

Fully Atomistic Modeling in Computational Spectroscopy: Tryptophan in Aqueous Solution as a Test Case

Published as part of *The Journal of Physical Chemistry A* special issue “Forty Years of Response Function Theory”.

Marco Trinari, Chiara Sepali, Tommaso Giovannini, and Chiara Cappelli*



Cite This: <https://doi.org/10.1021/acs.jpca.5c04511>



Read Online

ACCESS |



Metrics & More

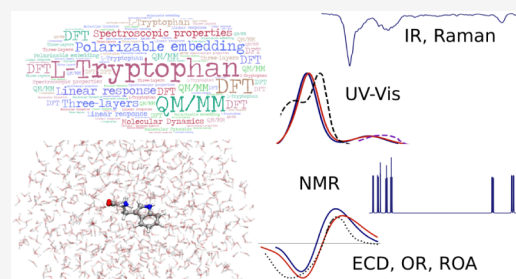


Article Recommendations



Supporting Information

ABSTRACT: We present a multiscale computational protocol for the simulation of a wide range of spectroscopic properties—electronic, magnetic, and vibrational—of zwitterionic L-tryptophan in aqueous solution. The approach combines density functional theory (DFT) for the solute with polarizable embedding models (QM/FQ and QM/FQF μ) for the solvent, and incorporates extensive conformational sampling via classical molecular dynamics. The protocol successfully reproduces UV–vis and ECD spectra, including the characteristic $S_0 \rightarrow S_1$ transition and chiroptical features, and captures the negative optical rotation at the sodium D-line with good agreement to experiment. NMR chemical shifts and spin–spin couplings are also computed, and a hybrid QM/FDE/FQF μ scheme is employed to improve the description of solvent-sensitive nuclei. Vibrational spectra (IR, Raman, and ROA) are calculated and analyzed, with all models yielding results consistent with experimental data where available. The comparison between QM/FQ and QM/FQF μ highlights the importance of accurate solvent treatment, especially for chiroptical and magnetic properties.



1. INTRODUCTION

Since its foundations, the field of computational chemistry has expanded in scope, and thanks to both technological and methodological advances, it has become an indispensable tool in most other areas of chemical research.¹ Among its many applications, a particularly important one is the development of computational methods for the modeling of spectroscopic properties of molecular systems.² Spectroscopic techniques provide direct access to molecular-level information, offering invaluable insights into structure, dynamics, and chemical interactions. As a result, computational spectroscopy has evolved into a synergistic counterpart to experiment, enabling the interpretation of spectral features in terms of molecular motifs and physical mechanisms.^{3–7}

Computational spectroscopy generally relies on a quantum mechanical (QM) description of the chemical system. In the gas phase, a variety of well-established approaches have been developed, ranging from derivative-based response methods² to real-time propagation techniques,^{8–10} that can predict spectroscopic observables across different regimes with varying degrees of accuracy and computational demand. However, to realistically describe most chemical and biochemical processes, one must consider the condensed phase, with aqueous solution being the most prominent environment. Here, the unique physical and chemical properties of water, such as its polarity, hydrogen bonding capabilities, and dielectric response, can

profoundly influence the spectroscopic signatures of solvated species.¹¹ Consequently, gas-phase models often fail unless an explicit treatment of the surrounding medium is considered.

To address this challenge, multiscale models have been developed in which the solute is treated quantum mechanically, and the solvent is represented using classical or semiclassical descriptions.^{11–17} Among these, polarizable embedding schemes stand out for their ability to capture the mutual polarization between solute and solvent.^{11,18–23} Polarizable QM/Molecular Mechanics (MM) embedding techniques belong to the family of “focused models”,¹¹ the latter represent a targeted approach in which computational accuracy is concentrated on the portion of the system directly involved in the chemical processes of interest, while the rest is treated at a more affordable level of theory. These models, in particular, allow for an accurate description of local electronic properties (such as electronic transitions, excited states, or chemical reactivity) by applying a refined quantum mechanical treatment to the active region (usually small), while

Received: June 30, 2025

Revised: October 16, 2025

Accepted: October 16, 2025

maintaining a realistic representation of the surrounding environment through polarizable embedding.^{24–26} It should be noted that the use of polarizable embedding approaches in the context of ultrafast excited-state dynamics is not trivial and may give rise to an unphysical description of time-resolved processes and signals.¹⁸ In the present work, however, we target the modeling of static spectral signals, for which the linear-response polarizable embedding adopted here is robust, as already demonstrated in several previous studies.^{27–31} Focused models are especially useful for simulating complex systems, such as solvated systems, where a full quantum mechanical treatment would be computationally prohibitive.

In this work, we employ two polarizable QM/MM models: the QM/Fluctuating Charges (QM/FQ)^{24,32} and QM/Fluctuating Charges and Fluctuating Dipoles (QM/FQF μ) methods,^{33,34} together with their extension based on frozen density embedding (FDE),^{35,36} QM/FDE/FQ(F μ).^{37,38} These methods enable the accurate and self-consistent inclusion of environmental effects by treating the solvent as a polarizable medium that dynamically responds to the solute's quantum mechanical charge distribution.³⁹ To demonstrate the versatility and effectiveness of these approaches, we apply them to the modeling of a broad range of spectroscopies—including UV/vis absorption,²⁷ vibrational (IR^{28,40} and Raman^{29,41,42}), electronic circular dichroism (ECD),^{37,43,44} optical rotation (OR),^{45–47} Raman optical activity (ROA),^{29,41,48} and nuclear magnetic resonance (NMR).^{38,49} Although individual methods have previously been employed—each to varying extents—to model specific spectroscopic properties,^{11,27,28,37,38,41,50,51} the present work introduces a comprehensive and unified framework capable of addressing a broad range of observables within a consistent multiscale strategy. By combining polarizable QM/(QM')/MM methodologies with molecular dynamics simulations in a single protocol, we accurately simulate electronic, vibrational, and chiroptical properties, while also highlighting the differences and specific challenges that arise when the framework is applied to diverse spectral signals. Other strategies for including solvation effects on the calculations of spectroscopic properties have been widely employed in the literature,⁵² ranging from continuum models (PCM/CPCM)^{53,54} to explicit solvent molecular dynamics based on MM non-polarizable fixed-charge force fields such as TIP3P,⁵⁴ or hybrid cluster-continuum approaches where selected solvent molecules are treated at the QM or MM level while the bulk is eventually described as a dielectric medium.^{53–60} These approaches have proven effective and computationally efficient,⁶¹ but they generally lack an explicit description of mutual solute–solvent polarization. In contrast, the polarizable QM/FQ(F μ) and QM/FDE/FQ(F μ) models adopted here explicitly include such effects in a self-consistent manner, offering a balanced compromise between accuracy and computational cost.

To highlight the potentialities and quality of the approaches, they are applied to *L*-tryptophan (TRP) in aqueous solution. TRP is an essential amino acid featuring a distinctive indole chromophore, which renders it highly responsive to various spectroscopic probes. Its relevance spans biochemical, pharmaceutical, and environmental contexts, and its interactions with water are known to significantly affect its structural and electronic characteristics. Moreover, the wealth of experimental data available for tryptophan makes it an ideal benchmark system for validating multiscale quantum/classical

methods. By applying our theoretical protocol across diverse spectroscopies, we establish a unified framework for accurately simulating the spectroscopic behavior of solvated systems, thereby enhancing our understanding of solvation effects and paving the way for applications to larger and more complex biomolecular environments.

2. THEORETICAL METHODOLOGY

The spectroscopic signals of tryptophan in aqueous solution are modeled by means of multiscale fully atomistic approaches, based on polarizable QM/FQ(F μ) embedding or three-layer polarizable QM/FDE/FQ(F μ). In both methods, the solute is treated at the QM level, specifically at the Density Functional Theory (DFT) and its time-dependent (TDDFT) extension. While in QM/FQ(F μ) the whole solvent is described at the classical FQ(F μ) level,³³ in QM/FDE/FQF μ the solvent molecules that are closest to the solute are treated at the FDE level,³⁵ and the remaining ones at the FQF μ level.³⁷ A graphical depiction of the partitioning of the two approaches is provided in Figure 1. The two methods mainly differ in the

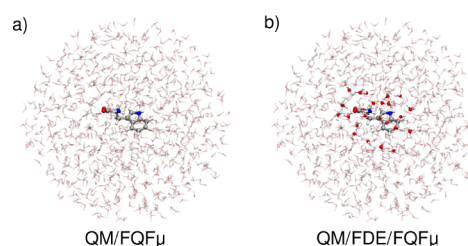


Figure 1. Graphical depiction of QM/FQF μ (a) and QM/FDE/FQF μ (b) partitioning for TRP in aqueous solution.

description of the short-range interactions. In particular, in polarizable QM/FQF μ , the solute–solvent interactions are approximated to electrostatics and polarization effects, under the assumption that nonelectrostatics, such as Pauli repulsion and dispersion, play a minor role, especially for polar solvents,¹⁴ such as water. Differently, in QM/FDE/FQ(F μ), part of the nonelectrostatic interactions, i.e., Pauli repulsion, are taken into account by the FDE layer.³⁷

The energy of the total system can be written as^{38,51}

$$\mathcal{E} = E_{\text{QM}/(\text{FDE})} + E_{\text{FQF}\mu} + E_{\text{QM}/(\text{FDE})/\text{FQF}\mu}^{\text{int}} \quad (1)$$

where $E_{\text{QM}/(\text{FDE})}$ and $E_{\text{FQF}\mu}$ are the energies of the QM(/FDE) and FQF μ layers, respectively, whereas $E_{\text{QM}/(\text{FDE})/\text{FQF}\mu}^{\text{int}}$ represents the interaction energy between the two parts. The QM/FQ and QM/FQF μ models account for mutual solute–solvent polarization effects by endowing each classical atom in the environment with charges q and dipoles μ that adjust to the QM density.^{11,33} Specifically, in the QM/FQ model, the polarization response of the classical region is described via fluctuating atomic charges determined through the Electronegativity Equalization Principle (EEP).^{62,63} In QM/FQF μ , additional dipoles are introduced to account for anisotropic polarization.³³ A set of Lagrangian multipliers λ is included in both force fields to ensure the charge neutrality of each solvent molecule.

The interaction energy $E_{\text{QM}/(\text{FDE})/\text{FQF}\mu}^{\text{int}}$ in eq 1 can then be written as the electrostatic interaction between the solvent multipolar moments and the total electric potential and field generated by the QM(/FDE) subsystems. By defining ρ_{tot} as

the total density of the QM (and FDE) portion, $E_{\text{QM}/\text{FDE}/\text{FQF}\mu}^{\text{int}}$ can be written as

$$\begin{aligned} E_{\text{QM}/\text{FDE}/\text{FQF}\mu}^{\text{int}} &= \sum_i^{\text{MM}} q_i V[\rho_{\text{tot}}](\mathbf{r}_i) - \boldsymbol{\mu}_i \cdot \mathbf{E}[\rho_{\text{tot}}](\mathbf{r}_i) \\ &= \mathbf{q}_\lambda^\dagger \mathbf{V}[\rho_{\text{tot}}] - \boldsymbol{\mu}^\dagger \mathbf{E}[\rho_{\text{tot}}] \end{aligned} \quad (2)$$

where i runs over the atoms of the MM part. $V[\rho_{\text{tot}}](\mathbf{r}_i)$ and $\mathbf{E}[\rho_{\text{tot}}](\mathbf{r}_i)$ are the QM potential and field evaluated at position \mathbf{r}_i of the i -th atomic MM site.

We can now rewrite eq 1 in terms of the energy functional of FQF μ force field (FF)³³ and the interaction energy defined in eq 2:

$$\begin{aligned} E_{\text{QM}/\text{FDE}/\text{FQF}\mu}[\rho_{\text{tot}}, \mathbf{q}, \boldsymbol{\lambda}, \boldsymbol{\mu}] &= E_{\text{QM}/\text{FDE}}[\rho_{\text{tot}}] \\ &\quad + \frac{1}{2} \mathbf{q}_\lambda^\dagger \mathbf{M} \mathbf{q}_\lambda + \mathbf{q}_\lambda^\dagger \mathbf{C}_Q \\ &\quad + \mathbf{q}^\dagger \mathbf{V}[\rho_{\text{tot}}] \\ &\quad + \mathbf{q}^\dagger \mathbf{T}^{qq} \boldsymbol{\mu} + \frac{1}{2} \boldsymbol{\mu}^\dagger \mathbf{T}^{\mu\mu} \boldsymbol{\mu} \\ &\quad - \boldsymbol{\mu}^\dagger \mathbf{E}[\rho_{\text{tot}}] \end{aligned} \quad (3)$$

where, \mathbf{q}_λ is a vector containing FQ charges \mathbf{q} and a set of Lagrange multipliers $\boldsymbol{\lambda}$, that fix the total charge Q on each FQ molecule. \mathbf{M} is a matrix collecting the charge–charge interaction kernel (\mathbf{T}^{qq}) and a set of Lagrangian blocks.⁶⁴ \mathbf{T}^{qq} diagonal elements describe charge self-interaction energy and are expressed in terms of atomic chemical hardnesses η . \mathbf{C}_Q is a vector containing atomic electronegativities χ and charge constraints \mathbf{Q} . The FQF μ FF additionally accounts for charge–dipole and dipole–dipole interactions, which are expressed in terms of $\mathbf{T}^{q\mu}$ and $\mathbf{T}^{\mu\mu}$ interaction kernels.⁶⁵ The diagonal elements of the $\mathbf{T}^{\mu\mu}$ tensor represent the dipole–dipole self-interaction and are written in terms of the atomic polarizabilities α . Notably, QM(/FDE)/FQ is obtained by retaining the terms exclusively depending on charges in eq 3.³⁷

In accordance with the EEP principle, the polarizable charges (and dipoles) are finally obtained by a constrained minimization of eq 3 with respect to charges (dipoles) and Lagrange multipliers, yielding the following linear system of equations:

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_\lambda & \mathbf{T}^{q\mu} \\ \mathbf{1}_\lambda^\dagger & \mathbf{0} & \mathbf{0} \\ (\mathbf{T}^{q\mu})^\dagger & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ \mathbf{Q} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} -\mathbf{V}[\rho_{\text{tot}}] \\ \mathbf{0} \\ \mathbf{E}[\rho_{\text{tot}}] \end{pmatrix} \quad (4)$$

where the second term on the right-hand side describes the mutual polarization between the QM(/FDE) and FQF μ regions. A functional differentiation of eq 3 with respect to the solute density function (ρ_i in QM/FDE/FQF μ ; ρ in QM/FQF μ) gives the Kohn–Sham (KS) Hamiltonian that defines the Self-Consistent Field (SCF) procedure. In both cases, this results in the addition of the following operator:^{24,33}

$$\hat{H}_{\text{QM}/\text{FDE}/\text{FQF}\mu} = \sum_i \frac{q_i}{|\mathbf{r}_i - \mathbf{r}|} - \sum_i \boldsymbol{\mu}_i \frac{\mathbf{r}_i - \mathbf{r}}{|\mathbf{r}_i - \mathbf{r}|^3} \quad (5)$$

where charges and dipoles are obtained by solving eq 4.

The computational description of molecular properties and spectra in (aqueous) solutions can be largely modified by the quality of the modeling of solute–solvent interactions, being particularly crucial the inclusion of solute–solvent polarization

effects (which are featured by QM/FQF μ).^{11,21,22,51} To further extend QM(/FDE)/FQF μ to response properties of solvated systems, the modification of the ground state (GS) molecular orbitals (MOs), which results from the SCF procedure (see eq 5), is not sufficient. A generic response property can be computed by solving the proper response equations,^{66,67} which must be modified in order to account for mutual polarization effects resulting from the perturbation of the QM density. In a DFT framework, this means that the embedding method has to be extended to electric and magnetic Coupled-Perturbed KS (CPKS) equations,⁶⁶ accounting for the adjustment of the polarization sources to the perturbed density, and specific terms due to the usage of Gauge—including atomic orbitals (GIAOs) in magnetic response.⁶⁸ Absorption spectra are instead computed by including specific polarization terms into Casida's equations,⁶⁹ in order to account for the dynamic response (linear regime) of polarization sources (we refer the reader to ref 27 for a detailed discussion on modeling absorption spectroscopy using polarizable QM/MM methods). For vibrational response, nuclear perturbation must also be considered, extending the method to the analytical evaluation of energy first derivatives (gradients) using the machinery of quantum chemistry.⁷⁰ Second and higher order derivatives can be computed by using analytical or numerical differentiation, opening the way to the simulation of IR,^{28,40} Raman,^{29,41} and Raman Optical Activity spectra.^{29,41}

3. RESULTS AND DISCUSSION

3.1. Computational Protocol. The application of QM(/FDE)/FQ(F μ) to the calculation of molecular properties and spectral signals of aqueous TRP requires its coupling to a specific multistep protocol, which has been proposed and refined in recent years.^{11,27} Similar strategies, based on the combination of MD sampling with quantum chemical calculations of several spectroscopic properties in solution, have also been employed in previous works.^{71–74} The protocol consists of the following steps:

- (1) *Definition of the system:* TRP (in its zwitterionic form) is treated at the QM level, whereas the surrounding water environment is treated classically, by using FQ and FQF μ .
- (2) *Conformational and configurational sampling:* TRP–water configurational space is explored through classical molecular dynamics (MD) simulations performed using the GROMACS package⁷⁵ and conducted on the nanosecond time scale. Both intramolecular and intermolecular interactions are treated using the General Amber Force Field (GAFF2),⁷⁶ while the standard TIP3P force field⁷⁷ is used to describe the solvent. TRP initial geometry is chosen to match the optimized structure at the B3LYP/aug-cc-pVDZ level of theory, and then solvated by preparing a cubic simulation box with a side length of 5.9 Å, containing 6864 water molecules. CM5 charges are employed.⁷⁸

To avoid atomic overlap and unrealistic interactions between molecules, the system's energy is minimized. Subsequently, before starting the production step, two equilibration stages are performed. First, a 1 ns NVT equilibration stage is run, maintaining the temperature constant at 298.15 K using a velocity-rescaling thermostat⁷⁹ with a coupling constant of 0.1 ps and an integration time step of 2.0 fs. Then, a 2 ns NPT

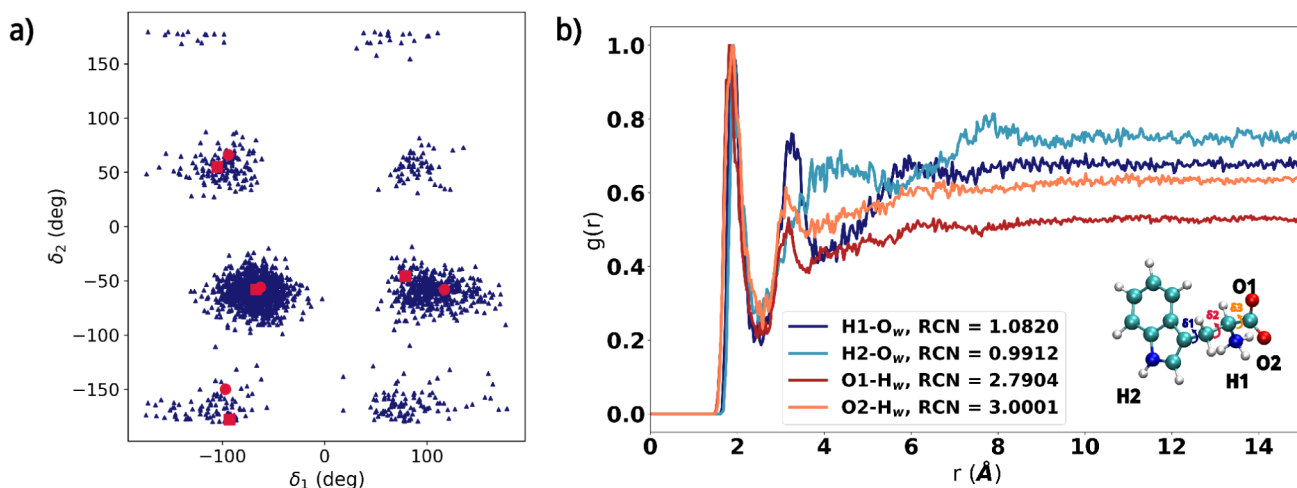


Figure 2. (a) Left: MD conformational analysis (blue triangles) for TRP zwitterionic form in aqueous solution as a function of δ_1 and δ_2 . The conformations of the 4 main clusters obtained through clustering analysis are also shown as red circles, while the red squares represent the dihedrals of the lowest energy conformers reported in the literature.¹⁰⁷ (b) Right: Radial distribution functions (RDFs) between TRP and water molecules, with the corresponding running coordination numbers (RCNs). The definition of δ_1 , δ_2 and δ_3 is also reported.

equilibration stage is carried out, maintaining the pressure at 1.0 bar using the Parrinello–Rahman barostat⁸⁰ with a coupling constant of 2.0 ps and an integration time step of 2.0 fs. Finally, an NPT production stage is run for 50 ns with an integration time step of 2.0 fs, using the LINCS algorithm⁸¹ to maintain all bond lengths constrained, allowing for stable and efficient simulations. The particle-mesh Ewald (PME)⁸² method is used to handle electrostatic interactions, with a grid spacing of 0.16 nm and a spline interpolation of order 4. Short-range Coulomb and van der Waals interactions are truncated using the Verlet cutoff scheme,⁸³ with a cutoff distance of 1.2 nm.

- (3) *Extraction of structures:* From the production stage of the MD simulation, several uncorrelated snapshots are extracted every 10 ps.^{84–91} Then, solute-centered spheres (droplets) of radius equal to 18 Å are cut. This dimension is large enough to retain solute–solvent interactions in a physically consistent way.
- (4) *QM/FQ(F μ) calculations of the target molecular properties/spectra:* Calculations of the target response/spectral properties are performed on a variable number of spherical droplets obtained at the previous step. The number of droplets is chosen to ensure numerical convergence of the results. UV and ECD spectra are simulated by processing respectively 300 and 950 snapshots at the TD-CAMY-B3LYP/TZ2P level of theory (see Figures S3–S6). The choice of this functional is consistent with previous studies demonstrating the reliability of range-separated hybrids for these properties, including applications to L-tryptophan.^{92,93} NMR, IR, Raman, ROA, and OR calculations are performed at the B3LYP/TZ2P level of theory, in line with several previous works where hybrid functionals have been successfully applied to vibrational and magnetic properties of L-tryptophan.^{94–96} While 200 snapshots are sufficient for NMR, IR, Raman, and ROA, a significantly larger number—4000 snapshots—is necessary to achieve convergence in OR (see Tables S1,S2, Figures S10, S11 and S13–S18).

FQ and FQF μ layers are modeled by exploiting the parametrization reported in refs 33 and 97.

- (5) *Data analysis and extraction of final properties/spectra:* For UV and ECD, the final spectra are obtained by applying a Gaussian convolution with a Full Width at Half Maximum (fwhm) of around 0.5 eV to averaging data. In the case of ¹H NMR, the full spectrum is computed by including nuclear spin–spin couplings under an external magnetic field of 600 MHz. To generate the final spectral profile, the resulting chemical shift values are convoluted using a Lorentzian function with a fwhm of 0.002 eV. In contrast, for ¹³C, ¹⁵N, and ¹⁷O NMR, spin–spin couplings are not considered. Accordingly, only average values obtained from the QM/FQ and QM/FQF μ calculations are reported. Specifically, chemical shifts are provided for ¹³C, whereas nuclear shielding constants are reported for ¹⁵N and ¹⁷O. Chemical shifts are obtained from computed shielding constants using tetramethylsilane (TMS) as reference, with $\sigma_{\text{ref}} = 31.7$ ppm for ¹H and $\sigma_{\text{ref}} = 182.852$ ppm for ¹³C, as employed in Amsterdam Modeling Suite (AMS).⁹⁸ For IR, Raman, and ROA spectra, stick spectra obtained from individual snapshots are averaged and convoluted using a Lorentzian function with a fwhm of 10 cm⁻¹. A Raman excitation energy of 634 nm is used to simulate the spectra. In addition to QM/FQ and QM/FQF μ , for NMR, a QM/FDE/FQF μ approach is employed. The FDE protocol consisted of three sequential steps: (i) a full-system QM/MM calculation with the solute and all solvent molecules, (ii) a QM/FQF μ single-point calculation on the water molecules within the FDE layer (treated at QM level), embedded in the remaining solvent molecules (described at FQF μ level), to generate a frozen electron density (ρ_2), and (iii) a final QM/FDE/FQF μ calculation where this frozen density is used to represent the inner solvent shell, while the remaining solvent is treated as a polarizable MM region. In step (iii), to calculate the FDE terms, the PW91K86 functional⁹⁹ is used to approximate the kinetic energy, while the nonadditive exchange-correlation terms are treated using the PBE

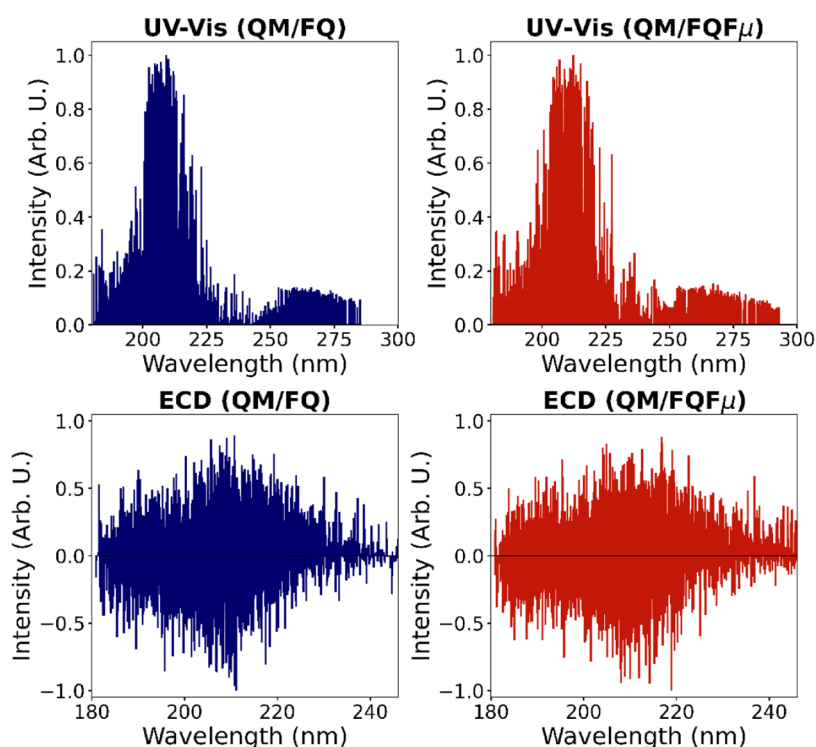


Figure 3. UV and ECD QM/FQ (left) and QM/FQF μ (right) stick spectra of L-tryptophan (TRP) in aqueous solution, derived from 300 (for UV) and 950 (for ECD) structures extracted from the MD simulation.

functional.³⁶ The selection of PW91K is based on previous studies that demonstrated its effectiveness in simulating ground-state properties and NMR spectra with FDE.^{38,100–102} QM/FDE/FQF μ calculations are performed using an FDE radius of 3 Å and incorporating freeze-and-thaw (FT) cycles.¹⁰³

As regards OR calculations, a single value corresponding to the sodium D-line energy (589.3 nm) is computed for each snapshot across both QM/FQ and QM/FQF μ methods, highlighting the frame to frame variation of that property.

All QM/FQ(F μ) and QM/FDE/FQ(F μ) calculations are performed with the AMS.^{104,105}

3.2. Conformational Analysis and Hydration Patterns.

TRP is a flexible molecule, and its conformation is expected to be affected by the presence of the aqueous environment, which can also interact with TRP by forming hydrogen bonding (HB).

Two main factors ensure a high-quality conformational sampling for the TRP aqueous system, i.e. (i) the sampling of the solute conformation in terms of the internal dihedral angles within the molecule that originate from intramolecular interactions (such as weak noncovalent interactions, HB, ...), and (ii) the sampling of the solvent configurations and the distribution and orientation of solvent molecules around the solute.¹¹

The MD trajectory is initially analyzed using clustering analysis¹⁰⁶ of GROMACS to identify the main configurations of TRP, employing a root-mean-square deviation (RMSD) threshold of 0.11 nm. Four primary conformers are identified, shown in Figure S1, with relative populations of 66%, 18%, 7%, and 4%, respectively. TRP is characterized by three key dihedral angles: δ_1 , δ_2 , and δ_3 —the definitions and specific values of the representative structures are detailed in Figure 2b

and S1. The dihedral distribution functions (DDFs) for δ_1 , δ_2 , and δ_3 are presented in Figure S2. δ_1 , the dihedral angle connecting the aromatic ring to the main chain, shows a dominant peak centered at approximately -60° and a secondary peak centered at about 100° , which indicates some rotational flexibility around this bond. Similarly, δ_2 , which links the main chain to the amino group, shows a dominant peak near -60° , along with minor peaks at around 60° and $\pm 180^\circ$. δ_3 , connecting the amino group to the carboxyl group, shows broad peaks centered at -60° and 130° , reflecting significant rotational flexibility. The δ_1 and δ_2 values obtained from the main clusters, identified through the clustering analysis, show good agreement with the representative conformers reported in the literature (see Figure 2a).¹⁰⁷ As shown in the figure, both our clusters and those from the literature fall within the high-density regions of the dihedral angle distribution, indicating that they represent the most populated conformational basins sampled during the simulation. No clusters fall within the low-density regions, which is expected since only the most representative structures are selected. The third dihedral angle, δ_3 , is not included in the 2D analysis due to its high flexibility, as it spans nearly the entire angular range (see Figure S2). Overall, these results support the consistency between different sampling and analysis strategies and confirm the reliability of our conformational model.

To analyze solute–solvent HB interactions, radial distribution functions $g(r)$ (RDFs) and running coordination numbers (RCNs) are extracted from the MD trajectory (see Figure 2b). Around oxygen atoms of the carboxyl group (O1 and O2), there is a first solvation shell at about 1.82 Å (O1) and 1.88 Å (O2), respectively, with corresponding RCNs of 2.79 and 3.00. This indicates that the carboxyl group forms, on average, HBs with three water molecules. In the case of the hydrogen atoms

of the amino group (H1) and the pyrrolic group (H2), the first solvation shell is placed at approximately 1.85 Å for H1 and 1.91 Å for H2, with RCNs of 1.08 and 0.99, reflecting the formation of an HB with a single water molecule. A similar analysis can be performed to investigate potential intramolecular HBs between the amino and the carboxyl groups. No significant peaks are detected, thus indicating the absence of persistent intramolecular HBs within TRP, and also suggesting that the solute primarily forms HBs with the solvent molecules.

3.3. UV and ECD Spectra. This section presents computed QM/FQ ($F\mu$) UV and ECD spectra of TRP in aqueous solution. Figure 3 displays the calculated stick spectra, representing the raw excitation energies and corresponding oscillator strengths obtained from individual snapshots extracted from the MD trajectory. Each snapshot produces a distinct spectral signal as a result of the varying configurations of the solute and different distributions of the solvent around it, effectively capturing solvent-induced inhomogeneous band broadening without the need for ad-hoc corrections.

QM/FQ ($F\mu$) UV and ECD spectra are shown in Figure 4. The results of the study of the convergence of computed

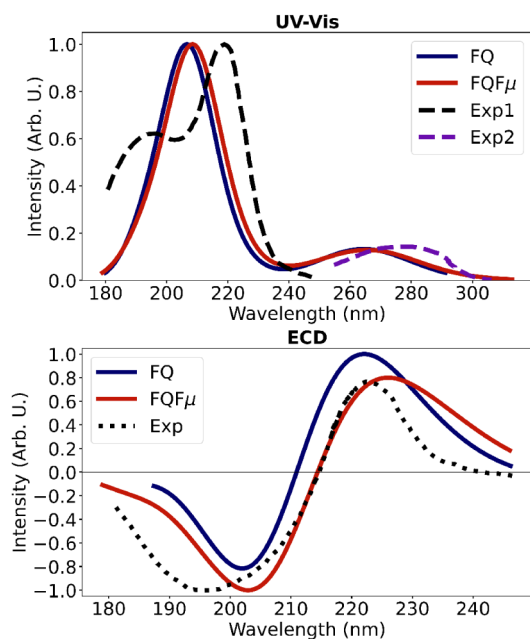


Figure 4. QM/FQ and QM/FQF μ UV (top) and ECD (bottom) spectra of L-tryptophan (TRP) in aqueous solution, derived from 300 (for UV) and 950 (for ECD) structures extracted from the MD simulation and computed at the CAMY-B3LYP/TZ2P level of theory. A Gaussian convolution having a fwhm of 0.5 eV is exploited to draw the spectra. Computed data are compared with experimental UV spectra^{108,109} and experimental ECD spectrum.¹⁰⁸

spectra with respect to the number of snapshots are reported in Figures S3–S6 in the SI. Convergence of averaged values is achieved employing 200 snapshots for UV and 900 for ECD, at both QM/FQ and QM/FQF μ level. The larger number of snapshots which are required to converge ECD spectra highlights the complex interplay of computed positive and negative rotational strengths (see Figure 3), and aligns with previous studies of some of us.^{44,51}

The averaged QM/FQ ($F\mu$) UV spectra (see Figure 4) exhibit an intense high-energy absorption peak accompanied

by a weaker peak at lower energy. The corresponding ECD spectra, also depicted in Figure 4, display a distinctive (–, +) sign pattern for both solvent models. Canonical molecular orbital (CMO) decomposition, based on natural bonding orbital (NBO) analysis,^{110,111} reveals that the lower-energy peak originates from the HOMO \rightarrow LUMO $\pi \rightarrow \pi^*$ electronic transition. CMOs involve orbitals localized in the aromatic ring of TRP, as shown in Figure S7. Notice that the NBO analysis is performed on the representative structure of the main conformer of TRP, obtained through the GROMACS cluster analysis, as described in Section 3.2. Solvent effects are implicitly considered by employing the COSMO continuum solvent model.^{112,113}

Only minor differences between QM/FQ and QM/FQF μ UV and ECD spectra are observed. The inclusion of fluctuating dipoles in QM/FQF μ leads to a slight redshift in peak positions, while the shape of the bands and the relative intensities remain largely unchanged. In UV spectra, the main peak appears at 206 nm for QM/FQ and 208 nm for QM/FQF μ , with a weaker band at 264–265 nm. Both ECD spectra display the characteristic (–, +) pattern, with bands at 201/223 nm for QM/FQ and 203/226 nm for QM/FQF μ .

Figure 4 reports also experimental UV and ECD spectra, enabling a direct comparison between theoretical predictions and experimental observations. The experimental UV data for L-tryptophan in aqueous solution are taken from two distinct measurements covering different spectral regions. One experiment reports the high-energy transitions,¹⁰⁸ which give rise to two bands centered at 190 and 220 nm, with the latter being the more intense. The other experiment focuses on the near-UV region,¹⁰⁹ showing a broad band centered at about 278 nm. As shown in Figure 4, the simulations predict only one band in the high-energy region, while the experimental spectrum resolves two contributions at 190 and 220 nm, and they slightly overestimate the position of the 278 nm band, possibly due to intrinsic limitations of DFT. Overall, the computed spectra reproduce pretty well the positions of the experimental peaks and the observed band broadening. The ECD experimental spectrum¹⁰⁸ shows a negative peak at around 200 nm, followed by a positive peak near 220 nm. Both computed spectra accurately reproduce the key features of the experimental spectrum, especially the sign pattern (–, +) that is the most important characteristic of chiral spectroscopies. Furthermore, the calculations yield excellent agreement in terms of peak positions and band broadening for both QM/FQ and QM/FQF μ methods, while the agreement in relative intensities is observed only with the QM/FQF μ approach. These results improve upon those previously reported in the literature using PCM TDDFT/CAM-B3LYP or TDDFT/B3LYP with the aug-cc-pVDZ basis set, where the final spectrum was obtained by Boltzmann averaging over the nine most stable conformers optimized by the authors.⁹³ In particular, our protocol not only reproduces the experimental alternation of signs (–, +), as already achieved in earlier studies, but also provides a more accurate description of both peak positions and band broadening.

3.4. NMR Spectra. Computed QM/FQ ($F\mu$) chemical shifts for TRP hydrogen and carbon nuclei, as well as shielding constants for nitrogen and oxygen, are reported in Table 1 and Figure 5 for atom labeling. Tabulated values are obtained by averaging computed values for 200 snapshots, to ensure convergence (see Tables S1 and S2).

Table 1. QM/FQ and QM/FQF μ Computed Chemical Shifts (ppm)(for Hydrogen and Carbon) and Magnetic Shielding Constants (for Nitrogen and Oxygen)^{a,b}

Atom	QM/FQ	QM/FQF μ
Hydrogens		
H1	3.7336	3.6629
H2	3.4384	3.4743
H3	4.3763	4.3996
H4	10.0063	10.4016
H5	6.0134	6.2879
H6	5.9277	6.1879
H7	6.3728	6.7483
H8	8.2038	8.2446
H9	7.6214	7.6119
H10	7.7782	7.7860
H11	8.0512	8.1204
H12	7.9716	8.0644
Carbons		
C1	63.9518	63.5953
C2	30.4370	30.4645
C3	176.2754	178.5201
C4	117.7629	117.2523
C5	135.2466	136.4871
C6	143.8707	143.8123
C7	132.6885	132.6688
C8	124.3865	124.5879
C9	127.0266	126.6992
C10	129.4505	129.1790
C11	121.6423	122.0687
Nitrogens		
N1	210.4982	211.7675
N2	104.0095	101.8629
Oxygens		
O1	35.1612	42.1765
O2	48.8983	55.5504

^aValues are Obtained by Averaging Over 200 Snapshots of Solvated TRP. ^bQM Level of Theory: B3LYP/TZ2P.

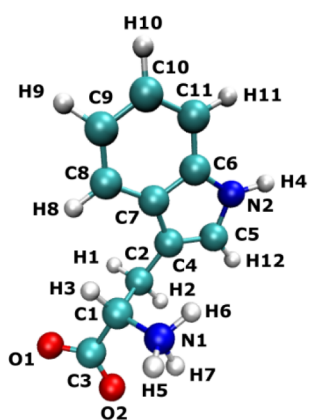


Figure 5. Atom labeling scheme of TRP used for chemical shift and shielding assignment.

For ^1H nuclei, the differences between QM/FQ and QM/FQF μ are small, as evidenced by the average absolute deviations reported in Table S3. The maximum intermodel deviation for hydrogen nuclei is 0.4 ppm, observed for the N–H proton of the pyrrole ring H4 (see Figure 5).

In comparison to QM/FQ, the inclusion of dipoles in the solvent description alters the electron density around the

nuclei, leading to upfield or downfield shifts. A noticeable deshielding of 0.3–0.4 ppm is observed for nuclei that strongly interact with the aqueous environment, such as H4, H5, H6, and H7, as a consequence of the enhanced solute–solvent electrostatic interaction described by QM/FQF μ compared to QM/FQ.

Consistently with the trends observed for ^1H , also in the case of ^{13}C nuclei the differences among the QM/FQ and QM/FQF μ models are generally low and range between 0.02 and 1.0 ppm for most nuclei, with the exception of carboxylic carbon (C3) and C5 (see Table S3).

The averaged data for oxygen and nitrogen nuclei are also presented in Table 1. Due to their strong interactions with the surrounding medium, oxygen nuclei exhibit high sensitivity to the choice of solvent model, with large variations across the different approaches, and a maximum difference of around 7 ppm observed between QM/FQ and QM/FQF μ . Conversely, nitrogen nuclei show a lower sensitivity, following a trend similar to hydrogen and carbon, with a maximum shielding variation of approximately 2 ppm for the pyrrolic nitrogen N2, as observed when comparing QM/FQ and QM/FQF μ .

NMR chemical shifts and shielding constants are local magnetic response properties, strongly influenced by the electronic environment around the nucleus. They are particularly sensitive to short-range effects such as hydrogen-bond directionality and local electronic polarization. To better account for these interactions—especially those involving the first solvation shell—we employed a hybrid QM/FDE/FQF μ model. In this scheme, solvent molecules within a certain radius of the solute are treated quantum mechanically using the FDE approach, while the remaining environment is described by a polarizable molecular mechanics layer (FQF μ). This enables a more accurate representation of the local electronic environment without requiring a full QM treatment of the entire solvent.³⁸ To assess the reliability and cost-efficiency of this setup, a series of tests is performed on 10 frames by systematically varying key parameters. Specifically, we evaluated the effect of the FDE radius (3, 4, and 5 Å), the use of a mixed basis set in step (ii), and the inclusion of FT cycles in step (iii) (see Section 3.1 for the steps). The influence of the FDE radius and the use of mixed basis sets is first examined (see Figure S8), revealing that the mixed basis has a negligible effect on the computed chemical shifts. Subsequently, the effect of the FDE radius was further investigated in combination with FT cycles (Figure S9), highlighting a non-negligible sensitivity for nuclei such as H4–H7, O1, O2, and C3. Among the tested setups, the configuration with a 3 Å FDE radius with FT cycles showed good agreement with the reference calculation (5 Å with FT), while keeping the computational cost low. These settings are therefore adopted for the NMR property calculations.

Similarly to QM/FQ and QM/FQF μ computed NMR spectra, also for QM/FDE/FQF μ model, the final chemical shifts and shieldings values are obtained by averaging over 200 snapshots, which ensures the convergence of the computed values (see Table S4). As shown in Table S5, the inclusion of the FDE layer has little impact on calculated chemical shifts of the hydrogen and carbon nuclei, with the exception of H4–H7 and C3, which show differences of approximately 0.5–0.6 and 4 ppm, respectively, between QM/FQF μ and QM/FDE/FQF μ (see Table S6). As expected, a much more pronounced effect is observed for the oxygen nuclei, with deviations up to 16 ppm compared to QM/FQF μ values. Moving to ^{13}C , the

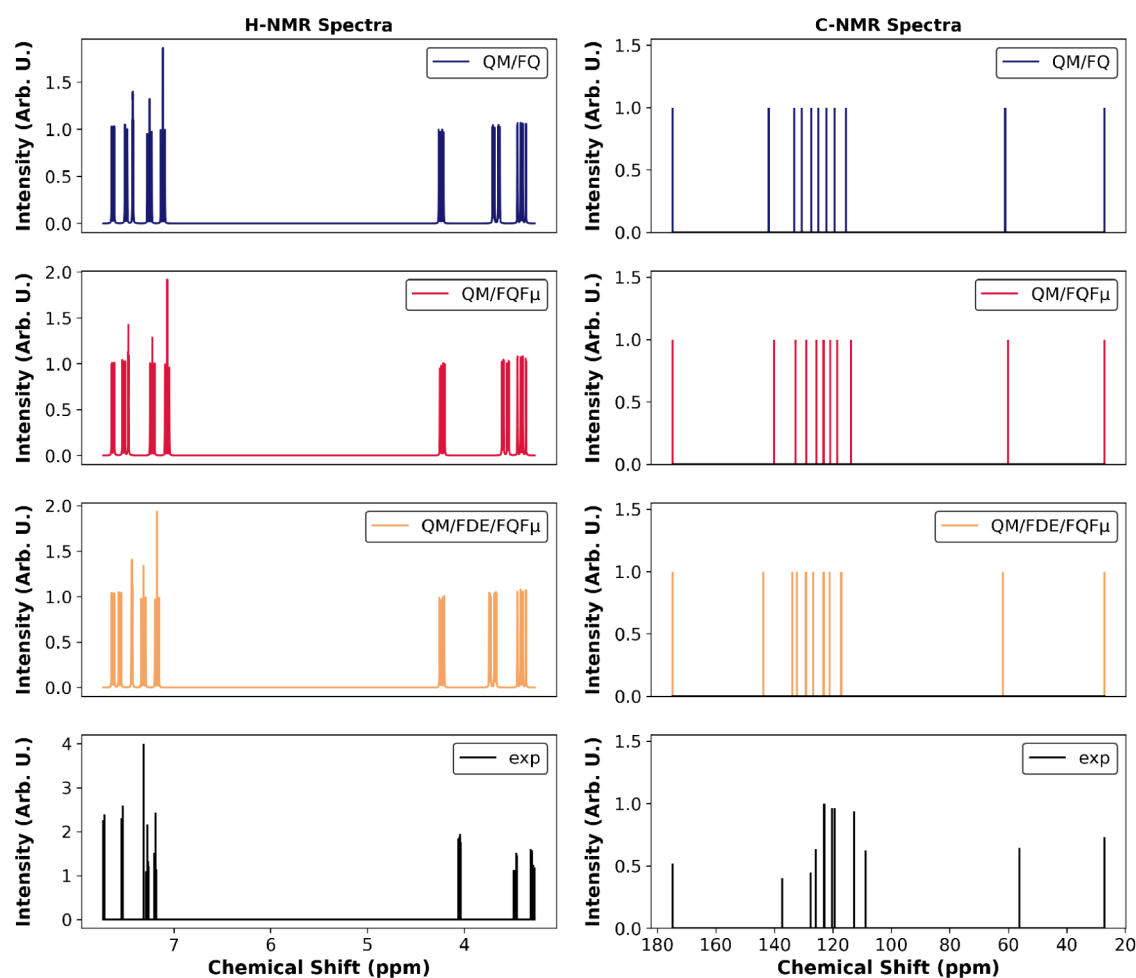


Figure 6. Computed coupled ^1H NMR and uncoupled ^{13}C NMR spectra of L-tryptophan (TRP) in aqueous solution. QM/FQ, QM/FQF μ and QM/FDE/FQF μ approaches are compared. QM level: B3LYP/TZ2P. Values are averaged over 200 snapshots. Chemical shifts are obtained from computed shielding constants using tetramethylsilane (TMS) as reference, with $\sigma_{\text{ref}} = 31.7$ ppm for ^1H and $\sigma_{\text{ref}} = 182.852$ ppm for ^{13}C , as employed in AMS.⁹⁸ Experimental NMR spectra are also shown for comparison.^{114,115}

subsequent inclusion of the FDE layer induces the same opposing trend observed for ^1H , resulting in a pronounced shielding effect of up to 4 ppm for C3.

To end the discussion, computed data are compared with available experimental values (see Figure 6). The experimental chemical shifts for hydrogen and carbon nuclei of L-tryptophan in aqueous solution are also reported in Table S13.^{114,115} To maintain consistency with experimental conditions, hydrogen atoms bound to nitrogen (H4, H5, H6, and H7) are excluded from the coupled spectra. For ^1H nuclei, NMR coupled spectra are simulated assuming a magnetic field strength corresponding to a proton resonance frequency of 600 MHz. The J-coupling constants are computed for all models, as reported in Tables S7–S9.

Hydrogen atoms H8–H9–H10–H11 of the benzene ring show ortho J-couplings (J_{ortho}) of approximately 7–8 Hz. Additionally, long-range couplings, such as ^5J (e.g., H8–H4), are observed for the indole ring, although J values are lower than 1 Hz, due to the extensive resonance delocalization that facilitates spin–spin interactions over greater distances.

For the aliphatic chain, the highest coupling constants correspond to geminal couplings (^2J) (H1–H2, H5–H6, H6–H7, H5–H7), ranging from 17 to 19 Hz, while vicinal

couplings (^3J) (H1–H3, H2–H3, H3–H5, H3–H6, H3–H7) fall within the 4–12 Hz range.

J-coupling constants computed with the three models exhibit minimal differences, as shown in Tables S10–S12, with a maximum difference of 0.4 Hz observed for the H3–H7 coupling moving from QM/FQF μ to QM/FDE/FQF μ .

Computed ^1H NMR spectra reveal two distinct chemical shift regions. Hydrogen atoms belonging to the indole moiety (H8–H12) resonate above 7 ppm, while aliphatic hydrogens (H1–H3) appear below 4 ppm. All computational models successfully reproduce the experimental ^1H NMR spectrum, with a good agreement in both chemical shifts and relative peak intensities. Although experimental J-coupling constants are not available for direct comparison, computed spectra capture the expected coupling patterns.

The computed decoupled ^{13}C NMR spectrum, shown in Figure 6, displays aromatic carbon signals between 110 and 150 ppm. Carboxylic carbon (C3), which is highly deshielded, resonates around 180 ppm, consistent with its electron-withdrawing environment. The two aliphatic carbons (C1 and C2) are found at approximately 30 and 60 ppm, respectively, reflecting their different chemical environments adjacent to the indole ring and the amino group. Theoretical models accurately reproduce the experimental peak distribution,

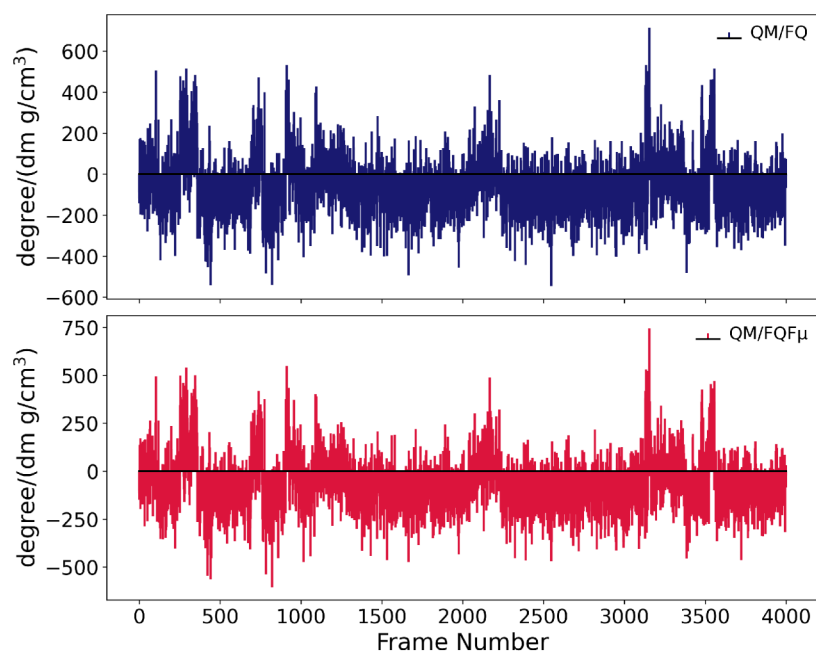


Figure 7. Computed QM/FQ and QM/FQF μ optical rotation at the sodium D-line (589.3 nm) of L-tryptophan in aqueous solution for each of the 4000 frames extracted from the molecular dynamics trajectory. Each stick represents the OR of a single solute–solvent configuration. QM level of theory: B3LYP/TZ2P.

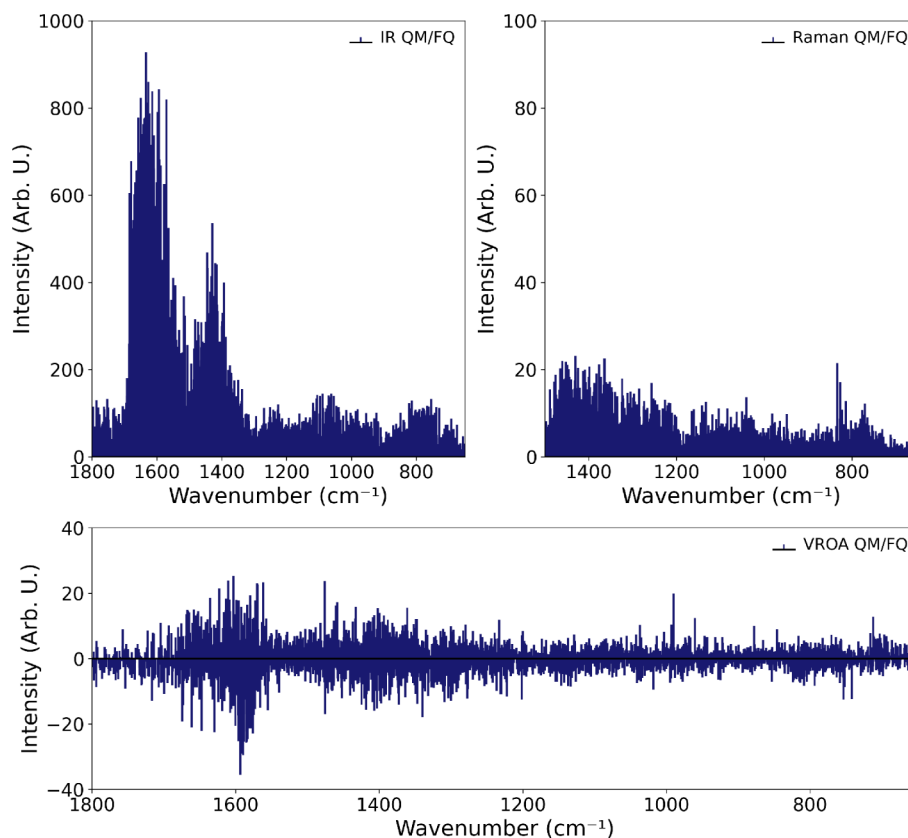


Figure 8. IR, Raman, and ROA stick spectra of L-tryptophan (TRP) in aqueous solution computed using the QM/FQ approach at the B3LYP/TZ2P level of theory. Spectra correspond to normal-mode analyses performed on 200 individual snapshots extracted from the molecular dynamics simulation. IR and Raman spectra are shown in the top panel, and the ROA spectrum in the bottom panel.

particularly within the aliphatic region (30–60 ppm), with deviations generally within a few ppm.

Because of the absence of unambiguous experimental assignments for the aromatic carbon signals, a direct

comparison with the computed spectra is not possible. Nevertheless, the spectra are qualitatively in good agreement, again supporting the validity of our computational approach for both hydrogen and carbon NMR spectral predictions.

3.5. Optical Rotation. To complement the analysis of ECD and further investigate the chiroptical properties of L-tryptophan in solution, OR is computed. Simulations are performed using both the QM/FQ and QM/FQF μ models at the B3LYP/TZ2P level of theory. A total of 4000 frames extracted from the MD trajectory are used to ensure statistical convergence (see Figures S10 and S11). The specific incident wavelength used in the calculations corresponds to the sodium D line (589.3 nm), which is commonly adopted for direct comparison with experimental measurements. To gain a deeper understanding of the frame-by-frame fluctuations in OR, stick spectra are generated by reporting the OR values for each of the 4000 snapshots extracted from the MD trajectory. The results are depicted in Figure 7 for both QM/FQ and QM/FQF μ calculations. The large fluctuations in both the sign and intensity of the OR values highlight how strongly this property depends on the specific mutual arrangement of the solute and surrounding water molecules. This behavior is typical of OR, which is highly sensitive to small structural and environmental variations.^{41,46} Despite the intrinsic complexity of OR, both QM/FQ and QM/FQF μ successfully reproduce the experimentally observed negative sign of the OR, measured as -31.5 degree/(dm \cdot g/cm³) at 23 °C.¹¹⁶ The computed values are -64 degree/(dm \cdot g/cm³) with QM/FQ and -71 degree/(dm \cdot g/cm³) with QM/FQF μ , reflecting a consistent qualitative agreement with experiment.

3.6. Vibrational Spectra. We now turn our attention to vibrational spectra, which involve a perturbation with respect to nuclear displacements. To stay within the "focused" approach that we have discussed in the Introduction, we employ the so-called Partial Hessian Vibrational Analysis (PHVA),^{117–119} which has been used in numerous studies^{21,28,48} and offers a reliable and computationally efficient strategy for vibrational analysis in complex systems.

Within the PHVA, vibrational spectra are computed for the QM active region of the system, in the present case TRP. The influence of the surrounding environment is still accounted for, but only in terms of its effect on the vibrational properties of the active site. This means that not all of the system's nuclei are included in the vibrational analysis: only the atoms belonging to the active region are used to construct the Hessian matrix, which in turn determines the vibrational modes. Vibrational analysis within the double-harmonic approximation must be performed on energy minima. For this reason, a partial optimization of the various snapshots extracted from the MD is performed, allowing only the QM degrees of freedom to relax and then keeping the solvent frozen. In this way, the sampling of the solvent configuration around the solute achieved in the MD is preserved. It is worth noting that alternative MD-based approaches extract vibrational spectra from time-correlation functions of dipole moments or polarizabilities computed along trajectories.^{120–123} These methods naturally include dynamic effects as well as explicit solute–solvent couplings, however they fully rely on a specific choice of the classical force field. The PHVA offers a harmonic but efficient description of solute vibrations at the full quantum-mechanical level.

The computed stick spectra for IR, Raman, and ROA, obtained with the QM/FQ and QM/FQF μ approaches, are shown respectively in Figures 8 and S12. The results refer to calculations on 200 snapshots, a number that is appropriate to reach convergence on averaged spectra (see Figures S13–S18). For IR and Raman, stick spectra closely reflect the final spectral

shapes, while the ROA stick spectrum is more complex to interpret because of the coexistence of positive and negative signals within the same spectral regions, which prevents immediate prediction of the final spectral profile.

The averaged computed IR spectrum (Figure 9) shows three dominant bands centered around 1600 cm⁻¹, 1420 cm⁻¹, and

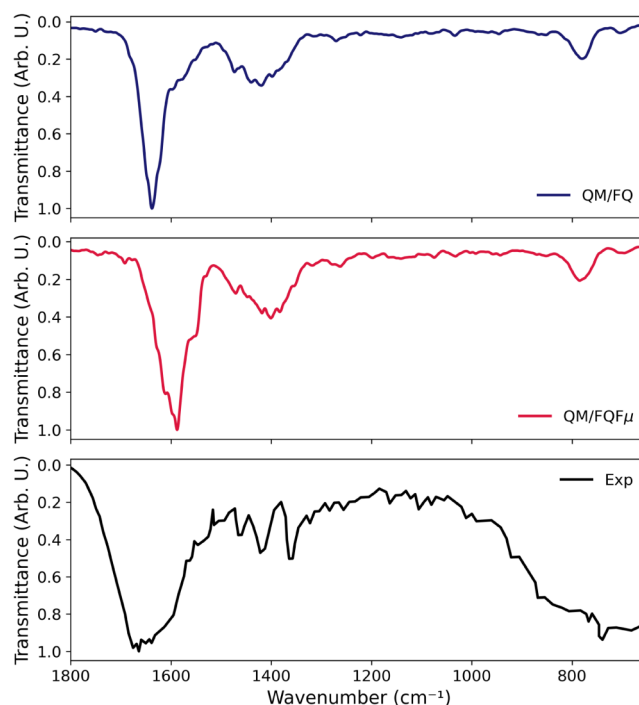


Figure 9. Computed IR spectra of L-tryptophan (TRP) in aqueous solution using QM/FQ and QM/FQF μ approaches at the B3LYP/TZ2P level of theory, averaged over 200 snapshots. The experimental IR spectrum⁹⁴ is also shown for comparison.

780 cm⁻¹. The first, and most intense band, is assigned to the asymmetric stretching of the carboxylate group. This mode is coupled with the symmetric deformation of the NH₃⁺ group, the out-of-plane bending of the pyrrolic N–H, CH₂ scissoring, and stretching of conjugated C–C and C–N bonds within the aromatic moiety. The normal modes corresponding to the most prominent vibrational bands are analyzed for the most representative MD frame in Figure S19. The second band, at approximately 1420 cm⁻¹, is mainly attributed to the symmetric stretching of the carboxylate group, with additional contributions from aromatic stretching and in-plane C–H bending. The third band at 780 cm⁻¹, involves CO₂ scissoring, out-of-plane bending of aromatic C–H bonds, and rocking motions of the aliphatic CH₂ and NH₃⁺ groups.

In addition to these main bands, several weaker features appear in the 1300–900 cm⁻¹ region, which arise from combinations of in-plane and out-of-plane bending modes of aromatic and aliphatic C–H groups, along with C–C and C–N stretching modes. The rocking of CH₂ and NH₃⁺ groups give minor contributions to this region. As illustrated in Figure 9, IR spectra obtained with the QM/FQ and QM/FQF μ methods show minimal differences. The only noticeable deviation concerns the dominant band assigned to the asymmetric stretching of the carboxylate ion: it is slightly red-shifted in the QM/FQF μ spectrum and appears broader and better resolved than at the QM/FQ level. Apart from this

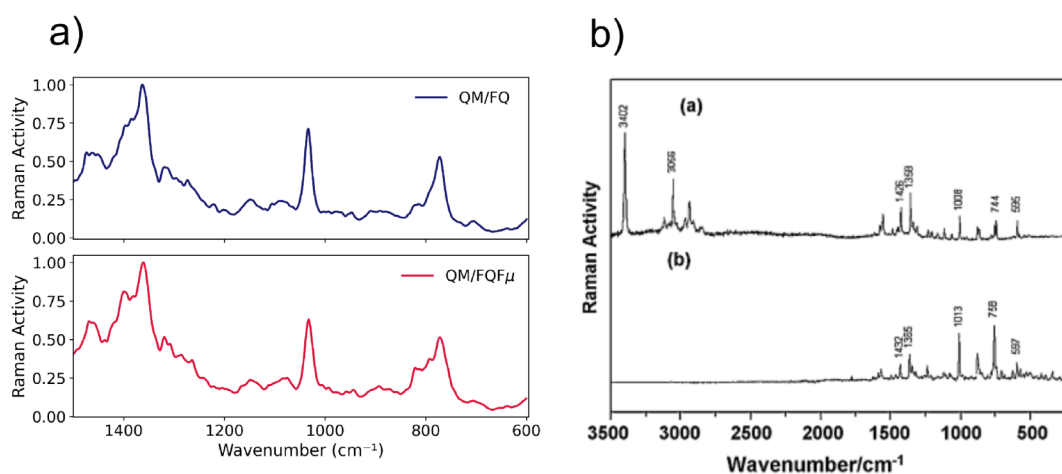


Figure 10. a) Computed Raman spectra of L-tryptophan (TRP) in aqueous solution using QM/FQ and QM/FQFμ approaches at the B3LYP/TZ2P level of theory, averaged over 200 snapshots. b) Experimental Raman spectra. Reproduced with permission from ref.⁹⁴ Copyright Elsevier 2012. The top panel corresponds to the solid phase, while the bottom panel shows the experimental Raman spectrum of zwitterionic L-tryptophan in aqueous solution, used here for comparison with our results.

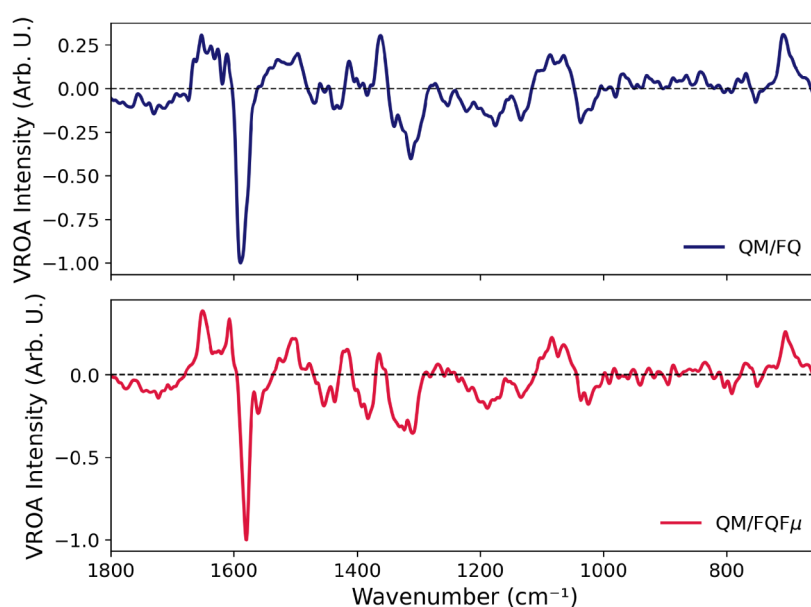


Figure 11. Computed ROA spectra of L-tryptophan (TRP) in aqueous solution using QM/FQ and QM/FQFμ approaches at the B3LYP/TZ2P level of theory, averaged over 200 snapshots.

minor difference, all remaining bands coincide in both positions and intensities, yielding two essentially identical spectra.

The computed Raman spectrum at 634 nm is reported in Figure 10. The dominant vibrational features substantially overlap with those observed in the IR spectrum. The band located at 1460 cm⁻¹ primarily involves symmetric stretching of the carboxylate group, coupled with aliphatic C–C stretching and out-of-plane C–H bending. The most intense band, located at 1360 cm⁻¹, is mainly attributed to aromatic C–C stretching, with additional contributions from in-plane C–H bending. Bands around 1200 cm⁻¹ and 900 cm⁻¹ are associated with mixed in-plane and out-of-plane bending, rocking, and stretching of aliphatic moieties. In particular, the intense band at approximately 1050 cm⁻¹ is associated with C–C and C–N stretching in combination with in plane aromatic C–H bending modes.

Finally, the 780 cm⁻¹ band, which is also observed in the IR spectrum, includes contributions from CO₂ deformation, aromatic out-of-plane bending, and rocking of CH₂ and NH₃⁺ groups. Similarly to IR, the Raman spectra reported in Figure 10 reveal only subtle differences between the QM/FQ and QM/FQFμ approaches. The main vibrational bands are well-aligned in both positions and relative intensities, with the QM/FQFμ spectrum displaying slightly sharper features.

As expected, the ROA spectrum (Figure 11) exhibits a pronounced alternation of positive and negative signals, underscoring the intrinsic complexity of chiroptical responses. This spectral variability reflects both the conformational flexibility of the solute and the dynamic nature of solvent organization, as sampled by MD simulations. The major ROA bands correspond closely to those in the IR and Raman spectra, with the most intense signal assigned to the asymmetric stretching of the carboxylate group, consistently

exhibiting a negative sign. Despite the intrinsic complexity of the ROA spectrum, the results reported in Figure 11 show a high level of consistency between the QM/FQ and QM/FQF μ models. Both approaches yield comparable band positions and reproduce the same sign pattern across the spectral range. In contrast to IR and Raman spectra—where the overall spectral shape remains essentially unchanged regardless of the number of snapshots, even though peak intensities and positions may vary—the ROA spectrum exhibits markedly slower convergence. As illustrated in Figures S15 and S18, using a limited number of snapshots can lead not only to intensity variations but also to incorrect sign predictions for key bands. This behavior further highlights the sensitivity of ROA signals to subtle conformational and environmental effects, reinforcing the notion that reliable chiroptical predictions require extensive configurational sampling. Similar predictive challenges have also been reported in previous studies. For instance, Blanch et al. demonstrated that the sign of the ROA band associated with the W3 vibration of the tryptophan indole ring (1550 cm⁻¹) is highly sensitive to the local stereochemistry and conformation of the side chain, leading to sign inversions across different protein environments.¹²⁴

We now move on to compare the computed results with the experimental data,⁹⁴ which are available only for IR and Raman spectra. Computed IR spectra obtained with both QM/FQ and QM/FQF μ approaches are in good agreement with the experimental spectrum reported in the literature.⁹⁴ The most intense band, corresponding to the asymmetric stretching of the carboxylate group, is centered around 1600 cm⁻¹ in both theoretical and experimental spectra, confirming the reliability of the simulations in capturing the dominant vibrational features. However, while both models slightly underestimate the width of the experimental band, the QM/FQF μ spectrum shows a broader and smoother profile, which is more similar to the experimental shape.

The band centered near 1400 cm⁻¹ in the simulated spectra also well reproduces the experimental profile. This region spans approximately from 1450 to 1350 cm⁻¹ and includes multiple resolved components, while in the computed spectra it appears as a broad band.

The experimental spectral region between 1300 and 900 cm⁻¹ is overall well reproduced by both computational approaches, with consistent band positions and relative intensities. The low-frequency region around 780 cm⁻¹, which is associated with out-of-plane bending and rocking motions, is also present in the experimental spectrum, though shifted to lower wavenumbers (744 cm⁻¹) and appearing with higher intensity. Despite the slight redshift, the computed band profile in this region is in good qualitative agreement with the experimental shape.

Simulated Raman spectra, both QM/FQ and QM/FQF μ , are in excellent agreement with experiments. The first band in the experimental spectrum, located at 1420 cm⁻¹, aligns well with the simulated bands. However, a key difference lies in the intensity: the experimental band is weaker, whereas in the simulated spectrum, this band is the most intense.

Similarly, the band at 1360 cm⁻¹ is present in both the experimental and computed spectra, with calculations accurately reproducing both the experimental band position and intensity.

Regarding the experimental bands at 1013 cm⁻¹ and 758 cm⁻¹, both are well reproduced in the computed spectra, both in terms of position and intensity. These bands are among the

most intense in the experimental Raman spectrum, and the simulations accurately reflect these features.

4. CONCLUSIONS

In this work, we have applied multiscale modeling^{12–15} techniques to compute a broad range of spectroscopic properties—electronic, magnetic, and vibrational—of zwitterionic L-tryptophan in aqueous solution. The solute is treated at a high level of theory using density functional theory (DFT), while solute–solvent interactions, including hydrogen bonding, are captured through atomistic polarizable embedding models, QM/FQ and QM/FQF μ .^{24,32,37,38} To account for the inherently dynamic nature of both solute and solvent in solution, extensive sampling is performed via classical molecular dynamics simulations by using GROMACS package.⁷⁵

We demonstrate that our computational protocol is not limited to a single spectroscopic observable but is broadly applicable across a variety of techniques. In particular, we have successfully simulated UV–vis and ECD spectra, OR values, NMR, IR, Raman, and ROA spectra. The UV spectroscopy results confirm that the bright S₀ → S₁ transition of tryptophan, which is observed experimentally¹⁰⁹ at 278 nm, is accurately reproduced by both QM/FQ and QM/FQF μ approaches, with a deviation of approximately 14 nm. NBO analysis,^{110,111} performed on the representative QM/COSMO structure, confirms the π → π^* character of the transition. The ECD spectrum is also successfully simulated, capturing the characteristic (–, +) sign pattern seen in the experimental data.¹²⁵ As a chiroptical property, ECD is highly sensitive to the environment, and this is reflected in the more pronounced differences between the QM/FQ and QM/FQF μ models compared to UV spectra. In particular, the peak intensities in the ECD spectrum are better reproduced by the QM/FQF μ approach, underscoring the importance of an accurate treatment of the solvent in simulating such environment-dependent properties.

We then extend our investigation to another fundamental chiroptical observable: OR. As with ECD, OR is highly sensitive to both the molecular environment and conformational fluctuations. Remarkably, despite the large frame-by-frame variability observed throughout the molecular dynamics trajectory, both QM/FQ and QM/FQF μ models are able to correctly capture the experimental negative sign of the optical rotation at the sodium D-line. The computed values, –64 degree/(dm · g/cm³) for QM/FQ and –71 degree/(dm · g/cm³) for QM/FQF μ , show good agreement with the experimental measurement of –31.5° at 23°C.¹¹⁶

To further broaden the scope of the analysis, magnetic properties are also investigated through the simulation of NMR spectra using both QM/FQ and QM/FQF μ models. The chemical shifts of hydrogen and carbon nuclei are computed and compared with available experimental data^{114,115} while for nitrogen and oxygen atoms, only theoretical shielding values are obtained due to the lack of experimental references. For hydrogen nuclei, the coupled ¹H NMR spectra are also simulated, including scalar spin–spin coupling constants. The resulting ¹H and ¹³C spectra show good agreement with experiment, and only minor differences are observed between the two solvation models, indicating a relatively low sensitivity of these nuclei to the specific environment description. In contrast, more significant variations are found for the oxygen shielding constants,

underscoring the importance of solvent treatment for accurately describing the magnetic responses of these nuclei.

Given the local nature of NMR observables and their strong dependence on the electronic density around each nucleus, a hybrid QM/FDE/FQF μ approach is also employed for this property.^{35,37,38} In this model, solvent molecules within a 3 Å radius of the solute are treated at the quantum level using a frozen density embedding (FDE) scheme, and freeze-and-thaw cycles are included to allow mutual polarization between the active and frozen regions. As expected, this refinement leads to a substantial improvement in the description of oxygen shielding, while only modest changes are observed for hydrogen and carbon, further confirming their lower sensitivity to the environment.

We also investigate vibrational properties, including IR, Raman, and ROA spectra, by applying QM/FQ and QM/FQF μ models, further confirming the general applicability of our protocol across a wide range of spectroscopic techniques. These spectroscopic techniques provide detailed insights into the vibrational structure and enable the characterization of both achiral and chiroptical responses. The analysis of the computed vibrational modes reveals that the main normal modes involved span a broad frequency range: at higher frequencies (1700–1300 cm⁻¹), we observe the symmetric and asymmetric stretching of the carboxylate group, C=C stretching, NH₃⁺ bending, and in-plane C–H bending. At lower frequencies (1300–700 cm⁻¹), the main contributions come from single bond stretching modes (C–C and C–N) and out-of-plane bending motions involving C–H and the carboxylate group. IR and Raman spectra are consistently reproduced by both QM/FQ and QM/FQF μ models, with only minor differences observed between them. For these two techniques, the computed results show good agreement with experimental data.⁹⁴ ROA spectra likewise show good consistency between the two models, particularly in terms of band positions and sign patterns; however, due to the absence of experimental ROA data for zwitterionic tryptophan in aqueous solution, we cannot verify whether the predicted sign alternation is accurate.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpca.5c04511>.

TRP configurations in aqueous solution, TRP dihedral distribution functions, convergence plots of spectroscopic properties, NBO analysis, computed TRP NMR chemical shifts, shieldings, and coupling constants, absolute differences between chemical shifts and shieldings obtained with different solvent models, study of FDE parameters, experimental NMR H and C chemical shifts, IR, Raman, and ROA stick spectra, and TRP normal modes. (PDF)

■ AUTHOR INFORMATION

Corresponding Author

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

Authors

Marco Trinari – *Scuola Normale Superiore, Pisa I-56126, Italy*

Chiara Sepali – *Scuola Normale Superiore, Pisa I-56126, Italy*

Tommaso Giovannini – *Department of Physics, University of Rome Tor Vergata, and INFN, Rome I-00133, Italy*;

orcid.org/0000-0002-5637-2853

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jpca.5c04511>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge the Center for High-Performance Computing (CHPC) at SNS for providing the computational infrastructure. MT thanks Sara Gomez and Piero Lafiosca for their guidance on MD simulations and QM/FDE calculations.

■ REFERENCES

- (1) Barone, V. *Computational strategies for spectroscopy: from small molecules to nano systems*; John Wiley & Sons, 2011.
- (2) Barone, V.; Alessandrini, S.; Biczysko, M.; Cheeseman, J. R.; Clary, D. C.; McCoy, A. B.; Di Risio, R. J.; Neese, F.; Melosso, M.; Puzzarini, C. Computational molecular spectroscopy. *Nat. Rev. Meth. Primers* **2021**, *1*, 38.
- (3) 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.
- (4) Maiuri, M.; Garavelli, M.; Cerullo, G. Ultrafast spectroscopy: State of the art and open challenges. *J. Am. Chem. Soc.* **2020**, *142*, 3–15.
- (5) Xie, F.; Mendolicchio, M.; Omarouyache, W.; Murugachandran, S. I.; Lei, J.; Gou, Q.; Sanz, M. E.; Barone, V.; Schnell, M. Structural and Electronic Evolution of Ethanolamine upon Microhydration: Insights from Hyperfine Resolved Rotational Spectroscopy. *Angew. Chem., Int. Ed.* **2024**, *63*, No. e202408622.
- (6) Cannelli, O.; Giovannini, T.; Baiardi, A.; Carlotti, B.; Elisei, F.; Cappelli, C. Understanding the interplay between the solvent and nuclear rearrangements in the negative solvatochromism of a push–pull flexible quinolinium cation. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32544–32555.
- (7) Cannelli, O.; Wiktor, J.; Colonna, N.; Leroy, L.; Puppini, M.; Bacellar, C.; Sadykov, I.; Krieg, F.; Smolentsev, G.; Kovalenko, M. V.; et al. Atomic-Level Description of Thermal Fluctuations in Inorganic Lead Halide Perovskites. *J. Phys. Chem. Lett.* **2022**, *13*, 3382–3391.
- (8) Provorse, M. R.; Isborn, C. M. Electron dynamics with real-time time-dependent density functional theory. *Int. J. Quantum Chem.* **2016**, *116*, 739–749.
- (9) Li, X.; Govind, N.; Isborn, C.; De Prince, A. E., III; Lopata, K. Real-time time-dependent electronic structure theory. *Chem. Rev.* **2020**, *120*, 9951–9993.
- (10) Lopata, K.; Govind, N. Modeling fast electron dynamics with real-time time-dependent density functional theory: Application to small molecules and chromophores. *J. Chem. Theory Comput.* **2011**, *7*, 1344–1355.
- (11) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664–5677.
- (12) 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.
- (13) 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.

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

(15) 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.

(16) Perrella, F.; Petrone, A.; Rega, N. Understanding charge dynamics in dense electronic manifolds in complex environments. *J. Chem. Theory Comput.* **2023**, *19*, 626–639.

(17) Wildman, A.; Donati, G.; Lipparini, F.; Mennucci, B.; Li, X. Nonequilibrium environment dynamics in a frequency-dependent polarizable embedding model. *J. Chem. Theory And Comput.* **2019**, *15*, 43–51.

(18) 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.

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

(20) Cammi, R. *Molecular Response Functions for the Polarizable Continuum Model: physical basis and quantum mechanical formalism*; Springer, 2013.

(21) Giovannini, T.; Cappelli, C. Continuum vs. atomistic approaches to computational spectroscopy of solvated systems. *Chem. Commun.* **2023**, *59*, 5644–5660.

(22) Olsen, J. M. H.; Kongsted, J. *Advances in quantum chemistry*; Elsevier, 2011, Vol. 61, pp. 107–143.

(23) 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.

(24) 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.

(25) 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.

(26) Hedegård, E. D.; Bast, R.; Kongsted, J.; Olsen, J. M. H.; Jensen, H. J. A. Relativistic polarizable embedding. *J. Chem. Theory Comput.* **2017**, *13*, 2870–2880.

(27) Gómez, S.; Giovannini, T.; Cappelli, C. Multiple facets of modeling electronic absorption spectra of systems in solution. *ACS Phys. Chem. Au* **2023**, *3*, 1–16.

(28) 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.

(29) Giovannini, T.; Olszówka, 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.

(30) Steinmann, C.; Reinholdt, P.; Norby, 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.

(31) 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.

(32) 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.

(33) 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.

(34) Gómez, S.; Lafiosca, P.; Giovannini, T. Modeling UV/Vis Absorption Spectra of Food Colorants in Solution: Anthocyanins and Curcumin as Case Studies. *Molecules* **2024**, *29*, 4378.

(35) Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.* **1993**, *97*, 8050–8053.

(36) Jacob, C. R.; Neugebauer, J.; Visscher, L. A flexible implementation of frozen-density embedding for use in multilevel simulations. *J. Comput. Chem.* **2008**, *29*, 1011–1018.

(37) 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.

(38) Lafiosca, P.; Rossi, F.; Egidi, F.; Giovannini, T.; Cappelli, C. Multiscale frozen density embedding/molecular mechanics approach for simulating magnetic response properties of solvated systems. *J. Chem. Theory Comput.* **2024**, *20*, 266–279.

(39) Quintela Rodriguez, F. E.; Lafiosca, P.; Giovannini, T.; Cappelli, C. Quantum dynamics of dissipative polarizable media. *Phys. Rev. B* **2025**, *111*, 235144.

(40) 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.

(41) Sepali, C.; Lafiosca, P.; Gómez, S.; Giovannini, T.; Cappelli, C. Effective fully polarizable QM/MM approaches to compute Raman and Raman Optical Activity spectra in aqueous solution. *Spectr. Acta A* **2024**, *305*, 123485.

(42) Giovannini, T.; Gómez, S.; Cappelli, C. Modeling Raman Spectra in Complex Environments: From Solutions to Surface-Enhanced Raman Scattering. *J. Phys. Chem. Lett.* **2025**, *16*, 3106–3121.

(43) 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.

(44) Gómez, S.; Ambrosetti, M.; Giovannini, T.; Cappelli, C. Close-Up Look at Electronic Spectroscopic Signatures of Common Pharmaceuticals in Solution. *J. Phys. Chem. B* **2024**, *128*, 2432–2446.

(45) 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.

(46) 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.

(47) Sepali, C.; Gómez, S.; Grifoni, E.; Giovannini, T.; Cappelli, C. Computational Spectroscopy of Aqueous Solutions: The Underlying Role of Conformational Sampling. *J. Phys. Chem. B* **2024**, *128*, 5083–5091.

(48) 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.

(49) 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.

(50) 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.

(51) 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.

(52) Csizi, K.-S.; Reiher, M. Universal QM/MM approaches for general nanoscale applications. *WIREs: Comput. Mol. Sci.* **2023**, *13*, No. e1656.

- (53) Palivec, V.; Pohl, R.; Kaminsky, J.; Martinez-Seara, H. Efficiently computing NMR ¹H and ¹³C chemical shifts of saccharides in aqueous environment. *J. Chem. Theory Comput.* **2022**, *18*, 4373–4386.
- (54) Palivec, V.; Kopecký, V.; Jungwirth, P.; Bouř, P.; Kaminský, J.; Martinez-Seara, H. Simulation of Raman and Raman optical activity of saccharides in solution. *Phys. Chem. Chem. Phys.* **2020**, *22*, 1983–1993.
- (55) Palivec, V.; Johannessen, C.; Kaminský, J.; Martinez-Seara, H. Use of Raman and Raman optical activity to extract atomistic details of saccharides in aqueous solution. *PLoS Comput. Biol.* **2022**, *18*, No. e1009678.
- (56) Mutter, S. T.; Zielinski, F.; Cheeseman, J. R.; Johannessen, C.; Popelier, P. L.; Blanch, E. W. Conformational dynamics of carbohydrates: Raman optical activity of D-glucuronic acid and N-acetyl-D-glucosamine using a combined molecular dynamics and quantum chemical approach. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6016–6027.
- (57) Hopmann, K. H.; Ruud, K.; Pecul, M.; Kudelski, A.; Dračinský, M.; Bour, P. Explicit versus implicit solvent modeling of Raman optical activity spectra. *J. Phys. Chem. B* **2011**, *115*, 4128–4137.
- (58) D’Cunha, R.; Crawford, T. D. Modeling complex solvent effects on the optical rotation of chiral molecules: a combined molecular dynamics and density functional theory study. *J. Phys. Chem. A* **2021**, *125*, 3095–3108.
- (59) Cheeseman, J. R.; Shaik, M. S.; Popelier, P. L.; Blanch, E. W. Calculation of Raman optical activity spectra of methyl- β -D-glucose incorporating a full molecular dynamics simulation of hydration effects. *J. Am. Chem. Soc.* **2011**, *133*, 4991–4997.
- (60) Kaminský, J.; Andrushchenko, V.; Bouř, P. Natural and magnetic circular dichroism spectra of nucleosides: effect of the dynamics and environment. *RSC Adv.* **2021**, *11*, 8411–8419.
- (61) Bensberg, M.; Türtscher, P. L.; Unsleber, J. P.; Reiher, M.; Neugebauer, J. Solvation free energies in subsystem density functional theory. *J. Chem. Theory Comput.* **2022**, *18*, 723–740.
- (62) Sanderson, R. An interpretation of bond lengths and a classification of bonds. *Science* **1951**, *114*, 670–672.
- (63) 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.
- (64) 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.
- (65) 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.
- (66) Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. Recent advances in wave function-based methods of molecular-property calculations. *Chem. Rev.* **2012**, *112*, 543–631.
- (67) Norman, P.; Ruud, K.; Saue, T. *Principles and practices of molecular properties: theory, modeling, and simulations*; John Wiley & Sons, 2018.
- (68) Helgaker, T.; Jaszunski, M.; Ruud, K. Ab initio methods for the calculation of NMR shielding and indirect spin-spin coupling constants. *Chem. Rev.* **1999**, *99*, 293–352.
- (69) Casida, M. E. *Recent Advances in Density Functional Methods*; World Scientific, 1995; pp. 155–192.
- (70) Frisch, M.; Head-Gordon, M.; Pople, J. Direct analytic SCF second derivatives and electric field properties. *Chem. Phys.* **1990**, *141*, 189–196.
- (71) Perrella, F.; Coppola, F.; Rega, N.; Petrone, A. An expedited route to optical and electronic properties at finite temperature via unsupervised learning. *Molecules* **2023**, *28*, 3411.
- (72) Petrone, A.; Cerezo, J.; Ferrer, F. J. A.; Donati, G.; Improta, R.; Rega, N.; Santoro, F. Absorption and emission spectral shapes of a prototype dye in water by combining classical/dynamical and quantum/static approaches. *J. Phys. Chem. A* **2015**, *119*, 5426–5438.
- (73) Mancini, G.; Del Galdo, S.; Chandramouli, B.; Pagliai, M.; Barone, V. Computational spectroscopy in solution by integration of variational and perturbative approaches on top of clusterized molecular dynamics. *J. Chem. Theory Comput.* **2020**, *16*, 5747–5761.
- (74) Zanetti-Polzi, L.; Aschi, M.; Daidone, I.; Amadei, A. Theoretical modeling of the absorption spectrum of aqueous riboflavin. *Chem. Phys. Lett.* **2017**, *669*, 119–124.
- (75) 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*, 19–25.
- (76) 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.
- (77) 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.
- (78) 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.
- (79) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- (80) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (81) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (82) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (83) Páll, S.; Hess, B. A flexible algorithm for calculating pair interactions on SIMD architectures. *Comput. Phys. Commun.* **2013**, *184*, 2641–2650.
- (84) Wiebeler, C.; Schapiro, I. QM/MM benchmarking of cyanobacteriochrome Slr1393g3 absorption spectra. *Molecules* **2019**, *24*, 1720.
- (85) Egdi, 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/PCM approach. *Opt. Mater. Express* **2015**, *5*, 196–209.
- (86) Gelabert, R.; Moreno, M.; Lluch, J. M. Predicting the electronic absorption band shape of azobenzene photoswitches. *Int. J. Mol. Sci.* **2023**, *24*, 25.
- (87) Bastida, A.; Zúñiga, J.; Fogolari, F.; Soler, M. A. Statistical accuracy of molecular dynamics-based methods for sampling conformational ensembles of disordered proteins. *Phys. Chem. Chem. Phys.* **2024**, *26*, 23213–23227.
- (88) Hudson, P. S.; Boresch, S.; Rogers, D. M.; Woodcock, H. L. Accelerating QM/MM free energy computations via intramolecular force matching. *J. Chem. Theory Comput.* **2018**, *14*, 6327–6335.
- (89) Pérez-Barcia, Á.; Cárdenas, G.; Nogueira, J. J.; Mandado, M. Effect of the QM size, basis set, and polarization on QM/MM interaction energy decomposition analysis. *J. Chem. Inf. Model.* **2023**, *63*, 882–897.
- (90) Pan, X.; Rosta, E.; Shao, Y. Representation of the QM subsystem for long-range electrostatic interaction in non-periodic ab initio QM/MM calculations. *Molecules* **2018**, *23*, 2500.
- (91) Pederson, J. P.; McDaniel, J. G. DFT-based QM/MM with particle-mesh Ewald for direct, long-range electrostatic embedding. *J. Chem. Phys.* **2022**, *156*, 174105.
- (92) Komjáti, B.; Urai, Á.; Hosztafi, S.; Kökösi, J.; Kováts, B.; Nagy, J.; Horváth, P. Systematic study on the TD-DFT calculated electronic circular dichroism spectra of chiral aromatic nitro compounds: A comparison of B3LYP and CAM-B3LYP. *Spectrochim. Acta, Part A* **2016**, *155*, 95–102.
- (93) Guillaume, M.; Ruud, K.; Rizzo, A.; Monti, S.; Lin, Z.; Xu, X. Computational study of the one-and two-photon absorption and

circular dichroism of (L)-tryptophan. *J. Phys. Chem. B* **2010**, *114*, 6500–6512.

(94) Leyton, P.; Brunet, J.; Silva, V.; Paipa, C.; Castillo, M. V.; Brandán, S. A. An experimental and theoretical study of L-tryptophan in an aqueous solution, combining two-layered ONIOM and SCRF calculations. *Spectrochim. Acta, Part A* **2012**, *88*, 162–170.

(95) Hssain, A. H.; Devi, T. G.; Salih, R. O.; Abdullah, N. R. Structural and spectroscopic study of L-tryptophan dimer state using DFT and MD: Computational and experimental analysis. *J. Mol. Struct.* **2025**, *1331*, 141582.

(96) Fatima, A.; Kumar, A.; Saral, A. K. A.; Muthu, S.; Afzal, M.; Haq, N.; Nazar, I.; Siddiqui, N.; Javed, S. Study of L-tryptophan (a neurotransmitter precursor): spectral, Hirshfeld surface, molecular docking and dynamics simulations. *Z. Phys. Chem.* **2023**.

(97) 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.

(98) Wolff, S.; Ziegler, T. Calculation of DFT-GIAO NMR shifts with the inclusion of spin-orbit coupling. *J. Chem. Phys.* **1998**, *109*, 895–905.

(99) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(100) Jacob, C. R.; Visscher, L. Calculation of nuclear magnetic resonance shieldings using frozen-density embedding. *J. Chem. Phys.* **2006**, *125*, 194104.

(101) Buló, R. E.; Jacob, C. R.; Visscher, L. NMR solvent shifts of acetonitrile from frozen density embedding calculations. *J. Phys. Chem. A* **2008**, *112*, 2640–2647.

(102) Götz, A. W.; Autschbach, J.; Visscher, L. Calculation of nuclear spin-spin coupling constants using frozen density embedding. *J. Chem. Phys.* **2014**, *140*, 104107.

(103) Wesolowski, T. A.; Weber, J. Kohn-Sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules. *Chem. Phys. Lett.* **1996**, *248*, 71–76.

(104) Vrije Universiteit AMS 2024.1, SCM, *Theoretical Chemistry*; Vrije Universiteit: Amsterdam, The Netherlands, 2024.

(105) Baerends, E. J.; Aguirre, N. F.; Austin, N. D.; Autschbach, J.; Bickelhaupt, F. M.; Buló, R.; Cappelli, C.; van Duin, A. C. T.; Egidi, F.; Fonseca Guerra, C.; et al. The Amsterdam Modeling Suite. *J. Chem. Phys.* **2025**, *162*, 162501.

(106) 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.

(107) Méndez-Hurtado, J.; Menéndez, M. I.; López, R.; Ruiz-López, M. F. An ab initio analysis of the structure of L-tryptophan tautomers in microhydrated environments, in water and in hydrophobic solvents. *Comput. Theor. Chem.* **2014**, *1034*, 17–25.

(108) Auer, H. E. Far-ultraviolet absorption and circular dichroism spectra of L-tryptophan and some derivatives. *J. Am. Chem. Soc.* **1973**, *95*, 3003–3011.

(109) Liu, H.; Zhang, H.; Jin, B. Fluorescence of tryptophan in aqueous solution. *Spectrochim. Acta, Part A* **2013**, *106*, 54–59.

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

(111) Glendening, E. D.; Landis, C. R.; Weinhold, F. NBO 6.0: Natural bond orbital analysis program. *J. Comput. Chem.* **2013**, *34*, 1429–1437.

(112) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

(113) Klamt, A. Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *J. Phys. Chem.* **1995**, *99*, 2224–2235.

(114) Wishart Research Group, *1H-NMR Spectrum of Tryptophan (HMDB)*. https://moldb.wishartlab.com/system/documents/files/000/031/283/original/041023_P00_05_HW_IDP_Assigned20121204-87231-1ne9758.png?1354661960, 2025; Accessed 9 April 2025.

(115) National Center for Biotechnology Information, *PubChem Compound Summary for CID 6305, Tryptophan*. <https://pubchem.ncbi.nlm.nih.gov/compound/Tryptophan>, 2025; Accessed 9 April 2025.

(116) Amino, *L-Tryptophan USP-39 Oral Grade*. <https://amino.de/fileadmin/mediadb/amino/produkte/specifications/aminoacids/L-Tryptophan-USP-39-oral-grade.pdf>, Accessed 26 June 2025.

(117) Jin, S.; Head, J. D. Theoretical investigation of molecular water adsorption on the Al (111) surface. *Surf. Sci.* **1994**, *318*, 204–216.

(118) 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.

(119) 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*, 1157.

(120) Strachan, A. Normal modes and frequencies from covariances in molecular dynamics or Monte Carlo simulations. *J. Chem. Phys.* **2004**, *120*, 1–4.

(121) Rega, N. Vibrational analysis beyond the harmonic regime from ab-initio molecular dynamics. *Theor. Chem. Acc.* **2006**, *116*, 347–354.

(122) Gaigeot, M.-P. Theoretical spectroscopy of floppy peptides at room temperature. A DFTMD perspective: gas and aqueous phase. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3336–3359.

(123) Thomas, M.; Brehm, M.; Fligg, R.; Vöhlinger, P.; Kirchner, B. Computing vibrational spectra from ab initio molecular dynamics. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608–6622.

(124) Blanch, E. W.; Hecht, L.; Day, L. A.; Pederson, D. M.; Barron, L. D. Tryptophan absolute stereochemistry in viral coat proteins from Raman optical activity. *J. Am. Chem. Soc.* **2001**, *123*, 4863–4864.

(125) Strickland, E. H.; Horwitz, J.; Billups, C. Fine structure in the near-ultraviolet circular dichroism and absorption spectra of tryptophan derivatives and chymotrypsinogen A at 77 K. *Biochemistry* **1969**, *8*, 3205–3213.