Computational Spectroscopy of Aqueous Solutions: The Underlying Role of Conformational Sampling

Chiara Sepali, Sara Gómez, Emanuele Grifoni [a], Tommaso Giovannini, and Chiara Cappelli*

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

E-mail: chiara.cappelli@sns.it

[a] present address: PQE Group - Località Prulli di Sopra, 103/c, 50066 Prulli di Sopra FI

Abstract

Fully atomistic multiscale polarizable quantum mechanics (QM)/molecular mechanics (MM) approaches, combined with techniques to sample the solute-solvent phase space, constitute the most accurate method to compute spectral signals in aqueous solution. Conventional sampling strategies, such as classical molecular dynamics (MD), may encounter drawbacks when the conformational space is particularly complex and transition barriers between conformers are high. This can lead to inaccurate sampling, which can potentially impact the accuracy of spectral calculations. For this reason, in this work, we compare classical MD with enhanced sampling techniques, i.e. Replica Exchange MD and Metadynamics. In particular, we show how the different sampling techniques affect computed UV, Electronic Circular Dichroism (ECD), Nuclear Magnetic Resonance (NMR) shieldings, and Optical Rotatory Dispersion (ORD) of *N*-acetylproline-amide in aqueous solution. Such a system is a model peptide characterized by complex conformational variability. Calculated values suggest that spectral properties are influenced by solute conformers, relative population, and solvent effects, therefore particular care needs to be paid for when choosing the sampling technique.

1 Introduction

The modeling of the spectral signals originating from a molecular system dissolved in an aqueous solution is extremely challenging, because of the presence of specific solute-solvent interactions, such as hydrogen bonding (HB).[1–5] The most reliable methods for these systems are quantum mechanics (QM)/molecular mechanics (MM) approaches, in which the solute is described at the QM level, while solvent molecules are modeled using a classical MM force field (FF).[1,6–10] These methods are generally coupled to a configurational sampling of the solute-solvent phase-space, allowing for the description of the dynamical aspects of the solvation phenomenon. The quality of computed spectral signals thus depends on the accuracy of the QM level, the quality of treatment of the QM/MM interaction, which should include polarization to get a physically consistent picture,[1,2,9–14] and the reliable sampling of the solute's configurational space. While the refinement of QM/MM interactions and the QM level has attracted much interest in the context of computational spectroscopy, the role of an accurate configurational sampling has generally been overlooked and performed employing classical Molecular Dynamics (MD) simulations.[15,16] The latter is a good compromise between computational cost and the accuracy of the sampling,[1,2,9,10,17] however, anytime a molecular system is characterized by multiple metastable states in equilibrium, the exploration of the whole phase space can result frustrated by the presence of kinetic bottlenecks.[18] Since reactive events may need high activation energies to happen, crossing these energy barriers can occur on a time scale that is hardly reachable with standard MD simulations, thus leading to incorrect sampling.[18]

A possible solution comes from *enhanced sampling* methods,[19–23] a class of algorithms specifically designed to accelerate the exploration of the configurational phase space in MD simulations. These methods can be classified into two major groups: (i) umbrella sampling[24] and metadynamics[25–27] that act by adding an extra term to the Hamiltonian of the system, designed to vanish any energy barrier placed in between the states of interest; (ii) methods such as Replica Exchange (RE)[19,28,29] that mix trajectories obtained from several parallel MD simulations of the same system at different temperatures. High-temperature MD simulations permit a faster barrier crossing and exploration of new states, while lower-temperature simulations ensure an accurate sampling within each state. Despite their potentialities, to the best of our knowledge, little attention has been paid so far to coupling *enhanced sampling* with QM/MM spectral calculations.

In this work, the configurational space is explored with three different sampling strategies, namely, classical MD, REMD parallel tempering (PT), and well-tempered metadynamics, by keeping the same FF for describing the solute-solvent system. To showcase the different conformational sampling arising from various searching techniques and how they can affect computed spectral properties, the peculiar case of aqueous *N*-acetylproline amide (NAP) is selected. NAP is a peptide model whose conformations are highly influenced by the solvent, the choice of the FF, and the searching technique, potentially affecting the accuracy of computed spectroscopic and response properties.[30–32] In line with previous studies on NAP,[31–33] spectral properties of increasing complexity from the theoretical point of view are considered: UV, Electronic Circular Dichroism (ECD), Nuclear Magnetic Resonance (NMR) shieldings, and Optical Rotatory Dispersion (ORD). In all spectroscopic calculations, NAP is treated at the QM level whereas the solvent is described employing the polarizable Fluctuating Charges (FQ) force field.[34]

2 Computational Details

2.1 Molecular dynamics (MD)

To sample the phase space of NAP in water solution, classical Molecular Dynamics (MD), temperature Replica Exchange or Parallel Tempering MD (REMD-PT), and Metadynamics are exploited. GROMACS 2021.1[35,36] is used to run MD simulations, together with a patch with PLUMED2[37] for enhanced sampling. The three strategies share parametrization, energy minimization, and thermalization stages but differ in the production phase. NAP is parametrized with the ff19SB force field[38] employing ACPYPE-Antechamber.[39,40] NAP and around 1600 TIP3P water molecules[41] are added to a cubic box of length 3.65 nm. As general simulation settings, trajectories are integrated via the leapfrog algorithm[42] with a time step of 2 fs applying periodic boundary conditions (PBC) in all directions. Short-range electrostatics and van der Waals interactions are cut off at 1.0 nm, and long-range electrostatics are calculated using the particle mesh Ewald (PME) method.[43] Hydrogen atoms are constrained using the LINCS algorithm.[44,45]

Energy minimization is done using the steepest descent minimizer[46] with a maximum step size of 0.01 nm and a tolerance value of 1000.0 kJ/mol/nm. Then, two short equilibrations are performed: an initial

NVT equilibration (V-rescale thermostat[47] with τ_T= 0.1 ps) at 300 K (or at the temperature of the respective replica in REMD-PT) for 100 ps, and a second NPT equilibration (V-rescale thermostat[47] with τ_T= 0.1 ps, Parrinello-Rahman barostat[48] with τ_P= 2 ps) at 300 K (or at the temperature of the respective replica in REMD-PT) and 1 bar for 100 ps. After equilibrating the systems, NPT production simulations are performed with common V-rescale thermostat,[47] τ_T= 0.1 ps, and Parrinello-Rahman barostat,[48] τ_P= 2 ps, P = 1 bar, for the three methods but changing conditions as follows:

- Classical MD: 300 ns NPT production simulations for both *cis* and *trans* NAP conformers.
- REMD-PT:[19,28,29] Multiple replicas or copies of the system are concurrently run. A total of 23 replicas are arranged on a ladder from 300 K to 415 K every 5 K and each simulation is performed for 100 ns at the corresponding temperature value. Atomic configurations are allowed to move up and down the ladder by periodically (here every 500 steps) attempting to exchange two neighboring configurations, with an acceptance probability given by the equations in Ref. 18.
- Metadynamics:[26,49–51] The dihedral angle Ψ is employed as the collective variable in 200 ns simulation. The next parameters are adopted for PLUMED2[37] settings: Gaussian hills heights: 1.2, Bias factor: 8, T= 300 K, Gaussian hills widths (s_d): 0.35, Hills deposition rate: 500.

All trajectories are analyzed using GROMACS tools, MDAnalysis[52] and TRAVIS[53,54] utilities.

2.2 Extraction of structures

To compute spectra, a set of uncorrelated configurations is extracted from the entire trajectories. The different snapshots are further cut in spheres of radius 15 Å. Unlike unbiased MD and REMD-PT, structures extracted from metadynamics trajectories are associated with a weight based on the conformation of the solute due to the biased potential added to the simulation. This weight is employed when calculating the spectral property. If just a few hundred snapshots are considered, big differences in those weights could occur (see Figure S1(a) in the Supporting Information (SI)), thus weighting the solvent distributions in distinct ways. To overcome that drawback, ten thousand snapshots are extracted and the spectra are calculated by using configurations with weight homogeneity, which in turn reassures homogeneity in the solvent arrangements (see Figure S1(b) in the SI), keeping the populations to those obtained in the metadynamics.

2.3 Spectral calculations and convergence

In the final stage of the protocol, QM/FQ[34] UV, ECD, ORD, and NMR calculations are performed on the extracted snapshots. The parametrization from Ref. 12 is exploited and the QM portion is treated at the CAMY-B3LYP/TZ2P level in the UV, ECD, ORD calculations and the B3LYP/DZP level in the NMR calculations. 10 excited states are converged in TD-DFT calculations. The raw ECD and UV data acquired from every snapshot are convoluted using a Gaussian function, with an FWHM of 0.20 eV. Subsequently, the individual spectra are averaged to get the final ones. The convergence of QM/FQ computed spectra is carefully examined by considering an increasing number of snapshots. All calculations are conducted using a modified version of the Amsterdam Modeling Suite (AMS), release 2020.202.[55]

3 Results and Discussion

NAP conformers, portrayed in Figure 1(a) as associated to ω, Φ, and Ψ dihedral angles, derive from *cis/trans* structures and the relative position of the proline and methyl amide fragments: *trans*–P_{II}, *trans*–3_{10}, *trans*–C_7, *cis*–P_{II}, *cis*–3_{10} (see also Table S1 in the SI). Extracted free-energy maps for ω and Ψ are shown in Figure 1(b). Starting from the *trans*–P_{II} conformer (the most likely to occur in aqueous solution, see Ref. 31), classical MD only identifies the three *trans* motifs in a production run of 300 ns. Such an atypically long MD for a small system serves the purpose of increasing the odds of overcoming the energy barriers and finding *cis* conformers. However, the results indicate that those events require a longer time scale to happen. Starting the classical MD with the *cis*–P_{II} conformer led to a richer landscape of the conformers, thus indicating that probably the *cis* to *trans* energy barrier is smaller than the one to go from the *trans* to the *cis* motifs and that classical MD-based conformers are thus dependent of the selected starting point (see Figure S2 in the SI). In contrast, the complete set of conformers is found when using REMD-PT (23 replicas from 300 to 415 K) and metadynamics (Ψ was the chosen collective variable, CV) with computational costs comparable to that of the classical MD.

The computed populations of all conformers as calculated from each MD simulation are listed in Table 1. All sampling techniques used find *trans*–P_{II} as the main conformer, with 89%, 73%, and 74% of the population for classical, REMD-PT, and metadynamics, respectively, in agreement with previous studies.[30,31] In terms

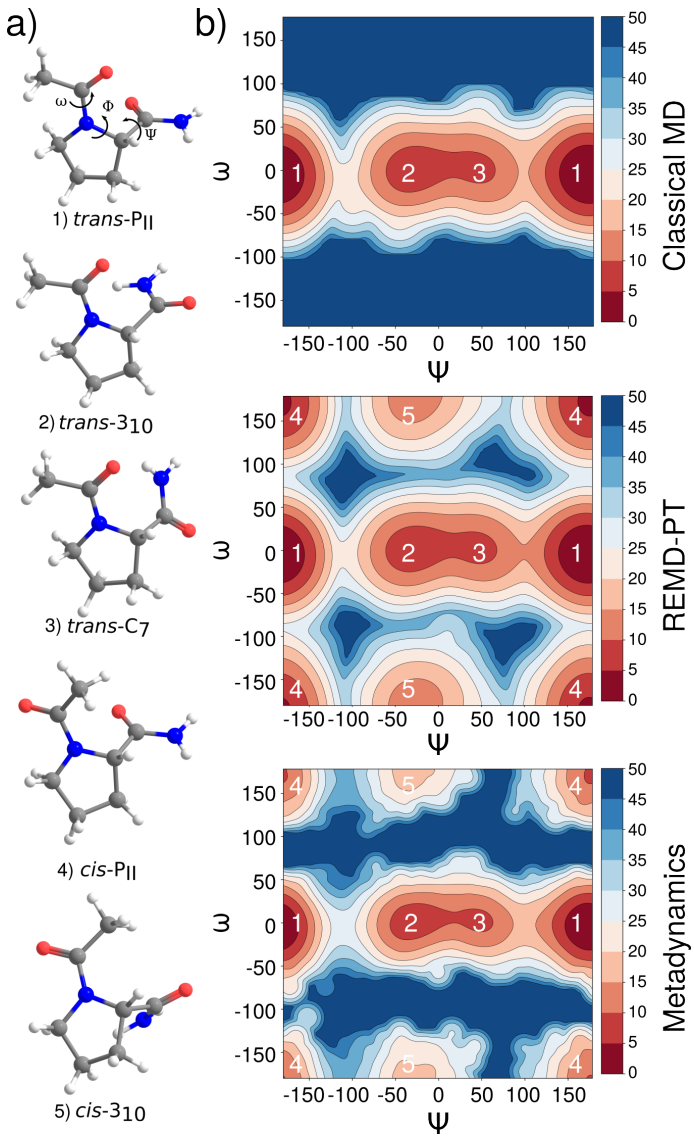

Figure 1: Conformational search of solvated NAP. (a) NAP conformers (1-5). (b) 2D free energy profiles as obtained with classical MD, REMD-PT, and metadynamics.

of the *cis* conformers, *cis*–P_{II} has a significant population and turns out to be the second most important conformer of NAP in water, as weighted by REMD-PT and metadynamics (16% and 15%). *cis*–3_{10} is instead sampled in both cases with a minor contribution. Overall, the similar populations described by REMD-PT and metadynamics give further reliability to the obtained values and highlight the inability of classical MD to provide a reliable sampling of the phase–space. Such a drastic difference can be also graphically appreciated in Figure 1(b), which highlights how classical MD does not sample specific regions of the solute phase-space.

Table 1: Relative populations of NAP conformers in aqueous solution obtained with three diverse sampling techniques

Conformer	Classical MD	Metadynamics	REMD-PT
trans–P_{II}	88.9	74.2	73.2
trans–3_{10}	7.5	6.5	6.6
trans–C_7	3.6	3.0	3.1
cis–P_{II}	-	15.4	16.3
cis–3_{10}	-	0.8	0.9

From the classical MD, metadynamics, and REMD-PT trajectories, three sets of uncorrelated structures are extracted and cut in spheres (see Ref. 1 for more details on the procedure). For each snapshot, the various spectral signals are computed at the QM/FQ level, which are then averaged to obtain the final spectra.

The UV, ECD, and OR (at 290 nm) stick spectra, as obtained from the raw QM/FQ computed data for the snapshots extracted from the three trajectories are depicted in Figure 2. Additional values for ^{1}H, ^{13}C, ^{15}N, and ^{17}O NMR chemical shieldings are reported in Tables S2 and S3 in the SI. Figure 2 reveals significant variability in the signals across different snapshots, emphasizing the dependence of the observed properties on both the solute arrangement and the distribution of the surrounding solvent molecules. Solvent-induced inhomogeneous broadening is particularly evident, and remarkably, it naturally arises from the computed spectrum without the need of any a-posteriori adjustment. Remarkably, ECD rotational strengths (see Figure 2) are both positive and negative, depending on both NAP conformation and the actual arrangement of water around it. Similarly, the majority of OR values at 290 nm are negative (see Figure 2), though some snapshots yield positive results. The change in sign across snapshots in both ECD and OR data reflects the complexity of these properties, in line with previous studies of some of us.[10,57–60] Also, while any variations in stick spectra across the three sample procedures are not discernible, differences emerge in the final averaged spectra (Figure 3), which, in case of ECD and ORD, result from a fine balance between +

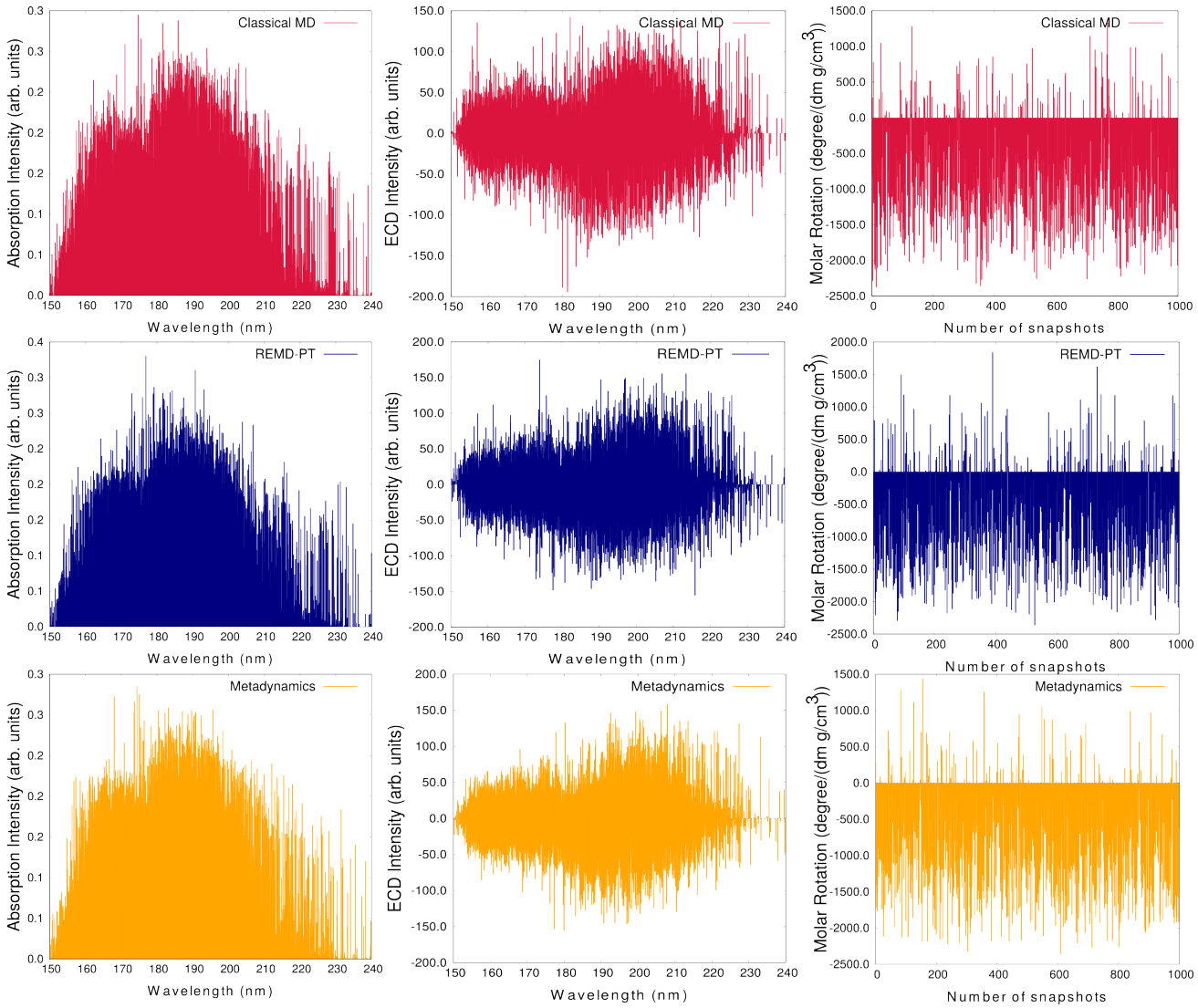

Figure 2: QM/FQ stick spectra (UV, ECD, and ORD data) as resulting from the three sampling methods: Classical MD, REMD-PT, and Metadynamics.

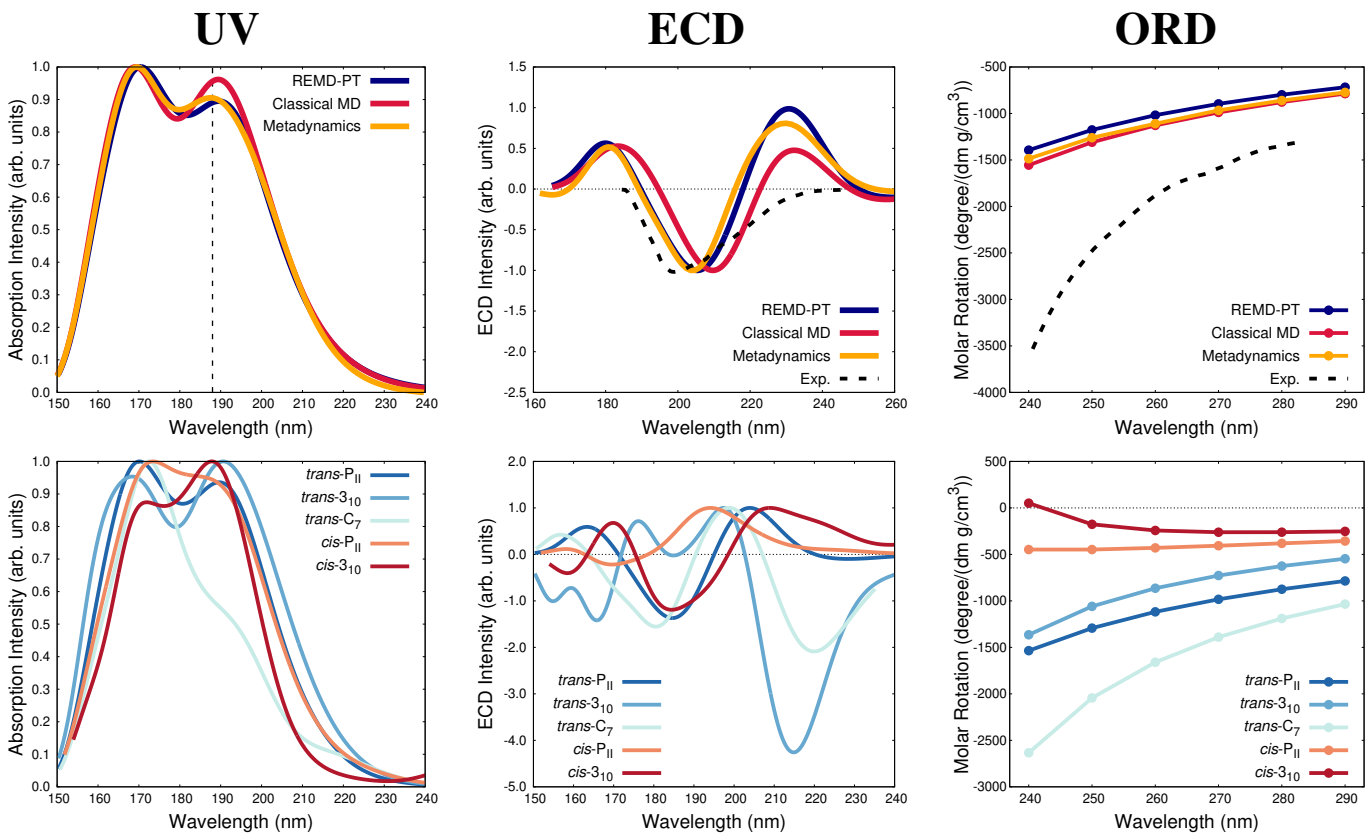

Figure 3: QM/FQ UV, ECD, and ORD of NAP in aqueous solution. Top panels: averaged convoluted spectra, as obtained with the three sampling techniques. Bottom panels: spectra for the various NAP conformers, as sampled by REMD-PT. Experimental data taken from Ref. 56 are also shown as dashed lines. The averaged ECD spectrum is scaled, so that the computed maximum absorption energy matches the experimental band.

and - values. As a consequence, an increasing number of snapshots is required to reach the convergence of computed spectra/values for UV (200 snapshots), NMR (400 snapshots), ECD (750 snapshots), and ORD (1000 snapshots) (see Figure S3 and Table S2 in the SI). It is also worth highlighting that to average solvent effects and guarantee a fair comparison between the three sampling techniques, for the snapshots extracted from metadynamics a homogeneous distribution of weightings is used for computing the property (further details are given in Section 2.2 in the Computational Details).

As it is shown in the top panel of Figure 3, computed UV spectra of NAP in aqueous solution are characterized by two main bands located at 168 and 190 nm, whereas ECD spectra are dominated by a (+,-,+) sign alternation with the negative band peaking at 210 nm. Canonical Molecular Orbital (MO) decomposition from Natural Bonding Orbital (NBO) analysis[61,62] indicates that the main bands are due to a $n \rightarrow \pi^*$ electronic transition, whose nature involves the orbitals displayed in Figure **??** in the SI, which

are mainly located in the carbonyl and amide regions of the molecule. ORD in the 240-290 nm range has a monotonic behavior until apparently reaching a plateau at different values depending on MD runs, but all above -1000 degree/(dm g/cm^3). Final computed spectra directly relate to the different depictions of NAP conformers as arising from the various MDs (see Figure 1(a)). To analyze the specific signals of each conformer, individual spectra are shown in Figure 3, bottom panel, for the structures sampled by REMD-PT. Since in REMD-PT and metadynamics trajectories *cis*–P_{II} and *trans*–3_{10} are the second and third most populated conformers, they are expected to have a major impact on final spectra/data, along with the *trans*–P_{II} structure, which is the most populated conformer in all simulations. Indeed, in the UV the presence of *cis*–P_{II} changes the relative intensities between the two bands with respect to *trans*–P_{II}, while redshifting the band at 168 nm, whereas the *trans*–3_{10} yields a blue-shift of the same band (see Figure 3). In ECD, the signal of *trans*–3_{10} counterbalances the intensity of *trans*–P_{II}. As a result, *cis*–P_{II} lowers the band intensity at 168 nm and 188 nm, while keeping the positive sign and intensity of the band at 205 nm. The latter is blue-shifted by both *trans*–3_{10} and *cis*–P_{II}. Additionally, in the ORD, *trans*–3_{10} and *cis*–P_{II} contributions drastically decrease the (absolute) value of the *trans*–P_{II} molar rotation, moving from -1555 (Classical MD) to -1394 (REMD-PT) for an incident wavelength of 240 nm. Thus, changing the sampling technique can yield differences raising up to 10 % in ORD calculations.

Conformer-dependent results, depicted in Figure 3 indicate that each NAP motif has a unique spectrum, especially for ECD and ORD. Indeed, each conformer affects both spectral shapes and sign patterns in ECD, relative intensities, and position of the absorption maximum wavelengths in UV, while sign and magnitude in ORD. Note that the *cis*–P_{II} conformer, which is the second most populated conformed in REMD-PT, shows the most drastic changes in sign for ECD and in shape for UV. Concerning ORD, the two *cis* conformers exhibit the least negative molar rotation values at the studied wavelengths.

The same conclusions can be drawn by taking conformers from the other sampling techniques as long as they have a common conformer set. In fact, for the *trans*–P_{II} conformer, UV, ECD, and ORD spectra from classical MD, metadynamics, and REMD-PT (see Figure **??** in the SI) very similar, keep the main features and resemble the final averaged spectra of Figure 3, top panel. Indeed, considering *trans*–P_{II} spectra obtained by the three sampling techniques, we note that slight differences in UV and ECD spectra are present (not affecting the main features, such as sign patterns), while ORD signals are almost superimposable (see

Figure **??**). Therefore, regardless of the sampling technique, the same conformer provides similar spectral profiles. From a physico-chemical point of view, it means that the three MDs can catch similarly both the variability of the solvent arrangement and the degrees of freedom of the target (e.g. the main three dihedral angles, the rotation of the methyl group and the flexibility of the ring part). Such findings are supported by the Radial Distribution Functions (RDFs, *g(r)*, Figure **??** in the SI) for those atoms potentially involved in HB networks with the solvent and the Dihedral Distribution Functions (DDFs, see Figure **??** in the SI). Indeed, the differences in the computed spectra obtained from the three MDs are intimately tied to the slight differences in RDFs and DDFs resulting from the three MDs.

Let us move to comparing our findings with available experiments and previous computational work.[31,32] As it can be noticed from Figure 3, computed spectra of NAP in water show a good agreement with experimental data. The actual accuracy depends on the complexity of the property and the sampling procedure. We note that the wavelength shift with respect to the experiment can be ascribed to the selected level of theory. For ORD, the lower values reported in our work, as compared with experiments, are mainly related to the blueshift in the ECD maximum reported by our calculations. With respect to continuum solvation approaches,[32] significant lowering of the OR at 280 nm is observed (-845 vs. -1976 degree/dm g/cm^3). This is due to the atomistic picture we are providing and the subsequent presence of the *cis* conformers.[32] In fact, the presence of *cis*–P_{II} and *cis*–3_{10} conformers, that are not found in Ref. 32, makes the OR values less negative, thus shifting the final value towards the experimental data.

Moving on to NMR shieldings (see Table S3 in the SI) ^{1}H-NMR, ^{13}C-NMR, and ^{15}N-NMR results averaged over the 400 snapshots extracted from the three MD simulations are very similar and in good agreement with the available experiments.[30] As expected, our data substantially differ from previous work[31] exploiting continuum solvation models in case of N and H atoms of the amide group, which mostly interact *via* HB with the water environment. This is primarily due to the atomistic picture provided by the QM/MM+MD methodology (see also RDFs in Figure **??** in the SI). Remarkably, for ^{17}O a huge discrepancy with respect to the continuum solvation picture[31] is reported. This is also in line with the RDFs reported in Figure **??** in the SI, which clearly show that the two oxygen atoms strongly interact with the surrounding water molecules. Notably, for the amide group oxygen atom, the classical MD sampling yields a smaller chemical shift (17.78 ppm) than metadynamics (22.52 ppm) and REMD-PT (21.50 ppm). This can be ascribed to the presence

of *cis* conformers in both metadynamics and REMD-PT samplings, clearly affecting the final computed chemical shift for this oxygen atom, which is involved in the dihedral rotation. Finally, the QM/FQ+MD approach can catch the similarity in H8, H9 and H10, which is instead missing by using continuum solvation. This is most probably due to a better description of the rotating methyl group achieved by employing atomistic simulations.

4 Conclusions

The data reported in this paper show that the choice of the configurational sampling strategy, which is generally overlooked, represents a delicate task in the modeling of spectroscopies of aqueous systems. Conventional MD has some shortcomings, such as being dependent on the starting point of the PES and unable to overcome highly energetic barriers. Enhanced-sampling methods are the solution to thoroughly explore the PES and find ''hidden" conformers. In the case of NAP, there are high energy barriers when going from the *trans* to the *cis* conformers. In addition, reliable modeling of the distribution of the solvent molecules around the solute is needed to grasp the dynamic aspect of the solvation. This can be reached in different ways, either by extracting snapshots with the same weight (as in classical MD and REMD-PT) or by associating a weight (inherited from the weight given to the solute) to each one (as in metadynamics). Despite the similarity in the conformers found with the enhanced sampling methodologies, this study reveals that REMD-PT is a more robust method because it does not need to use the corrections related to the weight of snapshots. Yet, metadynamics can help find the whole set of conformers to start subsequent classical MDs with each motif. By calculating the spectra at the QM/FQ level based on different sampling techniques, this work provides the first in-depth analysis of how specific sampling strategies can affect computed spectral signals of aqueous solutions. The structures extracted from the diverse MD runs are associated with diverse spectral signals, with discrepancies increasing with the complexity of the property under investigation. Overall, the extent of the effect in the final spectrum is mainly associated with the conformer population, which is drastically affected by the sampling strategy.

Finally, since the effect of specific conformers on a given spectral signal cannot be known *a-priori* applying enhanced methodologies from the beginning of the protocol to better explore the solute-solvent

phase space might be a solution. Building upon all these findings, further studies will be able to determine the impact of the enhanced sampling techniques on other systems.

5 Supporting Information

NAP conformers obtained from 300 ns classical MD runs with cis-P_{II} as a starting point, convergence plots, NPA orbitals, NAP trans-P_{II} computed properties, NAP-water RDFs and DDFs, NMR data.

Acknowledgement

We gratefully acknowledge the Center for High-Performance Computing (CHPC) at SNS for providing the computational infrastructure. CC acknowledges the support of the European Union by the Next Generation EU project ECS00000017 'Ecosistema dell'Innovazione' Tuscany Health Ecosystem (THE, PNRR, Spoke 4: Nanotechnologies for diagnosis and therapy).

References

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

(2) Gómez, S.; Giovannini, T.; Cappelli, C. Multiple facets of modeling electronic absorption spectra of systems in solution. *ACS Phys. Chem. Au.* **2022**, *3*, 1–16.

(3) Cramer, C. J.; Truhlar, D. G. Implicit solvation models: equilibria, structure, spectra, and dynamics. *Chem. Rev.* **1999**, *99*, 2161–2200.

(4) Orozco, M.; Luque, F. J. Theoretical methods for the description of the solvent effect in biomolecular systems. *Chem. Rev.* **2000**, *100*, 4187–4226.

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

(6) 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.

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

(8) 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.

(9) 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.

(10) 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.

(11) 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.

(12) 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.

(13) Giovannini, T.; Olszowka, 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.

(14) 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.

(15) Brooks, I., Charles L.; Case, D. A.; Plimpton, S.; Roux, B.; van der Spoel, D.; Tajkhorshid, E. Classical molecular dynamics. *J. Chem. Phys.* **2021**, *154*, 100401.

(16) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford university press, 2017.

(17) 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.

(18) Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annu. Rev. Phys. Chem.* **2016**, *67*, 159–184.

(19) Earl, D. J.; Deem, M. W. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910–3916.

(20) Bussi, G.; Branduardi, D. *Reviews in Computational Chemistry Volume 28*; John Wiley & Sons, Ltd, 2015; Chapter 1, pp 1–49.

(21) Yang, Y. I.; Shao, Q.; Zhang, J.; Yang, L.; Gao, Y. Q. Enhanced sampling in molecular dynamics. *J. Chem. Phys.* **2019**, *151*, 070902.

(22) Abrams, C.; Bussi, G. Enhanced Sampling in Molecular Dynamics Using Metadynamics, Replica-Exchange, and Temperature-Acceleration. *Entropy* **2014**, *16*, 163–199.

(23) Bernardi, R. C.; Melo, M. C.; Schulten, K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim. Biophys. Acta* **2015**, *1850*, 872–877.

(24) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.

(25) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. (USA)* **2002**, *99*, 12562.

(26) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 020603.

(27) Dama, J. F.; Parrinello, M.; Voth, G. A. Well-tempered metadynamics converges asymptotically. *Phys. Rev. Lett.* **2014**, *112*, 240602.

(28) Hansmann, U. H. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* **1997**, *281*, 140–150.

(29) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(30) Lee, K.-K.; Hahn, S.; Oh, K.-I.; Choi, J. S.; Joo, C.; Lee, H.; Han, H.; Cho, M. Structure of N-acetylproline amide in liquid water: experimentally measured and numerically simulated infrared and vibrational circular dichroism spectra. *J. Phys. Chem. B* **2006**, *110*, 18834–18843.

(31) Cappelli, C.; Mennucci, B. Modeling the Solvation of Peptides. The Case of (s)- N -Acetylproline Amide in Liquid Water. *J. Phys. Chem. B* **2008**, *112*, 3441–3450.

(32) Egidi, F.; Barone, V.; Bloino, J.; Cappelli, C. Toward an Accurate Modeling of Optical Rotation for Solvated Systems: Anharmonic Vibrational Contributions Coupled to the Polarizable Continuum Model. *J. Chem. Theory Comput.* **2012**, *8*, 585–597.

(33) Cappelli, C.; Monti, S.; Scalmani, G.; Barone, V. On the Calculation of Vibrational Frequencies for Molecules in Solution Beyond the Harmonic Approximation. *J. Chem. Theory Comput.* **2010**, *6*, 1660–1669.

(34) 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.

(35) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(36) 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.

(37) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604–613.

(38) Tian, C.; Kasavajhala, K.; Belfon, K. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q. ff19SB: Amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J. Chem. Theory Comput.* **2019**, *16*, 528–552.

(39) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE-Antechamber python parser interface. *BMC research notes* **2012**, *5*, 1–8.

(40) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247–260.

(41) 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.

(42) Hockney, R. W.; Goel, S.; Eastwood, J. Quiet high-resolution computer models of a plasma. *J. Comput. Phys.* **1974**, *14*, 148–158.

(43) 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.

(44) 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.

(45) Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116–122.

(46) Haug, E.; Arora, J.; Matsui, K. A steepest-descent method for optimization of mechanical systems. *J. Optim. Theory Appl.* **1976**, *19*, 401–424.

(47) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(48) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(49) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–12566.

(50) Iannuzzi, M.; Laio, A.; Parrinello, M. Efficient exploration of reactive potential energy surfaces using Car-Parrinello molecular dynamics. *Phys. Rev. Lett.* **2003**, *90*, 238302.

(51) Laio, A.; Rodriguez-Fortea, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. Assessing the accuracy of metadynamics. *J. Phys. Chem. B* **2005**, *109*, 6714–6721.

(52) 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.

(53) 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.

(54) Brehm, M.; Thomas, M.; Gehrke, S.; Kirchner, B. TRAVIS—A free analyzer for trajectories from molecular simulation. *J. Chem. Phys.* **2020**, *152*, 164105.

(55) Baerends, E.; Ziegler, T.; Atkins, A.; Autschbach, J.; Baseggio, O.; Bashford, D.; Bérces, A.; Bickelhaupt, F.; Bo, C.; Boerrigter, P.; et. al, ADF *(version 2020.202)*, Software for Chemistry & Materials, SCM. 2020; `http://www.scm.com`, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands.

(56) Madison, V.; Schellman, J. Location of proline derivatives in conformational space. I. Conformational calculations; optical activity and NMR experiments. *Biopolymers* **1970**, *9*, 511–567.

(57) 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.

(58) 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.

(59) Paolino, M.; Giovannini, T.; Manathunga, M.; Latterini, L.; Zampini, G.; Pierron, R.; Léonard, J.; Fusi, S.; Giorgi, G.; Giuliani, G.; et al., On the transition from a biomimetic molecular switch to a rotary molecular motor. *J. Phys. Chem. Lett.* **2021**, *12*, 3875–3884.

(60) 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 Explicit-Continuum approaches. *J. Phys. Chem. A* **2015**, *119*, 5396–5404.

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

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