond Critical Point. *J. Phys. Chem. A* **2003**, *107*, 272−284.

DOI: 10.1021/acs.jpcb.9b09705
J. Phys. Chem. B 2020, 124, 79−90

THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB Article

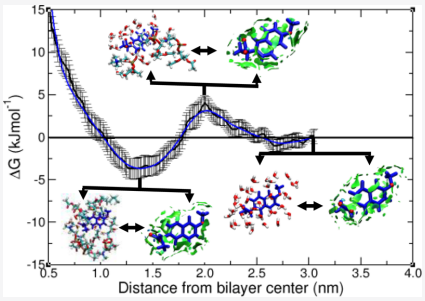
Thermodynamics and Intermolecular Interactions during the Insertion of Anionic Naproxen into Model Cell Membranes

Natalia Rojas-Valencia,* Sara Gómez, Francisco Núñez-Zarur, Chiara Cappelli, Cacier Hadad, and Albeiro Restrepo*

Cite This: J. Phys. Chem. B 2021, 125, 10383−10391 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The insertion process of Naproxen into model dimyristoylphosphatidylcholine (DMPC) membranes is studied by resorting to state-of-the-art classical and quantum mechanical atomistic computational approaches. Molecular dynamics simulations indicate that anionic Naproxen finds an equilibrium position right at the polar/nonpolar interphase when the process takes place in aqueous environments. With respect to the reference aqueous phase, the insertion process faces a small energy barrier of ≈5 kJ mol^{-1} and yields a net stabilization of also ≈5 kJ mol^{-1}. Entropy changes along the insertion path, mainly due to a growing number of realizable microstates because of structural reorganization, are the main factors driving the insertion. An attractive fluxional wall of noncovalent interactions is characterized by all-quantum descriptors of chemical bonding (natural bond orbitals, quantum theory of atoms in molecules, noncovalent interaction, density differences, and natural charges). This attractive wall originates in the accumulation of tiny transfers of electron densities to the interstitial region between the fragments from a multitude of individual intermolecular contacts stabilizing the tertiary drug/water/membrane system.



1. INTRODUCTION

Naproxen (NAP in what follows), (S)-(+)-6-methoxy-methyl-2-naphthaleneacetic acid, $C_{14}H_{14}O_3$, is a nonsteroidal anti-inflammatory drug (NSAID). The US Food and Drug Administration (FDA) approved its use without the need of a prescription in 1994, and it has since become one of the most popular drugs over the counter medications worldwide, sold as a sodium salt and commonly used to relieve pain, fever, menstrual cramps, and other aches.[1] NSAIDs work by nonselectively inhibiting the cyclooxygenase (COX) enzyme in the production of prostaglandins.[2] It is thought that inhibition of COX-2 leads to the desired therapeutic effect, while inhibition of COX-1 leads to unwanted side effects.[2]

Once ingested, all drugs face complex multicomponent chemical environments within the host organism. In the particular case of NSAIDs, their ability to permeate cell membranes is influenced by explicit intermolecular interactions and by configurational changes in the tertiary drug/water/membrane environment; however, one limitation of experimental techniques is that because of the structural and dynamic complexity of cell membranes, thermodynamic measurements alone do not provide a detailed picture of the insertion mechanism; moreover, the responses of the system as the drug traverses the changing chemical environments from the aqueous phase to the interior of the membrane cannot be extracted from thermodynamics. Notwithstanding the experimental limitations, a detailed knowledge of the insertion of drugs into cell membranes is highly desired. Fittingly, several recent studies have studied explicit interactions between the drugs and reduced models of phospholipid membranes under experimental and computational methodologies.[3−10]

Aiming at providing a detailed picture of the insertion process of anionic Naproxen into model dimyristoylphosphatidylcholine (DMPC) membranes (see Figure 1 for the structures of isolated anionic NAP and DMPC), we divide our work into three stages: first, we ran classical molecular dynamics (MD) trajectories in aqueous media under the proper conditions to simulate physiological environments. Second, in order to study the intricate and delicate interplay between the involved thermodynamic quantities, we decompose the Gibbs free energy profile resulting from structural changes in the drug/water/membrane tertiary system as a function of the position of Naproxen from the purely aqueous phase to the approximate center of the membrane. Third, we explore the evolution of

Received: July 30, 2021
Revised: August 13, 2021
Published: September 7, 2021

https://doi.org/10.1021/acs.jpcb.1c06766
J. Phys. Chem. B 2021, 125, 10383−10391

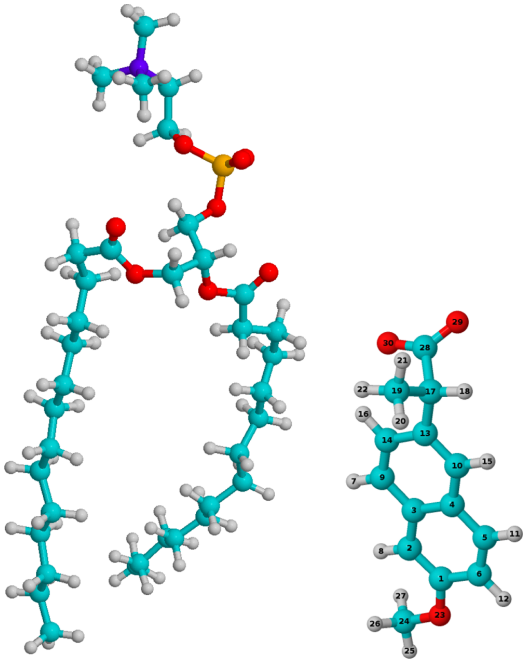


Figure 1. Isolated DMPC (left) and anionic Naproxen (right). Color code: P atoms in orange, C atoms in light blue, O atoms in red, N atoms in blue, and H atoms in light gray.

bonding interactions along the insertion path using rigorous theoretical tools firmly rooted in quantum mechanics.

2. COMPUTATIONAL METHODS

In order to obtain meaningful finite differences to derive thermodynamic quantities, MD simulations at 303, 317, and 331 K coupled to umbrella sampling methods were carried out following the protocols established by Rojas-Valencia and co-workers[9] and are summarized in Table S1 in Supporting Information The insertion process happens along a trajectory normal to the surface of the cell, which we define as the *z*-direction and set its origin at the center of the bilayer. A total of 64 lipids per layer were used to construct the membrane. The free energy profiles for this process were decomposed into the the corresponding enthalpic and entropic contributions. The forces driving the insertion process were obtained from the derivatives of the free energy profiles along the insertion coordinate. The CHARMM36 force field[11] for the membrane and the CHARMM general force field (CGenFF)[12,13] for Naproxen were used for all MD runs including 6152 TIP3P waters.[14] Bonding interactions were analyzed using natural bond orbitals[15−17] (NBOs), quantum theory of atoms in molecules[18−20] (QTAIM), and noncovalent interaction (NCI) indices[21,22] from randomly chosen snapshots from the last 15 nanoseconds of the trajectories.[10,2331] All MD calculations were carried out using Gromacs5.02.[32] NBO6[33] as implemented in Gaussian09,[34] the AIMall suite,[35] and NCIPLOT[36] fed with promolecular densities were used to calculate bonding descriptors.

3. RESULTS AND DISCUSSION

3.1. Energetics. The intricacies of the changing chemical environment that Naproxen has to overcome as it migrates from the aqueous phase to its equilibrium position, at the interface between the polar and nonpolar regions (≈1.35 nm from the center of the bilayer as seen in the bottom panel of Figure 2), are

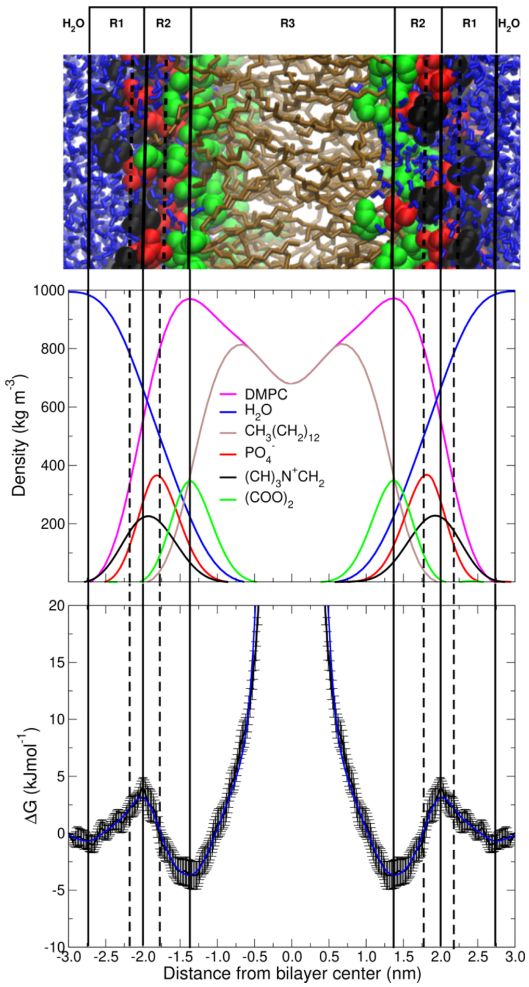


Figure 2. (Top) Snapshot of the geometry for the model DMPC membrane in an aqueous environment. (middle) Density of functional groups. (bottom) Potential of mean force. Solid lines mark the explicit boundaries between regions as defined by the points of vanishing net force. Dashed vertical lines mark the critical points of the net force. All plots derived from calculations at 317 K.

conveniently rationalized using a density of functional groups plot. We provide this plot in the top panel of Figure 2 for the equilibrium situation, dividing the entire system using the scheme suggested by Rojas-Valencia and co-workers.[9] This partition leads to a purely aqueous phase and to regions R1, which comprises an aqueous phase with the outer polar groups of the phospholipid, R2, including the majority of the polar

https://doi.org/10.1021/acs.jpcb.1c06766
J. Phys. Chem. B 2021, 125, 10383−10391

groups, and R3, containing the inner polar groups and the hydrophobic tails.

3.1.1. Potential of Mean Force (PMF). The first and second rows in Figure 3 unveil a molecular interaction picture behind both energy changes and the nature of the equilibrium points along the insertion path. After leaving the purely aqueous phase (3.0 → 2.5 nm), Naproxen faces a small energy barrier of ≈5 kJ mol^{-1} in the 2.5 → 2.0 nm interval, beautifully described by a negative, retarding force (fourth row plot). The density of the functional group plot clearly indicates that this barrier is explicitly due to a reduction in the number of Naproxen··· water contacts (detaching waters from Naproxen—as seen in the bottom panel—has an energy cost) and to a simultaneous increase in the presence of the phosphate and choline groups, which initially oppose the insertion of Naproxen. If the energy barrier is overcome, the 2.0 → 1.35 nm interval, taking Naproxen from the top of the barrier to its preferred equilibrium location at the interface between the polar and nonpolar regions of the phospholipid (1.35 nm, ≈5 kJ mol^{-1} below the aqueous phase), is dominated by a positive driving force. Interestingly, the only appreciable quantitative difference between insertion of anionic Naproxen and anionic ibuprofen[9] is the depth of the energy well: ≈5 and 16.5 kJ mol^{-1}, respectively, suggesting a larger structural affinity for the model membrane in favor of IBU$^-$. The density of functional groups indicates that right after the equilibrium position, that is, in the region increasingly populated by the hydrocarbon tail of the phospholipid, a retarding force prevents NAP$^-$ from penetrating further into the membrane. It is reassuring that the energy profile derived from our MD calculations is fully consistent with previous experimental[3,37,38] and computational[3,7,8,39,40] studies dealing with the insertion of NSAIDs into model membranes in two key aspects: the small energy barrier and the equilibrium position of the drug (to fully penetrate the membrane, transporting agents are needed).[41]

3.1.2. Entropy and Enthalpy. In order to obtain data for the finite difference analysis needed to decompose the free energy for the insertion process,[42] we used the PMF profiles at 303, 317, and 331 K and plotted the resulting enthalpy and entropy contributions in the third row of Figure 3. Notice that all quantities are calculated relative to the aqueous phase. The first observation is that there is a delicate interplay between the two terms which, in the two extreme cases, at the top of the barrier and at the equilibrium position, approaches a 5 kJ mol^{-1} difference. Increasing differences, leading to large positive free energies, are seen in the inaccessible nonpolar region, thus preventing further penetration into the membrane. Besides the reference purely aqueous phase, the two terms balance each other in at least two other points, one corresponding to a slight displacement of the maximum densities of the phosphate and choline groups (≈1.78 nm) and the other one located deep into the aliphatic chain (≈1.03 nm).

The evolution of the entropy and enthalpy terms reveals key thermodynamic aspects of the insertion process. Our decomposition of the PMF profile indicates that because of the signs of the contributing terms, the positive ΔH opposes insertion along the entire path, while the negative $-T\Delta S$ helps it. The entropy and enthalpy plots provide an exquisite molecular explanation for these behaviors: notice that along the entire insertion path, from the purely aqueous phase to the equilibrium position at the bottom of the energy well, the enthalpy curve runs opposite, while the entropy curve actually follows the curve quantifying the detachment of water molecules from NAP$^-$; thus, as stated

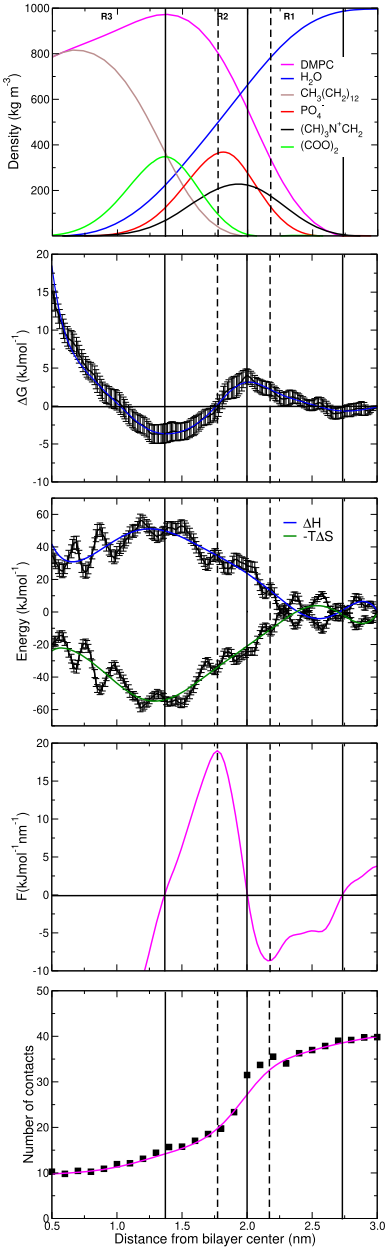


Figure 3. Dissection of the paths for Naproxen insertion into model DMPC membranes. Solid lines mark the explicit boundaries between regions as defined by the points of vanishing net force. Dashed vertical lines mark the critical points of the net force. (first row) Density of functional groups. (second row) Potential of mean force. (third row) Decomposition of the free energy into enthalpy and entropy contributions. (fourth row) Net resulting force. (fifth row) Number of water molecules within 3.0 Å of any given atom in Naproxen. All plots derived from calculations at 317 K.

https://doi.org/10.1021/acs.jpcb.1c06766
J. Phys. Chem. B 2021, 125, 10383−10391

above, removing water molecules from the microsolvated environment of NAP^- has an effective energy cost which is overcompensated by the $-T\Delta S$ term. The all-important entropy contributions are rationalized in two ways from a molecular perspective: first, detaching water molecules from NAP^- increases entropy by virtue of increasing the number of molecules in the purely aqueous region and thus increasing the number of degrees of freedom of the solvent. Second, insertion of the foreign NAP^- into the membrane distorts the packing order of the aliphatic region of the membrane, increasing the available local volume and leading to an increased freedom of motion, all of the abovementioned data being entropy-increasing factors.

A negative, retarding force opposes insertion of NAP^- into the membrane and originates a small barrier of circa 5 kJ mol^{-1}. $\Delta G > 0 \Rightarrow |\Delta H| > |-T\Delta S|$ at the top of the barrier, the point at which the difference between the enthalpy and entropy terms is at a maximum. The top of the barrier (2 nm from the center of the bilayer) is a pivotal point because as soon as NAP^- passes this point, a driving, positive force, characterized by an increasingly larger contribution from the entropy term, takes over and drives the drug all the way to its equilibrium position. The enthalpy and entropy terms actually balance each other at 1.77 nm; thus, in the trajectory for further penetration up to the equilibrium position (≈1.35 nm), the point at which the difference between the two terms is again at a maximum, entropy is the dominant factor because $\Delta G < 0 \Rightarrow |-T\Delta S| > |\Delta H|$.

It is worth summarizing at this point how the results extracted from our MD simulations match the available experimental evidence:[3−5,37,43] NAP^- finds an equilibrium position at the polar/nonpolar interface of the lipid membrane, as was established by Förster resonance energy transfer (FRET). A lack of heat transfer during isothermal titration calorimetry suggests that entropy plays a major role in the insertion process, and the precedent dissection of the interplay between the entropy and enthalpy terms provides solid evidence to conclude that entropy is indeed the major thermodynamic factor dominating the entire insertion path. Changes in phase transition temperatures and membrane distortions are observed from differential scanning calorimetry, IR, and X-ray studies. For the tertiary drug/water/membrane system, the changing density of functional groups and the loss of water molecules along the insertion path (top and bottom panels, Figure 3) with all the entropic considerations discussed above rationalize these experimental observations.

3.2. Evolution of Bonding Interactions. We provide next a detailed analysis of the evolution of the intermolecular interactions along the entire insertion path. Notice that in this work, intermolecular interactions are only assigned on the basis of the existence of well-defined bond paths connecting the fragments, as derived from the topology of the electron densities according to QTAIM. This analysis must be placed in the context of a highly dynamic system in which constant structural relaxations lead to rapid changes in both the strength and nature of individual intermolecular contacts. A very important point is that although molecular interactions constitute a sizeable part of the enthalpy contributions to the free energy, other factors are equally important. Among these enthalpy contributing factors, we mention, for example, the electronic energy with its many own components and its redistribution, configurational, and geometrical changes in individual molecules, thermal contributions from rotations, translations, and PV contributions. Clearly, under these circumstances, a generalized direct correlation between descriptors of chemical bonding and enthalpy would simply not be accurate; however, it should also be clear that these intermolecular interactions play a pivotal role in the calculated and measured macroscopic properties.

Following the established protocols,[10] in order to gain insights into the evolution of chemical bonds and intermolecular interactions during the insertion of anionic Naproxen into our model DMPC membranes, a total of 20 snapshots from each equilibrium position (aqueous phase, top of the barrier, and energy minimum) of the free energy profile shown in the second row of Figure 3 were chosen from the late stages of the corresponding MD simulations at 317 K. Table 1 lists key analysis variables, including the cutoff radius and the corresponding number of water molecules (n_w) and lipid fragments (n_L) contained within the interaction sphere.[10]

Table 1. Inventory of the Average Number of Water Molecules (n_W) and Lipid Residues (or Fragments) (n_L) within the Cutoff Radius in the 20 Snapshots Chosen from the MD Simulations of Each Equilibrium Position in the Free Energy Profile (Figure 3) for the Insertion of NAP^- into Model DMPC Membranes

	>2.74 nm aqueous phase	2.00 nm top of the barrier	1.35 nm energy minimum
cutoff radius (Å)	3.6	3.6	4.0
n_W	40	32	14
n_L	0	3	5

3.2.1. Electron Density and Charge Redistribution. Figure S1 in Supporting Information shows the redistribution of the electron density on NAP^- as it traverses the aqueous phase → top of the barrier → equilibrium position path. Green surfaces indicate the regions of total gain in electron density as the difference between the isolated molecule and the drug embedded in the chemical environment at the equilibrium position (1.35 nm), and red surfaces indicate regions of net loss of electron density. The region closer to the carboxylate group appears as the most responsive when changing chemical environments, with electrons flowing in both NAP^- ⇆ environment directions. A quantification of this electron redistribution is provided in the form of the difference in natural atom charges relative to isolated NAP^-, which shows only modest changes in individual atoms; however, the cumulative effect is significant: for the aqueous phase → top of the barrier → equilibrium position path, adding up all the charge gained by individual atoms with respect to isolated NAP^-, we obtained 0.25 → 0.25 → 0.15 $|e|$. Similarly, 0.36 → 0.37 → 0.31 $|e|$ were lost. Evidently, most of the electron flux occurs in going from isolated to solvated NAP^-, and removing the solvation waters (bottom panel, Figure 3 and Table 1) while surrounding NAP^- by a less-polar environment results in a net loss of electron density in NAP^- (≈0.36 − 0.31 = 0.05$|e|$) in going from the aqueous phase to the equilibrium position.

3.2.2. Noncovalent Interaction Index. The surfaces of noncovalent interactions plotted in Figure 4 uncover important aspects of the evolution of intermolecular interactions as the drug inserts into the model cell membrane:

1. The stronger interactions (blue surfaces according to the NCIPLOT[36] color code) are exclusively seen for the interactions between oxygen atoms in NAP^- and solvating waters.

https://doi.org/10.1021/acs.jpcb.1c06766
J. Phys. Chem. B 2021, 125, 10383−10391

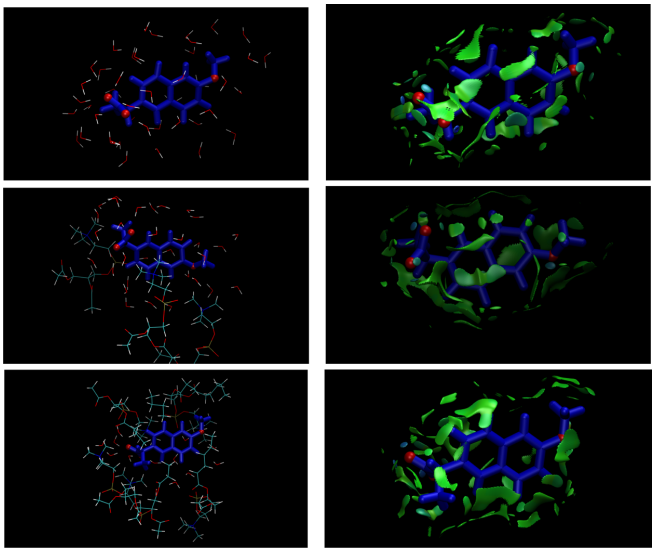


Figure 4. NCIs surrounding anionic Naproxen in the aqueous phase (top), top of the barrier (middle), and energy minimum (bottom) along the insertion path of anionic Naproxen into model DMPC membranes. All calculations using promolecular densities.

Table 2. Largest Donor → Acceptor NBO Energies, $-E^{(2)}_{d\to a}$, in kcal/mol and Electron Densities, $\rho(r_c)$, in a.u., at the Bond Critical Points for all Intermolecular Contacts during the Insertion of Anionic Naproxen into Model DMPC Membranes[a]

			$-E^{(2)}_{ij}$			$\approx 10^2 \times \rho(r_c)$			
label	fragments	orbitals involved	aqueous phase	top of the barrier	energy minimum	BCP	aqueous phase	top of the barrier	energy minimum
1	$-CO_2^-$ in Nap^- → water	$n_O \to \sigma^*_{O-H}$	29.08	12.96	27.58	O··· H	6	4	5
2	water ↔ water	$n_O \to \sigma^*_{O-H}$	22.74	24.57	24.17	O··· H	5	5	4
3	$-PO_2^-$ in DMPC → water	$n_O \to \sigma^*_{O-H}$	N/A	33.52	19.39	O··· H	N/A	7	5
4	−COOR in DMPC → water	$n_O \to \sigma^*_{O-H}$	N/A	2.49	10.22	O··· H	N/A	2	3
5	−choline in DMPC → water	$n_O \to \sigma^*_{C-H}$	N/A	13.28	6.38	O··· H	N/A	3	2
6	$-CO_2^-$ in Nap^- → choline in DMPC	$n_O \to \sigma^*_{C-H}$	N/A	N/A	3.03	O··· H	N/A	N/A	2
7	$-CO_2^-$ in Nap^- → choline in DMPC	$\pi_{C=O} \to \sigma^*_{C-H}$	N/A	N/A	3.89	O··· H	N/A	N/A	2
8	nonpolar Nap^- → nonpolar DMPC	$\sigma_{C-H} \to \sigma^*_{C-H}$	N/A	N/A	0.15	H··· H	N/A	N/A	<1
9	nonpolar Nap^- → nonpolar DMPC	$\pi_{C=C} \to \sigma^*_{C-H}$	N/A	N/A	1.00	C··· H	N/A	N/A	1
10	nonpolar DMPC → nonpolar Nap^-	$\sigma_{C-H} \to \pi^*_{C=C}$	N/A	0.20	0.06	C··· H	N/A	1	1
11	nonpolar Nap^- → choline in DMPC	$\pi_{C=C} \to \sigma^*_{C-H}$	N/A	1.07	1.29	C··· H	N/A	1	1
12	nonpolar Nap^- → water	$\pi_{C=C} \to \sigma^*_{O-H}$	0.12	0.19	0.57	C··· H	<1	<1	1
13	nonpolar Nap^- → water	$\sigma_{C-H} \to \sigma^*_{O-H}$	0.08	0.24	0.22	H··· H	1	1	1
14	water → nonpolar Nap^-	$n_O \to \sigma^*_{C-H}$	0.37	0.23	1.00	O··· H	1	1	1
15	water → nonpolar Nap^-	$n_O \to \sigma^*_{C-H}$	0.37	0.23	0.25	O··· H	1	1	1
16	CH_3O in Nap^- → water	$n_O \to \sigma^*_{O-H}$	5.61	5.76	8.43	O··· H	2	3	3

[a] Figure 5 shows the specific involved orbitals in each case. All calculations are at the CAM-B3LYP/6−31+G* level.

2. As water molecules are detached from the drug (bottom panel, Figure 3), the number of blue surfaces clearly diminishes.
3. Green surfaces, indicating weaker intermolecular contacts, change in the course of the insertion path. These changes are due to the diverse chemical environments the drug encounters, and since they are mostly in shape, they are consistent with the somewhat small energy differences among the stationary points along the PMF.
4. No matter the position of the drug along the insertion path, the green surface is always present. This noncovalent surface, which is the result of the accumulation of weak individual intermolecular contacts, acts as a stabilizing fluxional wall of attractive interactions. Very recent research has provided evidence in favor of fluxional attractive walls of noncovalent interactions in a wide variety of scenarios, including the stable conformations of methane and water cages,[44] the insertion of anionic Ibuprofen into model cell membranes,[10] the initial molecular recognition between the ACE2 receptor in eukaryotic cell membranes, and the receptor binding domain in the spike protein of SARS-COV-2,[45] and has been shown to play a pivotal role in the otherwise

https://doi.org/10.1021/acs.jpcb.1c06766
J. Phys. Chem. B 2021, 125, 10383−10391

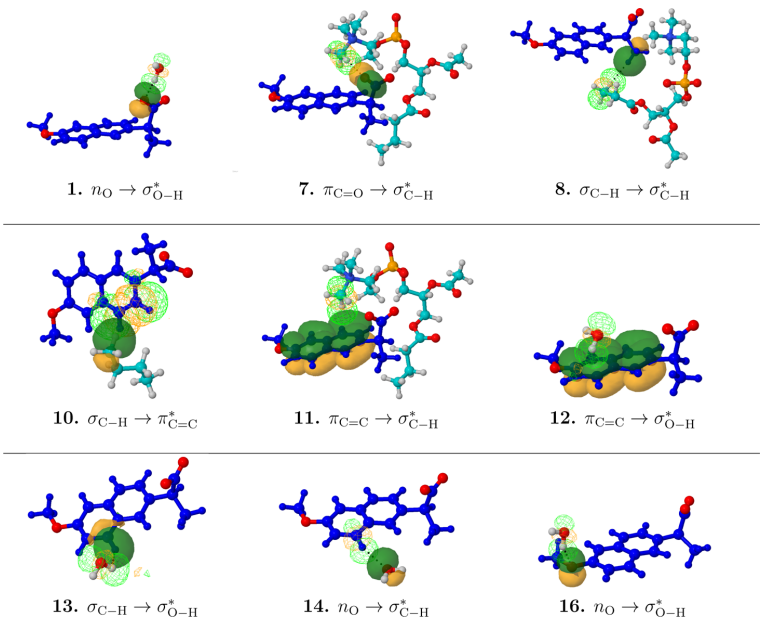


Figure 5. Orbital interactions within the NBO framework. The donor (solid surfaces) and acceptor (lined surfaces) are shown for a selected group of interactions listed in Table 2.

misunderstood fundamental nature of hydrophobic interactions.[46]

3.2.3. NBO Analysis. The fluxional walls of stabilizing interactions discussed above arise from the cumulative effect of a multitude of individual intermolecular contacts. We investigate next the quantum origin of these interactions under the donor → acceptor NBO model. Water ↔ water, water ↔NAP^-, water ↔ lipid, NAP^- ↔ lipid, and lipid ↔ lipid intermolecular interactions may be further divided into polar ↔ polar, polar ↔ nonpolar, and nonpolar ↔ nonpolar types. Among the myriad of individual contacts at the three equilibrium positions in the tertiary NAP^-/water/membrane system, Table 2 lists only the largest interaction energy for each one of the 16 types of $\phi_d \rightarrow \phi_a$ orbital interactions found in this work. Plots of the involved orbitals are shown in Figure 5. A few relevant observations are as follows:

1. Only four general types of intermolecular interactions are at play.
   (a) primary $n_O \rightarrow \sigma^*_{O-H}$ hydrogen bonds.
   (b) secondary $n_O \rightarrow \sigma^*_{C-H}$ hydrogen bonds.
   (c) $\sigma_{C-H} \rightarrow \sigma^*_{O-H}$ or $\sigma_{C-H} \rightarrow \sigma^*_{C-H}$ dihydrogen bonds.
   (d) $\pi_{C=C} \rightarrow \sigma^*_{C-H}$, $\pi_{C=C} \rightarrow \sigma^*_{O-H}$, or $\sigma_{C-H} \rightarrow \pi^*_{C=C}$, which we collectively characterize as $\pi \leftrightarrows \sigma$ interactions.
2. The DMPC model membrane and NAP^- have formal charges from the $-N(CH_3)_3^+$, $-PO_2^-$, and $-CO_2^-$ groups. Taking as a reference the 5.40 kcal/mol reported as the donor → acceptor energy in the $n_O \rightarrow \sigma^*_{O-H}$ orbital interaction responsible for the stabilization of the archetypal water dimer,[15] a strengthening effect of the formal charges on hydrogen bonds[47−58] is clearly observed for all chemical environments along the insertion path. The same $n_O \rightarrow \sigma^*_{O-H}$ interaction affords values as high as 22.74 kcal/mol for water··· water interactions in the aqueous phase (interaction 2 in Table 2) and 33.52 kcal/mol for the $-PO_2^-$···water (interaction 3 in Table 2) at the top of the barrier.
3. Each one of the individual donor··· acceptor orbital interactions taking apart in the nonpolar region of NAP^- is very small (0.06−1.29 kcal/mol); thus, the cumulative effect originating the large fluxional wall of noncovalent interactions stabilizing the tertiary drug/water/membrane system is rationalized using orbital interaction energies.
4. The $-OCH_3$ group in NAP^- carries interactions with water molecules all the way to the equilibrium position at the bottom of the energy profile. The associated $n_O \rightarrow \sigma^*_{O-H}$ orbital interactions actually grow as 5.61 (aqueous phase) → 5.76 (top of the barrier) → 8.43 (energy minimum) kcal/mol. Aside from a few water molecules, the bottom of the energy profile is a highly nonpolar environment; thus, as shown recently[46,59,60] in a changing paradigm of the understanding of the hydrophobic interactions, strengthening of the ether··· water interaction has its origin in the delicate balance between interactions among all competing groups.

3.2.4. Topology of the Electron Densities. QTAIM is a proven method to derive insights into the nature of bonding interactions. Analysis of the topology of the electron density at bond critical points (BCPs, r_c) yields a number of criteria fully discussed elsewhere.[18−20,61−63] In the context of this work, formal justification for the application of QTAIM-derived descriptors to study intermolecular interactions in structures that might be displaced from the stationary points of the quantum PES and a discussion of the equivalency between QTAIM and natural analysis of the topology of electron

https://doi.org/10.1021/acs.jpcb.1c06766
J. Phys. Chem. B 2021, 125, 10383−10391

densities is found in the work of Weinhold.[63] Accordingly, we draw information about the nature and evolution of intermolecular interactions along the insertion path of NAP^- into model DMPC membranes by focusing on the accumulation of electron density $\rho(r_c)$, the Laplacian of the electron density $\nabla^2\rho(r_c)$, and the total $\mathcal{H}(r_c)$, kinetic $\mathcal{G}(r_c)$, and potential $\mathcal{V}(r_c)$ energy densities at the well-defined bond critical points. The electron densities at bond critical points are listed in Table 2, and the plots associated to the following discussion are provided in Figure S2 in the ESI. These are the main points:

1. A hierarchy of interaction strength, fully consistent with the NBO results, may be established solely from the accumulation of electron densities around intermolecular BCPs (Table 2): primary $n_O \rightarrow \sigma^*_{O-H}$> secondary $n_O \rightarrow \sigma^*_{C-H} \approx$ dihydrogen $\sigma_{C-H} \rightarrow \sigma^*_{O-H} \approx \pi \leftrightarrows \sigma$> dihydrogen $\sigma_{C-H} \rightarrow \sigma^*_{C-H}$.
2. For all intermolecular interactions, without exceptions, $\nabla^2\rho(r_c) > 0$ describes local minima in the electron density in the region between the interacting atoms. This local depletion of the charge at BCPs is proper of noncovalent, long-range interactions, which in the context of the tertiary drug/water/membrane system is consistent with primary, secondary, and charge-assisted intermolecular hydrogen bonds. Further support for this characterization of the nature of the interactions is provided by the vast majority of contacts exhibiting $\mathcal{H}(r_c) > 0$ and $|\mathcal{V}(r_c)|/\mathcal{G}(r_c) < 1$.
3. The characterization of noncovalent, weak, and long-range intermolecular interactions is fully consistent with the few cases for which $\mathcal{H}(r_c) < 0$ with simultaneous $|\mathcal{V}(r_c)|/\mathcal{G}(r_c) > 1$. These specific cases neatly reflect the effect of the formal charges which strengthen the interactions without reaching the covalent status.[20] These few cases mostly comprise $H_2O\cdots$ H−O−H (top panel in Figure S2) and $CO_2^-\cdots$H−O−H (middle panel in Figure S2).
4. Physically meaningful exponential decays correctly describing the asymptotic behavior[64−68] of the electron densities as a function of the separation between the fragments, which are satisfied regardless of the position of NAP^- along the insertion path, are shown in the left plots of Figure S2. Notice that each intermolecular contact transfers a tiny amount of electron density to the corresponding BCP. These transferred densities no longer belong to either one of the groups; rather, the accumulation of these small electron densities in the region separating the fragments is the physical source of what we call fluxional attractive walls shown in Figure 4

4. SUMMARY AND CONCLUSIONS

The results of molecular dynamics studies of the insertion of anionic Naproxen into model cell membranes in aqueous environments are discussed here together with an analysis of the evolution of bonding interactions in the tertiary drug/water/membrane system along the insertion path.

In an otherwise spontaneous process, the drug faces a small barrier of ≈5 kJ mol^{-1} at 317 K. This barrier, represented by a retarding force, arises because of two main changes in the chemical environment as NAP^- proceeds from the purely aqueous phase to the exterior of the lipid bilayer: the loss of water molecules and the initial molecular recognition in the form of intermolecular interactions with the outer polar groups of the membrane. Once the barrier is overcome, a driving force originating in short-range favorable drug··· membrane interactions (primary and secondary hydrogen bonds and dihydrogen and $\pi \leftrightarrows \sigma$ interactions) and in an increase in the number of realizable microstates (increase in the number of solvent molecules in the aqueous phase and increase in the conformational freedom of the lipidic tails of the phospholipid bilayers) takes the drug to the equilibrium position at ≈1.35 nm from the center of the bilayer and ≈5 kJ mol^{-1} below the reference aqueous phase. Our calculations provide solid support to the experimental inference that entropy drives the entire insertion process: along the insertion path, the enthalpic contribution to the free energy is always destabilizing with respect to the reference aqueous phase, while the entropic contribution overcompensates to finally afford negative free energies at the preferred equilibrium position at the bottom of the energy profile.

Along the entire insertion path, individual intermolecular contacts result in tiny amounts of electron density being transferred to the interstitial region between fragments. Despite the small individual amounts, the cumulative effect of a large number of these contacts creates a fluxional wall of noncovalent interactions, providing stabilization for the tertiary drug/water/membrane system. The formal charge in anionic Naproxen strengthens the hydrogen bond network among the solvent molecules.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.1c06766.

Conditions for the MD trajectories and plots of the charge analysis, the density differences, and all QTAIM-related descriptors (PDF)

Cartesian coordinates in the xzy format for the quantum fragments taken from all the chosen frames (XYZ)

■ AUTHOR INFORMATION

Corresponding Authors

Natalia Rojas-Valencia − *Instituto de Química, Universidad de Antioquia UdeA, 050010 Medellín, Colombia; Facultad de Ciencias Básicas, Universidad de Medellín, 050026 Medellín, Colombia; Escuela de Ciencias y Humanidades, Departamento de Ciencias Básicas, Universidad Eafit, AA 3300 Medellín, Colombia*; Email: nandrea.rojas@udea.edu.co

Albeiro Restrepo − *Instituto de Química, Universidad de Antioquia UdeA, 050010 Medellín, Colombia*; orcid.org/0000-0002-7866-7791; Email: albeiro.restrepo@udea.edu.co

Authors

Sara Gómez − *Scuola Normale Superiore, Classe di Scienze, 56126 Pisa, Italy*; orcid.org/0000-0002-5430-9228

Francisco Núñez-Zarur − *Facultad de Ciencias Básicas, Universidad de Medellín, 050026 Medellín, Colombia*; orcid.org/0000-0002-9244-9328

Chiara Cappelli − *Scuola Normale Superiore, Classe di Scienze, 56126 Pisa, Italy*; orcid.org/0000-0002-4872-4505

Cacier Hadad − *Instituto de Química, Universidad de Antioquia UdeA, 050010 Medellín, Colombia*; orcid.org/0000-0002-2799-301X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.1c06766

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Partial funding for this project from H2020-MSCA-ITN-2017 European Training Network "Computational Spectroscopy In Natural sciences and Engineering" (COSINE), grant number 765739, is acknowledged. NRV thanks the Ministerio de Ciencias y Tecnología de Colombia, MINCIENCIAS, for a postdoctoral fellowship. FNZ acknowledges Universidad de Medellín for continuous support. Internal support from Universidad de Antioquia via "Estrategia para la sostenibilidad" is also acknowledged.

■ REFERENCES

(1) Wongrakpanich, S.; Wongrakpanich, A.; Melhado, K.; Rangaswami, J. A Comprehensive Review of Non-Steroidal Anti-Inflammatory Drug Use in The Elderly. *Aging Dis.* **2018**, *9*, 143.

(2) Vane, J. R.; Botting, R. M. Anti-inflammatory drugs and their mechanism of action. *J. Inflamm. Res.* **1998**, *47*, 78−87.

(3) Lichtenberger, L. M.; Zhou, Y.; Jayaraman, V.; Doyen, J. R.; O'Neil, R. G.; Dial, E. J.; Volk, D. E.; Gorenstein, D. G.; Boggara, M. B.; Krishnamoorti, R. Insight into NSAID-induced membrane alterations, pathogenesis and therapeutics: Characterization of interaction of NSAIDs with phosphatidylcholine. *Biochim. Biophys. Acta, Mol. Cell Biol. Lipids* **2012**, *1821*, 994−1002.

(4) Manrique-Moreno, M.; Villena, F.; Sotomayor, C. P.; Edwards, A. M.; Muñoz, M. A.; Garidel, P.; Suwalsky, M. Human cells and cell membrane molecular models are affected in vitro by the nonsteroidal anti-inflammatory drug ibuprofen. *Biochim. Biophys. Acta* **2011**, *1808*, 2656−2664.

(5) Manrique-Moreno, M.; Howe, J.; Suwalsky, M.; Garidel, P.; Brandenburg, K. Physicochemical Interaction Study of Non-Steroidal Anti-Inflammatory Drugs with Dimyristoylphosphatidylethanolamine Liposomes. *Lett. Drug Des. Discov.* **2010**, *7*, 50−56.

(6) Manrique-Moreno, M.; Suwalsky, M.; Villena, F.; Garidel, P. Effects of the nonsteroidal anti-inflammatory drug naproxen on human erythrocytes and on cell membrane molecular models. *Biophys. Chem.* **2010**, *147*, 53−58.

(7) Yousefpour, A.; Amjad Iranagh, S.; Nademi, Y.; Modarress, H. Molecular dynamics simulation of nonsteroidal antiinflammatory drugs, naproxen and relafen, in a lipid bilayer membrane. *Int. J. Quantum Chem.* **2013**, *113*, 1919−1930.

(8) Boggara, M. B.; Mihailescu, M.; Krishnamoorti, R. Structural Association of Nonsteroidal Anti-Inflammatory Drugs with Lipid Membranes. *J. Am. Chem. Soc.* **2012**, *134*, 19669−19676 PMID: 23134450.

(9) Rojas-Valencia, N.; Lans, I.; Manrique-Moreno, M.; Hadad, C. Z.; Restrepo, A. Entropy drives the insertion of ibuprofen into model membranes. *Phys. Chem. Chem. Phys.* **2018**, *20*, 24869−24876.

(10) Rojas-Valencia, N.; Gómez, S.; Montillo, S.; Manrique-Moreno, M.; Cappelli, C.; Hadad, C.; Restrepo, A. Evolution of Bonding during the Insertion of Anionic Ibuprofen into Model Cell Membranes. *J. Phys. Chem. B* **2020**, *124*, 79−90 PMID: 31790579.

(11) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **2010**, *114*, 7830−7843 PMID: 20496934.

(12) Vanommeslaeghe, K.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144−3154 PMID: 23146088.

(13) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155−3168 PMID: 23145473.

(14) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(15) Reed, A. E.; Curtiss, L. A.; Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **1988**, *88*, 899−926.

(16) Weinhold, F.; Landis, C. R.; Glendening, E. D. What is NBO analysis and how is it useful? *Int. Rev. Phys. Chem.* **2016**, *35*, 399−440.

(17) Weinhold, F.; Landis, C. R. *Discovering Chemistry with Natural Bond Orbitals*; Wiley-VCH: Hoboken NJ, 2012; p 319.

(18) Bader, R. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, 1990.

(19) Popelier, P. L. *Atoms in Molecules: An Introduction*; Prentice Hall: London, 2000.

(20) Grabowski, S. J. What Is the Covalency of Hydrogen Bonding? *Chem. Rev.* **2011**, *111*, 2597−2625.

(21) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498−6506.

(22) DiLabio, G. A.; Otero-de-la-Roza, A. Noncovalent Interactions in Density Functional Theory. *Reviews in Computational Chemistry*; John Wiley & Sons, Ltd, 2016; Chapter 1, pp 1−97.

(23) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664−5677.

(24) Cappelli, C. Integrated QM/polarizable MM/continuum approaches to model chiroptical properties of strongly interacting solute-solvent systems. *Int. J. Quant. Chem.* **2016**, *116*, 1532−1542.

(25) Etienne, T.; Very, T.; Perpète, E. A.; Monari, A.; Assfeld, X. A QM/MM Study of the Absorption Spectrum of Harmane in Water Solution and Interacting with DNA: The Crucial Role of Dynamic Effects. *J. Phys. Chem. B* **2013**, *117*, 4973−4980.

(26) Egidi, F.; Lo Gerfo, G.; Macchiagodena, M.; Cappelli, C. On the nature of charge-transfer excitations for molecules in aqueous solution: a polarizable QM/MM study. *Theor. Chem. Acc.* **2018**, *137*, 82.

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

(28) Giovannini, T.; Macchiagodena, M.; Ambrosetti, M.; Puglisi, A.; Lafiosca, P.; Lo Gerfo, G.; Egidi, F.; Cappelli, C. Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling. *Int. J. Quant. Chem.* **2019**, *119*, No. e25684.

(29) Puglisi, A.; Giovannini, T.; Antonov, L.; Cappelli, C. Interplay between conformational and solvent effects in UV-visible absorption spectra: curcumin tautomers as a case study. *Phys. Chem. Chem. Phys.* **2019**, *21*, 15504−15514.

(30) Cwiklik, L.; Aquino, A. J. A.; Vazdar, M.; Jurkiewicz, P.; Pittner, J.; Hof, M.; Lischka, H. Absorption and Fluorescence of PRODAN in Phospholipid Bilayers: A Combined Quantum Mechanics and Classical Molecular Dynamics Study. *J. Phys. Chem. A* **2011**, *115*, 11428−11437.

(31) Gómez, S.; Giovannini, T.; Cappelli, C. Absorption spectra of xanthines in aqueous solution: a computational study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5929−5941.

(32) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*−*2*, 19−25.

(33) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Landis, C. R.; Weinhold, F. *NBO 6.0* Theoretical Chemistry Institute, University of Wisconsin: Madison, 2013.

(34) 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. A. et al. *Gaussian 09* Revision E.01; Gaussian Inc.: Wallingford CT, 2009.
(35) Keith, T. *AIMALL* (version 19.10.12); TK Gristmill Software: Overland Park KS, USA, 2019.aim.tkgristmill.com.
(36) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. NCIPLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theory Comput.* **2011**, *7*, 625−632.
(37) Moreno, M. M.; Garidel, P.; Suwalsky, M.; Howe, J.; Brandenburg, K. The membrane-activity of Ibuprofen, Diclofenac, and Naproxen: A physico-chemical study with lecithin phospholipids. *Biochim. Biophys. Acta* **2009**, *1788*, 1296−1303.
(38) Pereira-Leite, C.; Figueiredo, M.; Burdach, K.; Nunes, C.; Reis, S. Unraveling the Role of Drug-Lipid Interactions in NSAIDs-Induced Cardiotoxicity. *Membranes* **2021**, *11*, 24.
(39) Sodeifian, G.; Razmimanesh, F. Diffusional interaction behavior of NSAIDs in lipid bilayer membrane using molecular dynamics (MD) simulation: Aspirin and Ibuprofen. *J. Biomol. Struct. Dyn.* **2019**, *37*, 1666−1684 PMID: 29695194.
(40) Liu, W.; Zhang, S.; Meng, F.; Tang, L. Molecular simulation of ibuprofen passing across POPC membrane. *J. Theor. Comput. Chem.* **2014**, *13*, 1450033.
(41) Wanat, K. Biological barriers, and the influence of protein binding on the passage of drugs across them. *Mol. Biol. Rep.* **2020**, *47*, 3221−3231.
(42) MacCallum, J. L.; Tieleman, D. P. Computer Simulation of the Distribution of Hexane in a Lipid Bilayer: Spatially Resolved Free Energy, Entropy, and Enthalpy Profiles. *J. Am. Chem. Soc.* **2006**, *128*, 125−130 PMID: 16390139.
(43) Micieli, D.; Giuffrida, M. C.; Pignatello, R.; Castelli, F.; Sarpietro, M. G. Interaction of naproxen amphiphilic derivatives with biomembrane models evaluated by differential scanning calorimetry and Langmuir-Blodgett studies. *J. Colloid Interface Sci.* **2011**, *360*, 359−369.
(44) Salazar-Cano, J.-R.; Guevara-García, A.; Vargas, R.; Restrepo, A.; Garza, J. Hydrogen bonds in methane-water clusters. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23508−23515.
(45) Gómez, S. A.; Rojas-Valencia, N.; Gómez, S.; Egidi, F.; Cappelli, C.; Restrepo, A. Binding of SARS-CoV-2 to Cell Receptors: A Tale of Molecular Evolution. *ChemBioChem* **2021**, *22*, 724−732.
(46) Gómez, S.; Rojas-Valencia, N.; Gómez, S. A.; Cappelli, C.; Merino, G.; Restrepo, A. A molecular twist on hydrophobicity. *Chem. Sci.* **2021**, *12*, 9233−9245.
(47) Hadad, C.; Florez, E.; Acelas, N.; Merino, G.; Restrepo, A. Microsolvation of small cations and anions. *Int. J. Quant. Chem.* **2019**, *119*, No. e25766.
(48) Flórez, E.; Acelas, N.; Ramírez, F.; Hadad, C.; Restrepo, A. Microsolvation of F−. *Phys. Chem. Chem. Phys.* **2018**, *20*, 8909−8916.
(49) Flórez, E.; Acelas, N.; Ibargüen, C.; Mondal, S.; Cabellos, J. L.; Merino, G.; Restrepo, A. Microsolvation of NO3−: structural exploration and bonding analysis. *RSC Adv.* **2016**, *6*, 71913−71923.
(50) Rojas-Valencia, N.; Ibargüen, C.; Restrepo, A. Molecular interactions in the microsolvation of dimethylphosphate. *Chem. Phys. Lett.* **2015**, *635*, 301−305.
(51) Zapata-Escobar, A.; Manrique-Moreno, M.; Guerra, D.; Hadad, C. Z.; Restrepo, A. A combined experimental and computational study of the molecular interactions between anionic ibuprofen and water. *J. Chem. Phys.* **2014**, *140*, 184312.
(52) Romero, J.; Reyes, A.; David, J.; Restrepo, A. Understanding microsolvation of Li+: structural and energetical analyses. *Phys. Chem. Chem. Phys.* **2011**, *13*, 15264−15271.
(53) Gonzalez, J. D.; Florez, E.; Romero, J.; Reyes, A.; Restrepo, A. Microsolvation of Mg2+, Ca2+: strong influence of formal charges in hydrogen bond networks. *J. Mol. Model.* **2013**, *19*, 1763−1777.
(54) Acelas, N.; Flórez, E.; Hadad, C.; Merino, G.; Restrepo, A. A Comprehensive Picture of the Structures, Energies, and Bonding in [SO4(H2O)n]2-, n = 1-6. *J. Phys. Chem. A* **2019**, *123*, 8650−8656.
(55) Ramírez-Rodríguez, F.; Restrepo, A. Structures, energies, and bonding in the microsolvation of Na+. *Chem. Phys.* **2021**, *544*, 111106.
(56) Chamorro, Y.; Flórez, E.; Maldonado, A.; Aucar, G.; Restrepo, A. Microsolvation of heavy halides. *Int. J. Quantum Chem.* **2021**, *121*, No. e26571.
(57) Velásquez, A.; Chamorro, Y.; Maldonado, A.; Aucar, G.; Restrepo, A. Microsolvation of Sr^{2+}, Ba^{2+}: Structures, energies, bonding, and nuclear magnetic shieldings. *Int. J. Quantum Chem.* **2021**, *121*, No. e26753.
(58) Uribe, L.; Gómez, S.; Giovannini, T.; Egidi, F.; Restrepo, A. An efficient and robust procedure to calculate absorption spectra of aqueous charged species applied to NO2−. *Phys. Chem. Chem. Phys.* **2021**, *23*, 14857−14872.
(59) Ben-Amotz, D. Water-Mediated Hydrophobic Interactions. *Annu. Rev. Phys. Chem.* **2016**, *67*, 617−638 PMID: 27215821.
(60) Tomar, D. S.; Paulaitis, M. E.; Pratt, L. R.; Asthagiri, D. N. Hydrophilic Interactions Dominate the Inverse Temperature Dependence of Polypeptide Hydration Free Energies Attributed to Hydrophobicity. *J. Phys. Chem. Lett.* **2020**, *11*, 9965−9970 PMID: 33170720.
(61) Espinosa, E.; Alkorta, I.; Elguero, J.; Molins, E. From weak to strong interactions: A comprehensive analysis of the topological and energetic properties of the electron density distribution involving X-H···F-Y systems. *J. Chem. Phys.* **2002**, *117*, 5529−5542.
(62) Cremer, D.; Kraka, E. A description of the chemical bond in terms of local properties of electron density and energy. *Croat. Chem. Acta* **1984**, *57*, 1259−1281.
(63) Weinhold, F. Natural bond critical point analysis: Quantitative relationships between natural bond orbital-based and QTAIM-based topological descriptors of chemical bonding. *J. Comput. Chem.* **2012**, *33*, 2440−2449.
(64) Ramírez, F.; Hadad, C. Z.; Guerra, D.; David, J.; Restrepo, A. Structural studies of the water pentamer. *Chem. Phys. Lett.* **2011**, *507*, 229−233.
(65) Farfán, P.; Echeverri, A.; Diaz, E.; Tapia, J. D.; Gómez, S.; Restrepo, A. Dimers of formic acid: Structures, stability, and double proton transfer. *J. Chem. Phys.* **2017**, *147*, 044312.
(66) Alkorta, I.; Rozas, I.; Elguero, J. Bond Length−Electron Density Relationships: From Covalent Bonds to Hydrogen Bond Interactions. *Struct. Chem.* **1998**, *9*, 243−247.
(67) Knop, O.; Rankin, K. N.; Boyd, R. J. Coming to Grips with N−H···N Bonds. 1. Distance Relationships and Electron Density at the Bond Critical Point. *J. Phys. Chem. A* **2001**, *105*, 6552−6566.
(68) Knop, O.; Rankin, K. N.; Boyd, R. J. Coming to Grips with N−H···N Bonds. 2. Homocorrelations between Parameters Deriving from the Electron Density at the Bond Critical Point. *J. Phys. Chem. A* **2003**, *107*, 272−284.

https://doi.org/10.1021/acs.jpcb.1c06766
J. Phys. Chem. B 2021, 125, 10383−10391

5.1.2. Electronic absorption spectra[10]

Water maintains the UV-Vis spectral features during the insertion of anionic Naproxen and Ibuprofen into model cell membranes

Natalia Rojas-Valencia,[*,†] Sara Gómez,[*,‡] Tommaso Giovannini,[‡] Chiara Cappelli,[‡] Albeiro Restrepo,[¶] and Francisco Núñez–Zarur[†]

†*Facultad de Ciencias Básicas, Universidad de Medellín, Carrera 87 No. 30-65, 050026, Medellín, Colombia*

‡*Scuola Normale Superiore, Classe di Scienze, Piazza dei Cavalieri 7, 56126, Pisa, Italy*

¶*Instituto de Química, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia*

E-mail: nandrea.rojas@udea.edu.co; sara.gomezmaya@sns.it

Abstract

UV-Vis spectra of anionic Ibuprofen and Naproxen in a model lipid bilayer of the cell membrane are investigated using computational techniques in combination with a comparative analysis of drug spectra in purely aqueous environments. The simulations aim at elucidating the intricacies behind the negligible changes in the maximum absorption wavelength in the experimental spectra. A set of configurations of the systems constituted by lipid, water, and drugs or just water and drugs are obtained from classical Molecular Dynamics simulations and UV-Vis spectra are computed in the framework of atomistic Quantum Mechanical /Molecular Mechanics (QM/MM) approaches together with Time-Dependent Density Functional Theory (TD-DFT). Our results suggest that the molecular orbitals involved in the electronic transitions are the same regardless of the chemical environment. A thorough analysis of the contacts between the drug and water molecules reveals that no significant changes in UV-Vis spectra arise as a consequence of Ibuprofen and Naproxen molecules being permanently microsolvated by water molecules despite the presence of lipid molecules.

1 Introduction

Nonsteroidal anti-inflammatory drugs (NSAIDs) are a family of chemical compounds widely used worldwide to relieve pain, fever, and swelling.[1] According to their chemical structure they are mainly classified as salicylates, sulfonanilides, and derivatives of acetic acid, enolic acid, and propionic acid.[2] The latter category includes Ibuprofen, (IBU), [(RS)-2-(4-(2-methylpropyl)phenyl propanoic acid)] and Naproxe, (NAP), [(2S)-2-(6-methoxynaphthalen-2-yl) propanoic acid]. Ibuprofen is used to treat pain in rheumatoid disorders and inflammatory diseases,[3] whereas Naproxen is mostly employed in the treatment of acute arthritis, osteoarthritis, musculoskeletal pain inflammation, and dysmenorrhea,[4] and it is thought to have a longer duration of action.[5]

IBU and NAP are weak acids with pK_a values of 5.2 and 4.2, respectively,[6] and thus they are found as deprotonated species, IBU^- and NAP^-, at physiological pH values. Another common characteristic of these molecules is their amphiphilic character due to the presence of one acetate and aromatic groups connected by a bridging carbon that is a chiral center. The S-conformations (see Figure 1) are reported to have greater therapeutic action.[7,8]

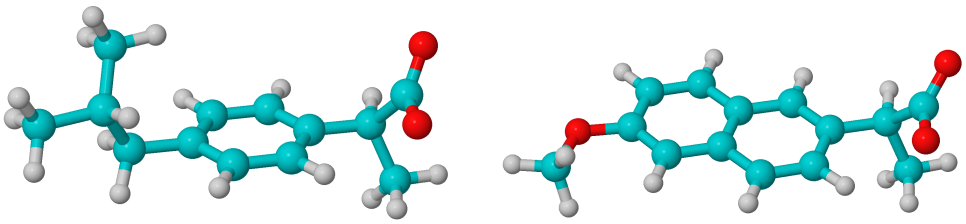


Figure 1: Isolated anionic forms of (S)–Ibuprofen (left) and (S)–Naproxen (right). Color code: C atoms in light blue, O atoms in red, and H atoms in light grey.

The main mechanism of action of IBU^- and NAP^-, as well as of the general set of NSAIDs, involves the inhibition of the cyclooxygenase (COX) enzyme during the production

of prostaglandins from the arachidonic acid.[9,10] Prostaglandins are mediators in the response of the body to a pathological or physiological stimulus.[9,11] In contrast to the NSAIDs therapeutic action, some experimental and computational studies suggest that they could affect cell membrane properties such as its fluidity and function due to their interaction with phospholipids which are the building blocks of the membrane.[2,12,13] Concerning those interactions, researchers have reached a consensus on the key role played by the NSAIDs lipid affinity in both the (undesired) toxic and therapeutic actions.[12]

The interaction between drugs and the lipidic environment of the cell membrane has been routinely investigated by means of the estimation of the drug partition coefficient in a system composed of octanol and water.[14] Notwithstanding, the complexity of the biological system is not well captured by this environment, and thus few reports have focused on more realistic models of cell membranes.[14–16] It is well known that the hydrophobic character of NSAIDs determines the extent to which they can be distributed in the membrane.[17] This fact has important implications for the lipid bilayer structure. For instance, Manrique-Moreno et al.[13,18] studied the effect of IBU^- and NAP^- on liposomes as model membranes under physiological conditions (pH 7.4). By using different experimental techniques such as Differential Scanning Calorimetry (DSC) and Fourier Transform Infrared Spectroscopy (FTIR), among others, the authors found changes both in the transition phase temperature and in the vibration of some polar groups of the lipid model when the drug was present. However, they did not find evidence of its insertion in the hydrophobic part of the membrane. Thereafter, several studies using Neutron Diffraction and Molecular Dynamics (MD) simulations were carried out to explore in more detail the location of the drug within the membrane.[19–21] Overall, those studies suggested that IBU and NAP can reside in the hydrophobic core and in the interface between the polar and nonpolar regions of the lipid bilayer in their protonated and deprotonated states, respectively. Without any exception, all the above studies agree that the anionic drug can produce changes in the lipid bilayer structure even though those drugs are not able to cross the membrane through a passive permeation mechanism. There-

fore, the simplistic view offered by the octanol/water partition coefficient is not good enough to gain reliable information on the insertion process. In the last two decades, experimental UV-Vis spectroscopy has been proposed as a new strategy to get the partition coefficient of drugs in lipid membranes due to the electronic changes of the molecule according to its chemical environment.[15,22]

Different experimental techniques have been used to get insight into the behavior of the complex tertiary systems constituted by phospholipids, drugs, and water. For the IBU^- and NAP^- cases, UV-Vis measurements have been performed in aqueous solutions and in lipid environments and the results did not show significant changes in the maximum absorption wavelength.[23,24] However, since the drug has interactions of a very distinct nature when is surrounded by water molecules or when immersed in a lipid environment, important band shifts should be expected as it happens when a generic solute is dissolved in polar *vs* in non-polar solvents.[25–27] To the best of our knowledge, no previous studies focus on understanding the intricacies involved in the electronic transitions of the drug as a response to the surrounding environment. Certainly, more detailed information about the system might be obtained using computational spectroscopy. In particular, the combination of MD simulations with Time-Dependent Density Functional Theory (TD-DFT) in the framework of Quantum Mechanics (QM)/Molecular Mechanics (MM) has proved to be a good strategy to study the absorption spectra of large-sized systems.[28–33]

In this work, we investigate, from a computational perspective, the reasons behind the small changes in experimental UV-Vis spectra of IBU^- and NAP^- when going from aqueous solution to the aqueous lipidic environment. To that end, we take reported configurations[21,34] from combined MD simulations performed with the umbrella sampling method[35] and calculate the electronic transitions involved in absorption spectra using QM/MM combined with TD-DFT. To account for the different chemical environments surrounding the drug,

the electronic transitions are characterized in a set of configurations belonging to the three key points where IBU^- or NAP^- could be located, namely, the aqueous environment, at the polar region of the lipid bilayer and at the polar/non–polar interface of the membrane, as pointed out in recent works.[21,34]

The paper is structured as follows: in the next section, to contextualize the problem, the results of the previously reported Gibbs free energy profiles for the insertion of IBU^- and NAP^- into model cell membranes are briefly summarized, thus explaining the key points from which the configurations are taken. Then, the computational details are given and in the section 3, UV-Vis and ECD spectra results are presented and discussed. Finally, in Section 4, some conclusions are drawn.

2 Methods

MD runs coupled with QM/MM have been shown to be particularly reliable because, on the one hand, MD simulations allow to capture the microscopic behavior of the molecules and have a good representation of macroscopic properties due to the ability of getting a set of different configurations of the equilibrated system, which on average represent the behavior of the system as a whole.[36–38] On the other hand, QM/MM approaches are relatively inexpensive because only a portion of the system to be excited is modeled at the QM level, while the environment is described classically with MM force fields.[39–42] By following such a computational strategy, it has been possible to get an outstanding reproduction of experimental UV-Vis spectra at reduced computational cost.[43–47] There are different conceptual ways of combining QM and MM parts, that differ from each other in the inclusion and method of describing the mutual interactions.[40,48,49] Possible approaches consist of Mechanical and Electrostatic Embedding. In the latter, the interaction term is formulated in terms of a set of fixed charges, whereas more sophisticated approaches make charges to adjust to the QM density,[40,50] however they may require specific parametrization for the considered

environments.[27] For example, QM/MM Electrostatic embedding was used in the calculation of the UV-Vis spectra of the residue pairs having persistent contacts in the attaching of the SARS-CoV-2 (and its variants) to the ACE2.[51,52] Also, a non-polarizable QM/MM approach was followed by Cwiklik et al.[33] to study a system aiming at reproducing the experimental absorption and emission spectra of PRODAN into a lipid bilayer, without elucidating the nature of the absorption electronic transitions.

In this work, we exploit a set of equilibrated configurations at 317 K reported in previous papers.[20,21] To summarize, during the MD simulations, each drug, anionic naproxen and ibuprofen, interacted with a lipid bilayer in an aqueous environment as a model of the cell membrane. The lipid bilayer was constructed with 64 Dimyristoylphosphatidylcholine (DMPC) molecules per layer. Explicit TIP3P water molecules were also included in the simulation box. A free energy profile for the affinity of the drug in different lipidic environments of the system was obtained following a reaction coordinate, running from the center of the bilayer to the aqueous environment. Three critical points were identified along the reaction coordinate: 1) a local minimum when the drug resides in the aqueous environment, 2) a maximum of the barrier when the drug is facing the polar part of the lipid, which is characterized by Choline and phospholipids groups and 3) the global minimum of the free energy when the drug is located in the interface between the polar/non-polar part of the lipids. These key points for IBU^- and NAP^- are displayed in Figure 2 along with a comparison of the free energy curves. It is worth mentioning that the global minima also represent the equilibrium position of the drugs and are in agreement with reported experimental results. More details about those studies can be found elsewhere.[20,21]

Based on the three critical points of the Gibbs free energy mentioned above (aqueous phase, top of the barrier, and energy minimum), we selected 200 random snapshots from the last 7 ns and 15 ns of the production steps of the reported MD simulations for each one of the three different chemical environments. In an effort to reduce the size of the system while

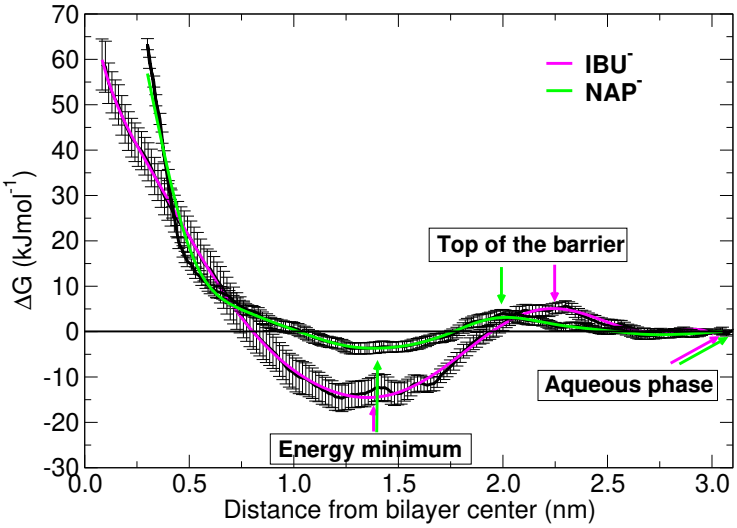


Figure 2: Free energy profiles for the insertion processes of the drugs into model dimyristoylphosphatidylcholine (DMPC) membranes at 317 K. The equilibrium position[20,21] is reached at ≈ 1.35 nm from the center of the lipid bilayer for both anions. Image adapted from Ref. 20. Copyright 2018 Royal Society of Chemistry, and from Ref. 21. Copyright 2021 American Chemical Society.

maintaining an accurate representation of the chemical environment during the simulation, only the water molecules within a cut-off radius of 3.6 Å from each atom of the drug were considered in the aqueous environment. 15 Å was the threshold for including lipids and water at the top of the barrier and in the energy minimum. A random snapshot of each point is shown in Figure 3 and the cut regions are highlighted. To determine the cut-off radius, we employed information from former studies[21,34] where radial distribution functions (RDFs) were exploited to define the boundaries and calculate quantum mechanical properties of the system –electron densities, natural charges, Non-covalent Interactions (NCI), Natural Bond Orbitals (NBOs)– in order to study the intermolecular interactions[53–55] taking place between the drug and the environment (water and phospholipids). As opposed to other works,[21,34] here, the first hydration sphere of the nonpolar part of the drug (first maximum of RDF) was explicitly considered, thus increasing the size of the regions. Finally, for the top of the barrier and the energy minimum, the coordinate corresponding to the emergence of a

plateau in the RDF was chosen as the last solvation sphere. Entire phospholipids are always included if present.

Non-polarizable, electrostatic embedding QM/MM calculations were performed to obtain the electronic absorption spectra. The first ten excited states were taken into account in each case and were enough to reproduce experimental findings. The drug (IBU^- or NAP^-) was included in the QM portion, whereas the remaining system was described through an electrostatic embedding using the same fixed charges employed during the MD runs. Vertical transition energies were computed for each configuration by means of TD-DFT calculations at the CAM-B3LYP/6-311++g(d,p) level of theory. This model chemistry has been chosen for two reasons: after benchmarking, (see Figure S1 in the Supplementary Information, SI) it showed the best reproduction of experimental UV-Vis absorption spectra when an anionic ibuprofen molecule was embedded by an aqueous environment, represented with the polarizable continuum model (PCM). Additionally, other studies with similar systems also used the same level of theory with a good reproduction of spectral properties.[56] Further analysis employing mechanical instead of electrostatic embedding yielded poor reproduction of the experimental data as can be seen in Figure S2 in the SI.

The obtained spectroscopic information was then convoluted using Gaussian functions with a full width half maximum (FWHM) of 0.4 eV to carry out a proper comparison between simulated and experimental absorption spectra. Despite using a large number of snapshots (600 in total, 200 in each environment) the spectra converged after considering only 140 snapshots (see Figure S3 in the SI). We also characterized all electronic transitions in terms of Canonical Molecular Orbitals (CMO) as described in the NBO framework.

All calculations were performed using the Gaussian 16 package.[57] NBO7[58] was employed for the CMO calculations.

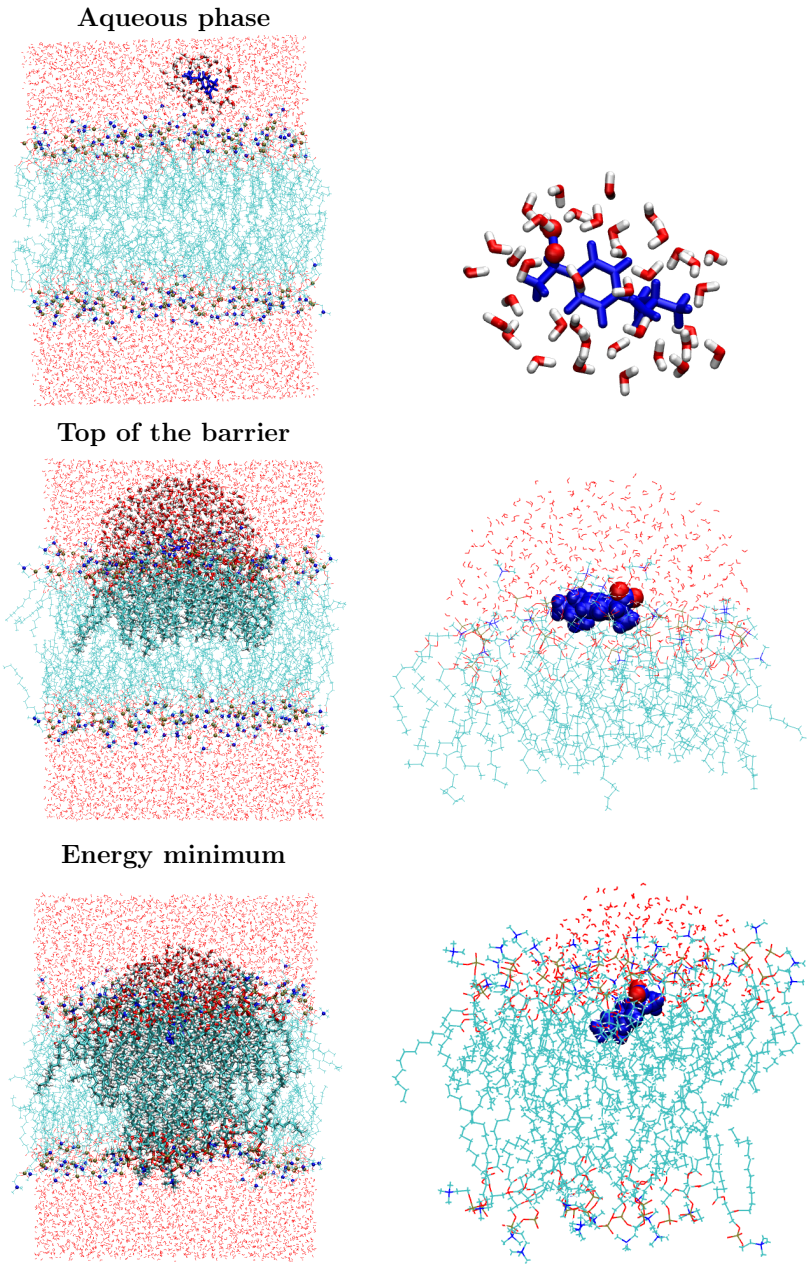


Figure 3: Representation of the selected system in the different environments: aqueous phase (top), top of the barrier (middle), and energy minimum (bottom). The highlighted regions in the left column, which are enlarged in the right column, correspond with the cut-off radius taken into account to calculate the UV-Vis spectra.

3 Results and discussion

In this section, the simulated UV-Vis absorption and ECD spectra of both Ibuprofen and Naproxen in three different chemical environments of the NSAID/DMPC/water system (aqueous phase, top of the barrier, and energy minimum) are presented. First, the results for the drugs in aqueous solution are analyzed and then the findings for each drug in an aqueous lipidic environment are shown. Later, an analysis of the small spectral differences is carried out.

3.1 UV-Vis spectra of ibuprofen and naproxen

3.1.1 Drugs in aqueous environments

The reported experimental UV-Vis spectrum of ibuprofen in aqueous environment has two absorption maxima at 222 and 190 nm.[24,59] As a matter of fact, it has been claimed that excitation energies occurring at wavelengths below 200 nm are difficult to interpret if the measure is not taken in *vacuum*.[60] For this reason, we will be focused on the 200-300 nm range. Experimental spectral features of naproxen in aqueous environment involve two bands, one in the 210-250 nm interval, centered at $\approx$ 230 nm, with a small shoulder in the vicinity of 270 nm, and a broad second band going from 300 to 340 nm.[23] These bands are usually associated to $\pi \rightarrow \pi^*$ transitions.[61] Figure 4 shows the QM/MM UV-Vis absorption spectra obtained from the convolution of a total of 200 snapshots when the drug is surrounded only by water molecules. Both simulated spectra (bottom panel of Figure 4) match the most important spectral characteristics observed in the experiments, with solvated IBU^- having an exceptional agreement. For solvated NAP^-, although the most intense band appears slightly shifted from the experimental position, the shoulder and the low-intensity band are mixed in a single spread band that covers all the region above 260 nm. Our computations are fully consistent with other measurements which do not distinguish the low-intensity band

from its onset.[62] These results confirm the accuracy of the level of theory used in this work.

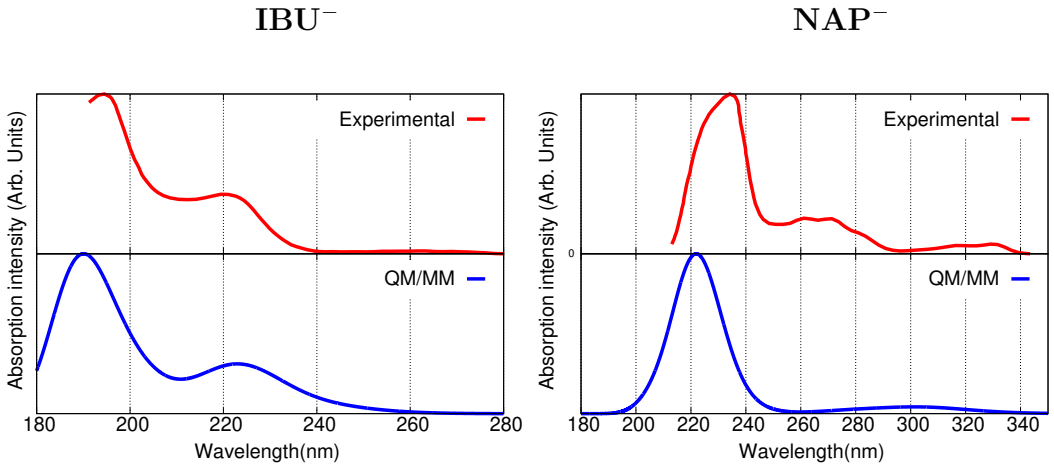


Figure 4: Experimental[23,24] (top) and simulated QM/MM CAM-B3LYP/6-311++g(d,p)/Electrostatic Embedding (bottom) UV-Vis spectra of IBU^- (left) and NAP^- (right) in aqueous solution. 200 snapshots were considered in the calculations. Averaged spectra were convoluted with an FWHM of 0.4 eV and normalized to the highest intensity band.

3.1.2 Drugs in aqueous lipidic environment

As mentioned in the Introduction, the experimental results show a negligible difference of about 2 nm in the maximum wavelength when the drug's surrounding environment is changed from purely aqueous to lipidic.[24] Computed spectra in the three different chemical environments, namely, aqueous solution, top of the barrier, and energy minimum (see Figures 2 and 3), are displayed in Figure 5 for both drugs. It is evident that beyond the different absolute intensities, there are no significant changes in the spectroscopic information related to the position of the main bands. Such an analogous behavior in diverse environments has been reported by Cwiklik et al.[33] who studied the changes in the absorption spectra for the insertion of PRODAN into palmitoyloleoylphosphatidylcholine (POPC) bilayers and pointed out a relatively weak or almost imperceptible dependence of the absorption energy on the type of environment. A better comparison of the convoluted spectra of IBU^- and NAP^-

can be seen in Figure S4 in the SI.

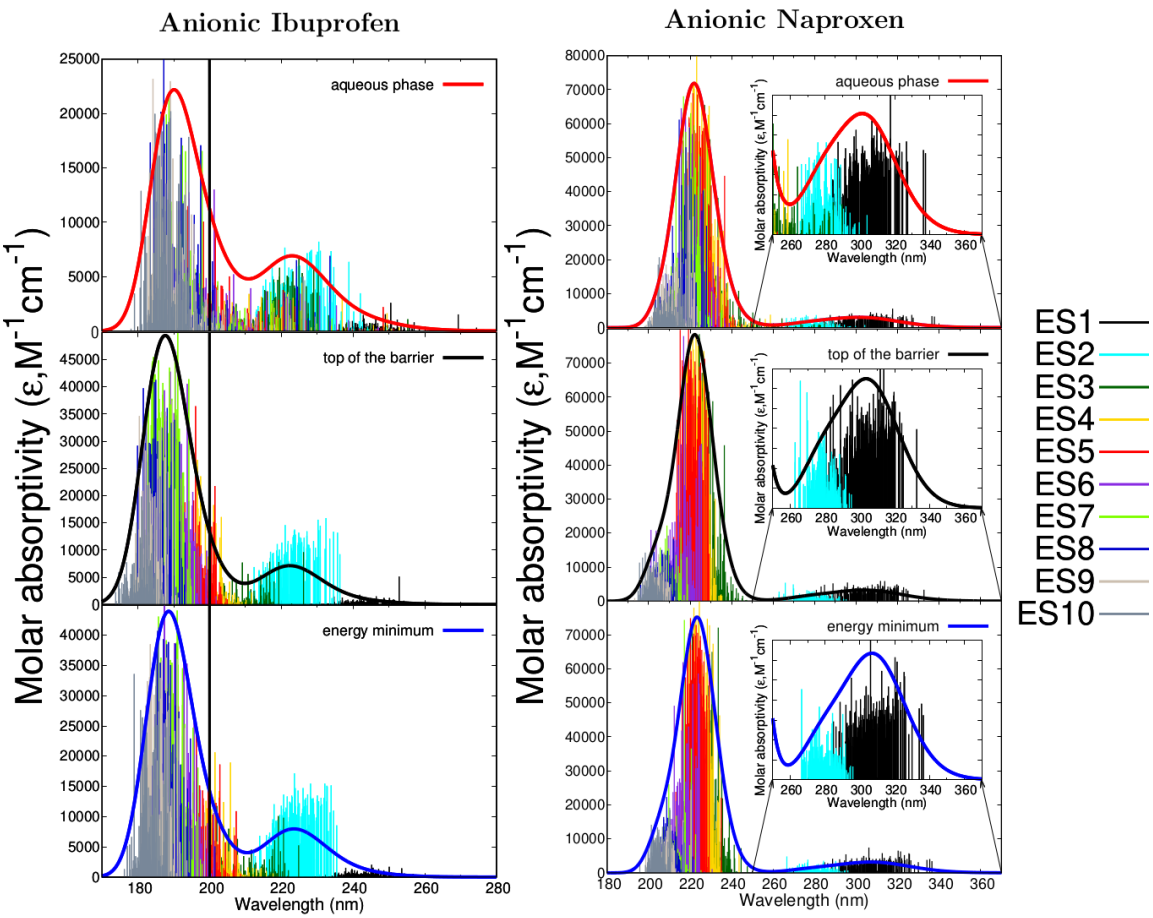


Figure 5: Stick-like UV-Vis spectra of Ibuprofen (left) and Naproxen (right) in the three different chemical environments considered in this work (see Figure 3). The shape of the convoluted spectrum is shown as a solid line. Each type of transition is color-coded, e. g., Excited state 1 (ES1) in black and Excited state 2 (ES2) in cyan. The vertical solid lines in the IBU^- case, mark the boundary up to the highest resolved experimental information. An inset showing the structure of the low-intensity band is also provided for NAP^-.

Figure 5 also reports the stick-like spectra of embedded IBU^- and NAP^- in all environments. It means that the raw data extracted from QM/MM calculations on each snapshot are reported as a position/oscillator strengths spectrum, which gives insight into the natural

spreading of the transition bands, both in wavelengths and intensities. In addition, each stick is colored depending on the associated transition (see labels at the right of the figure). In the IBU^- case, it is interesting to notice that for both the top of the barrier and the energy minimum, the band at 222 nm is the result of a single electronic transition to the excited state 2 (ES2). Conversely, in aqueous environment, although the major contribution comes from ES2, there are other electronic transitions leading to the appearance of such a band. These electronic transitions can be appreciated by taking a look at the diversity in the stick colors making part of the most intense absorption band (see Figure 5, top panel).

For NAP^-, the presence of the two fused aromatic rings makes the absorption spectra very distinct with respect to its ibuprofen counterpart which has only one ring. This structural feature has been documented to displace the position of the absorption maxima to longer wavelengths as in the benzene→naphthalene→anthracene series.[63] As displayed in the inset of the right panel of Figure 5, the first low-intensity band (between 250 and 340 nm) is basically made up of two transitions, i.e. ES1 (black color) and ES2 (cyan color). Additionally, the most intense band in NAP^- (240 nm) which comprises several excited states may overlap with the 222 nm band in IBU^- when IBU^-/NAP^- mixtures are considered, making it difficult to assign specific electronic transitions. For both anions in aqueous phase, these two bands, which are combinations of different electronic excited states, have a more spread distribution with respect to the other two environments. This observation may be attributed to the higher mobility of the drug in water solution, as compared to the limited freedom provided by the lipidic environment.

A deeper insight into the nature of the absorption bands can be attained by looking at the molecular orbitals (MOs) involved in the electronic transitions. IBU^- MOs in the three considered environments are depicted in Figure 6 along with their corresponding contributions to ES2. Table S1 in the SI lists the same information for the lowest-energy excited

state ES1. It can be observed that regardless of the environment, the excitation involves the HOMO and the first unoccupied MOs (LUMO, LUMO+1, LUMO+2 or LUMO+3 orbitals). Thus, the ES2 transition can be expressed as HOMO $\rightarrow$ 0.84(LUMO+1) + 0.55(LUMO+3), HOMO $\rightarrow$ LUMO and HOMO $\rightarrow$ 0.78LUMO + 0.62(LUMO+2) transitions for the drug into the aqueous phase, top of the barrier and energy minimum, respectively.

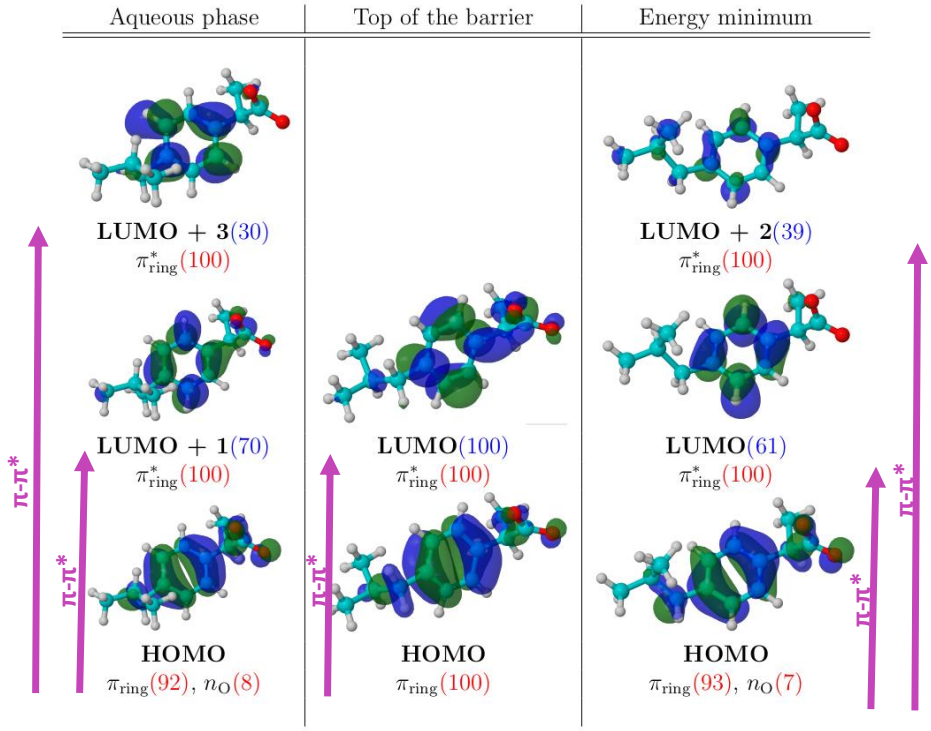


Figure 6: MOs involved in the electronic transition to ES2, responsible for the experimentally observed band of anionic Ibuprofen in different environments (see Figure 5, left panel). Blue and red numbers in parentheses correspond to the contributions of the MOs to the electronic transition and the contributions of the NBOs to the MOs after the CMO decomposition, respectively. Isosurface value: 0.03. These results correspond to a randomly chosen snapshot.

In addition, each MO can be characterized with the tools provided by NBO, specifically,

the Localized Analysis of CMOs, which allows for the assignment of the transition from/to specific portions of the chromophore. For instance, it is clearly seen that for IBU^-, the highest contributions to the HOMO orbitals come from both π bonds of the aromatic ring and from the lone pairs of the oxygen atoms always with a higher percentage of π_{ring}, independently of the environment. Furthermore, all unoccupied orbitals involve only π^* orbitals from the aromatic ring, validating the $\pi \rightarrow \pi^*$ assignment documented in the literature for the absorption band with a maximum at 222 nm when the drug is surrounded by aqueous environment.[56]

From the MO perspective, the NAP^- case is much more complex. Table S2 in the SI lists the orbitals having the largest coefficient for each electronic transition in one randomly chosen snapshot. It is important to stress that although we are showing results for one snapshot only, those findings are representative of the average of the electronic transitions for the entire group of configurations. The MOs involved in the ES1 and ES2 of NAP^- (giving rise to the band at 340 nm) are HOMO and HOMO-1. HOMO-1 has contributions of π orbitals from the aromatic ring and from the oxygen of the carboxylic group. Contributions from the ether oxygen are viewed in the case of the HOMO. Instead, excited states from ES3 to ES10 (leading to the highest intensity band) encompass an assorted set of occupied orbitals. Oscillator strengths in Table S2 in the SI indicate that ES5 and ES8 for aqueous phase, ES3 and ES4 for the top of the barrier, and ES4 and ES6 for energy minimum, are the primary excited states giving rise to the absorption intensities. Notice also that regardless of the environment, unoccupied orbitals (LUMO, LUMO+1,...) often involve the π^* from the aromatic rings. Several MOs involved in electronic transitions are displayed in Figure S5 as well as their CMO decomposition, which reinforces the $\pi \rightarrow \pi^*$ assignment link to that band.

It is worth mentioning that an implicit description of aqueous phase by means of PCM model yields similar MOs (and their contributions) as those obtained by using the atomistic

QM/MM approach. However, PCM cannot be exploited to describe the lipidic environment because a dielectric constant is not able to fairly represent the more complex top of the barrier and energy minimum environments. Thus, this highlights the need of using an atomistic QM/MM simulation when investigating electronic properties of large systems.

3.1.3 A short note on the electronic circular dichroism (ECD) spectra of anionic ibuprofen and naproxen

Chiroptical properties are particularly sensitive to the solute-environment interactions and to the instantaneous configuration of the system.[64] Further analysis of changes in the electronic structure of both drugs was done by calculating the ECD spectra. In Figure 7, the experimental (top) and convoluted (bottom) spectra for the drugs in aqueous environment are shown, exhibiting good agreement with the experimental positive and negative patterns for IBU^- and NAP^-, respectively. Simulated ECD spectra in the three critical points of the free energies profiles (see Figure 2) are included in Figure S6 in the SI and, as in the case of the UV-Vis spectra, there are no significant changes in the spectral curves in the different chemical environments.

For the particular case of naproxen, it has been revealed that the (R)- and (S) enantiomers had an unusual inversion in their ECD signals in the presence of ethanol and water when compared with polar aprotic solvents such as acetonitrile.[62] If we consider that this change in sign is only given by the polarity of the environment, the fact of maintaining the same ECD signs in our computational spectra demonstrates that there is not much change in the chemical environment as the drug goes from the aqueous phase to its location at the energy minimum, in the polar/non-polar interphase of the lipid bilayer.

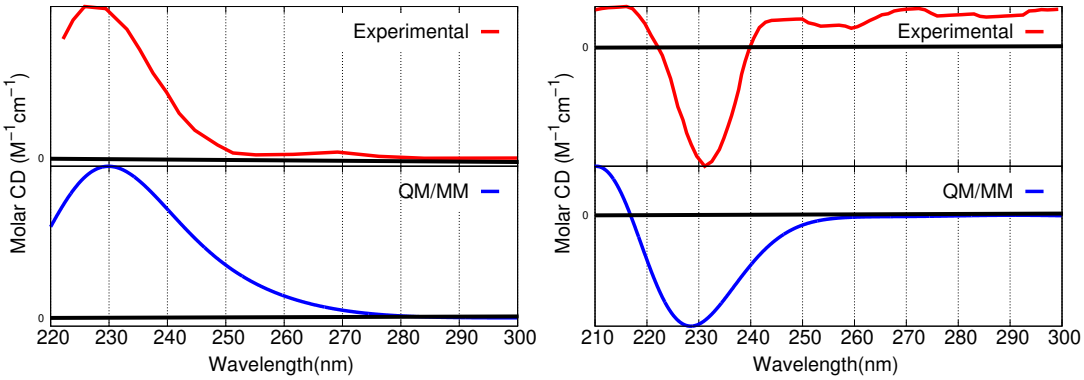


Figure 7: Experimental[62,65] (top) and simulated QM/MM CAM-B3LYP/6-311++g(d,p)/Electrostatic Embedding (bottom) Electronic Circular Dichroism spectra of IBU^- (left) and NAP^- (right) in aqueous solution. 200 snapshots were considered in the calculations. Averaged spectra were convoluted with an FWHM of 0.4 eV and normalized for comparison purposes.

Being clear that ECD properties of IBU^- and NAP^- do not depend on the environment, nor do their electronic absorption spectra, perhaps another spectroscopy like excited state emission spectra could be more fruitful to assess the insertion processes in cell membranes. Regarding emission spectra, for PRODAN in a membrane-type environment, a significant influence of the environment was observed in both experiments and simulations, in contrast to the weak dependence seen in the absorption processes.[33]

3.2 Analysis of spectral differences between all environments

From the above description of UV-Vis and ECD spectra, we have established a good agreement between experimental and calculated spectra for both drugs when they are embedded in diverse environments. From those results, it is clear the similarity between the UV-Vis and ECD spectra in the three regions, supported by almost the same molecular orbitals involved in the electronic transitions. Another piece of evidence to be added to that puzzling

situation is that there is no inversion of the ECD sign when polarity changes (as opposed to what was reported in Ref. 62). In order to give an answer to the key question of why there are no differences across the spectra in the several critical points of the free energy profile, it is necessary to analyze what is happening around the drug in its different locations in the system.

This fact can be explored by relying on the information reported in the works of Rojas-Valencia et al.,[21,34] In the first place, the authors analyzed the distribution of water molecules in the immediate vicinity (within 3.0 Å) of the entire drug and pointed out that no matter the environment, there are at least ten water molecules surrounding either IBU^- or NAP^-. Variations in the total number of drug-water interactions as a function of the distance from the bilayer center are collected in the left panel of Figure 8. Since our orbital analysis indicated that as a general rule the π orbitals are the most heavily involved ones in the electronic transitions, we went deeper and discriminate how many of those waters around the molecule were located close to the rings. The number of contacts between the solvent and the aromatic portion of each drug is plotted in the right panel of Figure 8. As expected, the majority of water molecules are placed near the carboxylate group because the formal charge is a strong attractor during the insertion process. Indeed, the microsolvation of IBU^- with up to three water molecules has been already studied[66] and it was clear that solvent molecules in direct contact with anionic Ibuprofen, preferred to cluster around the carboxylate oxygen atoms forming cyclic or bridged charge-assisted hydrogen bond networks. Here, we noticed that there are also persistent stabilizing interactions between the drug ring(s) and a few water molecules, about five and six for IBU^- or NAP^-, respectively. Therefore, it seems that the drug's local environment is quite similar in the three key points while it travels to the inner part of the bilayer, even if the global environment dictated by the presence or absence of lipids is changing.

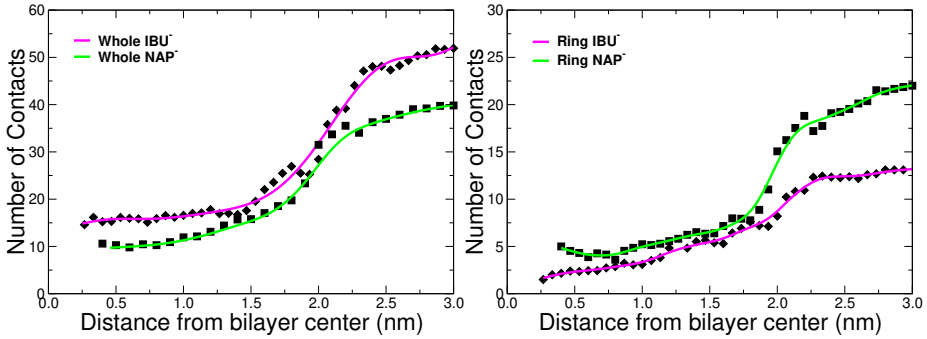


Figure 8: Number of water molecules within 3.0 Å of any given atom (left) or a specific atom of the ring(s) (right) in Naproxen and Ibuprofen during their insertion into the lipid bilayer.

Second, in Refs. 21,34, the same authors computed the NCI surfaces between the drug and its surroundings and found that there are many tiny collective non-covalent interactions throughout its insertion, with no significant qualitative changes. In fact, several of those interactions are associated with the ring$\cdots$solvent contacts mentioned above. These two aspects highlight the idea that the chemical environment around the NSAIDs stays approximately the same along its path from the aqueous environment up to the bottom of the energy profile. This translates into weak changes in the electronic structure reflected in turn in no large differences seen in the UV-Vis and ECD spectra.

4 Summary and conclusions

We have applied a multiscale QM/MM approach based on electrostatic embedding to simulate the UV-Vis absorption and ECD spectra of anionic naproxen and ibuprofen during their individual travel to the inner part of the cell membrane, by means of a series of TDDFT calculations. Spectra were calculated in three key points of the Gibbs free energy profile, namely, the aqueous phase, the top of the barrier, and the energy minimum, sampling configurations from MD trajectories previously reported. Having many accessible configurations leads to the arising of a natural broadening in the spectra. As a result of the application

of this methodology, all calculated spectra exhibit very good agreement when compared against available experimental data, being able to reproduce the main spectral features and the non-significant changes reported when the drug's spectra were measured in different environments. In all cases, electronic transitions were analyzed in terms of the orbitals involved. The CMO decomposition supports the $\pi \rightarrow \pi^*$ transitions assigned to the main bands. Our analysis of the excited states and orbitals, complemented by the quantification of the number of contacts between the drug and the solvent molecules along the insertion, allows us to gain insight into the origin of the small or almost imperceptible differences in both electronic absorption and circular dichroism spectra. In particular, we show that regardless of the environment there are always water molecules in the immediate vicinity of the ring/aromatic portion of IBU^- and NAP^-, thus maintaining similar local surroundings and not affecting to a large extent the electronic properties of these two compounds. Our calculations also indicate that the choice of the DFT functional and type of embedding were important steps, having large impacts on the accuracy of the results obtained. Finally, our findings are encouraging for using these methodologies when studying complex systems like drugs in lipidic environments and to understand where and how these or other spectroscopies could be eventually useful to distinguish between insertion or intercalation phenomena. In this respect, fluorescence has been proven to be useful in the case of other complex systems[33] and will be investigated in future works.

Acknowledgement

NRV thanks the Ministerio de Ciencias y Tecnología de Colombia, MINCIENCIAS, for a postdoctoral fellowship. FNZ acknowledges Universidad de Medellín for continuous support. Internal support from Universidad de Antioquia via "Estrategia para la sostenibilidad" is also acknowledged. We gratefully acknowledge the Center for High Performance Computing (CHPC) at SNS for providing the computational infrastructure.

Supporting Information Available

Plots for the analysis of the level of theory, the convergence of the UV-vis spectra according to the number of snapshots, and the UV-vis and ECD spectra for the three phases considered in this work. Tables including the main molecular orbitals involved in the ES1 for ibuprofen and all transitions for Naproxen. The oscillator strength values for naproxen in all environments were also included. Figure representing the most important MOs involved in the electronic transitions of naproxen in the three phases.

References

(1) Ghlichloo, I.; Gerriets, V. Nonsteroidal anti-inflammatory drugs (NSAIDs). **2019**,

(2) Pereira-Leite, C.; Nunes, C.; Reis, S. Interaction of nonsteroidal anti-inflammatory drugs with membranes: in vitro assessment and relevance for their biological actions. *Prog. Lipid Res* **2013**, *52*, 571–584.

(3) Ngo, V. T. H.; Bajaj, T. *Ibuprofen*; StatPearls Publishing, 2021.

(4) Brutzkus, J. C.; Shahrokhi, M.; Varacallo, M. *Naproxen*; StatPearls Publishing, 2022.

(5) Schiff, M.; Minic, M. Comparison of the analgesic efficacy and safety of nonprescription doses of naproxen sodium and Ibuprofen in the treatment of osteoarthritis of the knee. *J. Rheumatol.* **2004**, *31*, 1373–1383.

(6) Barbato, F.; La Rotonda, M. I.; Quaglia, F. Interactions of nonsteroidal antiinflammatory drugs with phospholipids: comparison between octanol/buffer partition coefficients and chromatographic indexes on immobilized artificial membranes. *J. Pharm. Sci.* **1997**, *86*, 225–229.

(7) Evans, A. M. Comparative pharmacology of S (+)-ibuprofen and (RS)-ibuprofen. *Clin. Rheumatol.* **2001**, *20*, 9–14.

(8) Kean, W. F.; Lock, C. J.; Rischke, J.; Butt, R.; Buchanan, W. W.; Howard-Lock, H. Effect of R and S enantiomers of naproxen on aggregation and thromboxane production in human platelets. *J. Pharm. Sci.* **1989**, *78*, 324–327.

(9) Vane, J.; Botting, R. Anti-inflammatory drugs and their mechanism of action. *Inflamm. Res.* **1998**, *47*, 78–87.

(10) Duggan, K. C.; Walters, M. J.; Musee, J.; Harp, J. M.; Kiefer, J. R.; Oates, J. A.; Marnett, L. J. Molecular basis for cyclooxygenase inhibition by the non-steroidal anti-inflammatory drug naproxen. *J. Biol. Chem.* **2010**, *285*, 34950–34959.

(11) Goodwin, G. *Prostaglandins: biochemisty, functions, types and roles*; Nova science publisher, New York, USA, 2010.

(12) Tsuchiya, H.; Mizogami, M. Membrane interactivity of non-steroidal anti-inflammatory drugs: a literature review. *J. adv. med* **2020**, *31*, 1–30.

(13) Manrique-Moreno, M.; Villena, F.; Sotomayor, C. P.; Edwards, A. M.; Muñoz, M. A.; Garidel, P.; Suwalsky, M. Human cells and cell membrane molecular models are affected in vitro by the nonsteroidal anti-inflammatory drug ibuprofen. *Biochim. Biophys. Acta - Biomembr.* **2011**, *1808*, 2656–2664.

(14) Seydel, J. K.; Wiese, M. *Drug-membrane interactions: analysis, drug distribution, modeling*; John Wiley & Sons, 2009.

(15) Magalhães, L. M.; Nunes, C.; Lúcio, M.; Segundo, M. A.; Reis, S.; Lima, J. L. High-throughput microplate assay for the determination of drug partition coefficients. *Nat. Protoc.* **2010**, *5*, 1823–1830.

(16) Bennion, B. J.; Be, N. A.; McNerney, M. W.; Lao, V.; Carlson, E. M.; Valdez, C. A.; Malfatti, M. A.; Enright, H. A.; Nguyen, T. H.; Lightstone, F. C., et al. Predicting a

drug's membrane permeability: A computational model validated with in vitro permeability assay data. *J. Phys. Chem. B* **2017**, *121*, 5228–5237.

(17) Lichtenberger, L. M.; Zhou, Y.; Jayaraman, V.; Doyen, J. R.; O'Neil, R. G.; Dial, E. J.; Volk, D. E.; Gorenstein, D. G.; Boggara, M. B.; Krishnamoorti, R. Insight into NSAID-induced membrane alterations, pathogenesis and therapeutics: characterization of interaction of NSAIDs with phosphatidylcholine. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2012**, *1821*, 994–1002.

(18) Manrique-Moreno, M.; Garidel, P.; Suwalsky, M.; Howe, J.; Brandenburg, K. The membrane-activity of Ibuprofen, Diclofenac, and Naproxen: A physico-chemical study with lecithin phospholipids. *Biochim. Biophys. Acta - Biomembr.* **2009**, *1788*, 1296–1303.

(19) Boggara, M. B.; Mihailescu, M.; Krishnamoorti, R. Structural association of nonsteroidal anti-inflammatory drugs with lipid membranes. *J. Am. Chem. Soc.* **2012**, *134*, 19669–19676.

(20) Rojas-Valencia, N.; Lans, I.; Manrique-Moreno, M.; Hadad, C. Z.; Restrepo, A. Entropy drives the insertion of ibuprofen into model membranes. *Phys. Chem. Chem. Phys.* **2018**, *20*, 24869–24876.

(21) Rojas-Valencia, N.; Gómez, S.; Núñez-Zarur, F.; Cappelli, C.; Hadad, C.; Restrepo, A. Thermodynamics and Intermolecular Interactions during the Insertion of Anionic Naproxen into Model Cell Membranes. *J. Phys. Chem. B* **2021**, *125*, 10383–10391.

(22) Santos, N. C.; Prieto, M.; Castanho, M. A. Quantifying molecular partition into model systems of biomembranes: an emphasis on optical spectroscopic methods. *Biochim. Biophys. Acta - Biomembr.* **2003**, *1612*, 123–135.

(23) Fernandes, E.; Soares, T. B.; Gonçalves, H.; Lúcio, M. Spectroscopic studies as a

toolbox for biophysical and chemical characterization of lipid-based nanotherapeutics. *Front. Chem.* **2018**, *6*, 323.

(24) Du, L.; Liu, X.; Huang, W.; Wang, E. A study on the interaction between ibuprofen and bilayer lipid membrane. *Electrochim. Acta* **2006**, *51*, 5754–5760.

(25) Raab, M. T.; Prỳmek, A. K.; Giordano, A. N. Estimation of the ground and excited state dipole moments for ibuprofen and naproxen sodium using the solvatochromic shift method. *J. Undergrad. Chem. Res.* **2021**, *20*, 68.

(26) Chen, X.; Qiao, W.; Miao, W.; Zhang, Y.; Mu, X.; Wang, J. the Dependence of implicit Solvent Model parameters and electronic Absorption Spectra and photoinduced charge transfer. *Scientific reports* **2020**, *10*, 1–8.

(27) Ambrosetti, M.; Skoko, S.; Giovannini, T.; Cappelli, C. Quantum Mechanics/Fluctuating Charge Protocol to Compute Solvatochromic Shifts. *J. Chem. Theory Comput.* **2021**, *17*, 7146–7156.

(28) Brunk, E.; Rothlisberger, U. Mixed quantum mechanical/molecular mechanical molecular dynamics simulations of biological systems in ground and electronically excited states. *Chem. Rev.* **2015**, *115*, 6217–6263.

(29) Zhang, K.; Ren, S.; Caricato, M. Multistate QM/QM Extrapolation of UV/Vis Absorption Spectra with Point Charge Embedding. *J. Chem. Theory Comput.* **2020**, *16*, 4361–4372.

(30) Zuehlsdorff, T. J.; Isborn, C. M. Modeling absorption spectra of molecules in solution. *Int. J. Quantum Chem.* **2019**, *119*, e25719.

(31) Giovannini, T.; Macchiagodena, M.; Ambrosetti, M.; Puglisi, A.; Lafiosca, P.; Lo Gerfo, G.; Egidi, F.; Cappelli, C. Simulating vertical excitation energies of sol-

vated dyes: From continuum to polarizable discrete modeling. *Int. J. Quantum Chem.* **2019**, *119*, e25684.

(32) Goletto, L.; Giovannini, T.; Folkestad, S. D.; Koch, H. Combining multilevel Hartree–Fock and multilevel coupled cluster approaches with molecular mechanics: a study of electronic excitations in solutions. *Phys. Chem. Chem. Phys.* **2021**, *23*, 4413–4425.

(33) Cwiklik, L.; Aquino, A. J.; Vazdar, M.; Jurkiewicz, P.; Pittner, J.; Hof, M.; Lischka, H. Absorption and fluorescence of PRODAN in phospholipid bilayers: a combined quantum mechanics and classical molecular dynamics study. *J. Phys. Chem. A* **2011**, *115*, 11428–11437.

(34) Rojas-Valencia, N.; Gómez, S.; Montillo, S.; Manrique-Moreno, M.; Cappelli, C.; Hadad, C.; Restrepo, A. Evolution of Bonding during the Insertion of Anionic Ibuprofen into Model Cell Membranes. *J. Phys. Chem. B* **2020**, *124*, 79–90.

(35) Kästner, J. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 932–942.

(36) Skoko, S.; Ambrosetti, M.; Giovannini, T.; Cappelli, C. Simulating Absorption Spectra of Flavonoids in Aqueous Solution: A Polarizable QM/MM Study. *Molecules* **2020**, *25*, 5853.

(37) Giovannini, T.; Grazioli, L.; Ambrosetti, M.; Cappelli, C. Calculation of ir spectra with a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles. *J. Chem. Theory Comput.* **2019**, *15*, 5495–5507.

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

(39) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

(40) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664–5677.

(41) Gómez, S.; Bottari, C.; Egidi, F.; Giovannini, T.; Rossi, B.; Cappelli, C. Amide Spectral Fingerprints are Hydrogen Bonding-Mediated. *J. Phys. Chem. Lett* **2022**, *13*, 6200–6207.

(42) Dohn, A. O. Multiscale electrostatic embedding simulations for modeling structure and dynamics of molecules in solution: a tutorial review. *Int. J. Quantum Chem.* **2020**, *120*, e26343.

(43) Gómez, S.; Giovannini, T.; Cappelli, C. Multiple Facets of Modeling Electronic Absorption Spectra of Systems in Solution. *ACS Physical Chemistry Au* **2022**, *0*, –, DOI: 10.1021/acsphyschemau.2c00050.

(44) Gómez, S.; Giovannini, T.; Cappelli, C. Absorption spectra of xanthines in aqueous solution: A computational study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5929–5941.

(45) Gómez, S.; Egidi, F.; Puglisi, A.; Giovannini, T.; Rossi, B.; Cappelli, C. Unlocking the power of resonance Raman spectroscopy: The case of amides in aqueous solution. *J. Mol. Liq.* **2022**, *346*, 117841.

(46) Uribe, L.; Gómez, S.; Giovannini, T.; Egidi, F.; Restrepo, A. An efficient and robust procedure to calculate absorption spectra of aqueous charged species applied to NO 2-. *Phys. Chem. Chem. Phys.* **2021**, *23*, 14857–14872.

(47) Uribe, L.; Gómez, S.; Egidi, F.; Giovannini, T.; Restrepo, A. Computational hints for the simultaneous spectroscopic detection of common contaminants in water. *J. Mol. Liq.* **2022**, *355*, 118908.

(48) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

(49) Mennucci, B.; Corni, S. Multiscale modelling of photoinduced processes in composite systems. *Nat. Rev. Chem.* **2019**, *3*, 315–330.

(50) Cappelli, C. Integrated QM/polarizable MM/continuum approaches to model chiroptical properties of strongly interacting solute–solvent systems. *Int. J. Quantum Chem.* **2016**, *116*, 1532–1542.

(51) Gómez, S. A.; Rojas-Valencia, N.; Gómez, S.; Egidi, F.; Cappelli, C.; Restrepo, A. Binding of SARS-CoV-2 to Cell Receptors: A Tale of Molecular Evolution. *ChemBioChem* **2021**, *22*, 724–732.

(52) Gómez, S. A.; Rojas-Valencia, N.; Gómez, S.; Cappelli, C.; Restrepo, A. The Role of Spike Protein Mutations in the Infectious Power of SARS-COV-2 Variants: A Molecular Interaction Perspective. *ChemBioChem* **2022**, *23*, e202100393.

(53) Gómez, S.; Rojas-Valencia, N.; Gómez, S. A.; Cappelli, C.; Merino, G.; Restrepo, A. A molecular twist on hydrophobicity. *Chem. Sci.* **2021**, *12*, 9233–9245.

(54) Rojas-Valencia, N.; Gómez, S.; Guerra, D.; Restrepo, A. A detailed look at the bonding interactions in the microsolvation of monoatomic cations. *Phys. Chem. Chem. Phys.* **2020**, *22*, 13049–13061.

(55) Gómez, S. A.; Rojas-Valencia, N.; Gómez, S.; Lans, I.; Restrepo, A. Initial recognition and attachment of the Zika virus to host cells: A molecular dynamics and quantum interaction approach. *ChemBioChem* **2022**, e202200351.

(56) Gómez, S.; Rojas-Valencia, N.; Giovannini, T.; Restrepo, A.; Cappelli, C. Ring Vi-

brations to Sense Anionic Ibuprofen in Aqueous Solution as Revealed by Resonance Raman. *Molecules* **2022**, *27*, 442.

(57) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian 16 Revision B.01. 2016; Gaussian Inc. Wallingford CT.

(58) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Karafiloglou, P.; Landis, C. R.; Weinhold, F. NBO 7.0. 2018; Theoretical Chemistry Institute, University of Wisconsin, Madison, WI.

(59) Choina, J.; Kosslick, H.; Fischer, C.; Flechsig, G.-U.; Frunza, L.; Schulz, A. Photocatalytic decomposition of pharmaceutical ibuprofen pollutions in water over titania catalyst. *Appl. Catal. B: Environ* **2013**, *129*, 589 – 598.

(60) Picollo, M.; Aceto, M.; Vitorino, T. UV-Vis spectroscopy. *Phys. Sci. Rev.* **2019**, *4*.

(61) Saji, R. S.; Prasana, J. C.; Muthu, S.; George, J.; Kuruvilla, T. K.; Raajaraman, B. Spectroscopic and quantum computational study on naproxen sodium. *Spectrochim. Acta - A: Mol. Biomol. Spectrosc.* **2020**, *226*, 117614.

(62) Ximenes, V. F.; Morgon, N. H.; Robinson de Souza, A. Solvent-dependent inversion of circular dichroism signal in naproxen: An unusual effect! *Chirality* **2018**, *30*, 1049–1053.

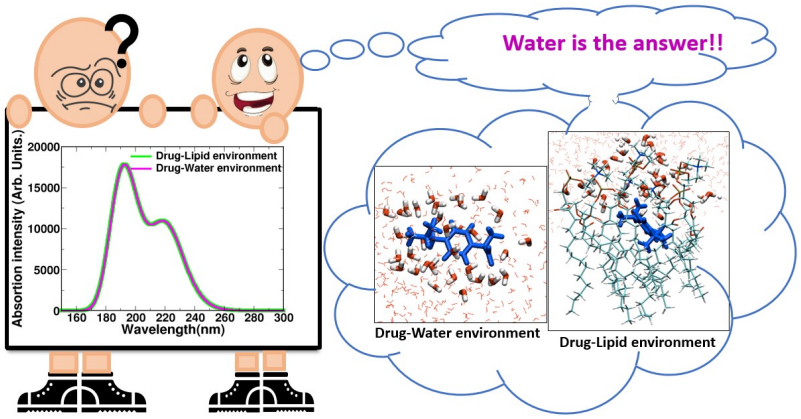
(63) Friedel, R. A.; Orchin, M. *Ultraviolet spectra of aromatic compounds*; Wiley New York, 1951; Vol. 40.

(64) Giovannini, T.; Egidi, F.; Cappelli, C. Theory and algorithms for chiroptical properties and spectroscopies of aqueous systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864–22879.

(65) Liu, M.; Chen, L.; Tian, T.; Zhang, Z.; Li, X. Identification and quantitation of enantiomers by capillary electrophoresis and circular dichroism independent of single enantiomer standard. *Anal. Chem.* **2019**, *91*, 13803–13809.

(66) Zapata-Escobar, A.; Manrique-Moreno, M.; Guerra, D.; Hadad, C. Z.; Restrepo, A. A combined experimental and computational study of the molecular interactions between anionic ibuprofen and water. *J. Chem. Phys.* **2014**, *140*, 184312.

TOC Graphic


Water is the answer!!
Absorption intensity (Arb. Units)
20000
15000
10000
5000
0
160 180 200 220 240 260 280 300
Wavelength(nm)
Drug-Lipid environment
Drug-Water environment
Drug-Water environment
Drug-Lipid environment

Chapter 6

Drugs and DNA

This chapter is related to the COSINE project in which the goal was the development of a multiscale QM/MM protocol for the calculation of resonance spectroscopies (RR and RROA) of DNA intercalated probes. After identifying one of the strongest binding DNA sequences of the widely used chemotherapy drug, DOX, several snapshots are extracted from MD trajectories and spectral properties are computed.

Initially, absorption spectra of DOX/Water and DOX/DNA/Water systems are calculated by exploiting the recently implemented coupling between the DFTB approach and the FQ model.[8] In the process of trying different approaches for modeling the MM portion in the QM/MM calculations of the DOX/DNA/Water systems, the DNA basis pairs are included in the QM portion and the solvent molecules are treated at the FQ level. The obtained TD-DFTB/FQ spectral profiles are perfectly comparable with those obtained at the pure DFT level.

Then, for the vibrational spectroscopies, in particular RR, an innovative methodology of projecting normal modes to avoid changes in the chemical environment of the DOX by doing partial or full optimizations, is presented. Diverse methods to obtain the normal modes of DOX are evaluated, ranging from optimizing conformations or rotating normal modes to projecting out a few of the soft coordinates. Once the projected normal modes are obtained, RR spectra are computed by means of the geometrical derivative of the complex electronic polarizability under resonance conditions via linear response theory. The results show a good agreement with experimental data, and they are better than those reported when the DNA or the aqueous environment are represented just by a dielectric constant, as is done in continuum approaches.

For the DOX-containing systems studied here, the good agreement between computed and experimental data allows us to easily identify the main features of RR spectra in both aqueous and DNA solutions, confirming the reliability of the methodology. Nevertheless, the protocol is general and can be used to study other drug-DNA complexes and several DNA sequences, as long as the conformational diversity of the target and the environment are taken into account by sampling the phase-space through MD simulations.

6.1. From MD sampling to UV-Vis and RR spectroscopies

6.1.1. DFTB/FQ to model absorption spectroscopies of large systems[8]

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

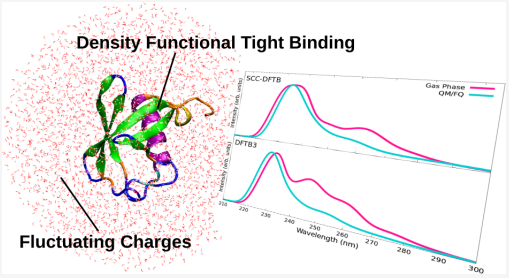
Absorption Properties of Large Complex Molecular Systems: The DFTB/Fluctuating Charge Approach

Piero Lafiosca, Sara Gómez, Tommaso Giovannini,* and Chiara Cappelli*

Cite This: *J. Chem. Theory Comput.* 2022, 18, 1765−1779 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We report on the first formulation of a novel polarizable QM/MM approach, where the density functional tight binding (DFTB) is coupled to the fluctuating charge (FQ) force field. The resulting method (DFTB/FQ) is then extended to the linear response within the TD-DFTB framework and challenged to study absorption spectra of large condensed-phase systems.



1. INTRODUCTION

The theoretical modeling of large molecular systems, with application in biological and technological fields, is one of the most challenging tasks for theoretical and computational chemistry.[1,2] In fact, the description of large molecular systems requires the treatment of a large number of degrees of freedom, from both nuclear and electronic points of view.[1,2] For this reason, high-level quantum mechanics (QM) methods are usually not applicable because they are usually associated with an unfavorable scaling with the number of atoms (and electrons).[3] Different strategies, usually based on chemical intuition, can be exploited to reduce the dimensionality of the system and to make high-level QM approaches applicable.[4−9] This is, for instance, the case for local excitations, which take place in a specific part of the considered molecule.[5,8,10] However, in many cases, especially for biomolecules, such an approximation may not be chemically justified because the phenomenon is the result of changes in the whole structure.

Semiempirical QM methods have been developed to treat in a realistic way this kind of system, which can be constituted by thousands of atoms.[11−14] Such methodologies introduce a set of integral approximations and parametrizations that make the computation particularly cheap. Clearly, the accuracy of each approach strongly depends on the quality of the parametrization. Among semiempirical methods, one of the most used is the density functional tight binding (DFTB) approach.[15−17] The theoretical starting point of such method is the density functional theory (DFT) energy in the Kohn−Sham (KS) framework, expressed by means of a linear combination of atomic orbitals (LCAO) over a minimal basis set. This quantity is then approximated by means of a Taylor expansion with respect to a reference density truncated at different orders by generating a hierarchy of DFTB methods.[15] In particular, the self-consistent charge DFTB approach (SCC-DFTB), which corresponds to a second-order expression of the KS energy, has been successfully applied to the calculation of energies, geometries, and vibrational frequencies of small organic molecules; its accuracy when compared with experimental values is comparable to that of full DFT calculations performed with a double-ζ plus polarization basis set.[17] Moreover, a time-dependent DFTB (TD-DFTB) approach has been developed to calculate excitation energies in a tight-binding fashion.[18,19] However, it has been shown that the standard pure or hybrid DFT functionals are not able to accurately treat charge-transfer excitations because their extension to the corresponding long-range-corrected versions is necessary.[20,21] In this context, the time-dependent long-range-corrected DFTB approach (TD-LC-DFTB)[22] has recently been proposed and explicitly designed for the treatment of charge-transfer states in large chromophores.

The DFTB approximation allows the boundaries of the systems treatable by most *ab initio* approaches to be pushed, making possible the QM description of large biomolecules, such as proteins.[19,23] However, most biomolecules are typically dissolved in an external environment, as water is the most common physiological solvent.[24] As for small organic molecules, also in such cases, the external aqueous solution may strongly

Received: October 22, 2021
Published: February 20, 2022

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

affect the properties of the biological system.[25] To take into account the solvent effect provided by water, the best compromise between computational cost and accuracy is to resort to the so-called focused models, in which the target system and the environment are described at different levels of theory based on the assumption that the phenomenon is carried on by the target and the environment just perturbs it.[26−28] Among the different focused models, the most accurate are the polarizable QM/molecular mechanics approaches.[29−44] In such methods, the environment molecules are classically and atomistically described by means of a polarizable force field, and mutual solute−solvent polarization is taken into account. In particular, excellent performances have been reported for the QM/fluctuating charge (QM/FQ) in the description of aqueous solutions[45] and recently for different solvents.[46] In such an approach, each solvent atom is endowed with a charge that is adjusted to the external potential generated by the solute density.[28,45,47] Such charges then polarize the QM density by entering the QM Hamiltonian in a mutual polarization fashion. In its basic formulation, the QM/FQ interaction is limited to electrostatics; however, nonelectrostatic interactions can also be considered.[48]

In this work, we have substantially extended the applicability of the polarizable QM/FQ approach by proposing a novel polarizable QM/FQ scheme based on the DFTB approach for the QM portion, allowing for the treatment of large, complex biomolecular systems. To the best of our knowledge, this is the first time that DFTB has been coupled to a polarizable MM approach. The newly developed DFTB/FQ approach has also been extended to the linear response regime by means of the time-dependent DFTB (TD-DFTB) approximation,[18,19] and it has been tested to reproduce the excitation energies of doxorubicin (DOX), an anticancer drug, in aqueous solution and intercalated in DNA and ubiquitin (UBI) protein dissolved in aqueous solution. The Article is organized as follows: In the next section, we briefly recall the DFTB approach, and we formulate the coupling between the DFTB and FQ portions for both ground-state and excitation energies calculations. DFTB/FQ is then applied to the calculation of the excitation energies of DOX, the DOX−DNA complex, and the UBI protein in aqueous solution. Conclusions and perspectives end the Article.

2. THEORY

In this section, the theoretical background of the DFTB/FQ approach is described. To this end, the fundamentals of DFTB and polarizable FQ approaches are briefly recapped, and the formulation of the DFTB/FQ coupling is presented. Then, the extension of the model to the linear response in a TD framework is discussed.

2.1. DFT Basis of TB Theory. In the general DFT framework, the energy functional in the KS picture reads[49]

$$E[\rho] = \sum_i^{\text{occ}} \int \mathrm{d}\mathbf{r}\, \psi_i(\mathbf{r})\left[-\frac{\Delta}{2} + V^{\text{ext}}(\mathbf{r}) + \frac{1}{2}\int \mathrm{d}\mathbf{r}'\, \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\right]\psi_i(\mathbf{r}) + E^{\text{xc}}[\rho] + V_{\text{NN}} \quad (1)$$

where ψ_i are occupied KS eigenstates, Δ is the Laplacian operator, V^{ext} is the external potential associated with the nuclei−electron interaction, E^{xc} is the exchange-correlation contribution, and V_{NN} is the nuclei−nuclei repulsion term.

In the DFTB theory, the electronic density ρ is expressed as $\rho = \rho_0 + \delta\rho$, where ρ_0 is a reference input density and $\delta\rho$ is a fluctuation, which is assumed to be small.[16,19,50] Within this assumption, the exchange-correlation energy contribution can be expanded in a Taylor expansion, and eq 1 becomes

$$E[\rho_0 + \delta\rho] = E^{\mathcal{H}^0} + E^{\text{rep}} + E^{\gamma} + E^{\Gamma} + \ldots \quad (2)$$

$$E^{\mathcal{H}^0} = \sum_i^{\text{occ}} \int \psi_i(\mathbf{r})\left[-\frac{\Delta}{2} + V^{\text{ext}}(\mathbf{r}) + \frac{1}{2}\int' \frac{\rho_0}{|\mathbf{r}-\mathbf{r}'|} + V^{\text{xc}}[\rho_0]\right]\psi_i(\mathbf{r}) \quad (3)$$

$$E^{\text{rep}} = -\frac{1}{2}\int'\int \frac{\rho_0\rho_0'}{|\mathbf{r}-\mathbf{r}'|} - \int V^{\text{xc}}[\rho_0]\rho_0 + E^{\text{xc}}[\rho_0] + V_{\text{NN}} \quad (4)$$

$$E^{\gamma} = \frac{1}{2}\int'\int \left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \left.\frac{\delta^2 E^{\text{xc}}}{\delta\rho\delta\rho'}\right|_{\rho_0}\right)\delta\rho\delta\rho' \quad (5)$$

$$E^{\Gamma} = \frac{1}{6}\int''\int'\int \left.\frac{\delta}{\delta\rho''}\frac{\delta^2 E^{\text{xc}}}{\delta\rho\delta\rho'}\right|_{\rho_0} \delta\rho\delta\rho'\delta\rho'' \quad (6)$$

where we have introduced the expectation value of the zeroth-order Hamiltonian $E^{\mathcal{H}^0}$ (which depends only on the reference density ρ_0) and the so-called repulsive energy contribution E^{rep}. E^{γ} and E^{Γ} collect the second- and third-order energy terms. Notice that in eqs 3−6, the usual shorthand notation such that $\int = \int \mathrm{d}\mathbf{r}$, $\delta\rho = \delta\rho(\mathbf{r})$, $\int' = \int \mathrm{d}\mathbf{r}'$, $\delta\rho' = \delta\rho(\mathbf{r}')$, and so on, is used.[16,19,50]

Different DFTB methods can be defined by truncating the Taylor expansion in eq 2 at different orders. The most basic approach consists of neglecting E^{γ} and E^{Γ} in eq 2. This gives rise to a set of non-self-consistent KS equations because the zeroth-order Hamiltonian depends only on the reference density, ρ_0. The repulsive energy E^{rep} is approximated as a sum of repulsive, short-ranged, two-body potentials, defined in terms of a set of parameters.[51] The Hamiltonian $\mathcal{H}^0_{\mu\nu}$ and overlap $S_{\mu\nu} = \langle\phi_\mu|\phi_\nu\rangle$ matrix elements are calculated at a set of relevant interatomic distances and are tabulated. By this, they do not need to be computed for each DFTB calculation, and this results in substantial computational savings as compared with standard DFT. Notice that various parametrizations for E^{rep} and the $\mathcal{H}^0_{\mu\nu}$ and $S_{\mu\nu}$ matrix elements have been proposed.[52]

A more sophisticated DFTB method, the SCC-DFTB, can be obtained by retaining E^{γ} in eq 2. The density fluctuation $\delta\rho$ is expressed as a sum of localized atomic contributions, $\delta\rho = \sum_\alpha \delta\rho_\alpha$, which are subsequently approximated through the monopolar term of a multipolar expansion,[18] that is,

$$\delta\rho_\alpha(\mathbf{r}) = \Delta q_\alpha F_\alpha(\mathbf{r}) \quad (7)$$

where $F_\alpha(\mathbf{r})$ is a normalized spherical density fluctuation centered on the αth atom, whereas the net charge $\Delta q_\alpha = q_\alpha - q^0_\alpha$ is computed through a Mulliken charge analysis. Within such an assumption, E^{γ} can be rewritten as

$$E^{\gamma} = \frac{1}{2}\int'\int \left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \left.\frac{\delta^2 E^{\text{xc}}}{\delta\rho\delta\rho'}\right|_{\rho_0}\right)\delta\rho\delta\rho' \approx \frac{1}{2}\sum_{\alpha\beta}\gamma_{\alpha\beta}\Delta q_\alpha \Delta q_\beta \quad (8)$$

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

where γ reads

$$\gamma_{\alpha\beta} = \int' \int \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta^2 E^{xc}}{\delta\rho\delta\rho'}\bigg|_{\rho_0} \right) F_\alpha(\mathbf{r}) F_\beta(\mathbf{r}') \tag{9}$$

Therefore, the total Hamiltonian matrix $\mathcal{H}_{\mu\nu}$ can be written as

$$\mathcal{H}_{\mu\nu} = \mathcal{H}^0_{\mu\nu} + \frac{1}{2} S_{\mu\nu} \sum_\xi (\gamma_{\alpha\xi} + \gamma_{\beta\xi}) \Delta q_\xi \tag{10}$$

Δq_ξ explicitly depends on MO coefficients through the density matrix, thus introducing a nonlinearity in the Hamiltonian. As a result, DFTB equations must be solved iteratively.

Last, the third-order term E^Γ in eq 2 may also be retained, such as in the DFTB3 approach.[53,54]

2.2. DFTB/FQ Approach. As stated in the Introduction, in this work, DFTB is coupled to the polarizable FQ force field, which represent each classical atom in terms of a charge q, which is allowed to "fluctuate" so as to fulfill the electronegativity equalization principle, which states that the instantaneous electronegativity χ of each atom must be the same at equilibrium. The total charge on each FQ molecule is fixed to a certain value Q by using Lagrangian multipliers λ. The FQ energy can be written as[47]

$$E_{FQ}[\mathbf{q}, \lambda] = \mathbf{q}_\lambda^\dagger \mathbf{C}_Q + \frac{1}{2} \mathbf{q}_\lambda^\dagger \mathbf{M} \mathbf{q}_\lambda \tag{11}$$

where $\mathbf{q}_\lambda$ is the vector of FQ charges and Lagrange multipliers, $\mathbf{C}_Q$ is a vector collecting atomic electronegativities and charge constraints Q, and the $\mathbf{M}$ matrix takes into account the interaction kernel between FQ charges and Lagrangian blocks. In particular, the diagonal elements of the FQ–FQ block of $\mathbf{M}$ account for the charge self-interaction by means of the chemical hardness η.[45] The minimization of the energy functional in eq 11 leads to a set of linear equations; their solution yields the FQ charges, that is

$$\mathbf{M}\mathbf{q}_\lambda = -\mathbf{C}_Q \tag{12}$$

Within a two-layer QM/MM scheme, the total energy of the DFTB/FQ system is written as

$$\mathcal{E} = E_{DFTB} + E_{DFTB/FQ} + E_{FQ} \tag{13}$$

where E_{DFTB} and E_{FQ} represent the energies of the DFTB and FQ portions and $E_{DFTB/FQ}$ is the interaction energy between the two layers. Here, similarly to most QM/MM approaches, a purely classical interaction term is considered; that is, the DFTB and FQ portions interact through the electrostatic potential generated on the FQ charges by the *total* DFTB density, that is, the reference density ρ_0 and the density fluctuation $\delta\rho$. Within the DFTB framework, the QM/FQ interaction can be approximated by only taking into account the potential generated by $\delta\rho$, similar to alternative DFTB/classical couplings.[55−57] Therefore, the corresponding approximated molecular electrostatic potential at the ith FQ charge placed at $\mathbf{r}_i$ can be written as

$$V_i(\rho) = V(\mathbf{r}_i) = -\int \frac{\delta\rho}{|\mathbf{r}_i - \mathbf{r}|} = -\sum_\alpha^{\text{nuclei}} \frac{\Delta q_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \tag{14}$$

where the implicit dependence of the electric potential on the density matrix through Mulliken charges is highlighted. Notice that in eq 14, the integration over the normalized spherical density fluctuations $F_\alpha(\mathbf{r})$ should be included. (See eq 7.) However, because the distance between FQ charges and QM atoms is typically larger than any intramolecular distance, we can safely assume that $F_\alpha(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{R}_\alpha)$. Therefore, density fluctuations can be described through a set of Mulliken point charges.

Moving back to the total DFTB/FQ energy functional, it can be rewritten as

$$\begin{aligned} \mathcal{E}[\delta\rho, \mathbf{q}, \lambda] &= E_{DFTB}[\delta\rho; \rho_0] + \mathbf{q}_\lambda^\dagger \mathbf{C}_Q \\ &\quad + \frac{1}{2}\mathbf{q}_\lambda^\dagger \mathbf{M} \mathbf{q}_\lambda - \sum_i^{FQs} \sum_\alpha^{\text{nuclei}} q_i \frac{\Delta q_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \\ &= E_{DFTB}[\delta\rho; \rho_0] + \mathbf{q}_\lambda^\dagger \mathbf{C}_Q + \frac{1}{2}\mathbf{q}_\lambda^\dagger \mathbf{M} \mathbf{q}_\lambda \\ &\quad + \mathbf{q}^\dagger \mathbf{V}(\delta\rho) \end{aligned} \tag{15}$$

Minimization of the energy functional in eq 15 with respect to both charges and Lagrangian multipliers yields the following linear system

$$\mathbf{M}\mathbf{q}_\lambda = -\mathbf{C}_Q - \mathbf{V}(\delta\rho) \tag{16}$$

The right-hand side of eq 16 collects both atomic electronegativities and the electric potential generated by the DFTB density. The latter term accounts for the mutual polarization among the DFTB and FQ portions of the system. In fact, KS equations need to be modified to include the DFTB/FQ contribution to the Hamiltonian matrix, which reads

$$\begin{aligned} \mathcal{H}^{DFTB/FQ}_{\mu\nu} &= -\frac{1}{2} \sum_i^{FQs} q_i S_{\mu\nu} \left(\frac{1}{|\mathbf{r}_i - \mathbf{R}_\alpha|} + \frac{1}{|\mathbf{r}_i - \mathbf{R}_\beta|} \right) \\ &= -\frac{1}{2} \sum_i^{FQs} q_i S_{\mu\nu} (\gamma^{FQ}_{\alpha i} + \gamma^{FQ}_{\beta i}) \end{aligned} \tag{17}$$

where the kernel γ^{FQ} that takes into account the interaction between the αth Mulliken charge (or the basis function μ) and the ith FQ charge. Notice that the DFTB/FQ term to the total energy and Hamiltonian matrix is the same for both the SCC-DFTB and DFTB3 methods because the Mulliken-based expansion for the density fluctuation $\delta\rho$ does not change. Note finally that the formulation presented above is not limited to FQ but can easily be extended to any kind of variational polarizable MM approach.[58]

2.3. Linear Response Regime. The extension of the approach to the linear response regime allows the calculation of some spectral signals and, in particular, vertical transition energies and absorption spectra. The TD-DFTB eigenproblem can be expressed in the Casida formalism as[18]

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{18}$$

where the eigenvalues ω correspond to excitation energies and the eigenvectors $\mathbf{X}$ and $\mathbf{Y}$ correspond to single-particle excitations and de-excitation amplitudes. Similarly to DFT/FQ,[34,59−63] to take into account the FQ layer, we need to modify the DFTB response matrices $\mathbf{A}$ and $\mathbf{B}$ as follows

$$\begin{aligned} A_{ai,bj} &= (\varepsilon_a - \varepsilon_i)\delta_{ab}\delta_{ij} + K_{ai,bj} + K^{FQ}_{ai,bj} \\ B_{ai,bj} &= K_{ia,bj} + K^{FQ}_{ai,jb} \end{aligned} \tag{19}$$

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

where the indices i, j and a, b run over the occupied and virtual molecular orbitals with energies ε. $K_{ia,jb}$ and $K^{FQ}_{ia,jb}$ are the DFTB and FQ coupling matrices, respectively. $K_{ia,jb}$ is usually simplified by exploiting the so-called γ-approximation,[18] similarly to the ground state in SCC-DFTB. (See eq 7.) In such an approximation, the transition density $p_{ia}(\mathbf{r}) = \psi_i(\mathbf{r})\psi_a(\mathbf{r})$ is decomposed as a sum of atomic contributions that, after a multipolar expansion, is approximated by means of the monopole term only. (See also eq 7.) Therefore, $K_{ia,jb}$ reads

$$K_{ia,jb} = \sum_{\alpha\beta}^{\text{nuclei}} q_\alpha^{ia} q_\beta^{jb} \overline{\gamma}_{\alpha\beta} \tag{20}$$

where q_α^{ia} and q_β^{jb} are Mulliken atomic transition charges. The $\overline{\gamma}$ functional is defined as in eq 9; however, the functional derivative of E_{xc} is evaluated on ρ. For systems with small charge-transfer effects, γ slightly depends on atomic charges, so that $\overline{\gamma}_{\alpha\beta}$ can be approximated with its ground-state counterpart $\gamma_{\alpha\beta}$.

By following refs 34, 45, and 47, the FQ contribution to the coupling matrix can be defined as

$$K^{FQ}_{ia,jb} = -\int' \int \psi_i(\mathbf{r})\psi_a(\mathbf{r}) \left[\sum_{pq}^{\text{FQs}} \frac{1}{|\mathbf{r} - \mathbf{r}_p|} M^{-1}_{pq} \frac{1}{|\mathbf{r}' - \mathbf{r}_q|} \right] \psi_j(\mathbf{r}')\psi_b(\mathbf{r}') \tag{21}$$

At this point, we can exploit the DFTB γ approximation, and eq 21 becomes

$$K^{FQ}_{ia,jb} = \sum_{\alpha\beta}^{\text{nuclei}} q_\alpha^{ia} q_\beta^{jb} \hat{\gamma}_{\alpha\beta} \tag{22}$$

where $\hat{\gamma}_{\alpha\beta}$ is defined as

$$\hat{\gamma}_{\alpha\beta} = \int' \int \left[\sum_{pq} \frac{1}{|\mathbf{r} - \mathbf{r}_p|} M^{-1}_{pq} \frac{1}{|\mathbf{r}' - \mathbf{r}_q|} \right] F_\alpha(\mathbf{r}) F_\beta(\mathbf{r}') \tag{23}$$

Similar to the ground state, we can assume $F_\alpha(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{R}_\alpha)$. Thus we obtain

$$\hat{\gamma}_{\alpha\beta} \approx \sum_{pq} \frac{1}{|\mathbf{r}_p - \mathbf{R}_\alpha|} M^{-1}_{pq} \frac{1}{|\mathbf{r}_q - \mathbf{R}_\beta|} = \gamma^{FQ}_{\alpha p} M^{-1}_{pq} \gamma^{FQ}_{\beta q} \tag{24}$$

where the interaction kernel defined in eq 17 is considered.

3. COMPUTATIONAL DETAILS

The equations presented in the previous section were implemented in a modified version of the Amsterdam Molecular Suite (AMS), release 2020.202 program.[23,64] TD-DFTB/FQ calculations were performed on 200 configurations extracted from MD simulations already reported in the literature.[65−67]

For DOX in the gas phase, we ran calculations on the same conformations coming from the MD after removing surrounding water molecules. Table 1 lists an inventory of the different systems/environments addressed in this Article and some other results that will be presented in the following discussion.

To explore diverse DFTB Hamiltonians, we relied on the Slater−Koster-based DFTB class and performed TD-DFTB calculations for the entire set of snapshots using both the second-order self-consistent charge extension SCC-DFTB (recently also called DFTB2) and the third-order extension known as DFTB3. These two DFTB schemes are thoroughly explained elsewhere (see, e.g., ref 70), but in short, SCC takes into account density fluctuations with improvements in the description of the polar bonds; likewise, DFTB3 describes hydrogen-bonded complexes and proton affinities, although at a little higher computational cost than SCC-DFTB calculations. SCC-DFTB and DFTB3 TD-DFTB calculations were performed by using mio-1-1[71] and 3ob-3-1[72] parameter sets, respectively.

Finally, to reduce the computational effort of TD-DFTB/FQ-based absorption spectra calculations, we tested the oscillator-strength-based truncation of the single-orbital transition space following the procedures introduced in a previous study.[23] A summary of the technical settings used in the TD-DFTB calculations can be found in Table S1 in the Supporting Information (SI). In all of the cases, absorption spectra profiles were obtained through a convolution of the TD-DFTB excitations by using Gaussian line shapes with a full width at half-maximum (fwhm) value of 0.3 eV, if not explicitly stated. A minimum number of 100 excited states were converged in each calculation. In all calculations, we exploited the FQ parameters proposed in ref 73.

Table 1. Inventory of the Number of Atoms, Water Molecules Described at the FQ Level, and Absorption Band Maxima (in Electronvolts) of All Systems under Study[a]

system	N_{atoms}	N_{FQwaters}	Abs_{SCC}	$\text{Abs}_{\text{DFTB3}}$
DOX in gas phase	69	0	2.41	2.43
DOX in water	3714	1215	2.33	2.35
DOX/water/DNA	9756	3100	2.37	2.36
UBI in gas phase[b]	1231	0	5.31, 4.75	5.37, 5.05, 4.77
UBI in water	11 731	3500	5.33, 4.80	5.43, 4.86

[a]Experimental results for the $\pi \rightarrow \pi^*$ transitions of DOX and UBI in aqueous solution are 2.58 and 4.51 eV, respectively.[68,69] [b]From the geometry reported in ref 23.

4. NUMERICAL RESULTS

In this section, we apply DFTB/FQ to describe absorption spectra. We analyze the effect that different choices of the DFTB Hamiltonian and the radius of the DFTB shell have on the spectra and test the accuracy of various intensity selection thresholds for the single orbital transition basis. Also, TD-DFTB/FQ spectra are compared with those obtained by using TD-DFTB calculations in the gas phase.

As test cases, we have chosen two biologically relevant, flexible organic molecules, namely, DOX and UBI, whose structures are depicted in the left panels of Figures 1 and 2, respectively.

4.1. Doxorubicin. DOX is an anticancer drug,[74] and it is commonly studied in the context of intercalation into DNA due to the proposed mechanism of action based on the insertion of its planar aromatic chromophore portion between sequential base pairs (BPs).[75−77] The drug in water is quite investigated as well given that it travels from the purely aqueous environment to penetrate DNA helices.[65,68,78−81] Regarding the DOX/DNA/water tertiary system, binding energy studies have shown that DOX affinity is sequence-dependent.[82] Although the preferential binding of DOX to double-stranded (ds) DNA is still a subject of debate, recent works reported that among some hexameric evaluated sequences, DOX prefers to bind to the d(CGATCG) in the case of the 1:1 complexes.[67] Therefore, we only discuss the intercalation complex of DOX with that DNA model.

Because of its importance as a chemotherapy medication, there are plenty of works dealing with the spectroscopic evidence of the insertion of a DOX molecule between pairs of

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

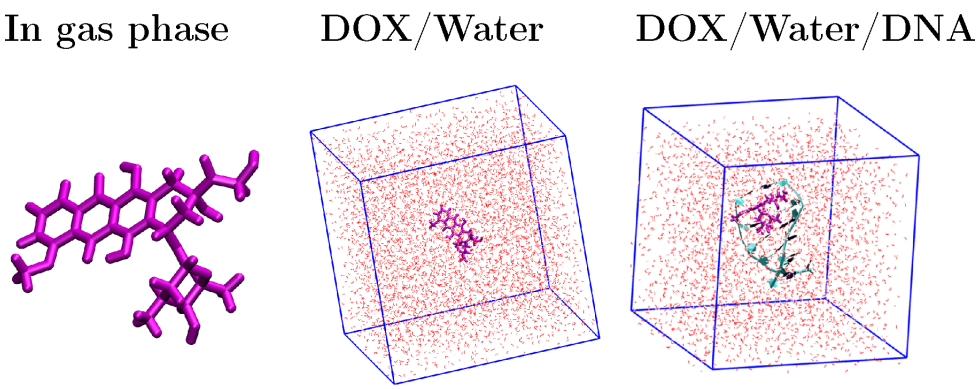


Figure 1. Three environments in which UV−vis spectra of doxorubicin were computed in this work. Left: Gas phase. Middle: Snapshot of the molecular dynamics of solvated DOX. Right: Snapshot of the molecular dynamics of DOX intercalated into DNA and surrounded by water molecules.

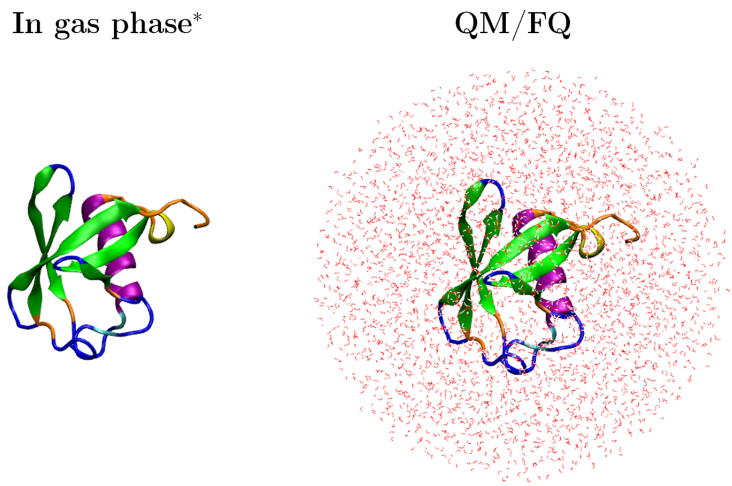


Figure 2. Environments in which the UV−vis spectra of ubiquitin were computed in this work. Left: Gas-phase conformation by using the same geometry reported in ref 23. Right: Snapshot of the molecular dynamics of solvated UBI is shown, which is treated with the QM/FQ approach.

nitrogen-containing nucleobases and for the spectral signatures of DOX in aqueous solution. Thus theoretical[65,80,81,83] and experimentally[84−88] obtained absorption spectra can be found in the literature for both environments. The main absorption band around 480 nm has been attributed to a $\pi \rightarrow \pi^*$ transition,[68,86] and some bathochromic and hypochromic effects are reported to occur upon intercalation.[86,89] Notwithstanding, it is difficult to observe those shifts because there is a vibronic component dominating the shape of the band.[65] The three environments in which we studied the absorption spectra of DOX are displayed in Figure 1.

From a set of snapshots (like those shown in Figure 1), the array of oscillator strengths obtained at their respective peak positions yields stick spectra (see Figure 3) with a natural broadening coming from the dynamical conformations of the chromophore and from the arrangements of the different molecules surrounding the system, that is, water molecules and the DNA basis. Computed stick spectra in the whole range of wavelengths are reported in Figure S1. It should be noted that the intensities of the sticks match the hypochromic effect reported to take place once the intercalation complex is formed.

Furthermore, considering that the quality of the results depends on whether there is a convergence of the desired property, some test computations on the UV−vis spectra of DOX in the more complex environment were also performed with an increasing number of snapshots extracted from the MD. Figure 4 shows the convergence of the energy and intensity of the first electronic transition with respect to the number of frames along with the associated 99% confidence intervals. The convergence behavior of the total spectra with respect to the number of frames is reported in Figure S2.

In addition, we evaluated the effect that solute−solvent nonelectrostatic interactions (neglected in the pure DFTB/FQ method) might have on the absorption spectra by adding the solute's closest water molecules to the QM portion and treating them with one of the DFTB Hamiltonians, whereas we described the remaining solvent molecules by means of FQ. This was done only for DOX in aqueous solution, and the results, together with the average number of water molecules (N_{QM}) for each radius threshold (R), are reported in Table 2. Clearly, the role of nonelectrostatic (mainly repulsion) effects is minimal.

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

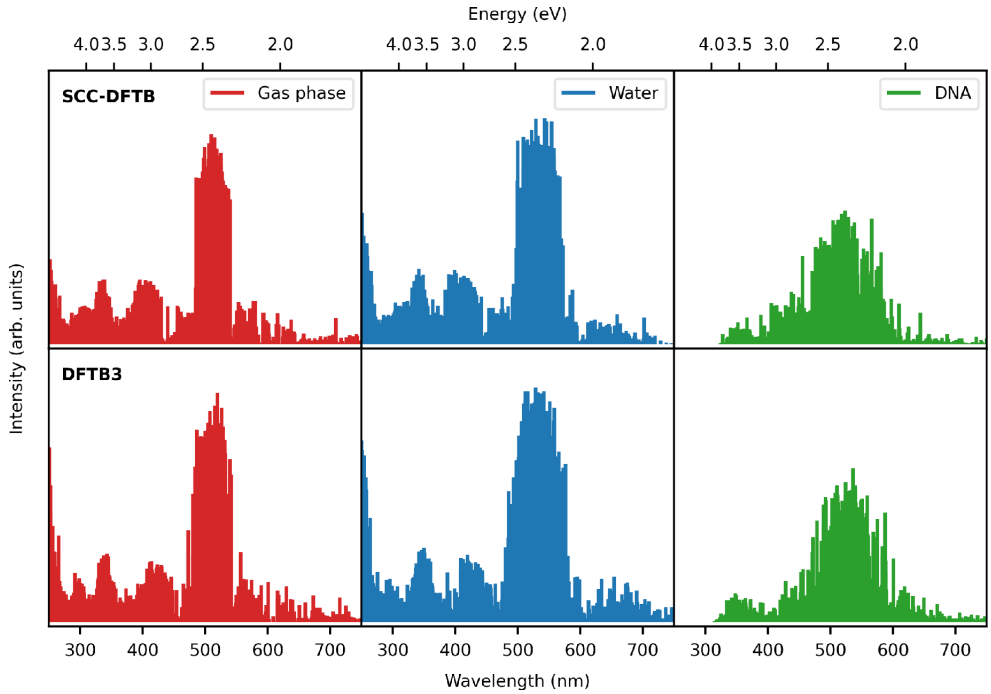


Figure 3. Stick spectra of doxorubicin *in vacuo* (red line), in water (blue line), and in DNA (green line) performed with different choices of the DFTB Hamiltonian.

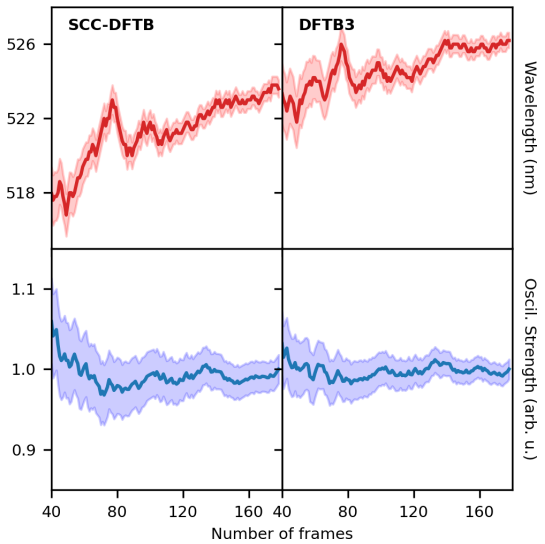


Figure 4. Convergence test for the absorption spectra of the tertiary DOX/water/DNA system. The position of the first excitation energy (top panel, red line) and the associated intensity (bottom panel, blue line) calculated with SCC-DFTB and DFTB3 model are reported. As a measure of the convergence, the 99% confidence intervals are reported.

4.1.1. DFTB Model Hamiltonians. As mentioned previously, we exploited two of the classic Slater−Koster-based DFTB Hamiltonians, SCC-DFTB and DFTB3. The resulting DFTB/FQ normalized absorption spectra obtained from the average of ∼180 structures of DOX in different environments are plotted in Figure 5. Table 1 also contains the maximum absorption energies, as obtained from the averaged spectra for both DFTB schemes and for all of the DOX environments under study. Interestingly, both DFTB Hamiltonians offer a similar description of the absorption spectra and the main band attributed to the $\pi \rightarrow \pi^*$ transition of the anthracycline chromophore. It should be emphasized that regardless of the DFTB model Hamiltonian and regardless of the environment, the HOMO and LUMO are, for the most part, the orbitals involved in the lowest energy transition. They are graphically depicted in Figure 6 along with other molecular orbitals belonging mainly to the rings of the DOX structure. These results are in line with those obtained by Olszówka et al.,[65] who reported a single HOMO → LUMO transition to be responsible for the appearance of the main band in the absorption spectra.

Table 2. Dependence of the Maximum Absorption Energies of Solvated DOX on the Size of the QM Shell[a]

R (Å)	N_{QM}	ΔVEE_{SCC} (eV)	ΔVEE_{DFTB3} (eV)
1	0	0.00	0.00
2	10	0.01	0.01
3	47	0.01	0.01
4	85	0.02	0.01
5	126	0.02	0.01
6	180	0.03	0.02

[a]ΔVEE is the energy difference with the maximum absorption calculated at QM/FQ level. N_{QM} is the average number of water molecules treated in the QM portion.

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

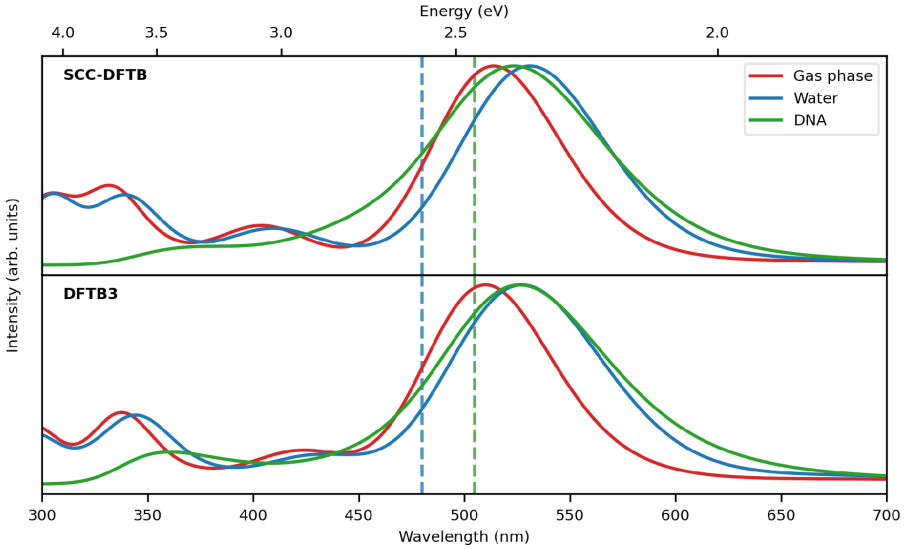


Figure 5. Absorption spectra of doxorubicin *in vacuo* (red line), in water (blue line), and in a water/DNA mix (green line) performed with different choices of the DFTB Hamiltonian. Experimental excitation energies from refs 68 and 86 in water and in water/DNA mix are reported with dashed lines.

As can be seen in Figure 5, the main environment effect is the red shift of the main band moving from the gas phase to water and water/DNA solutions. Overall, TD-DFTB/FQ reproduces the general shape of other published electronic absorption spectra,[81] although the main band is red-shifted by ∼0.2 eV when compared with the experimental maximum absorption energy of solvated DOX. Discrepancies between calculated and experimental results have already been reported for such systems and moderately corrected by using the vertical gradient (VG) or adiabatic Hessian (AH) approaches to vibronically resolve the spectra.[65] Also, as is reported in a recent paper,[90] it would be beneficial to consider different DOX tautomers to obtain a full description of the absorption spectra.

Going from water to water/DNA, there are just slight differences in the vertical excitation energies with both Hamiltonians; however, the spectral profile does exhibit some changes, including a broader main band when DOX is intercalated into DNA and also a different spectral shape at higher energies where the nucleotides are also involved in the electronic transitions. To understand the root causes of these differences in the spectra, especially in the main peak, we have plotted in Figure 7 the molecular orbitals that play a pivotal role in that particular excitation. By analysis and comparison of these orbitals with those displayed in Figure 6, it becomes clear that the aromatic rings of the DOX's nearest nucleotides are also participating in the transitions, although the majority of them include the anthraquinone rings (the portion that intercalates between two BPs of dsDNA) and the anchor domains of DOX, where the latter are responsible for stabilizing the DOX−DNA complex via hydrogen bonds with DNA bases. It is worth noting at this point that the HOMO of the ternary DOX/water/DNA system is not localized in the DOX molecule, unlike the situation in which DOX is in the aqueous solution environment.

4.1.2. Vertical Excitation Energy Dependence on the Size of the DFTB Shell. As shown in Figure 8, spectra obtained by varying the number of water molecules in the DFTB layer (QM/QM*w*/FQ) do not substantially differ each other, which is also confirmed by the data reported in Table 2, where ΔVEE, that is, the difference in vertical excitation energy (VEE), does not exceed 0.03 eV when compared with the QM/FQ result. It can therefore be argued that regardless of the Hamiltonian choice, the inclusion of the solvent does not play a pivotal role in the description of the bright $\pi \rightarrow \pi^*$ transition of solvated DOX; however, the spectral profiles look dissimilar at shorter wavelengths, with a more pronounced contrast in the SCC case.

4.1.3. Intensity Selection Thresholds. Figure 9 shows TD-DFTB calculated absorption spectra of DOX in the gas phase and in aqueous solution, obtained with intensity selection at different oscillator strength thresholds. It should be noticed that the reduced computational cost of the intensity-selected TD-DFTB leads to a loss in accuracy because there is a blue shift of the main band for larger thresholds. (See, for instance, $f > 0.1$ and $f > 0.01$.) Nevertheless, when a filter smaller than 0.001 is used, it is evident that the truncation of the basis in oscillator strength has a relatively small effect on the absorption spectrum. In fact, the relative intensities, number of peaks, and peak positions are kept, and the spectrum is practically unaltered compared with the nonfilter case, which is valid for both Hamiltonians. These findings indicate that a large part of the basis has a minor contribution to the spectra, as already reported for the simulation of the absorption spectra of C_{60} fullerene, $Ir(ppy)_3$, and UBI.[23]

4.2. Ubiquitin. UBI is a 76-amino acid polypeptide (1231 atoms) with diverse roles, mainly oriented to help in the regulation of the processes of other proteins in the body.[91−94] This small protein has been considered as a universal constituent of living cells.[95] Structurally speaking, UBI contains important chromophores like tyrosine and phenylalanine, with the former presenting higher absorbance. As a matter of fact, UBI has served as a model protein to study the sensitivity of UV−visible spectroscopy to environmental factors.[69,96]

DFTB/FQ is challenged in this section to compute the UV−vis absorption spectra of UBI in aqueous solution. The entire protein has been treated at the DFTB level, whereas water

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

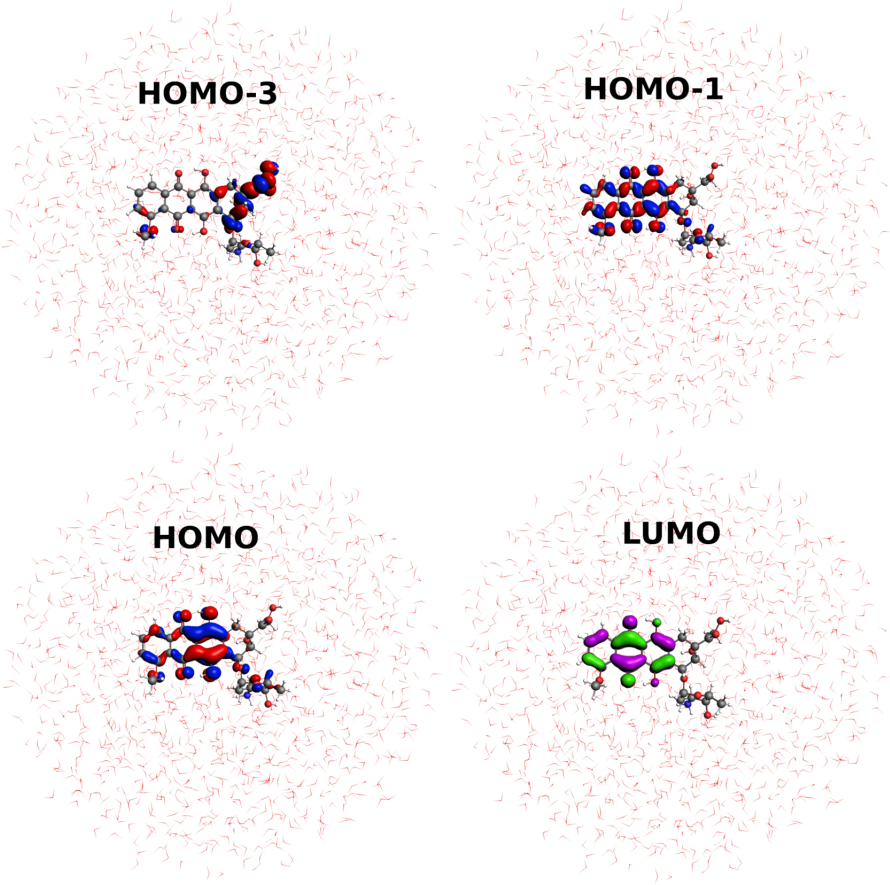


Figure 6. Most relevant MOs involved in the solvated doxorubicin absorption spectrum. For visualization purposes, virtual orbitals are depicted in different colors.

molecules are described by means of the FQ force field. Two major features are visible in the UBI experimental spectra in solution and in the gas phase, as reported by Bellina et al.:[69] (i) a broad band centered around 275−280 nm and (ii) an intense response at high energy with an onset at 250 nm. Indeed, it has been found that aromatic amino acids and proteins absorb UV light and show two main bands in UV−vis spectra, one centered on 280 nm that is the result of absorbance by the aromatic ring portion of their structure and a second one at lower wavelengths, which stems from the absorbance of peptide and carboxylic acid moieties. Because of this, it is not surprising that for UBI, a polypeptide containing tyrosine, the same bands are found. In particular, when Tyr is in aqueous solution, absorption maxima appear at ∼220 (higher absorbance) and 275 nm;[97−99] some authors have postulated that the two bands are probably arising from two well-separated $\pi \rightarrow \pi^*$ transitions.[100,101]

The influence of the environment on the absorption spectra of the UBI protein has already been demonstrated.[96] This effect can also be observed in Figure 10, which shows the spectra in the gas phase and in aqueous solution and a comparison between the results coming from the two model Hamiltonians. Such spectra have been obtained through a Gaussian convolution of the TD-DFTB excitations using an fwhm value of 0.2 eV. As a reference, the computed stick spectra of UBI in the gas phase and in aqueous solution over the entire spectral range are reported in Figure S3.

It can be observed that both SCC and DFTB3 yield the same spectral shape in the case of solvated UBI, whereas a different behavior is observed in the gas phase. Such discrepancies may be attributed to different parametrizations exploited in the two approaches.[102] Additionally, the presence of water has a tiny but non-negligible effect on the absorption spectra, overall red-shifting the main peaks that appear in gas phase spectra. The same calculations have also been performed by employing the nonpolarizable TIP3P force field to describe water molecules. (See Figure S4.) Absorption band maxima of UBI are also summarized in Table 1. To assign the transitions in the full protein, we applied a filter on the strongest oscillator strengths in the region of the maximum absorbance, and we quantified the contribution of single orbital transitions, thus identifying the leading contributing molecular orbitals and the part of the protein where they were situated. As has been indicated above, the band centered around 280 nm can be assigned to a $\pi \rightarrow \pi^*$ excitation, and even though orbitals from very distinct residues can participate in the transitions, the predominant ones are localized on the aromatic tyrosine and phenylalanine

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

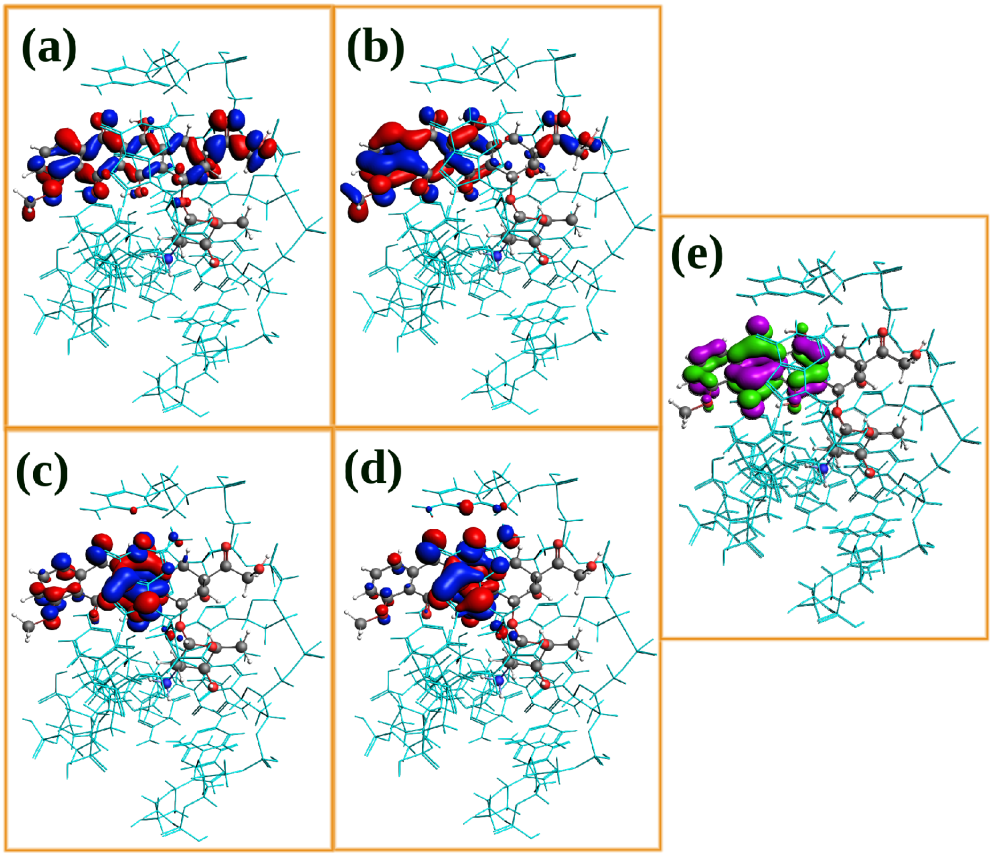


Figure 7. Most relevant MOs involved in the doxorubicin absorption spectrum when intercalated into DNA (represented by cyan sticks). (a−d) Occupied orbitals. (e) LUMO orbital. For visualization purposes, virtual orbitals are depicted in different colors.

chromophores, with tyrosine being responsible for most of the absorbance. A more in-depth analysis indicates that frontier molecular orbitals (HOMO, HOMO−1, LUMO, and LUMO +1) in the extended tyrosine residue resemble those involved in the absorption band at 280 nm of the UBI protein. The considered orbitals are presented in Figure 11 and are responsible for 90% of the state, having the greatest oscillator strengths in the region of the $\pi \rightarrow \pi^*$ excitation.

We move on to compare our results with experimental data on spectral shifts going from solution to the gas phase. It was previously determined that the $\pi \rightarrow \pi^*$ band in the gas phase is red-shifted as compared with the absorption in solution.[69] In addition, it is known that tyrosine is located at the surface of the protein[103] and thus has a strong effect on the environment, and the addition of water is also anticipated to cause a red shift of the tyrosine absorption spectrum.[104] In this context, it should be noted that although our results are in agreement with these observations, the experimental final shift is not fully reliable due to the fact that UV gas-phase spectra were measured for the isolated deprotonated protein and UBI can change its conformation going from the gas phase to solution.[69]

Finally, we remark that it is necessary to ensure the analysis to be performed on final converged spectra. Inspection of UBI spectra in aqueous solution, as obtained by averaging out a varying number of frames (Figure 12), reveals that increasing the number of frames has no impact on the final band shape; in fact, a sampling of 50 frames is sufficient to achieve convergence, with no missing features appearing in the spectra. Lastly, if single orbital transitions with an oscillator strength smaller than 0.001 are removed (results not shown here), then the spectra do not change, and as expected, the computational effort decreases.

5. SUMMARY, CONCLUSIONS, AND FUTURE PERSPECTIVES

We have presented a novel polarizable QM/MM approach where DFTB is coupled to the polarizable FQ force field. The model, which has been extended to TD-DFTB linear response, permits us to treat large systems in the condensed phase thanks to the favorable scaling of DFTB as compared with standard DFT and other *ab initio* methods. DFTB/FQ has been applied to the simulation of the electronic absorption spectra of DOX and UBI in different environments, showing that the inclusion of the FQ layer strongly affects spectral shapes and accounts for changes in both peaks' positions and relative intensities when going from the gas phase to condensed phase.

The low computational cost of DFTB has allowed for a detailed analysis of the role of nonelectrostatic effects in the case of DOX in aqueous solution. In particular, the size of the DFTB

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

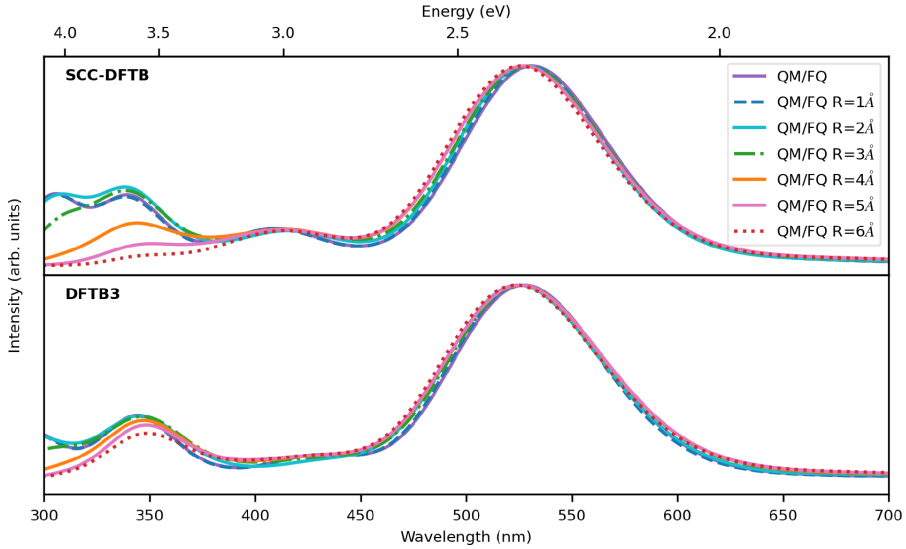


Figure 8. TD-DFTB/FQ absorption spectra of doxorubicin in water, varying the size of the QM portion and using two different DFTB Hamiltonians. The radius of the DFTB shell is reported in the key.

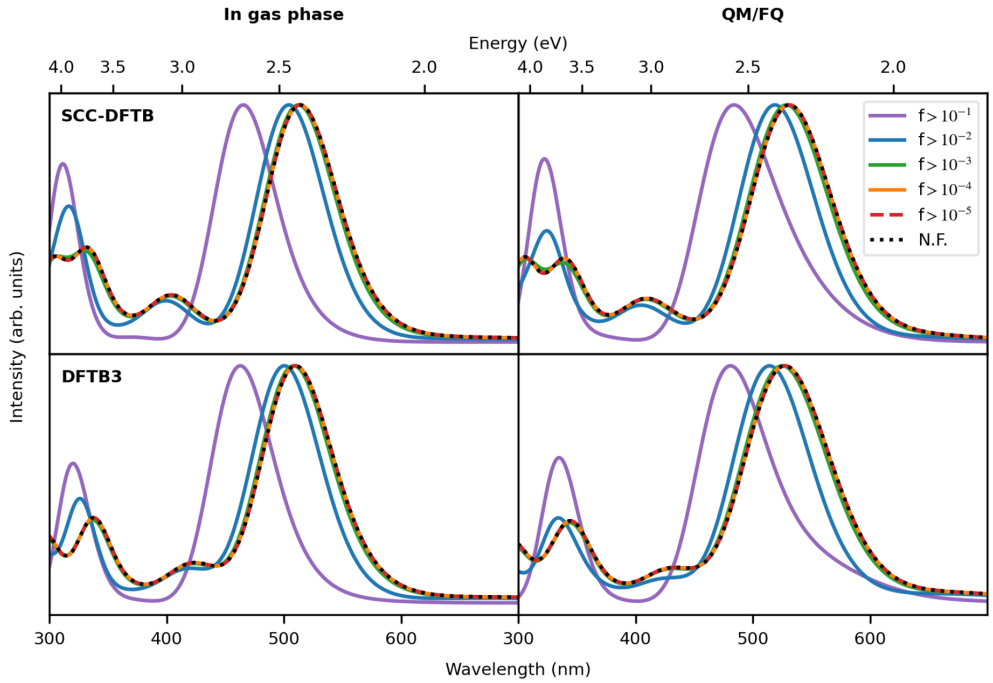


Figure 9. TD-DFTB and TD-DFTB/FQ absorption spectra of doxorubicin *in vacuo* (left panel) and in water (right panel) performed with different choices of the DFTB Hamiltonian and changing the intensity-selection thresholds for the single-orbital transitions basis. N.F. stands for no filter or that all single orbital transitions were considered.

portion has been varied by adding up a limited number of water molecules, showing that the DOX electronic response is almost unaffected by increasing the radius of the DFTB portion above 4 Å. This confirms the short-range nature of the nonelectrostatic interactions that are naturally included within the DFTB portion and that are dominated by Pauli repulsion. In addition, we show that by selecting the most intense spectral bands only, the accuracy of computed spectra is not particularly affected; however, the computational time of TD-DFTB/FQ calculations is substantially reduced. DOX and UBI spectral profiles have

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

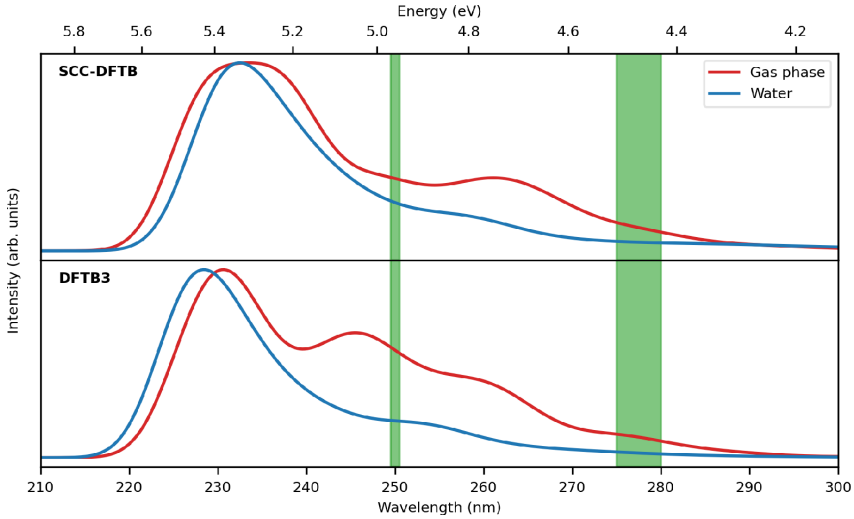


Figure 10. Comparison between calculated absorption spectra of UBI in the gas phase and aqueous solution, as obtained with different DFTB model Hamiltonians. The experimental data of UBI in the gas phase from ref 69 are reported with green blocks representing the experimental range of the two main absorption bands.

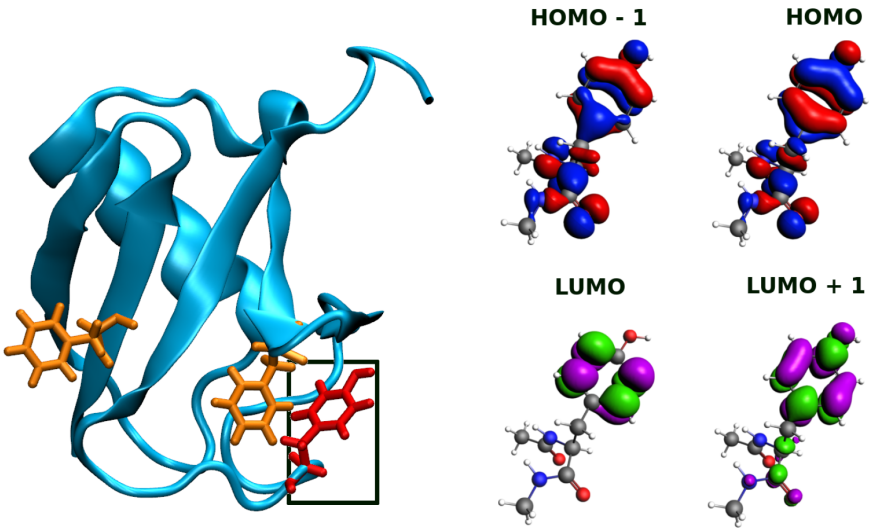


Figure 11. Left panel: Chromophores tyrosine (framed in black) and phenylalanine, which are the dominant sources of the UBI protein absorbance. Right panel: Molecular orbitals of extended Tyr mainly involved in the transitions. (See the text.) Extended Tyr means that in the calculations, the residue was capped with N-terminal acetyl and C-terminal N-Me amide capping groups to preserve the peptide bonds inside the protein. For visualization purposes, virtual orbitals are depicted in different colors.

been obtained by taking into account solute and solvent dynamics; the phase-space sampling and the consequent configurational variability in both solute and solvent moieties have brought up natural broadening in absorption bands, which is directly obtained from the signals arising on the different snapshots extracted from MD simulations.

Finally, the results reported in this work show that the combination of DFTB and FQ permits us to model absorption spectra of large molecules embedded in complex environments at a low computational cost and in nice agreement with experimental data. This opens up the opportunity to explore more challenging spectroscopies in different environments. In fact, DFTB/FQ (similarly to other QM/polarizable MM approaches) can simulate molecular properties in any kind of complex environment, pending appropriate parametrization.

To validate the accuracy of DFTB/FQ to describe solvatochromic effects, we have also computed vertical excitation energies at the TD-DFT/FQ level (see Table S2), in line with the preliminary calculations of some of the present authors.[80,81,105] DFTB/FQ and TD-DFT/FQ values are very similar, thus demonstrating that DFTB/FQ gives a correct description, from both a qualitative and quantitative point of

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

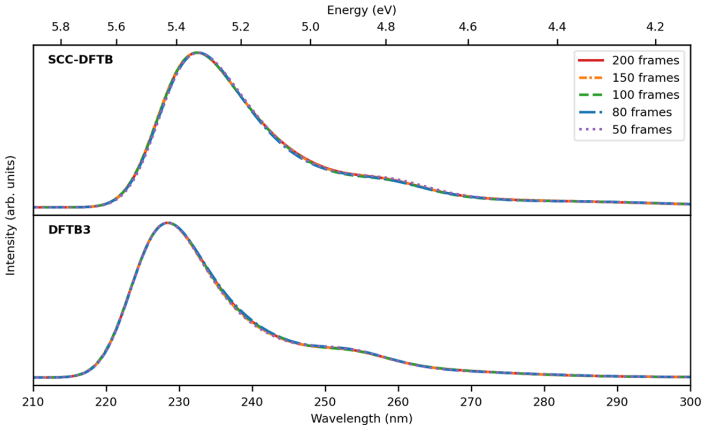


Figure 12. TD-DFTB/FQ absorption spectra of ubiquitin in aqueous solution, as obtained by averaging out an increasing number of snapshots.

view, for our systems. However, a more extended benchmark analysis on several systems would be required to finally validate the accuracy of the approach, as solvatochromic effects are strongly dependent on both the system and the nature of the excitation.

Improvement in the numerical performance of DFTB/FQ can be achieved by reparametrizing the FQ force field for DFTB calculations for both aqueous and nonaqueous solutions, in line with previous studies of some of the present authors.[46,106] Also, the description of the environment can be refined by adding polarizable dipole moments on MM atoms, such as in the recently developed FQFμ approach, which appropriately simulated anisotropic solvent effects.[35,37] In addition, in this Article, we have fully relied on a purely electrostatic model. Although electrostatics often dominates solvation effects, we have recently shown that nonelectrostatic interactions, in particular, Pauli repulsion, may strongly affect computed molecular properties.[107] Such terms can be included in the DFTB/FQ approach by following recent works of our group.[48,107]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c01066.

Technical settings for TD-DFTB/FQ calculations. Stick spectra of DOX *in vacuo*, in water, and in DNA/water computed at the TD-DFTB/FQ level. Convergence of the DOX/DNA/water absorption spectrum with the number of snapshots. Vertical excitation energies of DOX *in vacuo* and the DOX/water system computed at the TD-DFT/FQ and TD-DFTB/FQ levels. Stick spectra of UBI *in vacuo* and in aqueous solution. Absorption spectra of UBI/water computed at the TD-DFTB/TIP3P level (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Tommaso Giovannini − *Scuola Normale Superiore, Classe di Scienze, 56126 Pisa, Italy;* orcid.org/0000-0002-5637-2853; Email: tommaso.giovannini@sns.it

Chiara Cappelli − *Scuola Normale Superiore, Classe di Scienze, 56126 Pisa, Italy;* orcid.org/0000-0002-4872-4505; Email: chiara.cappelli@sns.it

Authors

Piero Lafiosca − *Scuola Normale Superiore, Classe di Scienze, 56126 Pisa, Italy*

Sara Gómez − *Scuola Normale Superiore, Classe di Scienze, 56126 Pisa, Italy;* orcid.org/0000-0002-5430-9228

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c01066

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Prof. Stefano Corni (Università di Padova) for providing the MD trajectory of ubiquitin in aqueous solution. This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 818064). We gratefully acknowledge the Center for High Performance Computing (CHPC) at SNS for providing the computational infrastructure.

■ REFERENCES

(1) Ratcliff, L. E.; Mohr, S.; Huhs, G.; Deutsch, T.; Masella, M.; Genovese, L. Challenges in large scale quantum mechanical calculations. *WIREs Comput. Mol. Sci.* **2017**, *7*, No. e1290.

(2) Dawson, W.; Degomme, A.; Stella, M.; Nakajima, T.; Ratcliff, L. E.; Genovese, L. Density functional theory calculations of large systems: Interplay between fragments, observables, and computational complexity. *WIREs Comput. Mol. Sci.* **2021**, e1574.

(3) Ochsenfeld, C.; Kussmann, J.; Lambrecht, D. S. Linear-scaling methods in quantum chemistry. *Rev. Comput. Chem.* **2007**, *23*, 1.

(4) Ma, Q.; Werner, H.-J. Explicitly correlated local coupled-cluster methods using pair natural orbitals. *WIREs Comput. Mol. Sci.* **2018**, *8*, No. e1371.

(5) Li, W.; Ni, Z.; Li, S. Cluster-in-molecule local correlation method for post-Hartree-Fock calculations of large systems. *Mol. Phys.* **2016**, *114*, 1447−1460.

(6) Folkestad, S. D.; Koch, H. Multilevel CC2 and CCSD Methods with Correlated Natural Transition Orbitals. *J. Chem. Theory Comput.* **2020**, *16*, 179.

(7) Marrazzini, G.; Giovannini, T.; Scavino, M.; Egidi, F.; Cappelli, C.; Koch, H. Multilevel density functional theory. *J. Chem. Theory Comput.* **2021**, *17*, 791−803.
(8) Bennie, S. J.; Curchod, B. F.; Manby, F. R.; Glowacki, D. R. Pushing the limits of EOM-CCSD with projector-based embedding for excitation energies. *J. Phys. Chem. Lett.* **2017**, *8*, 5559−5565.
(9) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F., III A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564−2568.
(10) Giovannini, T.; Koch, H. Energy-based molecular orbital localization in a specific spatial region. *J. Chem. Theory Comput.* **2021**, *17*, 139−150.
(11) Thiel, W. Semiempirical quantum-chemical methods. *WIREs Comput. Mol. Sci.* **2014**, *4*, 145−157.
(12) Thiel, W. *Theory and Applications of Computational Chemistry*; Elsevier, 2005; pp 559−580.
(13) Dral, P. O.; Wu, X.; Thiel, W. Semiempirical quantum-chemical methods with Orthogonalization and dispersion corrections. *J. Chem. Theory Comput.* **2019**, *15*, 1743−1760.
(14) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* **2021**, *11*, No. e1493.
(15) Seifert, G.; Joswig, J.-O. Density-functional tight binding—An approximate density-functional theory method. *WIREs Comput. Mol. Sci.* **2012**, *2*, 456−465.
(16) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B* **1998**, *58*, 7260.
(17) Gaus, M.; Cui, Q.; Elstner, M. Density functional tight binding: application to organic and biological molecules. *WIREs Comput. Mol. Sci.* **2014**, *4*, 49−61.
(18) Niehaus, T. A.; Suhai, S.; Della Sala, F.; Lugli, P.; Elstner, M.; Seifert, G.; Frauenheim, T. Tight-binding approach to time-dependent density-functional response theory. *Phys. Rev. B* **2001**, *63*, 085108.
(19) Elstner, M. The SCC-DFTB method and its application to biological systems. *Theor. Chem. Acc.* **2006**, *116*, 316−325.
(20) Jacquemin, D.; Perpète, E. A.; Scalmani, G.; Frisch, M. J.; Kobayashi, R.; Adamo, C. Assessment of the efficiency of long-range corrected functionals for some properties of large compounds. *J. Chem. Phys.* **2007**, *126*, 144105.
(21) Jacquemin, D.; Perpete, E. A.; Scuseria, G. E.; Ciofini, I.; Adamo, C. TD-DFT performance for the visible absorption spectra of organic dyes: conventional versus long-range hybrids. *J. Chem. Theory Comput.* **2008**, *4*, 123−135.
(22) Sokolov, M.; Bold, B. M.; Kranz, J. J.; Höfener, S.; Niehaus, T. A.; Elstner, M. Analytical Time-Dependent Long-Range Corrected Density Functional Tight Binding (TD-LC-DFTB) Gradients in DFTB+: Implementation and Benchmark for Excited-State Geometries and Transition Energies. *J. Chem. Theory Comput.* **2021**, *17*, 2266−2282.
(23) Rüger, R.; van Lenthe, E.; Lu, Y.; Frenzel, J.; Heine, T.; Visscher, L. Efficient Calculation of Electronic Absorption Spectra by Means of Intensity-Selected Time-Dependent Density Functional Tight Binding. *J. Chem. Theory Comput.* **2015**, *11*, 157−167.
(24) Paesani, F. Getting the right answers for the right reasons: Toward predictive molecular simulations of water with many-body potential energy functions. *Acc. Chem. Res.* **2016**, *49*, 1844−1851.
(25) Reichardt, C.; Welton, T. *Solvents and Solvent Effects in Organic Chemistry*; John Wiley & Sons, 2010.
(26) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999−3094.
(27) Mennucci, B. Polarizable continuum model. *WIREs Comput. Mol. Sci.* **2012**, *2*, 386−404.
(28) Cappelli, C. Integrated QM/polarizable MM/continuum approaches to model chiroptical properties of strongly interacting solute-solvent systems. *Int. J. Quantum Chem.* **2016**, *116*, 1532−1542.
(29) Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. Electronic energy transfer in condensed phase studied by a polarizable QM/MM model. *J. Chem. Theory Comput.* **2009**, *5*, 1838−1848.
(30) Olsen, J. M. H.; Kongsted, J. Molecular properties through polarizable embedding. *Adv. Quantum Chem.* **2011**, *61*, 107−143.
(31) Loco, D.; Polack, É.; Caprasecca, S.; Lagardère, L.; Lipparini, F.; Piquemal, J.-P.; Mennucci, B. A QM/MM approach using the AMOEBA polarizable embedding: from ground state energies to electronic excitations. *J. Chem. Theory Comput.* **2016**, *12*, 3654−3661.
(32) Loco, D.; Jurinovich, S.; Cupellini, L.; Menger, M. F.; Mennucci, B. The modeling of the absorption lineshape for embedded molecules through a polarizable QM/MM approach. *Photoch. Photobio. Sci.* **2018**, *17*, 552−560.
(33) 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.* **2012**, *8*, 4270−4278.
(34) Lipparini, F.; Cappelli, C.; Barone, V. Linear response theory and electronic transition energies for a fully polarizable QM/classical Hamiltonian. *J. Chem. Theory Comput.* **2012**, *8*, 4153−4165.
(35) Giovannini, T.; Puglisi, A.; Ambrosetti, M.; Cappelli, C. Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQFμ model. *J. Chem. Theory Comput.* **2019**, *15*, 2233−2245.
(36) Giovannini, T.; Grazioli, L.; Ambrosetti, M.; Cappelli, C. Calculation of ir spectra with a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles. *J. Chem. Theory Comput.* **2019**, *15*, 5495−5507.
(37) Giovannini, T.; Riso, R. R.; Ambrosetti, M.; Puglisi, A.; Cappelli, C. Electronic transitions for a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles: linear and corrected linear response regimes. *J. Chem. Phys.* **2019**, *151*, 174104.
(38) Egidi, F.; Giovannini, T.; Del Frate, G.; Lemler, P. M.; Vaccaro, P. H.; Cappelli, C. A combined experimental and theoretical study of optical rotatory dispersion for (R)-glycidyl methyl ether in aqueous solution. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3644−3655.
(39) Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. Excitation energies in solution: the fully polarizable QM/MM/PCM method. *J. Phys. Chem. B* **2011**, *115*, 3027−3037.
(40) Schwabe, T.; Olsen, J. M. H.; Sneskov, K.; Kongsted, J.; Christiansen, O. Solvation effects on electronic transitions: Exploring the performance of advanced solvent potentials in polarizable embedding calculations. *J. Chem. Theory Comput.* **2011**, *7*, 2209−2217.
(41) Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. Polarizable density embedding: A solution to the electron spill-out problem in multiscale modeling. *J. Phys. Chem. Lett.* **2017**, *8*, 5949−5958.
(42) Reinholdt, P.; Jørgensen, F. K.; Kongsted, J.; Olsen, J. M. H. Polarizable density embedding for large biomolecular systems. *J. Chem. Theory Comput.* **2020**, *16*, 5999−6006.
(43) Steinmann, C.; Reinholdt, P.; Nørby, M. S.; Kongsted, J.; Olsen, J. M. H. Response properties of embedded molecules through the polarizable embedding model. *Int. J. Quantum Chem.* **2019**, *119*, No. e25717.
(44) Boulanger, E.; Harvey, J. N. QM/MM methods for free energies and photochemistry. *Curr. Opin. Struct. Biol.* **2018**, *49*, 72−76.
(45) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664−5677.
(46) Ambrosetti, M.; Skoko, S.; Giovannini, T.; Cappelli, C. Quantum Mechanics/Fluctuating Charge Protocol to Compute Solvatochromic Shifts. *J. Chem. Theory Comput.* **2021**, *17*, 7146.
(47) Giovannini, T.; Egidi, F.; Cappelli, C. Theory and algorithms for chiroptical properties and spectroscopies of aqueous systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864−22879.
(48) Giovannini, T.; Lafiosca, P.; Cappelli, C. A general route to include Pauli repulsion and quantum dispersion effects in QM/MM approaches. *J. Chem. Theory Comput.* **2017**, *13*, 4854−4870.
(49) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133.
(50) Elstner, M. SCC-DFTB: what is the proper degree of self-consistency? *J. Phys. Chem. A* **2007**, *111*, 5614−5621.

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

(51) Foulkes, W. M. C.; Haydock, R. Tight-binding models and density-functional theory. *Phys. Rev. B* **1989**, *39*, 12520.
(52) DFTB. https://dftb.org/parameters/download (accessed on Feb 15, 2022).
(53) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. Extension of the self-consistent-charge density-functional tight-binding method: third-order expansion of the density functional theory total energy and introduction of a modified effective coulomb interaction. *J. Phys. Chem. A* **2007**, *111*, 10861−10873.
(54) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931−948.
(55) Xie, L.; Liu, H. The treatment of solvation by a generalized Born model and a self-consistent charge-density functional theory-based tight-binding method. *J. Comput. Chem.* **2002**, *23*, 1404−1415.
(56) Barone, V.; Carnimeo, I.; Scalmani, G. Computational Spectroscopy of Large Systems in Solution: The DFTB/PCM and TD-DFTB/PCM Approach. *J. Chem. Theory Comput.* **2013**, *9*, 2052−2071.
(57) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. *J. Phys. Chem. B* **2001**, *105*, 569−585.
(58) Nottoli, M.; Lipparini, F. General formulation of polarizable embedding models and of their coupling. *J. Chem. Phys.* **2020**, *153*, 224108.
(59) Gómez, S.; Giovannini, T.; Cappelli, C. Absorption spectra of xanthines in aqueous solution: A computational study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5929−5941.
(60) Puglisi, A.; Giovannini, T.; Antonov, L.; Cappelli, C. Interplay between conformational and solvent effects in UV-visible absorption spectra: Curcumin tautomers as a case study. *Phys. Chem. Chem. Phys.* **2019**, *21*, 15504−15514.
(61) Skoko, S.; Ambrosetti, M.; Giovannini, T.; Cappelli, C. Simulating Absorption Spectra of Flavonoids in Aqueous Solution: A Polarizable QM/MM Study. *Molecules* **2020**, *25*, 5853.
(62) Gómez, S.; Egidi, F.; Puglisi, A.; Giovannini, T.; Rossi, B.; Cappelli, C. Unlocking the power of resonance Raman spectroscopy: The case of amides in aqueous solution. *J. Mol. Liq.* **2022**, *346*, 117841.
(63) Gómez, S.; Rojas-Valencia, N.; Giovannini, T.; Restrepo, A.; Cappelli, C. Ring Vibrations to Sense Anionic Ibuprofen in Aqueous Solution as Revealed by Resonance Raman. *Molecules* **2022**, *27*, 442.
(64) Baerends, E.; et al. *DFTB*, version 2020.x; Theoretical Chemistry, Vrije Universiteit: Amsterdam, The Netherlands, 2020. http://www.scm.com (accessed on Feb 15, 2022).
(65) Olszówka, M.; Russo, R.; Mancini, G.; Cappelli, C. A computational approach to the resonance Raman spectrum of doxorubicin in aqueous solution. *Theor. Chem. Acc.* **2016**, *135*, 27.
(66) Brancolini, G.; Kokh, D. B.; Calzolai, L.; Wade, R. C.; Corni, S. Docking of Ubiquitin to Gold Nanoparticles. *ACS Nano* **2012**, *6*, 9863−9878.
(67) Jawad, B.; Poudel, L.; Podgornik, R.; Steinmetz, N. F.; Ching, W.-Y. Molecular mechanism and binding free energy of doxorubicin intercalation in DNA. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3877−3893.
(68) Hillig, K. W.; Morris, M. D. Pre-resonance Raman spectra of adriamycin. *Biochem. Biophys. Res. Commun.* **1976**, *71*, 1228−1233.
(69) Bellina, B.; Compagnon, I.; Joly, L.; Albrieux, F.; Allouche, A.-R.; Bertorelle, F.; Lemoine, J.; Antoine, R.; Dugourd, P. UV spectroscopy of entire proteins in the gas phase. *Int. J. Mass Spectrom.* **2010**, *297*, 36−40.
(70) Spiegelman, F.; Tarrat, N.; Cuny, J.; Dontot, L.; Posenitskiy, E.; Martí, C.; Simon, A.; Rapacioli, M. Density-functional tight-binding: basic concepts and applications to molecules and clusters. *Adv. Phys. X* **2020**, *5*, 1710252.
(71) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B* **1998**, *58*, 7260−7268.
(72) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* **2013**, *9*, 338−354.
(73) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* **1994**, *101*, 6141−6156.
(74) Lown, J. W. Discovery and development of anthracycline antitumour antibiotics. *Chem. Soc. Rev.* **1993**, *22*, 165−176.
(75) Chaires, J. B.; Satyanarayana, S.; Suh, D.; Fokt, I.; Przewloka, T.; Priebe, W. Parsing the Free Energy of Anthracycline Antibiotic Binding to DNA. *Biochemistry (Mosc.)* **1996**, *35*, 2047−2053.
(76) Gewirtz, D. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem. Pharmacol.* **1999**, *57*, 727−741.
(77) Barone, G.; Guerra, C. F.; Gambino, N.; Silvestri, A.; Lauria, A.; Almerico, A. M.; Bickelhaupt, F. M. Intercalation of Daunomycin into Stacked DNA Base Pairs. DFT Study of an Anticancer Drug. *J. Biomol. Struct. Dyn.* **2008**, *26*, 115−129.
(78) Poudel, L.; Wen, A. M.; French, R. H.; Parsegian, V. A.; Podgornik, R.; Steinmetz, N. F.; Ching, W.-Y. Electronic structure and partial charge distribution of doxorubicin in different molecular environments. *ChemPhysChem* **2015**, *16*, 1451−1460.
(79) Zhu, S.; Yan, L.; Ji, X.; Lu, W. Conformational diversity of anthracycline anticancer antibiotics: A density functional theory calculation. *J. Mol. Struct. THEOCHEM* **2010**, *951*, 60−68.
(80) Egidi, F.; Lo Gerfo, G.; Macchiagodena, M.; Cappelli, C. On the nature of charge-transfer excitations for molecules in aqueous solution: a polarizable QM/MM study. *Theor. Chem. Acc.* **2018**, *137*, 1−12.
(81) Giovannini, T.; Macchiagodena, M.; Ambrosetti, M.; Puglisi, A.; Lafiosca, P.; Lo Gerfo, G.; Egidi, F.; Cappelli, C. Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling. *Int. J. Quantum Chem.* **2019**, *119*, No. e25684.
(82) Jawad, B.; Poudel, L.; Podgornik, R.; Ching, W.-Y. Thermodynamic Dissection of the Intercalation Binding Process of Doxorubicin to dsDNA with Implications of Ionic and Solvent Effects. *J. Phys. Chem. B* **2020**, *124*, 7803−7818.
(83) Jia, M.; Song, X.; Zhang, Q.; Yang, D. A Theoretical Investigation About the Excited State Dynamical Mechanism for Doxorubicin Sensor. *J. Cluster Sci.* **2018**, *29*, 673−678.
(84) Manfait, M.; Bernard, L.; Theophanides, T. Resonance and pre-resonance Raman spectra of the antitumor drugs adriamycin and daunomycin. *J. Raman Spectrosc.* **1981**, *11*, 68−74.
(85) Manfait, M.; Alix, A. J.; Jeannesson, P.; Jardillier, J.-C.; Theophanides, T. Interaction of adriamycin with DNA as studied by resonance Raman spectroscopy. *Nucleic Acids Res.* **1982**, *10*, 3803−3816.
(86) Angeloni, L.; Smulevich, G.; Marzocchi, M. Absorption, fluorescence and resonance Raman spectra of adriamycin and its complex with DNA. *Spectrochim. Acta A-M* **1982**, *38*, 213−217.
(87) Smulevich, G.; Mantini, A. R.; Feis, A.; Marzocchi, M. P. Resonance Raman spectra and transform analysis of anthracyclines and their complexes with DNA. *J. Raman Spectrosc.* **2001**, *32*, 565−578.
(88) Lee, C.-J.; Kang, J.-S.; Kim, M.-S.; Lee, K.-P.; Lee, M.-S. The study of doxorubicin and its complex with DNA by SERS and UV-resonance Raman spectroscopy. *Bull. Korean Chem. Soc.* **2004**, *25*, 1211−1216.
(89) Yan, Q.; Priebe, W.; Chaires, J. B.; Czernuszewicz, R. S. Interaction of doxorubicin and its derivatives with DNA: Elucidation by resonance Raman and surface-enhanced resonance Raman spectroscopy. *Biospectroscopy* **1997**, *3*, 307−316.
(90) Florêncio e Silva, E.; Machado, E. S.; Vasconcelos, I. B.; Junior, S. A.; Dutra, J. D. L.; Freire, R. O.; da Costa, N. B. Are the Absorption Spectra of Doxorubicin Properly Described by Considering Different Tautomers? *J. Chem. Inf. Model.* **2020**, *60*, 513−521.
(91) Pickart, C. M.; Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *BBA - Mol. Cell. Res.* **2004**, *1695*, 55−72.
(92) Glickman, M. H.; Ciechanover, A. The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiol. Rev.* **2002**, *82*, 373−428.

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

(93) Schnell, J. D.; Hicke, L. Non-traditional Functions of Ubiquitin and Ubiquitin-binding Proteins. *J. Biol. Chem.* **2003**, *278*, 35857−35860.
(94) Mukhopadhyay, D.; Riezman, H. Proteasome-Independent Functions of Ubiquitin in Endocytosis and Signaling. *Science* **2007**, *315*, 201−205.
(95) Goldstein, G.; Scheid, M.; Hammerling, U.; Schlesinger, D. H.; Niall, H. D.; Boyse, E. A. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *P. Natl. Acad. Sci. USA* **1975**, *72*, 11−15.
(96) Antoine, R.; Dugourd, P. Visible and ultraviolet spectroscopy of gas phase protein ions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16494−16509.
(97) Creed, D. The photophysics and photochemistry of the near-UV absorbing amino acids-II. Tyrosyne and its simple derivatives. *Photochem. Photobiol.* **1984**, *39*, 563−575.
(98) Fornander, L. H.; Feng, B.; Beke-Somfai, T.; Nordén, B. UV Transition Moments of Tyrosine. *J. Phys. Chem. B* **2014**, *118*, 9247−9257.
(99) Del Galdo, S.; Mancini, G.; Daidone, I.; Zanetti Polzi, L.; Amadei, A.; Barone, V. Tyrosine absorption spectroscopy: Backbone protonation effects on the side chain electronic properties. *J. Comput. Chem.* **2018**, *39*, 1747−1756.
(100) Weber, G. Fluorescence-polarization spectrum and electronic-energy transfer in tyrosine, tryptophan and related compounds. *Biochem. J.* **1960**, *75*, 335.
(101) Longworth, J. *Excited States of Proteins and Nucleic Acids*; Plenum Press: New York, 1971; pp 319−484.
(102) Nishimoto, Y. Time-dependent density-functional tight-binding method with the third-order expansion of electron density. *J. Chem. Phys.* **2015**, *143*, 094108.
(103) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of ubiquitin refined at 1.8Å resolution. *J. Mol. Biol.* **1987**, *194*, 531−544.
(104) Wyer, J. A.; Ehlerding, A.; Zettergren, H.; Kirketerp, M.-B. S.; Brøndsted Nielsen, S. Tagging of Protonated Ala-Tyr and Tyr-Ala by Crown Ether Prevents Direct Hydrogen Loss and Proton Mobility after Photoexcitation: Importance for Gas-Phase Absorption Spectra, Dissociation Lifetimes, and Channels. *J. Phys. Chem. A* **2009**, *113*, 9277−9285.
(105) Cappelli, C. Unpublished.
(106) Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. Effective yet reliable computation of hyperfine coupling constants in solution by a QM/MM approach: Interplay between electrostatics and non-electrostatic effects. *J. Chem. Phys.* **2019**, *150*, 124102.
(107) Giovannini, T.; Ambrosetti, M.; Cappelli, C. Quantum confinement effects on solvatochromic shifts of molecular solutes. *J. Phys. Chem. Lett.* **2019**, *10*, 5823−5829.

https://doi.org/10.1021/acs.jctc.1c01066
J. Chem. Theory Comput. 2022, 18, 1765−1779

6.1.2. Modeling Resonance Raman spectroscopies of Doxorubicin in Complex Environments: Solution and DNA[9]

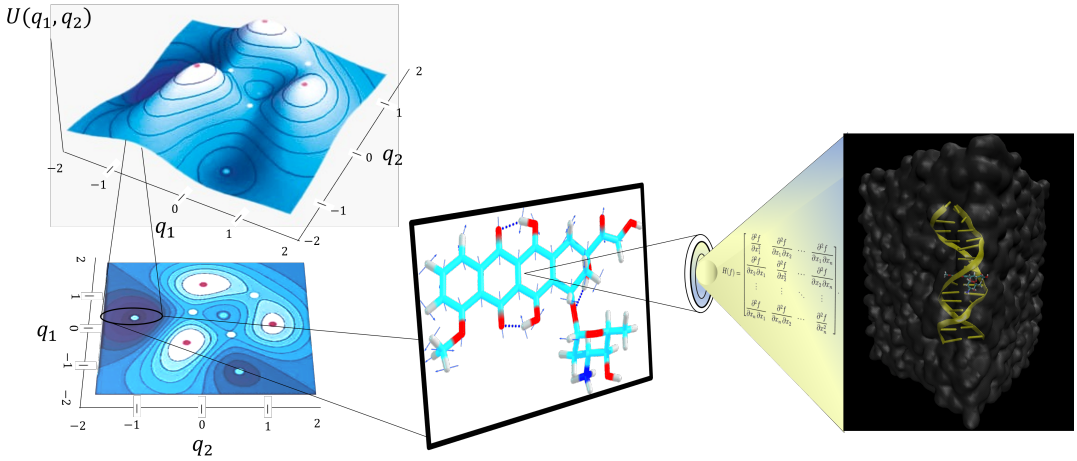


Figure 6.1. Graphical Abstract of paper 9.

UV-Resonance Raman Spectra of Systems in Complex Environments: A Multiscale Modeling applied to Doxorubicin intercalated into DNA

Sara Gómez,[*,†] Piero Lafiosca,[†] Franco Egidi,[‡] Tommaso Giovannini,[†] and Chiara Cappelli[*,†]

†*Scuola Normale Superiore, Classe di Scienze, Piazza dei Cavalieri 7, 56126, Pisa, Italy*

‡*Software for Chemistry & Materials BV, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands*

E-mail: sara.gomezmaya@sns.it; chiara.cappelli@sns.it

Abstract

UV-Resonance Raman (RR) spectroscopy is a valuable tool to study the binding of drugs to biomolecular receptors. The extraction of information at the molecular level from experimental RR spectra is made much easier and more complete thanks to the use of computational approaches, specifically tuned to deal with the complexity of the supramolecular system. In this paper we propose a protocol to simulate RR spectra of complex systems at different levels of sophistication, by exploiting a Quantum Mechanics/Molecular Mechanics (QM/MM) approach. The approach is challenged to investigate RR spectra of a widely used chemotherapy drug, Doxorubicin (DOX) intercalated into a DNA double strand. The computed results show a good agreement with experimental data, thus confirming the reliability of the computational protocol.

1 Introduction

UV- Resonance Raman (RR) spectroscopy is among the most powerful techniques used to investigate biological systems.[1] RR spectroscopy exploits the fact that during Raman measurements the incident frequency is tuned into an electronic absorption band, enhancing selected vibrational modes.[2] RR offers a unique selectivity as well as a high sensitivity to experimentally detect even traces of compounds, and thus it finds analytical applications in agriculture, life sciences, explosive detection, art, archaeology, and forensics, with additional current research in carbon nanotubes.[3] The key ingredient in the simulation of RR spectra of isolated systems is the transition polarizability tensor,[4–7] which can be obtained by exploiting a variety of approaches.[8–22]

When the system under investigation is in solution, the complexity of modeling increases and Quantum Mechanics/Molecular Mechanics (QM/MM) methods have been proven to be particularly successful, thanks to robust computational protocols developed in recent years.[23–26] Furthermore, in such systems, the partitioning between the QM and MM portions is generally straightforward because no covalent boundary exists between the solute and the solvent. Evidently, if the complexity of the system further increases (e.g. in the case of heterogeneous systems), the existing protocols for isolated or even solvated molecules must be adapted so that the new features can be properly considered in a physically-consistent way.

In this work, we extend a computational protocol, recently developed by us, which has been established as the state-of-the-art for the simulation of RR[26–28] and other diverse spectral signals in aqueous solutions, to target molecules embedded in biological matrices. The protocol involves a series of steps that start with a configurational sampling and from this, a number of structures are retrieved and used in subsequent quantum-classical calculations.[23,25] For systems in solution, basically, all kinds of spectroscopies have been covered in recently reported protocols[23,29,30] but when it comes to complex environments such as proteins, DNA,

and membranes, simulations usually focus on purely electronic spectroscopies.[31,32]

Given the combined electronic-vibrational nature of the RR signal, computing the property in a complex environment requires that the effects arising from electronic and vibrational parts be coherently inserted into the model. Accordingly, electronic transitions, normal modes, and polarizabilities are all ingredients that are influenced by the environment.

Indeed, the computational cost associated with the calculation of the vibrational responses of a complex system, such as a biological matrix, can be prohibitive. This is due to the fact that treating large systems implies including hundreds of vibrations in the calculation of the final spectra. Also, it is worth pointing out that, in order to obtain a reliable spectroscopic signal, the configurational phase-space of the target-environment system needs to be adequately sampled. This is usually done by resorting to a set of uncorrelated snapshots extracted from Molecular Dynamics (MD) simulations. However, when dealing with large, complex systems, this means that the vibrational analysis needs to be performed on each configuration, thus further increasing the computational complexity.

Here, we propose a series of strategies to compute the normal modes aiming at circumventing this problem. We apply the resulting protocol to Doxorubicin (DOX, trade name Adriamycin) intercalated into DNA. DOX is part of the anthracycline anticancer group and it has been claimed that this drug binds with DNA via an intercalation mechanism.[33,34] Figure 1 displays an intercalation complex between DOX and DNA. Drug–DNA complexes have been computationally studied and some works[35–47] are listed in Table S1 in the Supporting Information (SI).

Spectroscopy, and RR in particular, has been crucial for studying DOX in various environments, including DNA. Table S2 in the SI reports important contributions with a variety of techniques.[48–62] Indeed, many authors have pointed out spectroscopic consequences upon intercalation in the cases of absorption, fluorescence, Raman, and remarkably RR spectra.[51,56]

The manuscript is organized as follows: first, the methodology and a hierarchy of methods

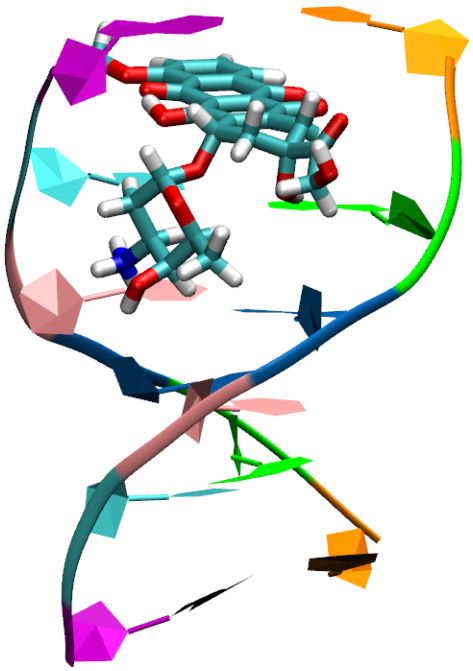


Figure 1: Graphical representation of doxorubicin bound to a d(CGATCG) sequence of DNA

to compute RR in complex environments are described along with computational details. Then, the different approaches are validated for the DOX/DNA systems or DOX dissolved in aqueous solution, by comparison with experimental results. Finally, conclusions and perspectives are drawn.

2 Methods

2.1 Computational protocol

In order to calculate the RR spectra of DOX dissolved in water and intercalated into DNA, we adapt an established approach proposed for aqueous solutions.[23,25] The protocol we propose is depicted step-by-step in Figure 2. A two-layer QM/MM system is defined for DOX (QM) in water (MM), whereas for the presence of DNA, we consider DOX as the target system

(QM) and DNA/Water as the environment (MM) unless otherwise stated. To sample the phase-space, we run a classical MD simulation for DOX in water using the same atom types and restrained electrostatic potential (RESP)-derived atomic charges as in Ref. 44. For the 1:1 DOX/DNA complex, we take the trajectories reported by Jawad et al.[44]: we select the sequence M3 d(CGATCG) because it is one of the strongest binding hexameric sequences of DOX after running Binding Free Energy (BFE) calculations by using MM-PBSA or MM-GBSA methods which gave BFE values of -12.74 and -9.10 kcal/mol, respectively.[44] From MD runs, 150 and 200 uncorrelated snapshots are taken for DOX/Water (one every 40 ps) and DOX/DNA/Water (one every 30 ps), respectively. Such snapshots contain free DOX or the intercalated DOX–DNA complex with surrounding water molecules within a cutoff distance of 18 and 8 Å, respectively (a selected snapshot is shown in the inset of Figure 2). To reduce the computational cost, clustering analysis of the MD runs is performed by following the methodology proposed in Ref. 63. 10 and 6 structural families are identified from the trajectories in water and in DNA respectively. Finally, on the representative structures, we computed RR spectra at the QM/MM level by resorting to both electrostatic (EE)[64,65] and polarizable embedding (PE) schemes,[23,25,66] the latter based on the fluctuating charges (FQ) force field.[67] In particular, while in EE fixed charges are assigned to MM atoms[68] (equal to those used in the MD runs), in FQ the charges may vary as a consequence of the interaction with the QM density (and vice-versa), thus introducing a mutual target-embedding polarization. Notice that in QM/FQ calculations, the QM region consists of the couple DOX/DNA, and the FQ region comprises all water molecules (described by using the parameters of Ref. 69). This combination will be referred to as QM/$\mathrm{QM_{DNA}}$/FQ in what follows.

RR spectra are calculated by resorting to a short-time dynamics approximation. RR intensities are directly computed from the geometrical derivatives of the frequency-dependent polarizability with respect to the normal coordinates.[18–21] The main advantage of that strategy is that all the electronic states are included in the polarizability, being also well-suited

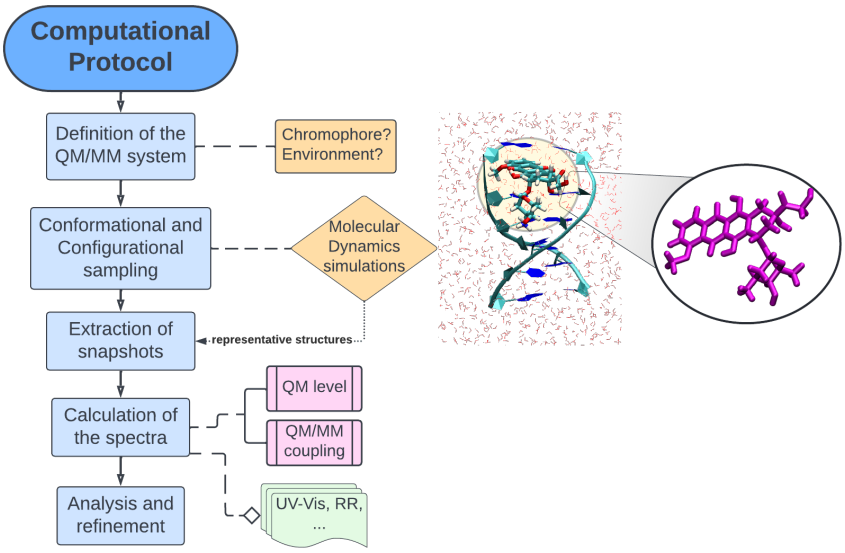


Figure 2: Flowchart of the protocol proposed in this work to model RR spectra of complex systems.

for dealing with large molecules or small molecules in complex environments.[18,22] We investigate different strategies to obtain DOX normal modes, by considering the variants listed in Table 1 and explained in the following paragraphs.

Table 1: Inventory of the different approaches employed in the calculation of normal modes. Normal modes serve to perform subsequent displacements and final calculations of the polarizability derivatives. See the text for a detailed explanation.

	Snapshots	QM/EE				QM/FQ		
		A0	PHVA	A1	A2	A0	PHVA	A1
DNAWater	150	✓						
	10	✓	✓	✓	✓	✓	✓	✓
	200	✓						
	6	✓	✓	✓	✓	✓	✓	✓

2.2 Strategies to obtain normal modes

To explore the effect of the quality of the normal modes on the final RR spectra, we systematically increase the level of sophistication in their acquisition. For this purpose, we try four

approaches.

2.2.1 The roughest approximation, A0

DOX normal modes are computed for each snapshot extracted from the MD simulation, without performing any optimization of the target system. In this way, DOX conformations are preserved in the complex, however, at the QM/MM level, DOX will not generally lie in its energy minimum. Although this approximation might be rather crude, it can give an initial insight into the vibrational spectral shape if the actual DOX geometry is not too different from the equilibrium one.

2.2.2 Partial Hessian Vibrational Approach, PHVA

Prior to the calculation of the normal modes, for each representative structure, DOX geometry is optimized by keeping water molecules or water and DNA base pairs (BPs) frozen in each reduced-size snapshot. This is consistent with the Partial Hessian Vibrational Approach (PHVA),[70–72] and DOX vibrational degrees of freedom are separated from those of the environment. Notably, this method preserves the environment in the configurations collected from MD simulations.

2.2.3 Transformation (Rotation) Matrix, A1

DOX normal modes are calculated on a reference structure optimized by using the Conductor Screening Model (COSMO)[73] to describe environmental effects. The dielectric constants for water and DNA given in Ref. 74 are used. The two geometries, i.e. the one of the reference DOX and the "distorted" DOX in each frame can be related to each other by means of a fitting (superimposing) procedure that uses rotations and translations transformations in order to minimize the root-mean-squared-deviation (RMSD) between the two lists of coordinates. Following that idea, for each frame, we construct a 3×3 transformation matrix providing the best alignment between DOX isolated optimized structure and DOX geometry

in the snapshot. To obtain the matrices, we use the *superpose3d* GitHub Repository[75] that implements the method outlined in Ref 76. Finally, the obtained transformation matrix is applied to the normal modes of the isolated optimized DOX to project them onto the extracted MD frame. The A1 strategy is tested on MD representative structures only.

2.2.4 Modification of the Adiabatic-Molecular Dynamics Generalized Vertical Hessian Ad-MD—gVH Approach, A2

It is an adaptation of the recently proposed mixed quantum-classical approach for the computation of electronic spectra of molecules characterized by a set of stiff (harmonic) modes and one or few internal large-amplitude (soft) motions.[77] For each MD snapshot, we take the reduced-dimensionality Hessian resulting from projecting out the soft coordinates from the ground state Hessian. This results in a new set of frequencies and normal modes, over which we later perform the corresponding displacements. The definition of flexible coordinates for DOX can be found in Section S3 in the SI. To obtain the reduced Hessian, we use the *FCclasses* code,[78] version 3.0.

All calculations are conducted by using a modified version of the Amsterdam Modeling Suite (AMS), release 2020.202.[79] In all cases, optimization (when it applies) and normal modes of DOX are calculated at the DFTB3 level, using the 3ob-3-1 parameter set.[80] Such normal modes are improved using the *Mode refinement* option[81] implemented in AMS. RR intensities are calculated *via* complex polarizability derivatives[18] at the B3LYP/DZP level with an incident frequency of 2.56 eV or 2.49 eV (477 nm or 497 nm) and a lifetime of 500 cm^{-1}. To compute the polarizability derivatives, the components of the polarizability tensor are obtained for two structures that have been displaced by 0.001 a.u. in two different directions along a vibrational mode. Details of the equations employed to compute cross-sections are described in Section S4 of the SI. Excitation wavelengths for RR are chosen based on DFTB/FQ UV-Vis spectra already reported by some of us.[46] Simulated spectra

are generated by convoluting RR peaks with a Lorentzian band shape with a half-width-at-the-half-maximum of 10 cm^{-1}. To obtain the final RR spectra, the resulting individual spectra are averaged. Persistence percentages (i.e the number of times each structure is sampled along the MD) are employed in the averaging process for those cases using only representative structures.

3 Results

3.1 MD results and choice of the representative structures

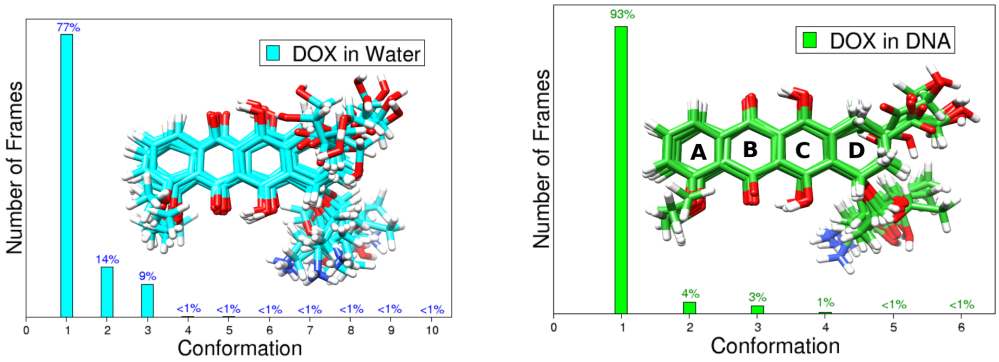


Figure 3: Percentage of frames exhibiting a similar conformation during MD simulations of DOX in water and of DOX intercalated into DNA (Model M3 shown in Figure 1). Differences in geometries for the most sampled conformers are also included as a superimposition.

DOX features three functional domains,[82] namely, the anthraquinone rings, the anchor, and the daunosamine region which contains an amino sugar group. During MD simulations, the stabilization of the drug via persistent intramolecular HBs between hydroxyl and carbonyl groups in the anthraquinone portion is common in both environments. However, the specificity and energetics of drug–DNA and drug-solvent interactions lead to different scenarios. For the complex formed between DOX and its target, the DNA conformational changes as well as the hydrogen bonding have already been analyzed in previous work.[44] Summarizing the principal remarks, five HBs between DOX and the chosen DNA sequence

are found, namely, two between the O12 (that linked to the ring D as an –OH group, see Figure 3 left) with H atoms bonded to N atoms of the Guanine 8 (G8) and three between the H–N^+ in the amino sugar and oxygen atoms of the Cytosine 5 (C5) and Thymine 4 (T4). Some of them are highly preserved throughout the simulation. As for DOX in aqueous solution, as a general rule, solute-solvent interactions are dictated by contacts involving O–H and C=O groups with hydrogen and oxygen atoms of the water molecules, in line with previous studies.[83,84]

DOX has many nuclear degrees of freedom and therefore a tremendous conformational diversity. Histograms of the RMSD calculated for each combination of DOX structures in the simulation (see Figure S2 in the SI) illustrate the flexibility of the DOX moiety along the entire trajectories. It can be seen that there is a wider distribution when DOX is free to move in solution. To group together a few conformers, we resort to a clustering methodology.[63] Figure 3 shows the predominant DOX conformers sampled during both MD simulations. The number of found clusters is 10 (solution) and 6 (DNA), supporting the fact that larger conformational freedom is present for solvated DOX. By looking at the *cluster size* we see that there is a predominance of two representative isomers, which are expected to contribute more to the calculated property. From the visual inspection of the conformations, it is clear that the anchor domain (glycolaldehyde) and the daunosamine region are the most flexible parts of the system. By observing the superimposition of the structures, the deformation of the last ring is readily appreciable. To apply the strategies explained in the methodology section, we use these sets of conformers. In the PHVA case, those geometries are energy minimized, which primarily changes the values of the dihedral angles involving the external OH groups.

3.2 RR spectra of DOX in Water and in DNA

Choice of incident frequency: The first necessary step when calculating RR is to set the incident wavelength to irradiate the sample. Early experimental UV-Vis of DOX in

water place the first electronic transition of the molecule at 480 nm (20800 cm^{-1})[48] (477 nm in Ref. 49) assigned to a $\pi \rightarrow \pi^*$ transition of the quinonoid compound.[49] According to Angeloni et al.[51] a bathochromic shift (≈10 nm) and a hypochromic effect are observed upon complexation with DNA. Computations by Egidi et al.[85] reveal that the firsts excited states have a strong charge transfer character involving HOMO → LUMO and HOMO-1 → LUMO transitions. Those molecular orbitals are all localized on the anthraquinone moiety.[86–88] Using CAM-B3LYP and QM/FQ, the vertical excitation energy for solvated DOX is reported to be 431 nm in Ref. 84. From simulations using DFTB/FQ there are some shifts in the calculated spectra with respect to the experimental ones.[46] By using an electrostatic embedding to treat the environment (see also Figure S3 in the SI) the absorption maximum of DOX intercalated in DNA lies at 497 nm and at 477 nm for DOX in water (B3LYP/DZP level of theory). To correctly reproduce the experimental conditions, we use shifted excitations as incident wavelengths in the simulation of the RR spectra.[16]

Analysis of the spectra: Figure 4 compares simulated QM/EE RR spectra for DOX in water and for DOX-DNA solution together with the experimental reports by Angeloni et al.[51] (in the 1000-2000 cm^{-1} region). Assignments of the major bands are given inside Figure 4a), and can also be found in Refs. 57,59. QM/EE RR spectra are computed by using the whole set of extracted snapshots (200 and 150 frames for aqueous and DNA solutions, respectively) and the representative structures provided by the clustering algorithm (10 and 6 frames for aqueous and DNA solutions, respectively). Although for DOX in water remarkable intensity changes are expected moving from pre-resonance to resonance conditions,[49] we point out that we are not interested in simulating enhancement factors, but in constructing a reliable computational tool to model RR spectra and examine the spectral changes when varying the environment.

RR of DOX is dominated by normal modes associated to the condensed aromatic rings, together with the typical signals of hydroxyanthraquinones.[48,49,59,60] The key point in those findings is that the vibrations of the aromatic ring modes (rings A, B, and C, see right

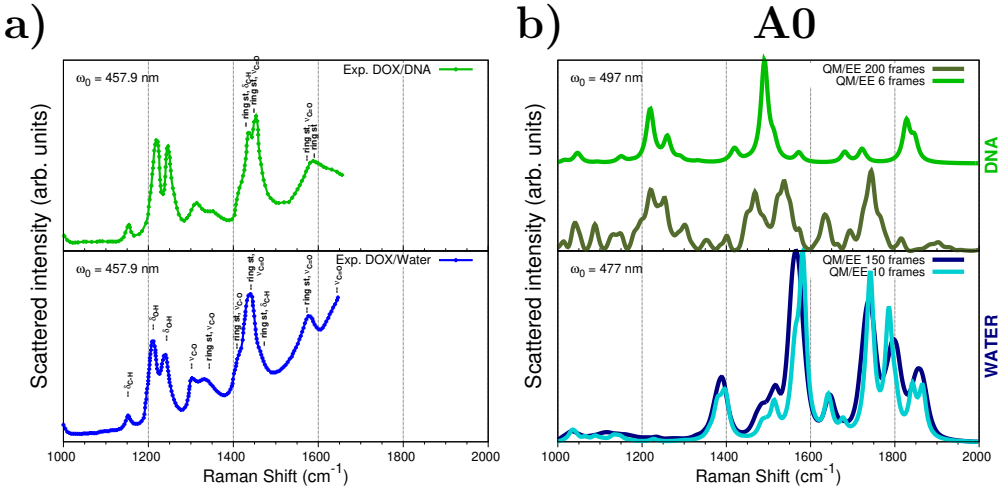


Figure 4: a) Experimental[51] Resonance Raman spectra of DOX in a DNA solution (top) and in water (bottom). b) Computed QM/MM EE Resonance Raman spectra after applying the approach A0 to displace geometries along the normal modes. RR intensities are calculated through complex polarizability derivatives using a damping factor of 500 cm^{-1}. A Lorentzian broadening with an FWHM = 20 cm^{-1} is used. QM level: B3LYP/DZP. All spectra are scaled such that the maximum intensity is unity.

panel of Figure 3) are coupled with the $\pi \rightarrow \pi^*$ electronic transition, therefore, it is not surprising that these modes are enhanced in RR spectra. If we compare the two experimental spectra in Figure 4a), we can see that there are some spectral perturbations suggesting a direct interaction between the drug and DNA. At this point of the discussion, it is useful to briefly summarize the effect of complexation on diverse spectroscopies as mentioned by some authors: (1) besides bathochromic and hypochromic effects on absorption spectra mentioned above and reported to occur upon the formation of the complex between DOX and DNA, there is also a reduction of fluorescence[51,56,89] (2) notable differences between RR spectra of the drug and of the complex are observed in the ranges near 450 cm^{-1} (not shown here) and 1400 cm^{-1}. In particular, Angeloni et al.[51] indicates that the bands at 430 and 1420 cm^{-1}, which are very weak for the pure drug, become prominent upon complexation. (3) Yan et al.[56] report that interactions between DOX and DNA primarily perturb the phenolic group and the π-system of the drug. Thus, from RR spectra it is apparent that once the

chromophoric rings are intercalated between adjacent DNA BPs, the intensity of the band at 444 cm^{-1} decreases and turns into a poorly resolved shoulder at 437 cm^{-1} (not shown here), accompanied by the splitting of the 1439 cm^{-1} band into two sub-bands at 1431 and 1449 cm^{-1} associated to the skeletal mode and the CCO stretching mode, respectively. Also, the bands at 1210 and 1241 cm^{-1} shift to 1213 and 1243 cm^{-1}, respectively. The latter band becomes sharper and moderately stronger in comparison with the former. Findings from Manfait et al.[50], Smulevich et al.[57] support such variations. (4) Similar prominent spectral changes in the same regions are claimed by Smulevich and Feis[53] on the basis of SERS experiments.

Spectra obtained with the EE and A0 schemes: Moving to QM/MM EE spectra presented in Figure 4b) and computed under the cheapest approach, A0, i.e. without optimizing the geometries extracted from MD trajectories, we see that for the intercalated DOX, position and shape of most Resonance Raman bands are in reasonably good agreement with the experimental observations. In contrast, by comparing computed and experimental RR spectra of aqueous DOX, it is clear that there are very diverse relative intensities and some important bands, as those associated with δ_{OH} (experimentally at around 1215 and 1245 cm^{-1}), are completely absent in simulated spectra. This is probably due to the occurrence of imaginary frequencies in both stiff and soft normal modes involving the OH group, which originate from the fact that the structures have not been reoptimized within the A0 approximation. Imaginary frequencies also appear in the complex with DNA, although in a smaller number (15 vs 3 on average as reported in Figure S4 of the SI): interestingly, this is related to the aforementioned rigidity of the structure when DOX is sandwiched within the DNA *via* stacking interactions. Furthermore, the out-of-equilibrium conditions yield a strong shift of the $\nu_{C=O}$ and ring stretching bands, which are wrongly predicted above 1800 cm^{-1}. This again indicates the inappropriateness and unsuitability of the A0 method. It can be finally noticed that RR spectra simulated by using the whole set of snapshots and the representative clustered ones are almost superimposable for DOX in water, while for DOX/DNA complex

some discrepancies, mainly related to relative intensities, are reported. Nevertheless, the main features of RR spectra are reproduced, and the discrepancies are reduced if DNA is included in the QM region in the QM/QM_{DNA}/FQ modeling (see also Figures S5 and S6 in the SI).

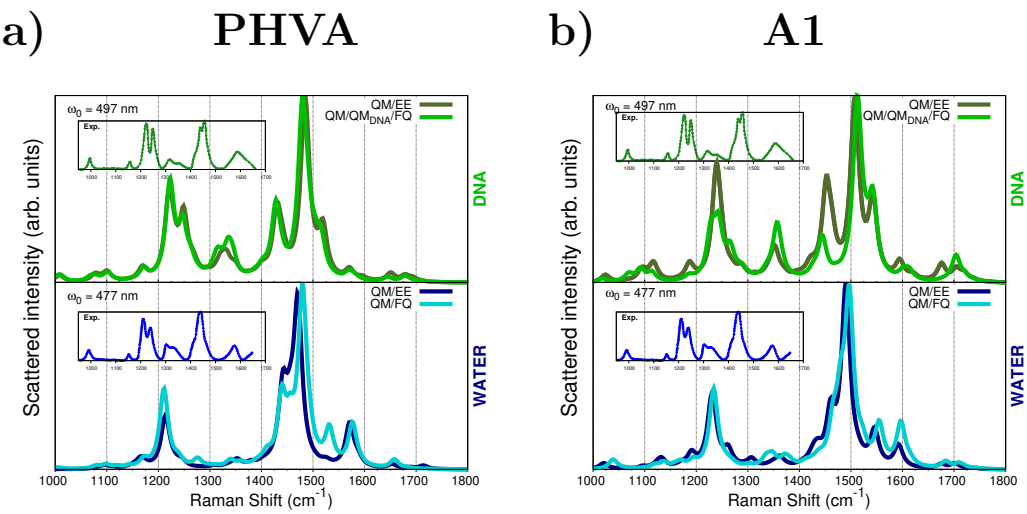


Figure 5: RR spectra of DOX in water and intercalated into DNA. RR intensities are calculated through complex polarizability derivatives using a damping factor of 500 cm^{-1}. Results from a) PHVA and b) A1 approaches are shown, see text. A Lorentzian broadening with an FWHM = 20 cm^{-1} is used. QM level: B3LYP/DZP. All spectra are scaled such that the maximum intensity is unity. Experimental RR spectra from Ref. 51 are given as inset.

Spectra obtained with the PHVA scheme: We now move to comment on the results obtained by applying the PHVA method to the representative structures, i.e., optimizing the solute in every MD snapshot while freezing the rest of the nuclear coordinates (see Figure 5a). We recall here that when the polarizable FQ scheme is used for the DNA solution, DNA is also included in the QM region, but only the DOX moiety is optimized (QM/QM_{DNA}/FQ in Figure 5). Remarkably, PHVA provides a better agreement with the experimental spectra than that found with the A0 approach. Overall, computed frequencies well reflect experimental values, while peaks' relative intensities exhibit appreciable discrepancies. For instance,

the δ_{OH} twofold band appears as a single signal accompanied by a shoulder in the computed spectra in water. Furthermore, the broadband measured at about 1600 cm^{-1} is barely visible in DNA. This is due to the fact that the ring stretching modes at 1575 and 1588 cm^{-1} mix to the modes belonging to the 1400-1500 cm^{-1} region, which are assigned to $\nu_{C=C} + \nu_{C-C}$ vibrations. As a matter of fact, a well-documented slight impact of the drug:DNA ratio[53] has been reported for such bands and for the $\nu_{C=O\cdots H}$ (hydrogen-bonded) vibration,[49,57,59] which is experimentally visible at around 1644 cm^{-1} (simulated at 1680 cm^{-1}).

Spectra obtained with the A1 scheme: In order to preserve the effects coming from the environment, RR spectra of DOX in water and DNA are computed by using the A1 scheme, which consists in taking the normal modes of one of the lowest energy DOX conformations and applying a rotation matrix (generated from an alignment procedure) to adjust them to the DOX actual geometry in each snapshot. Clearly, this procedure is much cheaper than PHVA, because a single geometry optimization is required. RR spectra computed for the representative structures are presented in Figure 5b. It can be observed that A1 spectra are especially similar to PHVA ones (see previous section), and are also in good agreement with the experiments. Only moderate differences in the position and shape of some bands are reported as compared to PHVA RR spectra (see also Figure 5a). As an example, the bands associated to the $\nu_{C-C} + \nu_{C-O}$ modes ($\sim$ 1330 cm^{-1}) are enhanced. However, few mismatches in relative intensities are still present, particularly for the most intense peaks, as the δ_{OH}, of which the intensity is significantly underestimated by A1 calculations.

As observed in Figure 5, for both approaches, namely PHVA and A1, RR spectra of the DOX-DNA complex are very similar to those calculated in solution. However, some substantial differences can be appreciated: the band at 1245 cm^{-1} is more enhanced for the DOX-DNA system, and a noticeable splitting of the band at 1441 cm^{-1} is observed. Such findings perfectly reproduce experimental features. Indeed, such bands are related to the two DOX C=O$\cdots$H–O groups, which are reported[53] to be involved in the interaction with DNA.

By comparing DOX RR spectra computed by using QM/EE and QM/FQ embedding schemes, it is possible to observe the marginal effects of mutual polarization. The same holds valid for the inclusion of the DNA in the QM region (QM/QM_{DNA}/FQ), highlighting the fact that non-electrostatic interactions between DOX and the DNA basis pairs also play a minor role as compare to electrostatics.

Spectra obtained with the A2 scheme: The last approach which is proposed is based on a modified version of the Ad-MD—gVH approach,[77] which has originally been designed as a general method for computing electronic spectra of flexible dyes in explicit environments. As explained above, in our adaptation, the DOX Hessian matrix has been constructed by removing all the soft coordinates, which involve the methyl and NH_3^+ torsions as well as the CCOH dihedral angle, the latter belonging to the DOX glycolaldehyde portion. We remark that this choice is indeed arbitrary and has been conducted by selecting all the normal modes that were involved in the computed imaginary frequencies. Figure 6 reports A2 RR spectra for DOX in water and DNA, which are computed at the QM/EE level only due to the observed similarities with more sophisticated approaches. In general, A2 frequencies well resemble the experiments, outperforming the other methods (see Table S3 in the SI). Also, all prominent spectral features are reproduced. This is due to the fact that the normal modes that have been eliminated in the construction of the Hessian matrix, are not associated with any enhanced peak in the selected region. Moreover, the DNA complexation basically alters 1400 – 1550 cm^{-1} spectral region, and it is also connected to a reduction in the intensity of the $\nu_{C=C}$ and $\nu_{C=O\cdots H}$ vibrations at 1587 and 1642 cm^{-1}. From a comparison between A2 RR spectra and those obtained by resorting to A0, PHVA, and A1 approximations (see also Figures S7, S8 in the SI), a small improvement is observed, regarding the enhancement of the peaks at 1308 and 1345 cm^{-1}, and the splitting of the peaks at 1215 and 1245 cm^{-1} in the aqueous solution. On the other hand, differently from the other approximations, such bands are mixed together when DOX interacts with DNA.

Continuum modeling: For the sake of completeness, we report RR spectra of DOX in water and DNA by describing environmental effects with an implicit, continuum description, as obtained by means of the COSMO model (see Figure S9 in the SI). As it has also been reported in previous studies,[74] a poor agreement with the experimental trends is obtained, as can be noticed in the largely overestimated intensity of the peaks located in the 1300–1400 cm^{-1} region. This further demonstrates the benefit of our atomistic modeling of the environment.

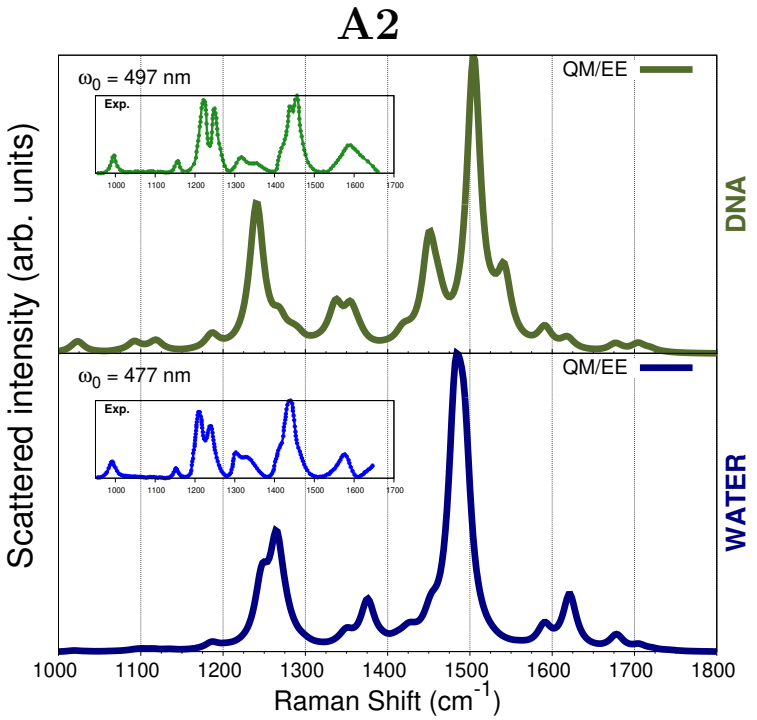


Figure 6: RR spectra of DOX in water and intercalated into DNA. RR intensities are calculated through complex polarizability derivatives using a damping factor of 500 cm^{-1}. Results from the A2 approach are shown, see text. A Lorentzian broadening with an FWHM = 20 cm^{-1} is used. QM level: B3LYP/DZP. All spectra are scaled such that the maximum intensity is unity. Experimental RR spectra from Ref. 51 are given as inset.

4 Conclusions and Final Remarks

In this work, we have proposed and analyzed different computational protocols for modeling RR spectra of Doxorubicin dissolved in aqueous solution and intercalated into DNA. The models are based on a multiscale QM/MM approach, which possibly accounts for mutual QM/MM polarization effects. RR spectra have been calculated by numerically differentiating the frequency-dependent complex polarizability with respect to normal vibrational coordinates. To correctly account for the configurational variability of the systems, several snapshots from MD trajectories have been extracted and possibly classified into different structural families based on the clustering analysis proposed in Ref. 63. RR spectra have subsequently been calculated for representative snapshots of each group. To obtain normal modes and proceed with geometry displacement, four approaches have been tested: (i) A0, i.e. the computationally cheapest, which involves the calculation of normal modes on raw, unoptimized, structures –by definition configurations out of equilibrium– and that, as expected, yields a poor reproduction of experimental values. (ii) Recalculating the normal modes from optimized DOX structures by applying the PHVA approach. This method improves results with respect to A0; (iii) avoiding the recalculation of normal modes for each snapshot by employing those of isolated DOX, and then using a linear transformation – different from one snapshot to another– to project the modes onto the actual DOX structures in each snapshot. This method yields computed values in good agreement with experimental observations while preserving the effects of the environment. Lastly, (iv) by borrowing the idea that flexible coordinates can be separated from stiff ones in each representative snapshot along the MD,[77] normal modes are recomputed from a reduced dimensionality Hessian. This last approach gives results in good agreement with experimental findings, but requires a prescreening procedure to analyze the occurrence of imaginary frequencies and to determine whether they can be reduced by moving such soft coordinates to what Cerezo et al.[77] called the "classical set". Indeed, the results reported in this paper show that reliable RR spectra

of doxorubicin in complex environments, treated atomistically, are obtained, in satisfactory agreement with experimental data. To conclude, it is amply documented that the specific DNA sequence plays an important role in the binding of a ligand to DNA strands. In that respect, the proposed methodology is general enough to treat different sequences and types of DNA binders, which opens up the possibility of computationally screening different drug candidates for any selected binding site. Finally, for large systems such as the one analyzed in this work, it would be interesting to compare in future studies our RR results with data computed with full vibronic approaches, such as those that have been proposed by some of us for simplest cases.[9,16]

Data Availability

DOX structure has been optimized by using the AMS code (version 2020.202 http://www.scm.com). MD simulations of DOX in aqueous solutions have been performed by using Gromacs version 2020.3 (https://www.gromacs.org). Structures of DOX intercalated into DNA have been provided by Professor Wai-Yim Ching and his research group at University of Missouri-Kansas City. QM/MM calculations have been performed by using a modified version of AMS release 2020.202 (http://www.scm.com).

Acknowledgement

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 818064). We thank Dr. Fabrizio Santoro for providing us with FCClasses 3.0, and Professor Wai-Yim Ching and his research group at the University of Missouri-Kansas City who graciously provided us with MD trajectories for DOX-DNA complexes. We also acknowledge the Center for High-Performance Computing (CHPC) at SNS for providing the computational infrastructure.

Supporting Information Available

Inventory of experimental and computational works on DOX and DOX-DNA complexes, definition of flexible coordinates, equations for RR intensities, RMSD distributions, additional spectra.

References

(1) Kelley, A. *Condensed-Phase Molecular Spectroscopy and Photophysics*; Wiley, 2022.

(2) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A. Elucidating peptide and protein structure and dynamics: UV resonance Raman spectroscopy. *J. Phys. Chem. Lett.* **2011**, *2*, 334–344.

(3) Chowdhury, J. In *Molecular and Laser Spectroscopy*; Gupta, V., Ed.; Elsevier, 2018; pp 147–164.

(4) Norman, P.; Ruud, K.; Saue, T. *Principles and Practices of Molecular Properties: Theory, Modeling, and Simulations*; Wiley, 2018.

(5) Ozaki, Y.; Wójcik, M.; Popp, J. *Molecular Spectroscopy, 2 Volume Set: A Quantum Chemistry Approach*; Wiley, 2019.

(6) Guthmuller, J. Comparison of simplified sum-over-state expressions to calculate resonance Raman intensities including Franck-Condon and Herzberg-Teller effects. *J. Chem. Phys.* **2016**, *144*, 064106.

(7) Guthmuller, J. *Molecular Spectroscopy: A Quantum Chemistry Approach*; John Wiley & Sons, Ltd, 2019; Chapter 17, pp 497–536.

(8) Xu, Q.; Aranda, D.; Yaghoubi Jouybari, M.; Liu, Y.; Wang, M.; Cerezo, J.; Improta, R.; Santoro, F. Nonadiabatic vibrational resonance Raman spectra from quantum dynamics

propagations with LVC models. Application to thymine. *J. Phys. Chem. A* **2022**, *126*, 7468–7479.

(9) Santoro, F.; Cappelli, C.; Barone, V. Effective Time-Independent Calculations of Vibrational Resonance Raman Spectra of Isolated and Solvated Molecules Including Duschinsky and Herzberg–Teller Effects. *J. Chem. Theory Comput.* **2011**, *7*, 1824–1839, PMID: 26596444.

(10) Heller, E. J.; Sundberg, R.; Tannor, D. Simple aspects of Raman scattering. *J. Phys. Chem.* **1982**, *86*, 1822–1833.

(11) Lee, S.-Y.; Heller, E. J. Time-dependent theory of Raman scattering. *J. Chem. Phys.* **1979**, *71*, 4777–4788.

(12) Baiardi, A.; Bloino, J.; Barone, V. A general time-dependent route to Resonance-Raman spectroscopy including Franck-Condon, Herzberg-Teller and Duschinsky effects. *J. Chem. Phys.* **2014**, *141*, 114108.

(13) Petrenko, T.; Neese, F. Analysis and prediction of absorption band shapes, fluorescence band shapes, resonance Raman intensities, and excitation profiles using the time-dependent theory of electronic spectroscopy. *J. Chem. Phys.* **2007**, *127*, 164319.

(14) Petrenko, T.; Neese, F. Efficient and automatic calculation of optical band shapes and resonance Raman spectra for larger molecules within the independent mode displaced harmonic oscillator model. *J. Chem. Phys.* **2012**, *137*, 234107.

(15) Horvath, R.; Gordon, K. C. Understanding excited-state structure in metal polypyridyl complexes using resonance Raman excitation profiles, time-resolved resonance Raman spectroscopy and density functional theory. *Coord. Chem. Rev.* **2010**, *254*, 2505–2518.

(16) Egidi, F.; Bloino, J.; Cappelli, C.; Barone, V. A robust and effective time-independent route to the calculation of resonance raman spectra of large molecules in condensed

phases with the inclusion of Duschinsky, Herzberg–Teller, anharmonic, and environmental effects. *J. Chem. Theory Comput.* **2014**, *10*, 346–363.

(17) Avila Ferrer, F. J.; Barone, V.; Cappelli, C.; Santoro, F. Duschinsky, Herzberg–Teller, and multiple electronic resonance interferential effects in resonance Raman spectra and excitation profiles. The case of pyrene. *J. Chem. Theory Comput.* **2013**, *9*, 3597–3611.

(18) Jensen, L.; Zhao, L.; Autschbach, J.; Schatz, G. Theory and method for calculating resonance Raman scattering from resonance polarizability derivatives. *J. Chem. Phys.* **2005**, *123*, 174110.

(19) Mohammed, A.; Ågren, H.; Norman, P. Time-dependent density functional theory for resonant properties: resonance enhanced Raman scattering from the complex electric-dipole polarizability. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4539–4548.

(20) Mohammed, A.; Ågren, H.; Norman, P. Resonance enhanced Raman scattering from the complex electric-dipole polarizability: A theoretical study on N2. *Chem. Phys. Lett.* **2009**, *468*, 119–123.

(21) Jensen, L.; Schatz, G. C. Resonance Raman scattering of rhodamine 6G as calculated using time-dependent density functional theory. *J. Phys. Chem. A* **2006**, *110*, 5973–5977.

(22) Al-Saidi, W.; Asher, S. A.; Norman, P. Resonance Raman spectra of TNT and RDX using vibronic theory, excited-state gradient, and complex polarizability approximations. *J. Phys. Chem. A* **2012**, *116*, 7862–7872.

(23) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664–5677.

(24) Gómez, S.; Giovannini, T.; Cappelli, C. Multiple Facets of Modeling Electronic Ab-

sorption Spectra of Systems in Solution. *ACS Physical Chemistry Au* **2022**, *0*, –, DOI: 10.1021/acsphyschemau.2c00050.

(25) Giovannini, T.; Egidi, F.; Cappelli, C. Theory and algorithms for chiroptical properties and spectroscopies of aqueous systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864–22879.

(26) Gómez, S.; Bottari, C.; Egidi, F.; Giovannini, T.; Rossi, B.; Cappelli, C. Amide Spectral Fingerprints are Hydrogen Bonding-Mediated. *J. Phys. Chem. Lett.* **2022**, *13*, 6200–6207.

(27) Gómez, S.; Egidi, F.; Puglisi, A.; Giovannini, T.; Rossi, B.; Cappelli, C. Unlocking the power of resonance Raman spectroscopy: The case of amides in aqueous solution. *J. Mol. Liq.* **2021**, 117841.

(28) Gómez, S.; Rojas-Valencia, N.; Giovannini, T.; Restrepo, A.; Cappelli, C. Ring Vibrations to Sense Anionic Ibuprofen in Aqueous Solution as Revealed by Resonance Raman. *Molecules* **2022**, *27*, 442.

(29) Gómez, S.; Giovannini, T.; Cappelli, C. Absorption spectra of xanthines in aqueous solution: A computational study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5929–5941.

(30) Di Remigio, R.; Giovannini, T.; Ambrosetti, M.; Cappelli, C.; Frediani, L. Fully polarizable QM/fluctuating charge approach to two-photon absorption of aqueous solutions. *J. Chem. Theory Comput.* **2019**, *15*, 4056–4068.

(31) Reinholdt, P.; Kjellgren, E. R.; Steinmann, C.; Olsen, J. M. H. Cost-effective potential for accurate polarizable embedding calculations in protein environments. *J. Chem. Theory Comput.* **2019**, *16*, 1162–1174.

(32) Reinholdt, P.; Jørgensen, F. K.; Kongsted, J.; Olsen, J. M. H. Polarizable density

embedding for large biomolecular systems. *J. Chem. Theory Comput.* **2020**, *16*, 5999–6006.

(33) Frederick, C. A.; Williams, L. D.; Ughetto, G.; Van der Marel, G. A.; Van Boom, J. H.; Rich, A.; Wang, A. H. J. Structural comparison of anticancer drug-DNA complexes: adriamycin and daunomycin. *Biochemistry* **1990**, *29*, 2538–2549, PMID: 2334681.

(34) Lown, J. W. Discovery and development of anthracycline antitumour antibiotics. *Chem. Soc. Rev.* **1993**, *22*, 165–176.

(35) Nakata, Y.; Hopfinger, A. Predicted mode of intercalation of doxorubicin with dinucleotide dimers. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 583–588.

(36) Barone, G.; Guerra, C. F.; Gambino, N.; Silvestri, A.; Lauria, A.; Almerico, A. M.; Bickelhaupt, F. M. Intercalation of Daunomycin into Stacked DNA Base Pairs. DFT Study of an Anticancer Drug. *J. Biomol. Struct. Dyn.* **2008**, *26*, 115–129, PMID: 18533732.

(37) Zhu, S.; Yan, L.; Ji, X.; Lu, W. Conformational diversity of anthracycline anticancer antibiotics: A density functional theory calculation. *J. Mol. Struct.: THEOCHEM* **2010**, *951*, 60–68.

(38) Lei, H.; Wang, X.; Wu, C. Early stage intercalation of doxorubicin to DNA fragments observed in molecular dynamics binding simulations. *J. Mol. Graph.* **2012**, *38*, 279–289.

(39) Airoldi, M.; Barone, G.; Gennaro, G.; Giuliani, A. M.; Giustini, M. Interaction of doxorubicin with polynucleotides. A spectroscopic study. *Biochemistry* **2014**, *53*, 2197–2207.

(40) Poudel, L.; Wen, A. M.; French, R. H.; Parsegian, V. A.; Podgornik, R.; Steinmetz, N. F.; Ching, W.-Y. Electronic structure and partial charge distribution of doxorubicin in different molecular environments. *ChemPhysChem* **2015**, *16*, 1451–1460.

(41) Jain, M.; Barthwal, S. K.; Barthwal, R.; Govil, G. Restrained molecular dynamics studies on complex of adriamycin with DNA hexamer sequence d-CGATCG. *Arch. Biochem.* **2005**, *439*, 12–24.

(42) Barthwal, R.; Agrawal, P.; Tripathi, A.; Sharma, U.; Jagannathan, N.; Govil, G. Structural elucidation of 4'-epiadriamycin by nuclear magnetic resonance spectroscopy and comparison with adriamycin and daunomycin using quantum mechanical and restrained molecular dynamics approach. *Arch. Biochem.* **2008**, *474*, 48–64.

(43) Agrawal, P.; Barthwal, S. K.; Barthwal, R. Studies on self-aggregation of anthracycline drugs by restrained molecular dynamics approach using nuclear magnetic resonance spectroscopy supported by absorption, fluorescence, diffusion ordered spectroscopy and mass spectrometry. *Eur. J. Med. Chem.* **2009**, *44*, 1437–1451.

(44) Jawad, B.; Poudel, L.; Podgornik, R.; Steinmetz, N. F.; Ching, W.-Y. Molecular mechanism and binding free energy of doxorubicin intercalation in DNA. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3877–3893.

(45) Jawad, B.; Poudel, L.; Podgornik, R.; Ching, W.-Y. Thermodynamic Dissection of the Intercalation Binding Process of Doxorubicin to dsDNA with Implications of Ionic and Solvent Effects. *J. Phys. Chem. B* **2020**, *124*, 7803–7818, PMID: 32786213.

(46) Lafiosca, P.; Gómez, S.; Giovannini, T.; Cappelli, C. Absorption Properties of Large Complex Molecular Systems: The DFTB/Fluctuating Charge Approach. *J. Chem. Theory Comput.* **2022**, *18*, 1765–1779.

(47) Rodrigues, E. S. B.; Macêdo, I. Y. L. d.; Silva, G. N. d. M. e.; de Carvalho e Silva, A.; Gil, H. P. V.; Neves, B. J.; Gil, E. d. S. DNA-Based Electrodes and Computational Approaches on the Intercalation Study of Antitumoral Drugs. *Molecules* **2021**, *26*, 7623.

(48) Hillig, K. W.; Morris, M. D. Pre-resonance Raman spectra of adriamycin. *Biochem. Biophys. Res. Commun.* **1976**, *71*, 1228–1233.

(49) Manfait, M.; Bernard, L.; Theophanides, T. Resonance and pre-resonance Raman spectra of the antitumor drugs adriamycin and daunomycin. *J. Raman Spectrosc.* **1981**, *11*, 68–74.

(50) Manfait, M.; Alix, A. J.; Jeannesson, P.; Jardillier, J.-C.; Theophanides, T. Interaction of adriamycin with DNA as studied by resonance Raman spectroscopy. *Nucleic Acids Res.* **1982**, *10*, 3803–3816.

(51) Angeloni, L.; Smulevich, G.; Marzocchi, M. Absorption, fluorescence and resonance Raman spectra of adriamycin and its complex with DNA. *Spectrochimica Acta Part A: Molecular Spectroscopy* **1982**, *38*, 213–217.

(52) Manfait, M.; Theophanides, T. Fourier Transform Infrared Spectra of cells treated with the drug adriamycin. *Biochem. Biophys. Res. Commun.* **1983**, *116*, 321–326.

(53) Smulevich, G.; Feis, A. Surface-enhanced resonance Raman spectra of adriamycin, 11-deoxycarminomycin, their model chromophores, and their complexes with DNA. *J. Phys. Chem.* **1986**, *90*, 6388–6392.

(54) Nonaka, Y.; Tsuboi, M.; Nakamoto, K. Comparative study of aclacinomycin versus adriamycin by means of resonance Raman spectroscopy. *Journal of Raman spectroscopy* **1990**, *21*, 133–141.

(55) Beljebbar, A.; Sockalingum, G.; Angiboust, J.; Manfait, M. Comparative FT SERS, resonance Raman and SERRS studies of doxorubicin and its complex with DNA. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **1995**, *51*, 2083–2090.

(56) Yan, Q.; Priebe, W.; Chaires, J. B.; Czernuszewicz, R. S. Interaction of doxorubicin

and its derivatives with DNA: Elucidation by resonance Raman and surface-enhanced resonance Raman spectroscopy. *Biospectroscopy* **1997**, *3*, 307–316.

(57) Smulevich, G.; Mantini, A. R.; Feis, A.; Marzocchi, M. P. Resonance Raman spectra and transform analysis of anthracyclines and their complexes with DNA. *J. Raman Spectrosc.* **2001**, *32*, 565–578.

(58) Lee, C.-J.; Kang, J.-S.; Kim, M.-S.; Lee, K.-P.; Lee, M.-S. The study of doxorubicin and its complex with DNA by SERS and UV-resonance Raman spectroscopy. *Bull Korean Chem Soc.* **2004**, *25*, 1211–1216.

(59) Das, G.; Nicastri, A.; Coluccio, M. L.; Gentile, F.; Candeloro, P.; Cojoc, G.; Liberale, C.; De Angelis, F.; Di Fabrizio, E. FT-IR, Raman, RRS measurements and DFT calculation for doxorubicin. *Microsc. Res. Tech.* **2010**, *73*, 991–995.

(60) Szafraniec, E.; Majzner, K.; Farhane, Z.; Byrne, H. J.; Lukawska, M.; Oszczapowicz, I.; Chlopicki, S.; Baranska, M. Spectroscopic studies of anthracyclines: structural characterization and in vitro tracking. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2016**, *169*, 152–160.

(61) Zhang, X.; Poniewierski, A.; Sozański, K.; Zhou, Y.; Brzozowska-Elliott, A.; Holyst, R. Fluorescence correlation spectroscopy for multiple-site equilibrium binding: a case of doxorubicin–DNA interaction. *Phys. Chem. Chem. Phys.* **2019**, *21*, 1572–1577.

(62) Zhang, R.; Zhu, J.; Sun, D.; Li, J.; Yao, L.; Meng, S.; Li, Y.; Dang, Y.; Wang, K. The Mechanism of Dynamic Interaction between Doxorubicin and Calf Thymus DNA at the Single-Molecule Level Based on Confocal Raman Spectroscopy. *Micromachines* **2022**, *13*, 940.

(63) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem. Int. Ed.* **1999**, *38*, 236–240.

(64) Bakowies, D.; Thiel, W. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

(65) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery Jr, J. A.; Morokuma, K.; Frisch, M. J. Combining quantum mechanics methods with molecular mechanics methods in ONIOM. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.

(66) Bondanza, M.; Nottoli, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. Polarizable embedding QM/MM: the future gold standard for complex (bio) systems? *Phys. Chem. Chem. Phys.* **2020**, *22*, 14433–14448.

(67) Cappelli, C. Integrated QM/Polarizable MM/Continuum Approaches to Model Chiroptical Properties of Strongly Interacting Solute-Solvent Systems. *Int. J. Quantum Chem.* **2016**, *116*, 1532–1542.

(68) Dohn, A. O. Multiscale electrostatic embedding simulations for modeling structure and dynamics of molecules in solution: a tutorial review. *Int. J. Quantum Chem.* **2020**, *120*, e26343.

(69) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* **1994**, *101*, 6141–6156.

(70) Jin, S.; Head, J. D. Theoretical investigation of molecular water adsorption on the Al (111) surface. *Surf. Sci.* **1994**, *318*, 204–216.

(71) Calvin, M. D.; Head, J. D.; Jin, S. Theoretically modelling the water bilayer on the Al (111) surface using cluster calculations. *Surf. Sci.* **1996**, *345*, 161–172.

(72) Biancardi, A.; Cammi, R.; Cappelli, C.; Mennucci, B.; Tomasi, J. Modelling vibrational coupling in DNA oligomers: a computational strategy combining QM and continuum solvation models. *Theor. Chem. Acc.* **2012**, *131*, 1–10.

(73) Pye, C. C.; Ziegler, T. An implementation of the conductor-like screening model of solvation within the Amsterdam density functional package. *Theo. Chem. Acc.* **1999**, *101*, 1432–2234.

(74) Russo, R. Modeling Structural and Spectroscopic Properties of Doxorubicin in Complex Environments. M.Sc. thesis, Università di Pisa, Pisa, Italy, 2015.

(75) Jewett, A. superpose3d. `https://github.com/jewettaij/superpose3d`, 2021.

(76) Diamond, R. A note on the rotational superposition problem. *Acta Crystallographica Section A* **1988**, *44*, 211–216.

(77) Cerezo, J.; Aranda, D.; Avila Ferrer, F. J.; Prampolini, G.; Santoro, F. Adiabatic-molecular dynamics generalized vertical hessian approach: a mixed quantum classical method to compute electronic spectra of flexible molecules in the condensed phase. *J. Chem. Theory Comput.* **2019**, *16*, 1215–1231.

(78) Santoro, F.; Cerezo, J. FCclasses 3.0, a Code for Vibronic Calculations. 2022; `http://www.iccom.cnr.it/en/fcclasses/`, Last consulted November 2022.

(79) Baerends, E.; et. al, ADF *(version 2020.202)*, Software for Chemistry & Materials, SCM. 2020; `http://www.scm.com`, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands.

(80) Gaus, M.; Goez, A.; Elstner, M. Parametrization and benchmark of DFTB3 for organic molecules. *J. Chem. Theory Comput.* **2013**, *9*, 338–354.

(81) Teodoro, T.; Koenis, M.; Galembeck, S.; Nicu, V.; Buma, W.; Visscher, L. Frequency range selection method for vibrational spectra. *J. Phys. Chem. Lett.* **2018**, *9*, 6878–6882.

(82) Chaires, J. B.; Satyanarayana, S.; Suh, D.; Fokt, I.; Przewloka, T.; Priebe, W. Parsing

the Free Energy of Anthracycline Antibiotic Binding to DNA. *Biochemistry* **1996**, *35*, 2047–2053, PMID: 8652545.

(83) Olszówka, M.; Russo, R.; Mancini, G.; Cappelli, C. A computational approach to the resonance Raman spectrum of doxorubicin in aqueous solution. *Theor. Chem. Acc.* **2016**, *135*, 1–15.

(84) Giovannini, T.; Macchiagodena, M.; Ambrosetti, M.; Puglisi, A.; Lafiosca, P.; Lo Gerfo, G.; Egidi, F.; Cappelli, C. Simulating vertical excitation energies of solvated dyes: From continuum to polarizable discrete modeling. *Int. J. Quantum Chem.* **2019**, *119*, e25684.

(85) Egidi, F.; Lo Gerfo, G.; Macchiagodena, M.; Cappelli, C. On the nature of charge-transfer excitations for molecules in aqueous solution: a polarizable QM/MM study. *Theor. Chem. Acc.* **2018**, *137*, 82.

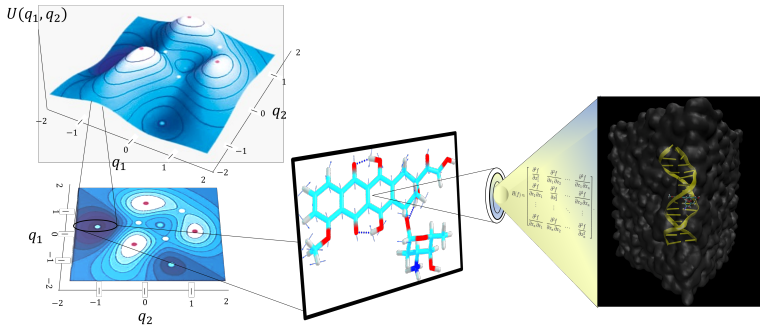
(86) Nicoli, L.; Giovannini, T.; Cappelli, C. Assessing the Quality of QM/MM Approaches to Describe Vacuo-to-water Solvatochromic Shifts. *J. Chem. Phys.* **2022**, DOI: 10.1063/5.0118664.

(87) Jia, M.; Song, X.; Zhang, Q.; Yang, D. A Theoretical Investigation About the Excited State Dynamical Mechanism for Doxorubicin Sensor. *J. Clust. Sci.* **2018**, *29*, 673–678.

(88) Florĕncio e Silva, E.; Machado, E. S.; Vasconcelos, I. B.; Junior, S. A.; L. Dutra, J. D.; Freire, R. O.; da Costa, N. B. Are the Absorption Spectra of Doxorubicin Properly Described by Considering Different Tautomers? *J. Chem. Inf. Model* **2019**, *60*, 513–521.

(89) Calendi, E.; Di Marco, A.; Reggiani, M. d.; Scarpinato, B.; Valentini, L. On physicochemical interactions between daunomycin and nucleic acids. *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1965**, *103*, 25–49.

TOC Graphic



Chapter 7

Summary, Conclusions and Future Perspectives

This thesis has presented the development and application of different multiscale protocols to describe molecular systems interacting with complex environments. The proposed protocols aim at obtaining a reliable description of experimental data and having predictive power. By means of illustrative applications, the crucial role of diverse aspects in the simulation of electronic and vibrational spectra, in particular, absorption, circular dichroism, Raman, and Resonance Raman, has been emphasized. The protocols integrate essential elements, such as phase-space sampling, solvation effects, and a physically sound calculation of the spectral signals, by using multiscale QM/MM methodologies. To obtain a representative set of configurations of the target and its environment, MD simulations with accurate FFs (desirably quantum-mechanically derived) have been proven to be an excellent tool. Regarding QM/MM calculations, although the Electrostatic Embedding does not take into account the polarization of the environment, is still useful in many applications, but the inclusion of mutual polarization effects as done in Polarizable Embedding schemes like QM/FQ is the most suitable way of dealing with complex systems.

For systems in aqueous solution, when QM/FQ in any of its flavors —QM/FQ or QM/FQFμ— has been challenged to reproduce absorption or RR spectra, results have found to be in a better agreement with experimental data than other approaches, which lack atomistic descriptions of the solvent molecules or discard polarization effects.

For systems in more complex systems, the protocol has been extended and the spectra have been modeled at different levels of sophistication, depending on how the MM portion is described and on how the normal modes are computed in the case of vibrational spectroscopies. Due to the size and complexity of the QM part, the problem is usually addressed by optimizing the target, but alternatively, simpler approaches may also be suitable for avoiding missing the dynamical aspects gained in the sampling step.

Since molecular properties and spectral signals arise as a consequence of the application of the protocol, it is worth mentioning that in all investigated cases, data for UV-Vis, ECD, Raman, and RR band shapes and intensities, theoretical and experimental, agree to a large extent. Discrepancies between the theory and experiment can be originated from one or more of the computational choices during the application of the computational protocol: selection of force fields in the MD runs, model chemistries in the computation of excited state energies,

transition densities, and polarizabilities, incident wavelength for the RR calculations, damping factors or excited state lifetimes, parametrizations of the electrostatic or polarizable schemes used in the QM/MM partitioning, etc.

Despite their achievements, the designed protocols can still stand to improve and the future prospects of the QM/FQ models are promising due to the existence of areas with room for improvement. Their wide parametrization for a variety of environments is the initial line of development, and it will enable them to tackle increasingly complicated systems, even biological environments. The parametrization effort includes quantum repulsion as well, and eventually quantum dispersion, which up until now have only been addressed for aqueous systems. The inclusion of charge transfer interactions (both between QM/MM or MM/MM moieties) also would help have a computational approach entirely coherent with the physics of the problem at hand. Indeed, more general methodologies to treat such effects are interesting topics for future investigations. Additionally, in order to obtain molecular structures that are adequate to be used in response property calculations, polarizable MD simulations would be advised to complement polarizable QM/MM calculations. Also, dealing with large systems necessitates the immediate application of enhanced MD sampling techniques and an appropriate examination of the system partitioning.

Finally, the demonstrated reliability of the protocol and its potential improvements make it suitable for promising applications to investigate various properties and spectroscopies of complex biosystems.

Appendix A

About Raman and Resonance Raman

This short section focuses on the general Raman scattering and on the vibrational RR spectroscopy, highlighting the higher sensitivity and usefulness of this latter in different fields. Most commonly used approximations to calculate RR Intensities and modeling a RR spectrum have been presented, with an emphasis on the calculation of the FC Overlap Integrals within and beyond the Independent Mode Displaced Harmonic Oscillator (IMDHO) model, by describing models as Adiabatic Hessian (AH), Vertical Hessian (VH), Vertical Gradient (VG) and Adiabatic Shift (AS). All these models differ in the treatment of the excited state and they become more or less computationally expensive depending on both the size of the system and the level of approximation. There are also other approaches in which the time-dependent formulation of RR intensities plays a pivotal role in the simulation of the spectra.

The discussion reported below on the basis of approaches to computing RR spectra is extracted from different sources found in the literature, in particular in Refs. 49,265,273,295.

Interactions of light and matter can be approached from different theoretical treatments including the purely classical one in which the radiation is described as an electromagnetic wave, and the material system as an assembly of independent classical rotors and vibrators, the purely quantized description in which both the radiation and the material system are treated from the quantum mechanical perspective, and a hybrid description where the classical treatment of the radiation is kept but the material system is represented by using the quantum mechanical theory. This latter treatment allows us to successfully examine the influence of the light on the molecular properties of the system.

There are some ways in which the light's path can be altered (Figure A.1), namely, absorption, reflection, refraction, scattering... Light-scattering phenomena have been widely studied once it was discovered by Raman and Krishnan in 1928 [296] and independently predicted by Smekal [297] in 1923. One of the principal and most used light-scattering phenomena are Raman scattering and Rayleigh scattering, which are also called inelastic (with change of frequency) and elastic (without change of frequency) scattering, respectively. Actually, more than 25 types of Raman spectroscopies are currently known, such as hyper-Rayleigh scattering, hyper-Raman scattering, coherent anti-Stokes Raman scattering, coherent Stokes Raman scattering, and stimulated Raman gain or loss spectroscopy, etc.

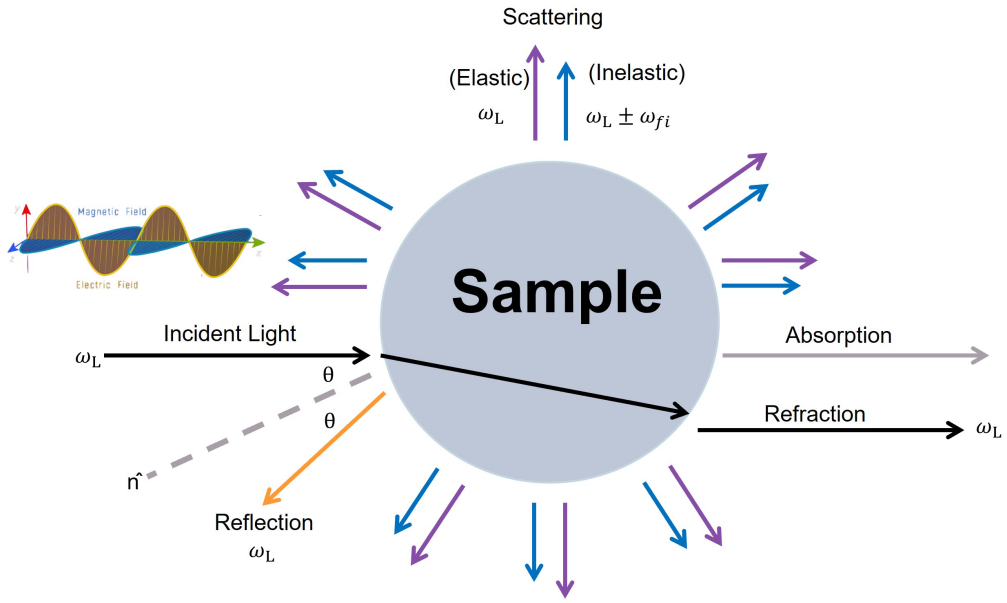


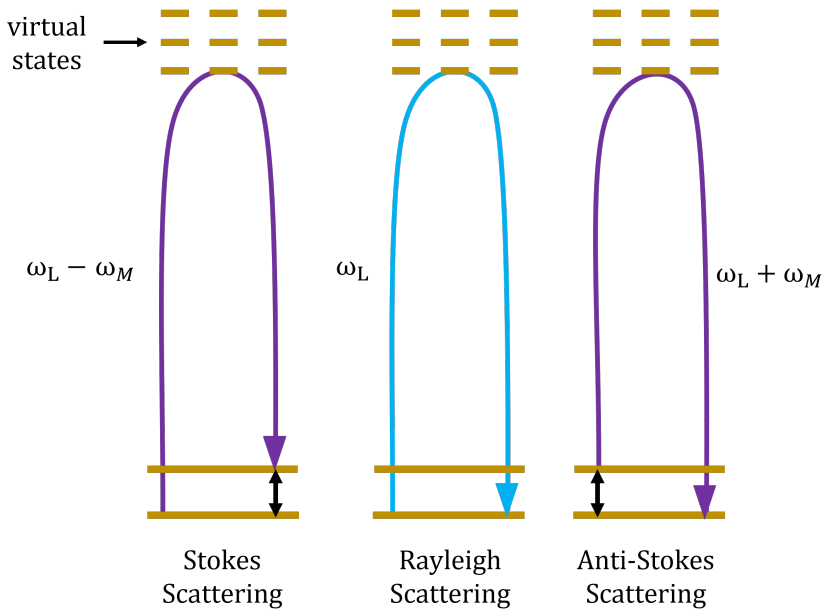
Figure A.1. Common interactions between light and matter. Taken and adapted from Ref. 298

A.1. Raman Basics

Raman spectroscopy studies the frequency change of light due to the interaction with matter. In the experimental set-up, monochromatic radiation of frequency ω_L is incident on systems such as dust-free, transparent gases and liquids, or optically perfect, transparent solids. In general, most of the incident light is transmitted without change but, in addition, some scattering of the radiation occurs, and some information about molecular properties can be obtained by analyzing the frequency content of the scattered radiation, ω_S.

From the Born-Oppenheimer Approximation (BOA), electronic and nuclear degrees of freedom can be separated. In turn, the nuclear motion can be approximately factored out in vibrational and rotational ones, whose coupling is minimized thanks to the Eckart conditions. In this manner, molecular states can be written as a product of the electronic, $|e\rangle$, vibrational, $|\nu\rangle$, and rotational, $|R\rangle$, wavefunctions, whereas the energy is obtained as the total sum of their corresponding energies. Given that most Raman scattering is observed under experimental conditions in which the rotational structure is not resolved, it is assumed a vibrational Raman scattering, which involves a transition from an initial vibrational state, $|\nu_i^g\rangle$ to a final vibrational state, $|\nu_f^g\rangle$ both belonging to the electronic ground state, $|e_g\rangle$, involving an intermediate state, $|\nu_m^e\rangle$, also known as *virtual state*, which belongs to the excited state, $|e_e\rangle$ [21]. Within this picture, the states of the system can be written as

$$\begin{aligned} |i\rangle &= |e_g\rangle\,|\nu_i^g\rangle \\ |m\rangle &= |e_e\rangle\,|\nu_m^e\rangle \\ |f\rangle &= |e_g\rangle\,|\nu_f^g\rangle \end{aligned} \tag{A.1}$$



Energy balance for the Stokes case

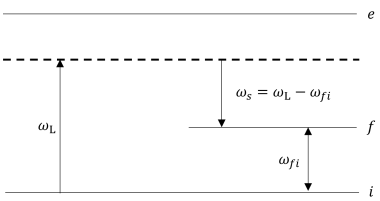
Initial state energy		Final state energy		Scattered frequency ω_S	Energy level diagram
Photons	Molecule	Photons	Molecule	determined by overall energy	
$n_1\hbar\omega_L$	E_i	$(n_1-1)\hbar\omega_L + \hbar\omega_S$	$E_f = E_i + \hbar\omega_{fi}$	$\omega_S = \omega_L - \omega_{fi}$	

Figure A.2. Representation of the Raman (Stokes and anti-Stokes) and Rayleigh scattering processes and energy balance for the Stokes case. The latter was taken and adapted from Ref. 21.

Note that the subscripts i, f and m indicate the associated vibrational quantum numbers, and in the case of inelastic scattering, i and f are different.

For Raman scattering, the energy of the final state can be higher or lower than the energy of the initial state. If $E_f > E_i$, and $\omega_S = \omega_L - \omega_M$, then Stokes Raman scattering is present, with $\omega_M = \omega_{fi} = \omega_f - \omega_i$ being the difference in the frequencies for the two vibrational states. Instead, if $E_f < E_i$, then $\omega_S = \omega_L + \omega_M$, which features anti-Stokes Raman scattering. In Figure A.2 both events are depicted along with the Rayleigh scattering wherein $|\nu_i^g\rangle = |\nu_f^g\rangle$. For the sake of clarity, a small table with the energy balance for the Raman Stokes process is also included in the bottom panel of Figure A.2. It is worth mentioning here that the energy $\hbar\omega_L$ does not match exactly any electronic transition energy and the photon is not absorbed in the strict spectroscopic sense, but rather the incident radiation is intended to perturb the molecule and opens the possibility of spectroscopic transitions distinct to direct absorption.

As a general rule, Raman shifts can be expressed as $\omega_L \pm \omega_M$ and since ω_M is related to the energy of a vibrational mode, it depends on molecular structure and environment, with atomic masses, bond lengths, molecular substituents and molecular geometry, all contributing. Stokes and anti-Stokes spectra only differ in intensity, which is directly proportional (10^{-6} times weaker) to the irradiance of the incident radiation and therefore such scattering can be described as a linear process.

A.1.1. Sum-Over-State (SOS) Formulation of the Vibrational Raman Intensities

Formal expressions for calculating Raman intensities have been provided by quantum mechanical treatments and by applying the time-dependent perturbation theory[263,299]. Overall, the radiation–molecule interaction is described by the electric dipole approximation and the first order response to the incident electric field is characterized by an induced transition electric dipole moment, with a component $(\mu_\rho^{(1)})_{i\to f}$ given as

$$(\mu_\rho^{(1)})_{i\to f} = (\alpha_{\rho\sigma})_{i\to f} E_\sigma \tag{A.2}$$

where $(\alpha_{\rho\sigma})_{i\to f}$ is the Raman polarizability tensor, $(\rho, \sigma = x, y, z)$, and E_σ is a component of the electric field amplitude. If $(\alpha_{\rho\sigma})_{i\to f}$ is known, it is possible to estimate the Raman scattering intensities. Thus, following the sum-over-state (SOS) expression originally derived by Kramers, Heisenberg, and Dirac[300], the Cartesian components of the transition polarizability tensor can be expressed as

$$(\alpha_{\rho\sigma})_{i\to f} = \frac{1}{\hbar} \sum_m \left\{ \frac{\langle f|\, \mu_\rho\, |m\rangle \langle m|\, \mu_\sigma\, |i\rangle}{\omega_{mi} - \omega_L - i\Gamma_m} + \frac{\langle f|\, \mu_\sigma\, |m\rangle \langle m|\, \mu_\rho\, |i\rangle}{\omega_{mf} + \omega_L + i\Gamma_m} \right\} \tag{A.3}$$

where μ_ρ is a component of the dipole moment operator, ω_{mi} is the energy separation between the initial and intermediate levels, Γ_m is the lifetime of the excited state m and ω_L is the frequency of the incident light. By assuming, besides the BOA (already applied in Eq. A.1), that the Eckart conditions are met, such as the vibrational part can be separated from the rotational and translational ones, some integrals for the Raman polarizability tensor become

$$\langle f|\, \mu_\rho\, |m\rangle = \langle \nu_f^g|\, \langle e_g|\, \mu_\rho\, |e_e\rangle\, |\nu_m^e\rangle = \langle \nu_f^g|\, (\mu_\rho)_{ge}\, |\nu_m^e\rangle \tag{A.4}$$

where $(\mu_\rho)_{ge}$ is a component of the electronic transition dipole moment between the electronic ground state (g) and an electronic excited state (e). Using these notations, and assuming that the lifetime Γ_m of the intermediate states is independent of the vibrational state, the Eq. A.3 can be further simplified by writing

$$(\alpha_{\rho\sigma})_{i\to f} = \frac{1}{\hbar} \sum_{e\neq g} \sum_m \left\{ \frac{\langle \nu_f^g|\, (\mu_\rho)_{ge} |\nu_m^e\rangle \langle \nu_m^e| (\mu_\sigma)_{eg}\, |\nu_i^g\rangle}{\omega_{eg} + \omega_{mi} - \omega_L - i\Gamma} + \frac{\langle \nu_f^g|\, (\mu_\sigma)_{ge} |\nu_m^e\rangle \langle \nu_m^e| (\mu_\rho)_{eg}\, |\nu_i^g\rangle}{\omega_{eg} + \omega_{mf} + \omega_L + i\Gamma} \right\} \tag{A.5}$$

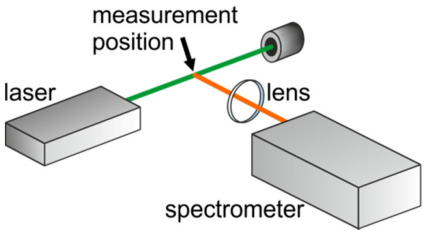


Figure A.3. Typical 90º-setup for Raman spectroscopy. Taken from Ref.[301]

where ω_{eg} and ω_{mi} are the Bohr frequencies associated to the electronic and vibrational energies, respectively. The previous equation gives the possibility of defining the cross section for a given experimental arrangement. In this respect, the initial consideration to take account of is that the relationship between the intensity, I, of Rayleigh and Raman scattered radiation from a single molecule and the irradiance, $\mathscr{I}$, of the incident radiation has the general form

$$I = \sigma' \mathscr{I} = \frac{\partial \sigma}{\partial \Omega} \mathscr{I} \tag{A.6}$$

where the quantities σ (total scattering cross-section per molecule) and σ' (first differential scattering cross-section with respect to a scattering solid angle Ω per molecule) have been included. It becomes clear in Eq. A.6 that I has a dependence on some experimental parameters. Explicitly, the polarization of the incident and scattered radiation and the illumination–observation geometry have to be taken into consideration. For example, it has been reported that in the case a set of N randomly oriented molecules irradiated with monochromatic radiation, irradiance $\mathscr{I}$ and polarization state p^i, the intensity of the radiation of polarization state p^S scattered along a general direction defined by θ (which arises from transitions from an initial state defined by the set of vibrational quantum numbers ν_i to a final state defined by the set of vibrational quantum numbers ν_f) may be written in the general form[260]

$$I(\theta; p^S; p^i) = \left(\frac{1}{16 \varepsilon_0^2 c_0^4 \pi^2} \right) N_{\nu_i} \omega_S^4 \left\{ \frac{1}{45} f(a^2, g^2, d^2, \theta) \right\} \mathscr{I} \tag{A.7}$$

where N_{ν_i} refers to the number of molecules in the initial vibrational state ν_i, c_0 is the speed of light, ε_0 is the electric permittivity in a vacuum, and $f(a^2, g^2, d^2, \theta)$ is a linear combination of the stated invariants[259] whose coefficients include any explicit dependence on θ.

In practice, the two most common experimental setups for Raman spectroscopy are the 90°-setup where the signal collection is arranged in a direction perpendicular to the laser propagation, and the back-scattering setup with a signal collection under 180º with respect to laser beam propagation[301]. The first one is illustrated in Figure A.3 and for such arrangement, the scattered intensity at an angle $\theta = 90$ for any polarization ($\perp^S + \parallel^S$) for incident light with perpendicular polarization ($\perp^L$) is

given as[2601]

$$I(\frac{\pi}{2};\perp^S+\|^S,\perp^L)=\left(\frac{N_{\nu_i}\omega_S^4}{16\varepsilon_0^2c_0^4\pi^2}\right)\left(\frac{45a^2+7g^2+5d^2}{45}\right)\mathscr{I} \tag{A.8}$$

where σ' can be recognized as the product of the terms in the first two parentheses (see Eq. A.6). Alternative definitions for the intensity employ the second differential cross section, σ'', which leads to having $\omega_S^3\omega_L$ rather than ω_S^4 in the numerator of Eq. A.8. Denoting $\alpha_{\rho\sigma}$ as the components of the Raman polarizability tensor of Eq. A.5, the invariants of the transition polarizability tensor, namely, a, g, and d take the following forms[259]

1. *The mean isotropic polarizability, a*

$$a=\frac{1}{3}(\alpha_{xx}+\alpha_{yy}+\alpha_{zz}) \tag{A.9}$$

2. *The antisymmetric anisotropy, d*

$$d^2=\frac{3}{4}\{|\alpha_{xy}-\alpha_{yx}|^2+|\alpha_{yz}-\alpha_{zy}|^2+|\alpha_{zx}-\alpha_{xz}|^2\} \tag{A.10}$$

3. *The symmetric anisotropy, g*

$$g^2=\frac{1}{2}\{|\alpha_{xx}-\alpha_{yy}|^2+|\alpha_{yy}-\alpha_{zz}|^2+|\alpha_{zz}-\alpha_{xx}|^2\}+\frac{3}{4}\{|\alpha_{xy}-\alpha_{yx}|^2+|\alpha_{xz}-\alpha_{zx}|^2+|\alpha_{yz}-\alpha_{zy}|^2\} \tag{A.11}$$

A.1.2. Normal Raman Scattering in the Double Harmonic Approximation

A widely used approximation in the study of vibrational spectroscopies is the double–harmonic approximation. Two well-known implications are contained, namely

(i) Molecular vibrations are described as harmonic oscillators so that the PES is restricted to the harmonic quadratic term in the mass-weighted normal coordinates Q_l. In this case, vibrational states in Eq. A.1 are direct products of one-dimensional states $|\chi\rangle$ for each vibrational mode l (collected initially in the subscripts i, f and m) from a total of M number of modes, such that

[1]The corresponding information for different observation geometries and polarization states of the scattered radiation is given in Tables 5.2(a) and 5.2(b) of Ref.[260].

$$
\begin{aligned}
|\nu_i^g\rangle &= |\chi_1^g\rangle \otimes |\chi_2^g\rangle \cdots \otimes |\chi_M^g\rangle = \prod_{l=1}^{M} |\chi_{i_l}^g\rangle \\
|\nu_f^g\rangle &= |\chi_1^g\rangle \otimes |\chi_2^g\rangle \cdots \otimes |\chi_M^g\rangle = \prod_{l=1}^{M} |\chi_{f_l}^g\rangle \\
|\nu_m^e\rangle &= |\chi_1^e\rangle \otimes |\chi_2^e\rangle \cdots \otimes |\chi_M^e\rangle = \prod_{l=1}^{M} |\chi_{m_l}^e\rangle
\end{aligned}
\tag{A.12}
$$

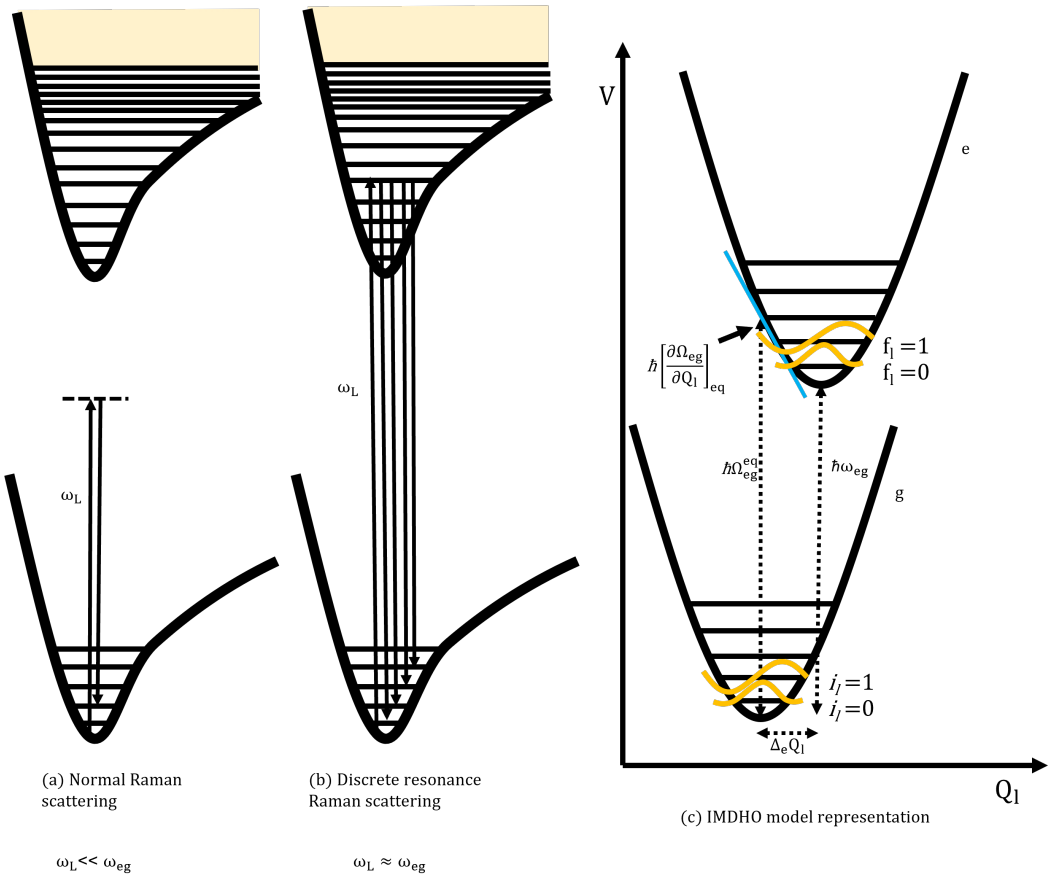


Figure A.4. (a-b) Two types of Raman scattering processes. Image taken and adapted from Ref. 299. (c) Pictorial view of the displaced harmonic oscillator model. Image taken from Ref. 263.

It is evident that the relative magnitudes of ω_L and $(\omega_{eg} + \omega_{mx}, x = i, f)$ will play an important role in the denominator of Eq. A.5 because their difference is involved. In the situation of Normal Raman scattering $\hbar\omega_{eg} \gg \hbar\omega_L$ (Figure A.4(a)), and if $\hbar\omega_L$ is assumed to be much larger than the vibrational transition energies (as it is in most experimental setups which use lasers in the visible spectrum), i.e. $\hbar\omega_L \gg \hbar\omega_{mi}, \hbar\omega_{mf}$, then, according to Placzek's approximation, the vibrational Bohr frequencies (written in red color in Eq. A.5) can be neglected so that $\omega_{eg} + \omega_{mi} \approx \omega_{eg}$, $\omega_{eg} + \omega_{mf} \approx \omega_{eg}$

in the denominators and the closure relation over the intermediate vibrational states, $\sum_m |\nu_m^e\rangle \langle \nu_m^e| = 1$, (written in blue color in Eq. A.5) can be applied in the numerator of such equation, arriving at

$$(\alpha_{\rho\sigma})_{i\to f} = \frac{1}{\hbar} \sum_{e \neq g} \left\{ \frac{\langle \nu_f^g | (\mu_\rho)_{ge} (\mu_\sigma)_{eg} | \nu_i^g \rangle}{\omega_{eg} - \omega_L - i\Gamma} + \frac{\langle \nu_f^g | (\mu_\sigma)_{ge} (\mu_\rho)_{eg} | \nu_i^g \rangle}{\omega_{eg} + \omega_L + i\Gamma} \right\} \quad \text{(A.13)}$$

or in a more compact expression, by defining $(\alpha_{\rho\sigma})_{i\to f} = \langle \nu_f^g | (\alpha_{\rho\sigma})_{gg} | \nu_i^g \rangle$ as the ground state electronic adiabatic polarizability:

$$(\alpha_{\rho\sigma})_{gg} = \frac{1}{\hbar} \sum_{e \neq g} \left\{ \frac{(\mu_\rho)_{ge} (\mu_\sigma)_{eg}}{\omega_{eg} - \omega_L - i\Gamma} + \frac{(\mu_\sigma)_{ge} (\mu_\rho)_{eg}}{\omega_{eg} + \omega_L + i\Gamma} \right\} \quad \text{(A.14)}$$

A more detailed mathematical treatment and analysis of the denominators (including simplifications) of Eq. A.5 can be found in Ref. 299.

(ii) The spectroscopic intensity associated with the excitation to a singly excited state of a particular normal mode, is determined by the magnitude of the first-order geometric derivative of the molecular electric dipole moment and polarizability[302]. In fact, if the tensor $(\alpha_{\rho\sigma})_{gg}$ in Eq. A.14 is developed as a Taylor series exploiting its dependency on the nuclear coordinates, this leads to

$$(\alpha_{\rho\sigma})_{gg} = (\alpha_{\rho\sigma})_{gg}^{eq} + \sum_l \left(\frac{\partial (\alpha_{\rho\sigma})_{gg}}{\partial Q_l} \right)_{eq} Q_l + \dots \quad \text{(A.15)}$$

which is subsequently restricted to the first order in Q_l to obtain for the ground state electronic adiabatic polarizability

$$(\alpha_{\rho\sigma})_{i\to f} = (\alpha_{\rho\sigma})_{gg}^{eq} \delta_{if} + \sum_l \left(\frac{\partial (\alpha_{\rho\sigma})_{gg}}{\partial Q_l} \right)_{eq} \langle \nu_f^g | Q_l | \nu_i^g \rangle \quad \text{(A.16)}$$

In Eq. A.15 and A.16, *eq* indicates that the terms are evaluated at the ground state equilibrium geometry, and the index l denotes a summation over all the mass-weighted normal coordinates Q_l of the electronic ground state. Interestingly, in Eq. A.16 one can recognize both an elastic scattering of light (i.e. Rayleigh scattering) in the first term because $\delta_{if} = 1$ only for $i = f$, and a description of the Raman scattering in the second term, provided that the initial vibrational quantum numbers i are different than the final vibrational quantum numbers f. By introducing the 1D harmonic oscillator states $|\chi\rangle$ from Eq. A.12 and considering only Raman scattering, the above equation becomes

$$(\alpha_{\rho\sigma})_{i\to f} = \sum_l \left(\frac{\partial (\alpha_{\rho\sigma})_{gg}}{\partial Q_l} \right)_{eq} \langle \chi_{f_l}^g | Q_l | \chi_{i_l}^g \rangle \prod_{l \neq l'} \delta_{f_{l'}, i_{l'}} \quad \text{(A.17)}$$

From the well-known identity for harmonic oscillators, namely

$$\langle \nu' | \, x \, | \nu \rangle = \left(\frac{\hbar}{2m\omega} \right)^{1/2} \left[\sqrt{\nu + 1} \delta_{\nu',\nu+1} + \sqrt{\nu} \delta_{\nu',\nu-1} \right] \tag{A.18}$$

it is clear that only transitions involving a modification of ± 1 in the vibrational quantum number i_l are non-zero (selection rules), because

$$\langle \chi^g_{f_l} | \, Q_l \, | \chi^g_{i_l} \rangle = \begin{cases} 0, & f_l \neq i_l \pm 1 \\ \sqrt{\frac{\hbar i_l}{2\omega_l}}, & f_l = i_l - 1 \quad \rightsquigarrow \text{anti-Stokes scattering} \\ \sqrt{\frac{\hbar (i_l+1)}{2\omega_l}}, & f_l = i_l + 1 \quad \rightsquigarrow \text{Stokes scattering} \end{cases} \tag{A.19}$$

Notice that by using the results of Eq. A.19 in Eq. A.17 it is possible to get the Raman polarizability tensor for a fundamental transition (in the double harmonic approximation, the overtone and combination transitions are forbidden) of the type $0 \rightarrow 1_n$

$$(\alpha_{\rho\sigma})_{g0 \rightarrow g1_n} = \sqrt{\frac{\hbar}{2\omega_n}} \left(\frac{\partial (\alpha_{\rho\sigma})_{gg}}{\partial Q_n} \right)_{eq} \tag{A.20}$$

where ω_n is the harmonic frequency for the *n-th* vibrational normal mode (normal coordinate Q_n) of the electronic ground state. The notation $g0$ means $i_l = 0 \ \forall l$ and the notation $g1_n$ means that $f_n = 1$ and $f_l = 0 \ \forall l \neq n$. By combining Eqs. A.20 and A.8 an expression for calculating the Raman differential cross section for a fundamental transition can be found and the Raman activity (S_n) of each vibrational mode Q_n can be defined as

$$S_n = 45a'^2 + 7g'^2 \tag{A.21}$$

where a' and g' (Eqs. A.9 and A.11) contain instead the derivatives of the $\alpha_{\rho\sigma}$ components and the rotational invariant d^2 is equal to zero due to the symmetry of the Raman polarizability tensor.

A.1.3. Signal Enhancement: Resonance Raman

As previously mentioned, photons are not absorbed in Raman scattering, but when $\hbar\omega_L$ approaches an electronic transition energy, an enhancement of the scattered intensity is observed leading to a RR signal. The following is a list of advantages/strengths that have been reported[49,50,269,293,303,304] for this particular kind of scattering:

- RR is a combined electronic-vibrational spectroscopy, whence it is much easier to extract structural information about the system
- Only the molecular vibrations that are affected by the electronic transitions (vibrations involving atoms close to the chromophore) will be visible in the spectrum, making RR spectra easier to interpret than non-resonant Raman ones

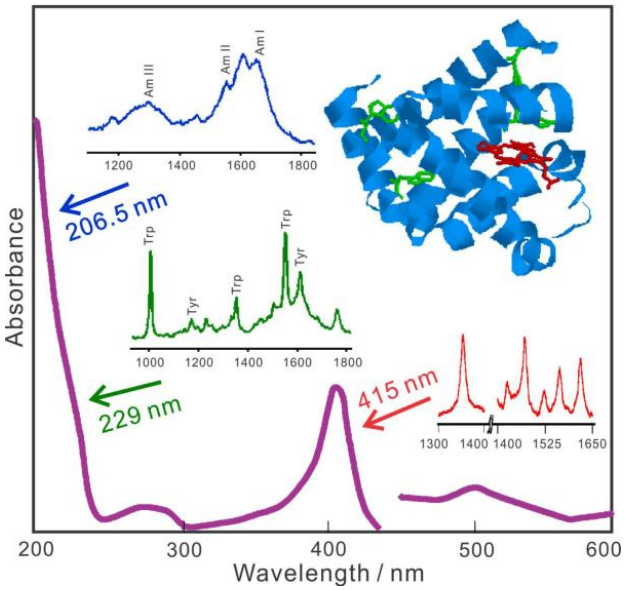


Figure A.5. Selectivity of RR spectral measurements of myoglobin showing the protein absorption spectrum and the different RR spectra obtained with different excitation wavelengths. Image taken from Ref.[303]. Horizontal axes in insets in cm^{-1}.

- The intensity of the scattered radiation is $10^3 - 10^6$ higher than the intensity of the regular Raman signal.
- A RR spectrum carries information about the excited state(s)
- RR has applications in many areas such as analytical chemistry and the study of metal complexes and biological systems, but with higher sensitivity than non-RR, allowing the detection of diluted species in solution, of complex sample mixtures, and of chromophores present in large molecular systems.
- The bandwidth only depends on the initial and final states, both belonging to the ground state PES, therefore the peaks' width is comparable to the case of a non-resonant Raman spectrum.

Figure A.5 shows an example of the UVRR selectivity in the study of a particular protein[303]. The visible wavelength absorption bands of this protein result from an in-plane $\pi \rightarrow \pi^*$ electronic transitions of a specific group, and the UVRR excitation of the protein at 415 nm (the strong absorption band), produce an intense UVRR spectrum which contains only the in-plane specific vibrations, being free from any "contamination" [50] coming from the non-resonant Raman signal.

A.2. Theory of Resonance Raman Scattering and some Approximations

A.2.1. Resonance Raman Intensities

Considering the situation where an incident frequency matches the transition energy or in a more general way $\omega_L \approx \omega_{eg} + \omega_{mi}$ (Figure A.4(b)), the first "resonant" term in Eq. A.5 will become dominant with respect to the second "non-resonant" term, that can be neglected, leading to the following definition for the RR polarizability tensor

$$(\alpha_{\rho\sigma})^{\mathrm{RR}}_{i\to f} = \frac{1}{\hbar}\sum_{e\neq g}\sum_{m}\left\{\frac{\langle\nu_f^g|\,(\mu_\rho)_{ge}\,|\nu_m^e\rangle\,\langle\nu_m^e|\,(\mu_\sigma)_{ge}\,|\nu_i^g\rangle}{\omega_{eg}+\omega_{mi}-\omega_L-i\Gamma}\right\} \tag{A.22}$$

A.2.2. The Time-Independent (TI) Formulation

The main problem with the above equation is that no analytic formulae for the electronic transition dipole moment, $(\mu_\sigma)_{ge}$, are known[269]. For the purpose of overcoming this limitation and taking into account that $(\mu_\sigma)_{ge}$ depends on the nuclear coordinates, this can be approximated as a Taylor series of the normal coordinates of the initial-ground (or intermediate[50]) electronic state about its respective equilibrium geometry, thus

$$(\mu_\rho)_{ge} = (\mu_\rho)^{eq}_{ge} + \sum_l \left(\frac{\partial(\mu_\rho)_{ge}}{\partial Q_l}\right)_{eq} Q_l + \ldots \tag{A.23}$$

where $(\mu_\rho)^{eq}_{ge}$ and $(\frac{\partial(\mu_\rho)_{ge}}{\partial Q_l})_{eq}$ are, respectively, the electronic transition dipole moment and its derivative, both evaluated at the ground state equilibrium geometry. If Eq. A.23 is truncated after the linear term with respect to Q_l and the result is reported in Eq. A.22, an approximated expression for the RR polarizability tensor can be reached, in the commonly referred terms of Franck-Condon (FC) and Herzberg-Teller (HT). It reads

$$(\alpha_{\rho\sigma})^{\mathrm{RR}}_{i\to f} = \frac{1}{\hbar\sqrt{2}}\sum_{e\neq g}\{(\mathrm{FC})^{\mathrm{RR}}_{\mathrm{e}} + (\mathrm{FC/HT})^{\mathrm{RR}}_{\mathrm{e}} + (\mathrm{HT})^{\mathrm{RR}}_{\mathrm{e}}\} \tag{A.24}$$

with the three terms inside the summation standing for the contributions of the electronic excited state (e) to the scattering, and taking the following forms

$$(\mathrm{FC})^{\mathrm{RR}}_{\mathrm{e}} = \sqrt{2}(\mu_\rho)^{eq}_{ge}(\mu_\sigma)^{eq}_{ge}\sum_m \frac{\langle\nu_f^g|\nu_m^e\rangle\,\langle\nu_m^e|\nu_i^g\rangle}{\omega_{eg}+\omega_{mi}-\omega_L-i\Gamma} \tag{A.25}$$

$$(\mathrm{FC/HT})^{\mathrm{RR}}_{\mathrm{e}} = \sqrt{2}\sum_l\left\{(\mu_\rho)^{eq}_{ge}\left(\frac{\partial(\mu_\sigma)_{ge}}{\partial Q_l}\right)_{eq}\sum_m\frac{\langle\nu_f^g|\nu_m^e\rangle\,\langle\nu_m^e|\,Q_l\,|\nu_i^g\rangle}{\omega_{eg}+\omega_{mi}-\omega_L-i\Gamma} + \left(\frac{\partial(\mu_\rho)_{ge}}{\partial Q_l}\right)_{eq}(\mu_\sigma)^{eq}_{ge}\sum_m\frac{\langle\nu_f^g|\,Q_l\,|\nu_m^e\rangle\,\langle\nu_m^e|\nu_i^g\rangle}{\omega_{eg}+\omega_{mi}-\omega_L-i\Gamma}\right\} \tag{A.26}$$

$$(\mathrm{HT})_{\mathrm{e}}^{\mathrm{RR}} = \sqrt{2}\sum_{l,l'}\left(\frac{\partial(\mu_\rho)_{ge}}{\partial Q_l}\right)_{eq}\left(\frac{\partial(\mu_\sigma)_{ge}}{\partial Q_{l'}}\right)_{eq}\sum_{m}\frac{\langle\nu_f^g|\,Q_l\,|\nu_m^e\rangle\,\langle\nu_m^e|\,Q_{l'}\,|\nu_i^g\rangle}{\omega_{eg}+\omega_{mi}-\omega_L-i\Gamma} \quad \text{(A.27)}$$

where $\langle\nu_f^g|\nu_m^e\rangle$ denotes an FC integral describing the overlap between the vibrational wave-functions of the ground state (g) and excited state (e). Since these FC overlap integrals are real quantities within the harmonic oscillator approximation (see Section A.1.2, item (i)), i.e. $\langle\nu_i^g|\nu_m^e\rangle = \langle\nu_m^e|\nu_i^g\rangle$, the *antisymmetric anisotropy, d,* in Eq. A.10 vanishes for pure FC transitions[49]. However, d is generally different from zero when the HT effect is considered.

As an inventory from Eq. A.25, A.26 and A.27 it is clear that the FC contribution depends on the product $(\mu_\rho)_{ge}^{eq}(\mu_\sigma)_{ge}^{eq}$, whereas the FC/HT term involves the products between the transition dipole moment $(\mu_\rho)_{ge}^{eq}$ and the derivatives $\left(\frac{\partial(\mu_\sigma)_{ge}}{\partial Q_l}\right)_{eq}$ with respect to all normal coordinates, and the HT contribution involves all products of the type $\left(\frac{\partial(\mu_\rho)_{ge}}{\partial Q_l}\right)_{eq}\left(\frac{\partial(\mu_\sigma)_{ge}}{\partial Q_{l'}}\right)_{eq}$. Again, by using the identity for harmonic oscillators (Eq. A.18) the FC/HT and HT terms of Eq. A.26 and A.27 can be expressed as a function of FC overlap integrals. In practice, to compute the RR intensities, quantum chemistry calculations should provide the next quantities

1. Electronic and vibrational transition frequencies, ω_{eg} and ω_{mi}
2. Components of the transition dipole moment $(\mu_\rho)_{ge}^{eq}$ and their derivatives $\left(\frac{\partial(\mu_\rho)_{ge}}{\partial Q_l}\right)_{eq}$
3. FC overlap integrals $\langle\nu_f^g|\nu_m^e\rangle$ between the ground (g) and the excited states (e).

A.2.3. The Time-Dependent (TD) Formulation

The time-dependent approach is an alternative method for the calculation of RR intensities. It was originally developed by Heller and coworkers, and is based on wave packet dynamics.[267–272,305] An equivalent expression for the RR polarizability shown in Eq. A.22 can be obtained in the time domain:

$$(\alpha_{\rho\sigma})_{i\to f}^{\mathrm{RR}} = \frac{i}{\hbar}\sum_{e\neq g}(\mu_\rho)_{ge}^{eq}(\mu_\sigma)_{ge}^{eq}\int_0^{+\infty}\sum_m\langle\nu_f^g|\nu_m^e\rangle\,\langle\nu_m^e|\nu_i^g\rangle\,e^{-i(\omega_{eg}+\omega_{mi}-\omega_L-i\Gamma)t}dt \quad \text{(A.28)}$$

Entering the following definitions in the above equation:

(i) the vibrational energies E_i and E_m of the ground (g) and excited (e) states, respectively, defining the vibrational Bohr frequency $\omega_{mi} = (E_m - E_i)/\hbar$

(ii) the vibrational Hamiltonian $\mathrm{H}_e^{\mathrm{Vib}}$ of the electronic excited state (e), verifying the eigenvalue equation $\mathrm{H}_e^{\mathrm{Vib}}\,|\nu_m^e\rangle = E_m\,|\nu_m^e\rangle$

(iii) the time evolution operator (i.e. propagator) $e^{-\frac{i}{\hbar}H_e^{\mathrm{Vib}}t} = \sum_m e^{-\frac{i}{\hbar}E_m t}\,|\nu_m^e\rangle\,\langle\nu_m^e|$

Then, the RR polarizability tensor will take the form:

$$(\alpha_{\rho\sigma})_{i\to f}^{\mathrm{RR}} = \frac{i}{\hbar}\sum_{e\neq g}(\mu_\rho)_{ge}^{eq}(\mu_\sigma)_{ge}^{eq}\int_0^{+\infty}\langle f|i_e(t)\rangle\, e^{-i(\omega_{eg}-E_m/\hbar-\omega_L-i\Gamma)t}dt \tag{A.29}$$

where $\langle i| \equiv |\nu_m^e\rangle$ and $\langle f| \equiv |\nu_f^g\rangle$ are the initial and final vibrational states, respectively, and $\langle i_e(t)| \equiv e^{-\frac{i}{\hbar}H_e^{\mathrm{Vib}}t}\,|i\rangle$ defines the time-dependent vibrational wavefunction of the excited state (e). The key piece in the TD formulation is that the RR polarizability tensor is expressed as a half-Fourier transform of a time-dependent overlap between the final vibrational state $\langle f|$ and the initial vibrational wave packet propagated by the excited state vibrational Hamiltonian $\mathrm{H}_e^{\mathrm{Vib}}$.

In contrast to the TI formulation that requires the evaluation of infinite summations over vibronic states, the TD method needs a numerical integration of an unbounded integral. What they have in common is that their expressions involve the necessity of a model for describing the ground and excited PESs, and the transition electric dipole moment.[50]

A.2.4. Description of the ground and excited states

Broadly speaking, the harmonic approximation for the BO PES of an electronic state can be used to collect vibrational frequencies and normal modes in both the ground and the excited state cases, but in particular for the excited state, it is computationally demanding. Consequently, several methods and algorithms have been proposed in the literature.

1. Independent Mode Displaced Harmonic Oscillator (IMDHO) model: It is the simplest and most widely used approximation. It is assumed that the ground and excited states have the same normal coordinates and the same vibrational frequencies, and they differ only in the equilibrium geometry. In fact, within this model, the difference between the ground and excited state PESs is only determined by the geometrical displacements $\Delta_e Q_l$ along the normal coordinates Q_l, and, as shown in Figure A.4(c), the dimensionless displacements, $\Delta_{e,l}$, can be obtained from the derivatives of the vertical transition frequency Ω_{eg} with respect to the normal coordinates Q_l (also called excited state gradients), evaluated at the ground state (g) geometry, according to the following

$$\Delta_{e,l} = \sqrt{\frac{\omega_l}{\hbar}}\Delta_e Q_l = -\frac{\sqrt{\hbar}}{\omega_l^{3/2}}\left(\frac{\partial\Omega_{eg}}{\partial Q_l}\right)_{eq} \tag{A.30}$$

   This approach is known as the **Vertical Gradient (VG)** model.

Another choice is an adiabatic approach, known as **Adiabatic Shift (AS)**, in which the excited-state geometry is optimized, but the shape of the PES is then assumed to be the same as in the ground state. Displacements, $\Delta_{e,l}$, are then calculated from the geometry difference between the equilibrium structures of the ground and excited states. AS model is, of course, computationally more expensive than the VG one, because it requires an optimization of the excited state geometry.

The definition of the displacements in the framework of the IMDHO model has an implication for the overlap integral, $\langle f|i_e(t)\rangle$, of Eq. A.29 in the TD formalism, and it is the fact that the dimensionless oscillator displacements are the only parameters appearing in the expression:[270]

$$\langle f|i_e(t)\rangle = \prod_l \left\{ \frac{(-1)^{k_l}\Delta_{e,l}^{k_l}}{2^{k_l/2}k_l!}(1-e^{-i\omega_l t})^{k_l} \right\} e^{-\frac{\Delta_{e,l}^2}{2}(1-e^{-i\omega_l t})} e^{-i(\omega_{eg}+\omega_{mi})t} \quad \text{(A.31)}$$

Where the k_l denote the excitation number of mode l in final vibrational state $|f\rangle$.

In addition, by taking advantage of the IMDHO model and exploiting the connection between absorption and RR spectroscopies, it is possible to derive simplified expressions for the RR polarizability tensor. For instance, an auxiliary function, $\Phi_e(\omega_L)$, that describes the Resonance Raman Excitation Profile (RREP) or the dependency of the RR intensity with respect to the incident frequency ω_L can be defined as[290]

$$\Phi_e(\omega_L) = \sum_m \frac{\langle \nu_0^g|\nu_m^e\rangle^2}{\omega_{eg}+\omega_{m0}-\omega_L-i\Gamma} \quad \text{(A.32)}$$

In the particular case of fundamental transitions $0 \rightarrow 1_n$, the evaluation of the FC contribution (Eq. A.25) requires the estimation of FC overlap integrals of the type $\langle \nu_0^g|\nu_m^e\rangle$ and $\langle \nu_{1_n}^g|\nu_m^e\rangle$ $\forall n, \forall \nu$. Since in the IMDHO model the recursive relations can be used to express the multidimensional FC overlap integrals as a product of 1D overlap integrals[263], $\langle \nu_0^g|\nu_m^e\rangle^2$ in Eq. A.32 can be written as

$$\langle \nu_0^g|\nu_m^e\rangle^2 = \prod_{l=1}^{M} \langle \chi_0^g|\chi_{m_l}^e\rangle^2 = \prod_{l=1}^{M} \frac{\Delta_{e,l}^{2m_l}}{2^{m_l}m_l!} e^{-\frac{\Delta_{e,l}^2}{2}} \quad \text{(A.33)}$$

which shows that the calculation of the FC factors of this type only requires the knowledge of the displacements $\Delta_{e,l}$ and of the quantum numbers m_l. Making use of those recursive relations and employing the $\Phi_e(\omega_L)$ definition, the FC term (Eq. A.25) for $0 \rightarrow 1_n$ transitions can be written as

$$(\mathrm{FC})_{\mathrm{e}}^{\mathrm{RR}} = (\mu_\rho)_{\mathrm{ge}}^{\mathrm{eq}}(\mu_\sigma)_{\mathrm{ge}}^{\mathrm{eq}}\Delta_{\mathrm{e,n}}\{\Phi_{\mathrm{e}}(\omega_{\mathrm{L}}) - \Phi_{\mathrm{e}}(\omega_{\mathrm{L}}-\omega_{\mathrm{n}})\} \quad \text{(A.34)}$$

where a specific relationship between the RR intensity and the geometrical displacements is shown. It is worth noticing that Eq. A.34 involves just products between the transition dipole moments, the dimensionless displacements, and a linear combination of the Φ_e function evaluated at different frequencies. Also, $\Phi_e(\omega_L)$ demands only the calculation of FC overlap integrals of the type $\langle \nu_0^g | \nu_0^e \rangle$.

By assuming small values for the displacements ($\Delta_{e,n} \ll 1$), even more simplified equations can be obtained. In those cases, fundamental transitions are expected to be the strongest ones and the largest FC factor is given by $\langle \nu_0^g | \nu_0^e \rangle \approx 1$ and $\langle \nu_0^g | \nu_m^e \rangle \approx 0$ for $m \neq 0$. In addition, the vibrational Bohr frequency $\omega_{00} = 0$ within the IMDHO model. Hence, Eq. A.32 and A.34 become, respectively:

$$\Phi_e(\omega_L) = \frac{1}{\omega_{eg} - \omega_L - i\Gamma} \tag{A.35}$$

$$(\mathrm{FC})_{\mathrm{e}}^{\mathrm{RR}} = (\mu_\rho)_{\mathrm{ge}}^{\mathrm{eq}} (\mu_\sigma)_{\mathrm{ge}}^{\mathrm{eq}} \Delta_{\mathrm{e,n}} \left(\frac{1}{\omega_{\mathrm{eg}} - \omega_{\mathrm{L}} - \mathrm{i}\Gamma} \right) \left(\frac{\omega_{\mathrm{n}}}{\omega_{\mathrm{eg}} - \omega_{\mathrm{L}} + \omega_{\mathrm{n}} - \mathrm{i}\Gamma} \right) \tag{A.36}$$

The latter equation is known as the **small-shift approximation**.

Further simplifications can be achieved assuming a pre-resonance situation with a single electronic excited state (e), i.e. $\omega_{eg} - \omega_L \gg \omega_n$. Thus, the vibrational frequency ω_n can be neglected in the previous denominator, and replacing the whole result in the FC contribution of Eq. A.24, we have

$$(\alpha_{\rho\sigma})_{i\to f}^{\mathrm{RR}} = \frac{1}{\hbar\sqrt{2}} \sum_{e \neq g} \left\{ (\mu_\rho)_{ge}^{eq} (\mu_\sigma)_{ge}^{eq} \Delta_{e,n} \left(\frac{1}{\omega_{eg} - \omega_L - i\Gamma} \right) \left(\frac{\omega_n}{\omega_{eg} - \omega_L - i\Gamma} \right) \right\} \tag{A.37}$$

Moreover, by using Eq. A.37 and $\omega_S = \omega_L - \omega_n \approx \omega_L$ altogether in Eq. A.8 (sequence $(\mathrm{FC})_{\mathrm{e}}^{\mathrm{RR}} \to (\alpha_{\rho\sigma})_{\mathrm{i}\to\mathrm{f}}^{\mathrm{RR}} \to (\alpha_{\rho\sigma})^2 \to$ RR Intensity), it follows that the relative RR intensities for fundamental transitions can be approximated by

$$I_{g0\to g1_n} \propto \omega_n^2 \Delta_{e,n}^2 = \frac{\omega_n^2 \hbar}{\omega_n^3} \left(\frac{\partial \Omega_{eg}}{\partial Q_n} \right)_{eq}^2 = \frac{\hbar}{\omega_n} \left(\frac{\partial \Omega_{eg}}{\partial Q_n} \right)_{eq}^2 \tag{A.38}$$

The first relation of proportionality is known as the Savin formula, whereas the final result is generally known as the **gradient method or the short-time approximation (STA)**, which is only applicable when a single electronic excited state is in resonance with the incident frequency, but does not require a harmonic PES for the excited state.

Finally, under short-time dynamics conditions, it is possible to write the ratio between the RR intensities of two modes j and l as[268,287]

$$\frac{I_j}{I_l} = \frac{\omega_j^2 \Delta_{e,j}^2}{\omega_l^2 \Delta_{e,l}^2} = \frac{\left(\frac{\partial \Omega_{eg}}{\partial Q_j}\right)_{eq}^2}{\left(\frac{\partial \Omega_{eg}}{\partial Q_l}\right)_{eq}^2} \tag{A.39}$$

2. Methods beyond the IMDHO model: They are employed as a generalization of the treatment in which the ground and excited states have different equilibrium geometries, different vibrational frequencies, and different normal coordinates. Even though the excited state's normal modes are calculated, the integrals appearing in the polarizability expressions cannot be directly computed since the harmonic wave functions of the two PESs are expressed in different basis sets[50]. If the molecule does not undergo too large distortions (semi-rigid molecule) between the equilibrium geometries of both states, the relation between the normal coordinates of the ground and excited states can be expressed as a linear transformation proposed by Duschinsky[3062]

$$\mathbf{Q}_g = \mathbf{J}\mathbf{Q}_e + \mathbf{k}_e \tag{A.40}$$

where $\mathbf{Q}_g$ and $\mathbf{Q}_e$ are column matrices containing the mass-weighted normal coordinates of the ground (Q_l) and excited ($Q_{e,l}$) states, respectively. $\mathbf{k}_e$ is the column matrix (shift vector) of the displacements, i.e. $(\mathbf{k}_e)_l = \sqrt{\frac{\hbar}{\omega_l}}\Delta_{e,l}$, and $\mathbf{J}$ is the rotation or Duschinsky matrix defined by

$$\mathbf{J} = (\mathbf{L}_g)^{-1}\mathbf{L}_e \tag{A.41}$$

where $\mathbf{L}_g$ and $\mathbf{L}_e$ are the transformation matrices connecting the mass-weighted normal coordinates to the mass-weighted Cartesian coordinates for the ground and excited states, respectively. $\mathbf{L}_g$ and $\mathbf{L}_e$ are obtained by solving the ground and excited state vibrational normal mode eigenvalue problem in the harmonic approximation[263].

It should be noticed here that the inclusion of Duschinsky rotation effects gives rise to an improvement of the AS model, often called the **Adiabatic Hessian (AH)** model. It consists in performing a geometry optimization followed by a vibrational analysis (computation of the Hessian matrix) for both ground and excited states. Therefore, it takes into account the changes in vibrational frequencies and normal modes between both electronic states. Furthermore, if the excited state Hessian matrix is calculated at the ground state equilibrium geometry, and the shift vector can be extrapolated from the gradient, the **Vertical Hessian (VH)** model is attained. Finally, if the

[2]IMDHO model neglects Duschinsky rotation

Table A.1. Computations Required to Generate Input Data for Simulation of RR Spectra with VG, AS, VH, and AH Models. Table adapted from Ref. 24

Approximations for the PES				
Calculation	VG	AS	VH	AH
Initial state				
Cartesian coordinates of atoms (equilibrium structure)	✓	✓	✓	✓
Energy at the minimum of PES (equilibrium geometry)	✓	✓	✓	✓
Frequencies	✓	✓	✓	✓
Normal modes, expressed by atomic displacements	✓	✓	✓	✓
Final state				
Cartesian coordinates of atoms at minimum of PES (equilibrium structure)		✓		✓
Energy at equilibrium geometry of initial state	✓		✓	
Energy at minimum of PES (equilibrium geometry)		✓		✓
Forces at equilibrium geometry of initial state	✓		✓	
Frequencies at equilibrium geometry of initial state			✓	
Frequencies at minimum of PES (equilibrium geometry)				✓
Normal modes, expressed by atomic displacements				✓

Duschinsky rotation is ignored and excited-state frequencies are not computed, the simplified, already explained, VG and AS methods are obtained. A practical summary of the data required as input by each model (VG, AS, VH, AH) is listed in Table A.1.

It is worth stressing here that for the calculation of the Franck-Condon overlap integrals between the harmonic PESs of the ground and excited states, several methods and algorithms have been proposed in the literature using either analytic or recursive formulae. One result is for example that one employed in Eq. A.33. However, in a sizable molecule, the number of vibrational states of the intermediate electronic state is huge, as is the number of the possible final vibrational states (belonging to the ground electronic state), then a prescreening scheme is needed to select *a priori* the vibrational levels that give the most important contributions to the RR spectrum.[49,50]

A.2.5. Transform Theory (TT) and Simplified Φ_e Approximation

Transform Theory is based on the relationship between the polarizability and the absorption spectrum. In general, the simulation of RR spectra is accompanied by the calculation of the absorption spectrum in order to identify the bands and vibronic transitions that are in resonance with the incident photon energy $\hbar\omega_L$. In the most common version, TT derives relative RR intensities from the differences in the equilibrium structures between the ground and the resonant excited states and a lineshape function Φ accessible from the experimental absorption spectrum[49]. In short, the experimental absorption band shape is related to the imaginary part of Φ_e, which by means of the Kramers–Kronig relations leads to the real part of Φ_e, that is finally the same function defined in Eq. A.32, from which RR intensities and RREPs can be simulated once the displacements $\Delta_{e,n}$ are obtained. Other contributions in the literature deal with computations of Φ from sum over states approaches and

not from the experimental data. The small-shift approximation referred to in Eqs. A.35 and A.36 is a particular case of the simplified Φ_e approximation, defined in the framework of the TT.

A.2.6. Resonance Polarizability Derivatives

A computationally different approach for the calculation of RR intensities is based on the geometrical derivatives of the frequency-dependent resonance polarizabilities with respect to the normal coordinates, calculated by including a finite lifetime (damping factor) of the electronic excited states[274–277]. This method relies on a short-time dynamics approximation and is similar to the simple excited-state gradient approximation method (see Eq. A.38) if only one electronic excited state is important.

It was already shown with the expression in Eq. A.20 that within the double-harmonic approximation, the Raman scattering cross sections are proportional to the derivatives of $\alpha_{\rho\sigma}$ with respect to vibrational normal modes. In the case of the Resonance Polarizability Derivatives, such a relationship holds but $\alpha_{\rho\sigma}$ is referred to the complex electronic polarizability:

$$(\alpha_{\rho\sigma})_{g0\rightarrow g1_n} = \sqrt{\frac{\hbar}{2\omega_n}} \left(\frac{\partial(\alpha_{\rho\sigma})_{gg}}{\partial Q_n} \right)_{eq} \tag{A.42}$$

where ω_n is the harmonic frequency for the *n-th* vibrational normal mode (normal coordinate Q_n) of the electronic ground state. The main advantage of that strategy is that all the electronic states are included in the polarizability, being also well-suited for dealing with large molecules or small molecules in complex environments.[274,278] However, it is not widely implemented because it requires the solution of response equations in the complex formalism.

Concerning Resonance Polarizability Derivatives, Rappoport et al.[307] derived a simplified expression including FC and HT contributions, by differentiating the SOS expansion for $\alpha(\omega_L)$ (see Eq. A.34) with respect to the vibrational normal mode Q. The proposed SOS approach uses only quantities calculated at the ground state geometry, namely, electronic excitation energies and their derivatives and transition dipole moments and their derivatives. Also, it has been demonstrated that within the IMDHO model, such expression is an approximation of the simplified Φ_e function approach.[290]

Bibliography

[1] Gómez, S.; Giovannini, T.; Cappelli, C. Multiple Facets of Modeling Electronic Absorption Spectra of Systems in Solution. *ACS Physical Chemistry Au* **2023**, *3*, 1–16.

[2] Gómez, S.; Giovannini, T.; Cappelli, C. Absorption spectra of xanthines in aqueous solution: a computational study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5929–5941.

[3] Gómez, S.; Rojas-Valencia, N.; Giovannini, T.; Restrepo, A.; Cappelli, C. Ring Vibrations to Sense Anionic Ibuprofen in Aqueous Solution as Revealed by Resonance Raman. *Molecules* **2022**, *27*.

[4] Gómez, S.; Egidi, F.; Puglisi, A.; Giovannini, T.; Rossi, B.; Cappelli, C. Unlocking the power of resonance Raman spectroscopy: The case of amides in aqueous solution. *J. Mol. Liq.* **2021**, *346*, 117841.

[5] Gómez, S.; Bottari, C.; Egidi, F.; Giovannini, T.; Rossi, B.; Cappelli, C. Amide Spectral Fingerprints are Hydrogen Bonding-Mediated. *J. Phys. Chem. Lett.* **2022**, *13*, 6200–6207.

[6] Rojas-Valencia, N.; Gómez, S.; Montillo, S.; Manrique-Moreno, M.; Cappelli, C.; Hadad, C.; Restrepo, A. Evolution of Bonding during the Insertion of Anionic Ibuprofen into Model Cell Membranes. *J. Phys. Chem. B* **2020**, *124*, 79–90.

[7] Rojas-Valencia, N.; Gómez, S.; Núñez-Zarur, F.; Cappelli, C.; Hadad, C.; Restrepo, A. Thermodynamics and Intermolecular Interactions during the Insertion of Anionic Naproxen into Model Cell Membranes. *J. Phys. Chem. B* **2021**, *125*, 10383–10391.

[8] Lafiosca, P.; Gómez, S.; Giovannini, T.; Cappelli, C. Absorption Properties of Large Complex Molecular Systems: The DFTB/Fluctuating Charge Approach. *J. Chem. Theory Comput.* **2022**, *18*, 1765–1779.

[9] Gómez, S.; Lafiosca, P.; Egidi, F.; Giovannini, T.; Cappelli, C. UV-Resonance Raman Spectra of Systems in Complex Environments: A Multiscale Modeling applied to Doxorubicin intercalated into DNA. *Just Accepted in J. Chem. Inf. Model.* **2023**,

[10] Rojas-Valencia, N.; Gómez, S.; Giovannini, T.; Cappelli, C.; Restrepo, A.; Núñez Zarur, F. Water maintains the UV-Vis spectral features during the insertion of anionic Naproxen and Ibuprofen into model cell membranes. *Submitted* **2023**,

[11] Gómez, S.; Rojas-Valencia, N.; Gómez, S. A.; Cappelli, C.; Merino, G.; Restrepo, A. A molecular twist on hydrophobicity. *Chem. Sci.* **2021**, *12*, 9233–9245.

[12] Goletto, L.; Gómez, S.; Andersen, J.; Koch, H.; Giovannini, T. Linear Response Properties of Solvated Systems: A Computational Study. *Phys. Chem. Chem. Phys.* **2022**, *24*, 27866–27878.

[13] Gómez, S. A.; Rojas-Valencia, N.; Gómez, S.; Egidi, F.; Cappelli, C.; Restrepo, A. Binding of SARS-CoV-2 to Cell Receptors: A Tale of Molecular Evolution. *ChemBioChem* **2021**, *22*, 724–732.

[14] Gomez, S. A.; Rojas-Valencia, N.; Gómez, S.; Cappelli, C.; Restrepo, A. The role of spike protein mutations in the infectious power of SARS-COV-2 variants: A molecular interaction perspective. *ChemBioChem* **2021**, *23*, e202100393.

[15] Uribe, L.; Gómez, S.; Giovannini, T.; Egidi, F.; Restrepo, A. An efficient and robust procedure to calculate absorption spectra of aqueous charged species applied to NO_2^-. *Phys. Chem. Chem. Phys.* **2021**, *23*, 14857–14872.

[16] Uribe, L.; Gómez, S.; Egidi, F.; Giovannini, T.; Restrepo, A. Computational hints for the simultaneous spectroscopic detection of common contaminants in water. *J. Mol. Liq.* **2022**, *355*, 118908.

[17] Arias, A.; Gómez, S.; Rojas-Valencia, N.; Núñez-Zarur, F.; Cappelli, C.; Murillo-López, J. A.; Restrepo, A. Formation and evolution of C–C, C–O, C=O and C–N bonds in chemical reactions of prebiotic interest. *RSC Adv.* **2022**, *12*, 28804–28817.

[18] Gómez, S.; Gómez, S.; David, J.; Guerra, D.; Cappelli, C.; Restrepo, A. Dissecting Bonding Interactions in Cysteine Dimers. *Molecules* **2022**, *27*.

[19] Sepali, C.; Lafiosca, P.; Gómez, S.; Giovannini, T.; Cappelli, C. Computational Protocol to simulate solvent effects on Raman and Raman Optical Activity spectra of systems in aqueous solution. *Submitted* **2022**,

[20] Gómez, S.; Ambrosetti, M.; Giovannini, T.; Cappelli, C. A Close-up Look at the Electronic Spectroscopic Signatures of Common Pharmaceuticals in Solution. *Submitted* **2022**,

[21] Long, D. *The Raman Effect*; John Wiley & Sons, Ltd, 2002; Chapter 1, pp 3–18.

[22] Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664–5677.

[23] Barone, V.; Alessandrini, S.; Biczysko, M.; Cheeseman, J. R.; Clary, D. C.; McCoy, A. B.; DiRisio, R. J.; Neese, F.; Melosso, M.; Puzzarini, C. Computational molecular spectroscopy. *Nat. Rev. Dis. Primers* **2021**, *1*, 1–27.

[24] Barone, V. *Computational Strategies for Spectroscopy: from Small Molecules to Nano Systems*; Wiley, 2011.

[25] Cappelli, C. Integrated QM/Polarizable MM/Continuum Approaches to Model Chiroptical Properties of Strongly Interacting Solute-Solvent Systems. *Int. J. Quantum Chem.* **2016**, *116*, 1532–1542.

[26] Tomasi, J.; Persico, M. Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent. *Chem. Rev.* **1994**, *94*, 2027–2094.

[27] Mennucci, B. Polarizable Continuum Model. *WIREs Comput. Mol. Sci.* **2012**, *2*, 386–404.

[28] Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

[29] Warshel, A.; Karplus, M. Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization. *J. Am. Chem. Soc.* **1972**, *94*, 5612–5625.

[30] Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981**, *55*, 117–129.

[31] Orozco, M.; Luque, F. J. Theoretical methods for the description of the solvent effect in biomolecular systems. *Chem. Rev.* **2000**, *100*, 4187–4226.

[32] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3094.

[33] Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

[34] Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700–733.

[35] Singh, U. C.; Kollman, P. A. A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the CH3Cl+ Cl- exchange reaction and gas phase protonation of polyethers. *J. Comput. Chem.* **1986**, *7*, 718–730.

[36] Olsen, J. M. H.; Kongsted, J. *Advances in quantum chemistry*; Elsevier, 2011; Vol. 61; pp 107–143.

[37] Dohn, A. O. Multiscale electrostatic embedding simulations for modeling structure and dynamics of molecules in solution: a tutorial review. *Int. J. Quantum Chem.* **2020**, *120*, e26343.

[38] Bondanza, M.; Nottoli, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. Polarizable embedding QM/MM: the future gold standard for complex (bio) systems? *Phys. Chem. Chem. Phys.* **2020**, *22*, 14433–14448.

[39] Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. The effective fragment potential method: A QM-based MM approach to modeling environmental effects in chemistry. *J. Phys. Chem. A* **2001**, *105*, 293–307.

[40] Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. The effective fragment potential: a general method for predicting intermolecular interactions. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177–193.

[41] Thompson, M. A. QM/MMpol: A Consistent Model for Solute/Solvent Polarization. Application to the Aqueous Solvation and Spectroscopy of Formaldehyde, Acetaldehyde, and Acetone. *J. Phys. Chem.* **1996**, *100*, 14492–14507.

[42] Loco, D.; Polack, É.; Caprasecca, S.; Lagardere, L.; Lipparini, F.; Piquemal, J.-P.; Mennucci, B. A QM/MM approach using the AMOEBA polarizable embedding: from ground state energies to electronic excitations. *J. Chem. Theory Comput.* **2016**, *12*, 3654–3661.

[43] Giovannini, T.; Puglisi, A.; Ambrosetti, M.; Cappelli, C. Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQFμ model. *J. Chem. Theory Comput.* **2019**, *15*, 2233–2245.

[44] Giovannini, T.; Egidi, F.; Cappelli, C. Theory and algorithms for chiroptical properties and spectroscopies of aqueous systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864–22879.

[45] Steinmann, C.; Reinholdt, P.; Nørby, M. S.; Kongsted, J.; Olsen, J. M. H. Response properties of embedded molecules through the polarizable embedding model. *Int. J. Quantum Chem.* **2019**, *119*, e25717.

[46] Loco, D.; Gelfand, N.; Jurinovich, S.; Protti, S.; Mezzetti, A.; Mennucci, B. Polarizable QM/Classical Approaches for the Modeling of Solvation Effects on UV–Vis and Fluorescence Spectra: An Integrated Strategy. *J. Phys. Chem. A* **2018**, *122*, 390–397.

[47] Cerezo, J.; Aranda, D.; Avila Ferrer, F. J.; Prampolini, G.; Santoro, F. Adiabatic-molecular dynamics generalized vertical hessian approach: a mixed quantum classical method to compute electronic spectra of flexible molecules in the condensed phase. *J. Chem. Theory Comput.* **2019**, *16*, 1215–1231.

[48] Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A. Elucidating peptide and protein structure and dynamics: UV resonance Raman spectroscopy. *J. Phys. Chem. Lett.* **2011**, *2*, 334–344.

[49] Santoro, F.; Cappelli, C.; Barone, V. Effective Time-Independent Calculations of Vibrational Resonance Raman Spectra of Isolated and Solvated Molecules Including Duschinsky and Herzberg–Teller Effects. *J. Chem. Theory Comput.* **2011**, *7*, 1824–1839.

[50] Egidi, F.; Bloino, J.; Cappelli, C.; Barone, V. A Robust and Effective Time-Independent Route to the Calculation of Resonance Raman Spectra of Large Molecules in Condensed Phases with the Inclusion of Duschinsky, Herzberg–Teller, Anharmonic, and Environmental Effects. *J. Chem. Theory Comput.* **2014**, *10*, 346–363.

[51] Olszówka, M.; Russo, R.; Mancini, G.; Cappelli, C. A computational approach to the resonance Raman spectrum of doxorubicin in aqueous solution. *Theor. Chem. Acc.* **2016**, *135*, 27.

[52] Reinholdt, P.; Kjellgren, E. R.; Steinmann, C.; Olsen, J. M. H. Cost-effective potential for accurate polarizable embedding calculations in protein environments. *J. Chem. Theory Comput.* **2019**, *16*, 1162–1174.

[53] Reinholdt, P.; Jørgensen, F. K.; Kongsted, J.; Olsen, J. M. H. Polarizable density embedding for large biomolecular systems. *J. Chem. Theory Comput.* **2020**, *16*, 5999–6006.

[54] Kessler, J.; Kapitán, J.; Bouř, P. First-principles predictions of vibrational Raman optical activity of globular proteins. *J. Phys. Chem. Lett.* **2015**, *6*, 3314–3319.

[55] Kessler, J.; Bouř, P. Classical Trajectory of Molecules in Electromagnetic Field: A Handy Method to Simulate Molecular Vibrational Spectra. *J. Chem. Theory Comput.* **2022**, *18*, 1780–1787.

[56] Kessler, J.; Andrushchenko, V.; Kapitán, J.; Bouř, P. Insight into vibrational circular dichroism of proteins by density functional modeling. *Phys. Chem. Chem. Phys.* **2018**, *20*, 4926–4935.

[57] Frenkel, D.; Smit, B. *Understanding molecular simulation: from algorithms to applications*, 2nd ed.; Academic Press: San Diego, 2002.

[58] Pérez, J.; Restrepo, A. ASCEC V02:Annealing Simulado con Energía Cuántica. 2008; Property, development, and implementation: Grupo de Química–Física Teórica, Instituto de Química, Universidad de Antioquia, Medellín, Colombia.

[59] Pérez, J. F.; Hadad, C. Z.; Restrepo, A. Structural studies of the water tetramer. *Int. J. Quantum Chem.* **2008**, *108*, 1653–1659.

[60] Pérez, J. F.; Florez, E.; Hadad, C. Z.; Fuentealba, P.; Restrepo, A. Stochastic Search of the Quantum Conformational Space of Small Lithium and Bimetallic Lithium–Sodium Clusters. *J. Phys. Chem. A* **2008**, *112*, 5749–5755.

[61] Vainio, M. J.; Johnson, M. S. Generating conformer ensembles using a multiobjective genetic algorithm. *J. Chem. Inf. Model.* **2007**, *47*, 2462–2474.

[62] Ishikawa, Y. A script for automated 3-dimentional structure generation and conformer search from 2-dimentional chemical drawing. *Bioinformation* **2013**, *9*, 988.

[63] Poli, G.; Seidel, T.; Langer, T. Conformational sampling of small molecules with iCon: Performance assessment in comparison with OMEGA. *Front. Chem.* **2018**, *6*, 229.

[64] Gao, J. Monte Carlo Quantum Mechanical-Configuration Interaction and Molecular Mechanics Simulation of Solvent Effects on the n $\rightarrow$ pi* Blue Shift of Acetone. *J. Am. Chem. Soc.* **1994**, *116*, 9324–9328.

[65] Hollingsworth, S. A.; Dror, R. O. Molecular dynamics simulation for all. *Neuron* **2018**, *99*, 1129–1143.

[66] Kemp, M. *An Introduction to Molecular Dynamics*; Physics research and technology; Nova Science Publishers, 2019.

[67] Zhou, K.; Liu, B. *Molecular Dynamics Simulation: Fundamentals and Applications*; Elsevier Science, 2022.

[68] Chen, G.; Huang, K.; Miao, M.; Feng, B.; Campanella, O. H. Molecular dynamics simulation for mechanism elucidation of food processing and safety: State of the art. *Compr. Rev. Food Sci. Food Saf.* **2019**, *18*, 243–263.

[69] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

[70] Burley, S. K.; Bhikadiya, C.; Bi, C.; Bittrich, S.; Chen, L.; Crichlow, G. V.; Christie, C. H.; Dalenberg, K.; Di Costanzo, L.; Duarte, J. M. et al. RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res.* **2020**, *49*, D437–D451.

[71] Nian, B.; Xu, Y.-J.; Liu, Y. Molecular dynamics simulation for mechanism revelation of the safety and nutrition of lipids and derivatives in food: State of the art. *Food Res. Int.* **2021**, *145*, 110399.

[72] Renger, T.; Müh, F. Understanding photosynthetic light-harvesting: a bottom up theoretical approach. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3348–3371.

[73] Jurinovich, S.; Viani, L.; Curutchet, C.; Mennucci, B. Limits and potentials of quantum chemical methods in modelling photosynthetic antennae. *Phys. Chem. Chem. Phys.* **2015**, *17*, 30783–30792.

[74] Rosnik, A. M.; Curutchet, C. Theoretical characterization of the spectral density of the water-soluble chlorophyll-binding protein from combined quantum mechanics/molecular mechanics molecular dynamics simulations. *J. Chem. Theory Comput.* **2015**, *11*, 5826–5837.

[75] Curutchet, C.; Mennucci, B. Quantum chemical studies of light harvesting. *Chem. Rev.* **2017**, *117*, 294–343.

[76] Padula, D.; Lee, M. H.; Claridge, K.; Troisi, A. Chromophore-dependent intramolecular exciton–vibrational coupling in the FMO complex: Quantification and importance for exciton dynamics. *J. Phys. Chem. B* **2017**, *121*, 10026–10035.

[77] Maity, S.; Bold, B. M.; Prajapati, J. D.; Sokolov, M.; Kubar Kubař, T.; Elstner, M.; Kleinekathöfer, U. DFTB/MM molecular dynamics simulations of the FMO light-harvesting complex. *J. Phys. Chem. Lett.* **2020**, *11*, 8660–8667.

[78] Chaillet, M. L.; Lengauer, F.; Adolphs, J.; Müh, F.; Fokas, A. S.; Cole, D. J.; Chin, A. W.; Renger, T. Static disorder in excitation energies of the Fenna–Matthews–Olson protein: Structure-based theory meets experiment. *J. Phys. Chem. Lett.* **2020**, *11*, 10306–10314.

[79] Andreussi, O.; Prandi, I. G.; Campetella, M.; Prampolini, G.; Mennucci, B. Classical force fields tailored for QM applications: Is it really a feasible strategy? *J. Chem. Theory Comput.* **2017**, *13*, 4636–4648.

[80] Cacelli, I.; Prampolini, G. Parametrization and Validation of Intramolecular Force Fields Derived from DFT Calculations. *J. Chem. Theory and Comput.* **2007**, *3*, 1803–1817.

[81] 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.* **2013**, *15*, 3736–3751.

[82] Horton, J. T.; Allen, A. E.; Dodda, L. S.; Cole, D. J. QUBEKit: Automating the derivation of force field parameters from quantum mechanics. *J. Chem. Inf. Model* **2019**, *59*, 1366–1381.

[83] Tuckerman, M. *Statistical Mechanics: Theory and Molecular Simulation*; Oxford Graduate Texts; OUP Oxford, 2010.

[84] Brooks, C. L.; Case, D. A.; Plimpton, S.; Roux, B.; van der Spoel, D.; Tajkhorshid, E. Classical molecular dynamics. *J. Chem. Phys.* **2021**, *154*, 100401.

[85] Iftimie, R.; Minary, P.; Tuckerman, M. E. Ab initio molecular dynamics: Concepts, recent developments, and future trends. *P. Natl. Acad. Sci.* **2005**, *102*, 6654–6659.

[86] Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press, 2009.

[87] Marrink, S. J.; De Vries, A. H.; Mark, A. E. Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **2004**, *108*, 750–760.

[88] Baron, R.; Trzesniak, D.; de Vries, A. H.; Elsener, A.; Marrink, S. J.; van Gunsteren, W. F. Comparison of thermodynamic properties of coarse-grained and atomic-level simulation models. *ChemPhysChem* **2007**, *8*, 452–461.

[89] Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The power of coarse graining in biomolecular simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 225–248.

[90] Qi, Y.; Ingólfsson, H. I.; Cheng, X.; Lee, J.; Marrink, S. J.; Im, W. CHARMM-GUI martini maker for coarse-grained simulations with the martini force field. *J. Chem. Theory Comput.* **2015**, *11*, 4486–4494.

[91] Isralewitz, B.; Baudry, J.; Gullingsrud, J.; Kosztin, D.; Schulten, K. Steered molecular dynamics investigations of protein function. *J. Mol. Graph. Model.* **2001**, *19*, 13–25.

[92] Isralewitz, B.; Gao, M.; Schulten, K. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.* **2001**, *11*, 224–230.

[93] Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120*, 5946–5961.

[94] Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K. Steered Molecular Dynamics. Computational Molecular Dynamics: Challenges, Methods, Ideas. Berlin, Heidelberg, 1999; pp 39–65.

[95] Daidone, I.; Amadei, A. Essential dynamics: foundation and applications. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 762–770.

[96] Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

[97] Bernardi, R. C.; Melo, M. C.; Schulten, K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim. Biophys. Acta - Gen. Subj.* **2015**, *1850*, 872–877.

[98] Voter, A. F. Hyperdynamics: Accelerated molecular dynamics of infrequent events. *Phys. Rev. Lett.* **1997**, *78*, 3908.

[99] Bal, K. M.; Neyts, E. C. Merging metadynamics into hyperdynamics: accelerated molecular simulations reaching time scales from microseconds to seconds. *J. Chem. Theory Comput.* **2015**, *11*, 4545–4554.

[100] Laio, A.; Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.* **2008**, *71*, 126601.

[101] Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 826–843.

[102] So/rensen, M. R.; Voter, A. F. Temperature-accelerated dynamics for simulation of infrequent events. *J. Chem. Phys.* **2000**, *112*, 9599–9606.

[103] Bader, R. F. W. A quantum theory of molecular structure and its applications. *Chem. Rev.* **1991**, *91*, 893–928.

[104] Bader, R. F. W. The Quantum Mechanical Basis of Conceptual Chemistry. *Monatshefte für Chemie / Chemical Monthly* **2005**, *136*, 819–854.

[105] Popelier, P. L. *Atoms in Molecules: An Introduction*; Prentice Hall, London, 2000.

[106] Espinosa, E.; Alkorta, I.; Elguero, J.; Molins, E. From weak to strong interactions: A comprehensive analysis of the topological and energetic properties of the electron density distribution involving X-H$\cdots$F-Y systems. *J. Chem. Phys.* **2002**, *117*, 5529–5542.

[107] Grabowski, S. J. What Is the Covalency of Hydrogen Bonding? *Chem. Rev.* **2011**, *111*, 2597–2625.

[108] Becke, A. *The quantum theory of atoms in molecules: from solid state to DNA and drug design*; John Wiley & Sons, 2007.

[109] Reed, A. E.; Curtiss, L. A.; Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **1988**, *88*, 899–926.

[110] Weinhold, F.; Landis, C. R. *Discovering Chemistry with Natural Bond Orbitals*; Wiley-VCH, Hoboken NJ, 319pp, 2012.

[111] Glendening, E. D.; Landis, C. R.; Weinhold, F. Natural bond orbital methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 1–42.

[112] Weinhold, F.; Landis, C.; Glendening, E. What is NBO analysis and how is it useful? *Int. Rev. Phys. Chem.* **2016**, *35*, 399–440.

[113] Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.

[114] Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. NCIPLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theory Comput.* **2011**, *7*, 625–632.

[115] DiLabio, G. A.; Otero-de-la Roza, A. *Rev. Comput. Chem.*; John Wiley & Sons, Ltd, 2016; Chapter 1, pp 1–97.

[116] Abrahama, M. J.; Murtola, T.; Schulz, R.; Pálla, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1-2*, 19–25.

[117] Brehm, M.; Kirchner, B. TRAVIS - A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories. *J. Chem. Inf. Model* **2011**, *51*, 2007–2023.

[118] Brehm, M.; Thomas, M.; Gehrke, S.; Kirchner, B. TRAVIS—A free analyzer for trajectories from molecular simulation. *J. Chem. Phys.* **2020**, *152*, 164105.

[119] Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J.; Melo, M. N.; Seyler, S. L.; Domanski, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M., et al. MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. Proceedings of the 15th python in science conference. 2016; p 105.

[120] Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.

[121] Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

[122] Keith, T. AIMALL *(version 19.10.12)*. 2019; TK Gristmill Software, Overland Park KS, USA, aim.tkgristmill. com.

[123] Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Karafiloglou, P.; Landis, C. R.; Weinhold, F. NBO 7.0. 2018; Theoretical Chemistry Institute, University of Wisconsin, Madison, WI.

[124] Boto, R. A.; Peccati, F.; Laplaza, R.; Quan, C.; Carbone, A.; Piquemal, J.-P.; Maday, Y.; Contreras-Garcia, J. NCIPLOT4: A new step towards a fast quantification of noncovalent interactions. 2020; https://github.com/juliacontrerasgarcia/nciplot.

[125] Rojas-Valencia, N.; Gómez, S.; Guerra, D.; Restrepo, A. A detailed look at the bonding interactions in the microsolvation of monoatomic cations. *Phys. Chem. Chem. Phys.* **2020**, *22*, 13049–13061.

[126] Gómez, S.; Ramírez-Malule, H.; Cardona-G, W.; Osorio, E.; Restrepo, A. Double-ring epimerization in the biosynthesis of clavulanic acid. *J. Phys. Chem. A* **2020**, *124*, 9413–9426.

[127] Gómez, S. A.; Rojas-Valencia, N.; Gómez, S.; Lans, I.; Restrepo, A. Initial recognition and attachment of the Zika virus to host cells: A molecular dynamics and quantum interaction approach. *ChemBioChem* **2022**, e202200351.

[128] Wiberg, K. Application of the Pople-Santry-Segal CNDO method to the cyclopropylcarbinyl and cyclobutyl cation and to bicyclobutane. *Tetrahedron* **1968**, *24*, 1083 – 1096.

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

[130] Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; Van Gunsteren, W. F.; Mark, A. E. Peptide folding: when simulation meets experiment. *Angew. Chem. Int. Ed.* **1999**, *38*, 236–240.

[131] Kratz, E. G.; Walker, A. R.; Lagardère, L.; Lipparini, F.; Piquemal, J.-P.; Andrés Cisneros, G. LICHEM: A QM/MM program for simulations with multipolar and polarizable force fields. *J. Comput. Chem.* **2016**, *37*, 1019–1029.

[132] Lipparini, F.; Scalmani, G.; Lagardère, L.; Stamm, B.; Cancès, E.; Maday, Y.; Piquemal, J.-P.; Frisch, M. J.; Mennucci, B. Quantum, classical, and hybrid QM/MM calculations in solution: General implementation of the ddCOSMO linear scaling strategy. *J. Chem. Phys.* **2014**, *141*, 184108.

[133] Vennelakanti, V.; Nazemi, A.; Mehmood, R.; Steeves, A. H.; Kulik, H. J. Harder, better, faster, stronger: Large-scale QM and QM/MM for predictive modeling in enzymes and proteins. *Curr. Opin. Struct. Biol.* **2022**, *72*, 9–17.

[134] Mennucci, B.; Cammi, R. *Continuum Solvation Models in Chemical Physics: From Theory to Applications*; Wiley, New York, 2008.

[135] Cances, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

[136] Mennucci, B.; Cancès, E.; Tomasi, J. Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: Theoretical bases, computational implementation, and numerical applications. *J. Phys. Chem. B* **1997**, *101*, 10506–10517.

[137] Mennucci, B. Continuum Solvation Models: What Else Can We Learn from Them? *J. Phys. Chem. Lett.* **2010**, *1*, 1666–1674.

[138] Barone, V.; Cossi, M.; Tomasi, J. A new definition of cavities for the computation of solvation free energies by the polarizable continuum model. *J. Chem. Phys.* **1997**, *107*, 3210–3221.

[139] Barone, V.; Cossi, M. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem A* **1998**, *102*, 1995–2001.

[140] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem* **2003**, *24*, 669–681.

[141] Cossi, M.; Barone, V. Analytical second derivatives of the free energy in solution by polarizable continuum models. *J. Chem. Phys.* **1998**, *109*, 6246–6254.

[142] Cammi, R.; Cossi, M.; Tomasi, J. Analytical derivatives for molecular solutes. III. Hartree–Fock static polarizability and hyperpolarizabilities in the polarizable continuum model. *J. Chem. Phys.* **1996**, *104*, 4611–4620.

[143] Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. *J. Chem. Phys.* **2002**, *117*, 43–54.

[144] Scalmani, G.; Frisch, M. J. Continuous surface charge polarizable continuum models of solvation. I. General formalism. *J. Chem. Phys.* **2010**, *132*, 114110.

[145] Lipparini, F.; Scalmani, G.; Mennucci, B.; Cancès, E.; Caricato, M.; Frisch, M. J. A variational formulation of the polarizable continuum model. *J. Chem. Phys.* **2010**, *133*, 014106.

[146] Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Pairwise solute descreening of solute charges from a dielectric medium. *Chem. Phys. Lett* **1995**, *246*, 122 – 129.

[147] Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* **1996**, *100*, 19824–19839.

[148] Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Self-Consistent Reaction Field Model for Aqueous and Nonaqueous Solutions Based on Accurate Polarized Partial Charges. *J. Chem. Theory Comput.* **2007**, *3*, 2011–2033.

[149] Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. Extension of a Temperature-Dependent Aqueous Solvation Model to Compounds Containing Nitrogen, Fluorine, Chlorine, Bromine, and Sulfur. *J. Phys. Chem. B* **2008**, *112*, 3024–3039.

[150] Liu, J.; Kelly, C. P.; Goren, A. C.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; Zhan, C.-G. Free Energies of Solvation with Surface, Volume, and Local Electrostatic Effects and Atomic Surface Tensions to Represent the First Solvation Shell. *J. Chem. Theory Comput.* **2010**, *6*, 1109–1117.

[151] Mennucci, B. Modeling environment effects on spectroscopies through QM/classical models. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6583–6594.

[152] Frediani, L.; Cammi, R.; Corni, S.; Tomasi, J. A polarizable continuum model for molecules at diffuse interfaces. *J. Chem. Phys.* **2004**, *120*, 3893–3907.

[153] Lin, H.; Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2007**, *117*, 185–199.

[154] Lipparini, F.; Cappelli, C.; Barone, V. Linear response theory and electronic transition energies for a fully polarizable QM/classical hamiltonian. *J. Chem. Theory Comput.* **2012**, *8*, 4153–4165.

[155] Lipparini, F.; Egidi, F.; Cappelli, C.; Barone, V. The optical rotation of methyloxirane in aqueous solution: a never ending story? *J. Chem. Theory Comput.* **2013**, *9*, 1880–1884.

[156] Egidi, F.; Carnimeo, I.; Cappelli, C. Optical rotatory dispersion of methyloxirane in aqueous solution: assessing the performance of density functional theory in combination with a fully polarizable QM/MM/PCMapproach. *Opt. Mater. Express* **2015**, *5*, 196–209.

[157] Egidi, F.; Russo, R.; Carnimeo, I.; D'Urso, A.; Mancini, G.; Cappelli, C. The Electronic Circular Dichroism of Nicotine in Aqueous Solution: A Test Case for Continuum and Mixed Explicit-Continuum Solvation Approaches. *J. Phys. Chem. A* **2015**, *119*, 5396–5404.

[158] Giovannini, T.; Olszowka, M.; Cappelli, C. Effective Fully Polarizable QM/MM Approach To Model Vibrational Circular Dichroism Spectra of Systems in Aqueous Solution. *J. Chem. Theory Comput.* **2016**, *12*, 5483–5492.

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

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

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

[162] Loco, D.; Cupellini, L. Modeling the absorption lineshape of embedded systems from molecular dynamics: A tutorial review. *Int. J. Quantum Chem.* **2019**, *119*, e25726.

[163] Jurinovich, S.; Curutchet, C.; Mennucci, B. The Fenna–Matthews–Olson Protein Revisited: A Fully Polarizable (TD) DFT/MM Description. *ChemPhysChem* **2014**, *15*, 3194–3204.

[164] Morzan, U. N.; Alonso de Armino, D. J.; Foglia, N. O.; Ramirez, F.; Gonzalez Lebrero, M. C.; Scherlis, D. A.; Estrin, D. A. Spectroscopy in complex environments from QM–MM simulations. *Chem. Rev.* **2018**, *118*, 4071–4113.

[165] Bakowies, D.; Thiel, W. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

[166] Ambrosetti, M.; Skoko, S.; Giovannini, T.; Cappelli, C. Quantum Mechanics/Fluctuating Charge Protocol to Compute Solvatochromic Shifts. *J. Chem. Theory Comput.* **2021**, *17*, 7146–7156.

[167] Lamoureux, G.; MacKerell Jr, A. D.; Roux, B. A simple polarizable model of water based on classical Drude oscillators. *J. Chem. Phys.* **2003**, *119*, 5185–5197.

[168] Lamoureux, G.; Roux, B. Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.* **2003**, *119*, 3025–3039.

[169] Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. Determination of electrostatic parameters for a polarizable force field based on the classical Drude oscillator. *J. Chem. Theory Comput.* **2005**, *1*, 153–168.

[170] Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell Jr, A. D. An empirical polarizable force field based on the classical drude oscillator model: development history and recent applications. *Chem. Rev.* **2016**, *116*, 4983–5013.

[171] Geerke, D.; van Gunsteren, W. F. The performance of non-polarizable and polarizable force-field parameter sets for ethylene glycol in molecular dynamics simulations of the pure liquid and its aqueous mixtures. *Mol. Phys.* **2007**, *105*, 1861–1881.

[172] Boulanger, E.; Thiel, W. Solvent boundary potentials for hybrid QM/MM computations using classical drude oscillators: a fully polarizable model. *J. Chem. Theory Comput.* **2012**, *8*, 4527–4538.

[173] Geerke, D. P.; Thiel, S.; Thiel, W.; van Gunsteren, W. F. Combined QM/MM molecular dynamics study on a condensed-phase SN2 reaction at nitrogen: The effect of explicitly including solvent polarization. *J. Chem. Theory Comput.* **2007**, *3*, 1499–1509.

[174] Boulanger, E.; Thiel, W. Toward QM/MM simulation of enzymatic reactions with the drude oscillator polarizable force field. *J. Chem. Theory Comput.* **2014**, *10*, 1795–1809.

[175] Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.* **1981**, *59*, 341–350.

[176] Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. Excitation energies in solution: the fully polarizable QM/MM/PCM method. *J. Phys. Chem. B* **2011**, *115*, 3027–3037.

[177] Caprasecca, S.; Jurinovich, S.; Viani, L.; Curutchet, C.; Mennucci, B. Geometry optimization in polarizable QM/MM models: the induced dipole formulation. *J. Chem. Theory Comput.* **2014**, *10*, 1588–1598.

[178] Donati, G.; Wildman, A.; Caprasecca, S.; Lingerfelt, D. B.; Lipparini, F.; Mennucci, B.; Li, X. Coupling Real-Time Time-Dependent Density Functional Theory with Polarizable Force Field. *J. Phys. Chem. Lett.* **2017**, *8*, 5283–5289.

[179] Jurinovich, S.; Pescitelli, G.; Di Bari, L.; Mennucci, B. A TDDFT/MMPol/PCM model for the simulation of exciton-coupled circular dichroism spectra. *Phys. Chem. Chem. Phys.* **2014**, *16*, 16407–16418.

[180] Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. Electronic energy transfer in condensed phase studied by a polarizable QM/MM model. *J. Chem. Theory Comput.* **2009**, *5*, 1838–1848.

[181] Curutchet, C.; Kongsted, J.; Muñoz-Losa, A.; Hossein-Nejad, H.; Scholes, G. D.; Mennucci, B. Photosynthetic Light-Harvesting Is Tuned by the Heterogeneous Polarizable Environment of the Protein. *J. Am. Chem. Soc.* **2011**, *133*, 3078–3084.

[182] Loco, D.; Jurinovich, S.; Cupellini, L.; Menger, M. F.; Mennucci, B. The modeling of the absorption lineshape for embedded molecules through a polarizable QM/MM approach. *Photochem. Photobiol. Sci.* **2018**, *17*, 552–560.

[183] Caprasecca, S.; Cupellini, L.; Jurinovich, S.; Loco, D.; Lipparini, F.; Mennucci, B. A polarizable QM/MM description of environment effects on NMR shieldings: from solvated molecules to pigment–protein complexes. *Theor. Chem. Acc.* **2018**, *137*, 84.

[184] Ren, S.; Lipparini, F.; Mennucci, B.; Caricato, M. Coupled cluster theory with induced dipole polarizable embedding for ground and excited states. *J. Chem. Theory Comput.* **2019**, *15*, 4485–4496.

[185] Ren, P.; Ponder, J. W. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

[186] Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.

[187] Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable atomic multipole-based AMOEBA force field for proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046–4063.

[188] Dziedzic, J.; Mao, Y.; Shao, Y.; Ponder, J.; Head-Gordon, T.; Head-Gordon, M.; Skylaris, C.-K. TINKTEP: A fully self-consistent, mutually polarizable QM/MM approach based on the AMOEBA force field. *J. Chem. Phys.* **2016**, *145*, 124106.

[189] Loco, D.; Buda, F.; Lugtenburg, J.; Mennucci, B. The Dynamic Origin of Color Tuning in Proteins Revealed by a Carotenoid Pigment. *J. Phys. Chem. Lett.* **2018**, *9*, 2404–2410.

[190] Mao, Y.; Shao, Y.; Dziedzic, J.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. Performance of the AMOEBA water model in the vicinity of QM solutes: A diagnosis using energy decomposition analysis. *J. Chem. Theory Comput.* **2017**, *13*, 1963–1979.

[191] Nottoli, M.; Mennucci, B.; Lipparini, F. Excited state Born–Oppenheimer molecular dynamics through coupling between time dependent DFT and AMOEBA. *Phys. Chem. Chem. Phys.* **2020**, *22*, 19532–19541.

[192] Marini, A.; Muñoz-Losa, A.; Biancardi, A.; Mennucci, B. What is Solvatochromism? *J. Phys. Chem. B* **2010**, *114*, 17128–17135.

[193] Jacobson, L. D.; Herbert, J. M. A one-electron model for the aqueous electron that includes many-body electron-water polarization: Bulk equilibrium structure, vertical electron binding energy, and optical absorption spectrum. *J. Chem. Phys.* **2010**, *133*, 154506.

[194] Jensen, L.; van Duijnen, P. T.; Snijders, J. G. A discrete solvent reaction field model for calculating frequency-dependent hyperpolarizabilities of molecules in solution. *J. Chem. Phys.* **2003**, *119*, 12998–13006.

[195] Poulsen, T. D.; Ogilby, P. R.; Mikkelsen, K. V. Linear response properties for solvated molecules described by a combined multiconfigurational self-consistent-field/molecular mechanics model. *J. Chem. Phys.* **2002**, *116*, 3730–3738.

[196] Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. Linear response functions for coupled cluster/molecular mechanics including polarization interactions. *J. Chem. Phys.* **2003**, *118*, 1620–1633.

[197] Jensen, L.; Swart, M.; van Duijnen, P. T. Microscopic and macroscopic polarization within a combined quantum mechanics and molecular mechanics model. *J. Chem. Phys.* **2005**, *122*, 034103.

[198] Jensen, L.; Van Duijnen, P. T. Refractive index and third-order nonlinear susceptibility of C60 in the condensed phase calculated with the discrete solvent reaction field model. *Int. J. Quantum Chem.* **2005**, *102*, 612–619.

[199] Jensen, L.; van Duijnen, P. T. The first hyperpolarizability of p-nitroaniline in 1, 4-dioxane: A quantum mechanical/molecular mechanics study. *J. Chem. Phys.* **2005**, *123*, 074307.

[200] Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. Polarizable density embedding: A new QM/QM/MM-based computational strategy. *J. Phys. Chem. A* **2015**, *119*, 5344–5355.

[201] Nåbo, L. J.; Olsen, J. M. H.; Holmgaard List, N.; Solanko, L. M.; Wüstner, D.; Kongsted, J. Embedding beyond electrostatics—The role of wave function confinement. *J. Chem. Phys.* **2016**, *145*, 104102.

[202] Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. Polarizable Density Embedding: A Solution to the Electron Spill-Out Problem in Multiscale Modeling. *J. Phys. Chem. Lett.* **2017**, *8*, 5949–5958.

[203] Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation methods: A route to accurate calculations on large systems. *Chem. Rev.* **2012**, *112*, 632–672.

[204] Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* **1994**, *101*, 6141–6156.

[205] Rick, S. W.; Stuart, S. J.; Bader, J. S.; Berne, B. Fluctuating charge force fields for aqueous solutions. *J. Mol. Liq.* **1995**, *65*, 31–40.

[206] Rick, S. W.; Berne, B. J. Dynamical Fluctuating Charge Force Fields: The Aqueous Solvation of Amides. *J. Am. Chem. Soc.* **1996**, *118*, 672–679.

[207] Lipparini, F.; Cappelli, C.; Barone, V. A gauge invariant multiscale approach to magnetic spectroscopies in condensed phase: General three-layer model, computational implementation and pilot applications. *J. Chem. Phys.* **2013**, *138*, 234108.

[208] 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.* **2012**, *8*, 4270–4278.

[209] Carnimeo, I.; Cappelli, C.; Barone, V. Analytical gradients for MP2, double hybrid functionals, and TD-DFT with polarizable embedding described by fluctuating charges. *J. Comput. Chem.* **2015**, *36*, 2271–2290.

[210] Egidi, F.; Lo Gerfo, G.; Macchiagodena, M.; Cappelli, C. On the nature of charge-transfer excitations for molecules in aqueous solution: a polarizable QM/MM study. *Theor. Chem. Acc.* **2018**, *137*, 82.

[211] Di Remigio, R.; Giovannini, T.; Ambrosetti, M.; Cappelli, C.; Frediani, L. Fully Polarizable QM/Fluctuating Charge Approach to Two-Photon Absorption of Aqueous Solutions. *J. Chem. Theory Comput.* **2019**, *15*, 4056–4068.

[212] Giovannini, T.; Ambrosetti, M.; Cappelli, C. Quantum Confinement Effects on Solvatochromic Shifts of Molecular Solutes. *J. Phys. Chem. Lett.* **2019**, *10*, 5823–5829.

[213] Giovannini, T.; Grazioli, L.; Ambrosetti, M.; Cappelli, C. Calculation of IR spectra with a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles. *J. Chem. Theory Comput.* **2019**, *15*, 5495–5507.

[214] Giovannini, T.; Riso, R. R.; Ambrosetti, M.; Puglisi, A.; Cappelli, C. Electronic transitions for a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles: linear and corrected linear response regimes. *J. Chem. Phys.* **2019**, *151*, 174104.

[215] Marrazzini, G.; Giovannini, T.; Egidi, F.; Cappelli, C. Calculation of linear and non-linear electric response properties of systems in aqueous solution: A polarizable quantum/classical approach with quantum repulsion effects. *J. Chem. Theory Comput.* **2020**, *16*, 6993–7004.

[216] Mortier, W. J.; Van Genechten, K.; Gasteiger, J. Electronegativity equalization: application and parametrization. *J. Am. Chem. Soc.* **1985**, *107*, 829–835.

[217] Rappe, A. K.; Goddard III, W. A. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem* **1991**, *95*, 3358–3363.

[218] Sanderson, R. An interpretation of bond lengths and a classification of bonds. *Science* **1951**, *114*, 670–672.

[219] Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual density functional theory. *Chem. Rev.* **2003**, *103*, 1793–1874.

[220] Moens, J.; Geerlings, P.; Roos, G. A conceptual DFT approach for the evaluation and interpretation of redox potentials. *Chem. Eur. J.* **2007**, *13*, 8174–8184.

[221] Geerlings, P.; De Proft, F. Conceptual DFT: the chemical relevance of higher response functions. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3028–3042.

[222] Geerlings, P.; Ayers, P. W.; Toro-Labbé, A.; Chattaraj, P. K.; De Proft, F. The Woodward–Hoffmann rules reinterpreted by conceptual density functional theory. *Acc. Chem. Res.* **2012**, *45*, 683–695.

[223] Geerlings, P.; Fias, S.; Boisdenghien, Z.; De Proft, F. Conceptual DFT: Chemistry from the linear response function. *Chem. Soc. Rev.* **2014**, *43*, 4989–5008.

[224] Geerlings, P.; Chamorro, E.; Chattaraj, P. K.; De Proft, F.; Gázquez, J. L.; Liu, S.; Morell, C.; Toro-Labbé, A.; Vela, A.; Ayers, P. Conceptual density functional theory: status, prospects, issues. *Theor. Chem. Acc.* **2020**, *139*, 1–18.

[225] Mayer, A. Formulation in terms of normalized propagators of a charge-dipole model enabling the calculation of the polarization properties of fullerenes and carbon nanotubes. *Phys. Rev. B* **2007**, *75*, 045407.

[226] Lennard-Jones, J. E. Cohesion. *Proc. Phys. Soc.* **1931**, *43*, 461.

[227] Gokcan, H.; Kratz, E. G.; Darden, T. A.; Piquemal, J.-P.; Cisneros, G. A. QM/MM Simulations with the Gaussian Electrostatic Model, A Density-Based Polarizable Potential. *J. Phys. Chem. Lett.* **2018**, *9*, 3062–3067.

[228] Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. Intermolecular electrostatic energies using density fitting. *J. Chem. Phys.* **2005**, *123*, 044109.

[229] Piquemal, J.-P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. Towards a force field based on density fitting. *J. Chem. Phys.* **2006**, *124*, 104101.

[230] Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. Generalization of the Gaussian electrostatic model: Extension to arbitrary angular momentum, distributed multipoles, and speedup with reciprocal space methods. *J. Chem. Phys.* **2006**, *125*, 184101.

[231] Tkatchenko, A.; Scheffler, M. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **2009**, *102*, 073005.

[232] Tkatchenko, A.; Romaner, L.; Hofmann, O. T.; Zojer, E.; Ambrosch-Draxl, C.; Scheffler, M. Van der Waals interactions between organic adsorbates and at organic/inorganic interfaces. *MRS bulletin* **2010**, *35*, 435–442.

[233] Tkatchenko, A.; DiStasio Jr, R. A.; Car, R.; Scheffler, M. Accurate and efficient method for many-body van der Waals interactions. *Phys. Rev. Lett.* **2012**, *108*, 236402.

[234] Hermann, J.; DiStasio, R. A.; Tkatchenko, A. First-Principles Models for van der Waals Interactions in Molecules and Materials: Concepts, Theory, and Applications. *Chem. Rev.* **2017**, *117*, 4714–4758.

[235] Curutchet, C.; Cupellini, L.; Kongsted, J.; Corni, S.; Frediani, L.; Steindal, A. H.; Guido, C. A.; Scalmani, G.; Mennucci, B. Density-Dependent Formulation of Dispersion–Repulsion Interactions in Hybrid Multiscale Quantum/Molecular Mechanics (QM/MM) Models. *J. Chem. Theory Comput.* **2018**, *14*, 1671–1681.

[236] Giovannini, T.; Lafiosca, P.; Cappelli, C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. *J. Chem. Theory Comput.* **2017**, *13*, 4854–4870.

[237] Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. Effective yet Reliable Computation of Hyperfine Coupling Constants in Solution by a QM/MM Approach: Interplay Between Electrostatics and Non-electrostatic Effects. *J. Chem. Phys.* **2019**, *150*, 124102.

[238] McWeeny, R. *Methods of molecular quantum mechanics*; Academic press: London, 1992.

[239] Amovilli, C.; McWeeny, R. A matrix partitioning approach to the calculation of intermolecular potentials. General theory and some examples. *Chem. Phys.* **1990**, *140*, 343–361.

[240] Kitaura, K.; Morokuma, K. A new energy decomposition scheme for molecular interactions within the Hartree-Fock approximation. *Int. J. Quantum Chem.* **1976**, *10*, 325–340.

[241] Morokuma, K.; Kitaura, K. *Chemical applications of atomic and molecular electrostatic potentials*; Springer, 1981; pp 215–242.

[242] Su, P.; Li, H. Energy decomposition analysis of covalent bonds and intermolecular interactions. *J. Chem.Phys.* **2009**, *131*, 014102.

[243] Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. Accurate first principles model potentials for intermolecular interactions. *Annu. Rev. Phys. Chem.* **2013**, *64*, 553–578.

[244] Sun, Q.; Chan, G. K.-L. Quantum embedding theories. *Acc. Chem. Res.* **2016**, *49*, 2705–2712.

[245] Knizia, G.; Chan, G. K.-L. Density matrix embedding: A strong-coupling quantum embedding theory. *J. Chem. Theory. Comput.* **2013**, *9*, 1428–1432.

[246] Chulhai, D. V.; Goodpaster, J. D. Projection-based correlated wave function in density functional theory embedding for periodic systems. *J. Chem. Theory Comput.* **2018**, *14*, 1928–1942.

[247] Chulhai, D. V.; Goodpaster, J. D. Improved accuracy and efficiency in quantum embedding through absolute localization. *J. Chem. Theory Comput.* **2017**, *13*, 1503–1508.

[248] Wen, X.; Graham, D. S.; Chulhai, D. V.; Goodpaster, J. D. Absolutely Localized Projection-Based Embedding for Excited States. *J. Chem. Theory Comput.* **2020**, *16*, 385–398.

[249] Ding, F.; Manby, F. R.; Miller III, T. F. Embedded mean-field theory with block-orthogonalized partitioning. *J. Chem. Theory Comput.* **2017**, *13*, 1605–1615.

[250] Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller III, T. F. Density functional theory embedding for correlated wavefunctions: Improved methods for open-shell systems and transition metal complexes. *J. Chem. Phys.* **2012**, *137*, 224113.

[251] Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller III, T. F. Accurate and systematically improvable density functional theory embedding for correlated wavefunctions. *J. Chem. Phys.* **2014**, *140*, 18A507.

[252] Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller III, T. F. A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.

[253] Goodpaster, J. D.; Ananth, N.; Manby, F. R.; Miller III, T. F. Exact nonadditive kinetic potentials for embedded density functional theory. *J. Chem. Phys.* **2010**, *133*, 084103.

[254] Zhang, K.; Ren, S.; Caricato, M. Multi-state QM/QM Extrapolation of UV/Vis Absorption Spectra with Point Charge Embedding. *J. Chem. Theory Comput.* **2020**, *16*, 4361–4372.

[255] Ramos, P.; Papadakis, M.; Pavanello, M. Performance of frozen density embedding for modeling hole transfer reactions. *J. Phys. Chem. B* **2015**, *119*, 7541–7557.

[256] Pavanello, M.; Neugebauer, J. Modelling charge transfer reactions with the frozen density embedding formalism. *J. Chem. Phys.* **2011**, *135*, 234103.

[257] Bennie, S. J.; Curchod, B. F.; Manby, F. R.; Glowacki, D. R. Pushing the limits of EOM-CCSD with projector-based embedding for excitation energies. *J. Phys. Chem. Lett.* **2017**, *8*, 5559–5565.

[258] Lee, S. J.; Welborn, M.; Manby, F. R.; Miller III, T. F. Projection-based wavefunction-in-DFT embedding. *Acc. Chem. Res.* **2019**, *52*, 1359–1368.

[259] Long, D. *The Raman Effect*; John Wiley & Sons, Ltd, 2002; Chapter 24, pp 471–495.

[260] Long, D. *The Raman Effect*; John Wiley & Sons, Ltd, 2002; Chapter 5, pp 85–152.

[261] Chowdhury, J. In *Molecular and Laser Spectroscopy*; Gupta, V., Ed.; Elsevier, 2018; pp 147–164.

[262] Norman, P.; Ruud, K.; Saue, T. *Principles and Practices of Molecular Properties: Theory, Modeling, and Simulations*; Wiley, 2018.

[263] Ozaki, Y.; Wójcik, M.; Popp, J. *Molecular Spectroscopy, 2 Volume Set: A Quantum Chemistry Approach*; Wiley, 2019.

[264] Guthmuller, J. Comparison of simplified sum-over-state expressions to calculate resonance Raman intensities including Franck-Condon and Herzberg-Teller effects. *J. Chem. Phys.* **2016**, *144*, 064106.

[265] Guthmuller, J. *Molecular Spectroscopy: A Quantum Chemistry Approach*; John Wiley & Sons, Ltd, 2019; Chapter 17, pp 497–536.

[266] Xu, Q.; Aranda, D.; Yaghoubi Jouybari, M.; Liu, Y.; Wang, M.; Cerezo, J.; Improta, R.; Santoro, F. Nonadiabatic vibrational resonance Raman spectra from quantum dynamics propagations with LVC models. Application to thymine. *J. Phys. Chem. A* **2022**, *126*, 7468–7479.

[267] Heller, E. J.; Sundberg, R.; Tannor, D. Simple aspects of Raman scattering. *J. Phys. Chem.* **1982**, *86*, 1822–1833.

[268] Lee, S.-Y.; Heller, E. J. Time-dependent theory of Raman scattering. *J. Chem. Phys.* **1979**, *71*, 4777–4788.

[269] Baiardi, A.; Bloino, J.; Barone, V. A general time-dependent route to Resonance-Raman spectroscopy including Franck-Condon, Herzberg-Teller and Duschinsky effects. *J. Chem. Phys.* **2014**, *141*, 114108.

[270] Petrenko, T.; Neese, F. Analysis and prediction of absorption band shapes, fluorescence band shapes, resonance Raman intensities, and excitation profiles using the time-dependent theory of electronic spectroscopy. *J. Chem. Phys.* **2007**, *127*, 164319.

[271] Petrenko, T.; Neese, F. Efficient and automatic calculation of optical band shapes and resonance Raman spectra for larger molecules within the independent mode displaced harmonic oscillator model. *J. Chem. Phys.* **2012**, *137*, 234107.

[272] Horvath, R.; Gordon, K. C. Understanding excited-state structure in metal polypyridyl complexes using resonance Raman excitation profiles, time-resolved resonance Raman spectroscopy and density functional theory. *Coord. Chem. Rev.* **2010**, *254*, 2505–2518.

[273] Avila, F. F.; Barone, V.; Cappelli, C.; Santoro, F. Duschinsky, Herzberg-Teller, and Multiple Electronic Resonance Interferential Effects in Resonance Raman Spectra and Excitation Profiles. The Case of Pyrene. *J. Chem. Theory Comput.* **2013**, *9*, 3597–3611.

[274] Jensen, L.; Zhao, L.; Autschbach, J.; Schatz, G. Theory and method for calculating resonance Raman scattering from resonance polarizability derivatives. *J. Chem. Phys.* **2005**, *123*, 174110.

[275] Mohammed, A.; Ågren, H.; Norman, P. Time-dependent density functional theory for resonant properties: resonance enhanced Raman scattering from the complex electric-dipole polarizability. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4539–4548.

[276] Mohammed, A.; Ågren, H.; Norman, P. Resonance enhanced Raman scattering from the complex electric-dipole polarizability: A theoretical study on N2. *Chem. Phys. Lett.* **2009**, *468*, 119–123.

[277] Jensen, L.; Schatz, G. C. Resonance Raman scattering of rhodamine 6G as calculated using time-dependent density functional theory. *J. Phys. Chem. A* **2006**, *110*, 5973–5977.

[278] Al-Saidi, W.; Asher, S. A.; Norman, P. Resonance Raman spectra of TNT and RDX using vibronic theory, excited-state gradient, and complex polarizability approximations. *J. Phys. Chem. A* **2012**, *116*, 7862–7872.

[279] Caricato, M.; Mennucci, B.; Tomasi, J.; Ingrosso, F.; Cammi, R.; Corni, S.; Scalmani, G. Formation and relaxation of excited states in solution: A new time dependent polarizable continuum model based on time dependent density functional theory. *J. Chem. Phys.* **2006**, *124*, 124520.

[280] Rinkevicius, Z.; Li, X.; Sandberg, J. A.; Mikkelsen, K. V.; Ågren, H. A hybrid density functional theory/molecular mechanics approach for linear response properties in heterogeneous environments. *J. Chem. Theory Comput.* **2014**, *10*, 989–1003.

[281] Olsen, J. M.; Aidas, K.; Kongsted, J. Excited states in solution through polarizable embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721–3734.

[282] Schwabe, T. General theory for environmental effects on (vertical) electronic excitation energies. *J. Chem. Phys.* **2016**, *145*, 154105.

[283] Schröder, H.; Schwabe, T. Corrected Polarizable Embedding: Improving the Induction Contribution to Perichromism for Linear Response Theory. *J. Chem. Theory Comput.* **2018**, *14*, 833–842.

[284] Casida, M. E. In *Recent Advances in Density Functional Methods Part I*; Chong, D. P., Ed.; World Scientific, Singapore, 1995; pp 155–192.

[285] Guido, C. A.; Chrayteh, A.; Scalmani, G.; Mennucci, B.; Jacquemin, D. Simple Protocol for Capturing Both Linear-Response and State-Specific Effects in Excited-State Calculations with Continuum Solvation Models. *J. Chem. Theory Comput.* **2021**, *17*, 5155–5164.

[286] Barron, L. D. *Molecular light scattering and optical activity*; Cambridge University Press, 2004.

[287] Tannor, D. J.; Heller, E. J. Polyatomic Raman scattering for general harmonic potentials. *J. Chem. Phys.* **1982**, *77*, 202–218.

[288] Myers, A. B. Resonance Raman intensity analysis of excited-state dynamics. *Accounts of chemical research* **1997**, *30*, 519–527.

[289] Tripathi, G. Electronic structure of para aminophenoxyl radical in water. *J. Chem. Phys.* **2003**, *118*, 1378–1391.

[290] Guthmuller, J.; Champagne, B. Time dependent density functional theory investigation of the resonance Raman properties of the julolidinemalononitrile push-pull chromophore in various solvents. *J. Chem. Phys.* **2007**, *127*, 164507.

[291] Yarasi, S.; Brost, P.; Loppnow, G. R. Initial excited-state structural dynamics of thymine are coincident with the expected photochemical dynamics. *J. Phys. Chem. A* **2007**, *111*, 5130–5135.

[292] Yarasi, S.; Ng, S.; Loppnow, G. R. Initial excited-state structural dynamics of uracil from resonance Raman spectroscopy are different from those of thymine (5-methyluracil). *J. Phys. Chem. B* **2009**, *113*, 14336–14342.

[293] Barclay, M. S.; Quincy, T. J.; Williams-Young, D. B.; Caricato, M.; Elles, C. G. Accurate Assignments of Excited-State Resonance Raman Spectra: A Benchmark Study Combining Experiment and Theory. *J. Phys. Chem. A* **2017**, *121*, 7937–7946.

[294] Macaluso, V.; Cupellini, L.; Salvadori, G.; Lipparini, F.; Mennucci, B. Elucidating the role of structural fluctuations, and intermolecular and vibronic interactions in the spectroscopic response of a bacteriophytochrome. *Phys. Chem. Chem. Phys.* **2020**, *22*, 8585–8594.

[295] Myers Kelley, A. Resonance Raman and resonance hyper-Raman intensities: Structure and dynamics of molecular excited states in solution. *J. Phys. Chem. A* **2008**, *112*, 11975–11991.

[296] Raman, C. V.; Krishnan, K. S. A New Type of Secondary Radiation. *Nature* **1928**, *121*, 501–502.

[297] Smekal, A. Zur quantentheorie der dispersion. *Naturwissenschaften* **1923**, *11*, 873–875.

[298] O'Donell, D. Seminar: Scattering of light: Raman Spectroscopy; Informal P-Chem Review. 2009; `https://www.slideshare.net/dodonne1/intro-to-raman-scattering`.

[299] Long, D. *The Raman Effect*; John Wiley & Sons, Ltd, 2002; Chapter 4, pp 49–84.

[300] Dirac, P. A. M.; Fowler, R. H. The quantum theory of dispersion. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1927**, *114*, 710–728.

[301] Kiefer, J. Recent advances in the characterization of gaseous and liquid fuels by vibrational spectroscopy. *Energies* **2015**, *8*, 3165–3197.

[302] Cornaton, Y.; Ringholm, M.; Louant, O.; Ruud, K. Analytic calculations of anharmonic infrared and Raman vibrational spectra. *Phys. Chem. Chem. Phys.* **2016**, *18*, 4201–4215.

[303] Xiong, K. UV Resonance Raman Spectroscopy: A Highly Sensitive, Selective and Fast Technique for Environmental Analysis. *J Environ Anal Chem* **2014**, *2*, e107.

[304] Lopez-Pena, I.; Leigh, B. S.; Schlamadinger, D. E.; Kim, J. E. Insights into protein structure and dynamics by ultraviolet and visible resonance Raman spectroscopy. *Biochemistry* **2015**, *54*, 4770–4783.

[305] Long, D. *The Raman Effect*; John Wiley & Sons, Ltd, 2002; Chapter 7, pp 221–270.

[306] Duschinsky, F. On the interpretation of electronic spectra of polyatomic molecules. *Acta Physicochim. URSS* **1937**, *7*, 551.

[307] Rappoport, D.; Shim, S.; Aspuru-Guzik, A. Simplified sum-over-states approach for predicting resonance raman spectra. Application to nucleic acid bases. *J. Phys. Chem. Lett.* **2011**, *2*, 1254–1260.