



HAL
open science

Vibrational circular dichroism spectroscopy with a classical polarizable force field: alanine in the gas and condensed phases

Jessica Bowles, Sascha Jähnigen, Federica Agostini, Rodolphe Vuilleumier,
Anne Zehnacker, Florent Calvo, Carine Clavaguéra

► **To cite this version:**

Jessica Bowles, Sascha Jähnigen, Federica Agostini, Rodolphe Vuilleumier, Anne Zehnacker, et al.. Vibrational circular dichroism spectroscopy with a classical polarizable force field: alanine in the gas and condensed phases. *ChemPhysChem*, 2024, e202300982, 10.1002/cphc.202300982 . hal-04489317

HAL Id: hal-04489317

<https://universite-paris-saclay.hal.science/hal-04489317>

Submitted on 7 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Vibrational circular dichroism spectroscopy with a classical polarizable force field: alanine in the gas and condensed phases

Jessica Bowles,^[a] Sascha Jähnigen,^[b] Federica Agostini,^[a] Rodolphe Vuilleumier,^[b] Anne Zehnacker,^[c] Florent Calvo,^{*[d]} Carine Clavaguéra,^{*[a]}

Polarizable force fields are an essential component for the chemically accurate modeling of complex molecular systems with a significant degree of fluxionality, beyond harmonic or perturbative approximations. In this contribution we examine the performance of such an approach for the vibrational spectroscopy of the alanine amino acid, in the gas and condensed phases, from the Fourier transform of appropriate time correlation functions generated along molecular dynamics (MD) trajectories. While the infrared (IR) spectrum only requires the electric dipole moment, the vibrational circular dichroism (VCD) spectrum further requires knowledge of the magnetic dipole moment, for which we provide relevant expressions to be used under polarizable force fields. The AMOEBA force field was employed here to model alanine in the neutral and zwitterionic isolated forms, solvated by water or nitrogen, and as a crystal. Within this framework, comparison of the electric and magnetic dipole moments to those obtained with nuclear velocity perturbation theory based on density-functional theory for the same MD trajectories are found to agree well with one another. The statistical convergence of the IR and VCD spectra is examined and found to be more demanding in the latter case. Comparisons with experimental frequencies are also provided for the condensed phases.

Introduction

Vibrational circular dichroism (VCD) is the differential absorption between left- and right-circularly polarized light in

the infrared (IR) range. It is an increasingly popular non-invasive technique to determine the absolute configuration of chiral molecules,^[1] including polypeptides,^[2–9] but also provides a wealth of information about solvation^[10] and is also useful for crystals.^[11–16]

The popularity of VCD as an experimental technique is largely due to progress in the sensitivity of the measurements, but also to the availability of accurate theories to compare and assess the observations. Historically, this was achieved with density-functional theory (DFT) and the implementation of rotational activity in popular quantum chemistry codes based on magnetic field perturbation theory.^[17,18] Such static calculations have become relatively routine, recent developments being essentially tied to the accuracy of the underlying quantum chemical method (improvements in the density functional itself, dispersion corrections, etc.) as well as corrections to account for the environment^[19–22] or anharmonicities in a perturbative way.^[23,24]

Highly fluxional systems are still a challenge for quantum chemistry, and one way to account for fluxionality, or the likely coexistence between multiple conformers, consists of summing their contributions by appropriate weighting, typically through Boltzmann factors.^[25] Such an approach is particularly relevant for solvated systems, in which fluxionality also arises due to the motion of the solvent molecules around the solute. Treating explicitly the first solvation shell this way, and more implicitly the remainder of the solvent in a continuum fashion, has led to the so-called cluster-in-a-liquid model.^[26] This approach remains limited in its approximate treatment of intrinsic anharmonicities, and struggles to describe very floppy molecules or systems that interact strongly with their environment.^[10] Molecular dynamics (MD) simulations carried out directly at the first-principles level of electronic structure can alleviate such difficulties and naturally account for anharmonic and temperature effects, the nuclear dynamics being treated classically.^[27,28]

In the context of VCD spectroscopy, nuclear velocity perturbation theory (NVPT) has been developed to be used with first principles MD (FPMD), the resulting method^[29–31] being successfully applied in the gas^[31] and condensed phases.^[32,33] One inherent limitation of FPMD methods is their computational cost, which can be reduced using hybrid schemes such as QM/MM, the force field (FF) part being usually of the nonpolarizable type, and a spectator to the problem of computing the infrared spectra of the solute.^[19,20,22] The use of a polarizable FF has occasionally been reported in the context of VCD spectroscopy at the QM/MM level.^[21,22]

Alternatively, polarizable force fields have become

[a] Dr. J. Bowles, Dr. F. Agostini, Dr. C. Clavaguéra*
Université Paris-Saclay, CNRS, Institut de Chimie Physique
UMR8000, 91405, Orsay, France
E-mail: carine.clavaguera@universite-paris-saclay.fr

[b] Dr. S. Sascha Jähnigen, Prof. R. Vuilleumier
PASTEUR Laboratory, Département de Chimie, Ecole Normale
Supérieure, PSL University, Sorbonne Université, CNRS, 75005
Paris, France

[c] Dr. A. Zehnacker
Université Paris-Saclay, CNRS, Institut des Sciences Moléculaires
d'Orsay UMR8214, 91405, Orsay, France

[d] Dr. F. Calvo*
Université Grenoble Alpes, CNRS, LIPhy, F-38000 Grenoble,
France
E-mail: florent.calvo@univ-grenoble-alpes.fr

sufficiently reliable to allow for the realistic modeling of entire infrared spectra, both in the gas^[34–37] and condensed^[34,35,38–40] phases. However, so far only non-polarizable FFs have been employed to determine VCD spectra,^[19,41,42] whereas suitably parametrized polarizable FFs can handle both large systems, long sampling times, still remaining chemically accurate. The purpose of the present article is to extend the framework of polarizable force fields to the calculation of VCD spectra. The key ingredient to such VCD simulations lies in the determination of the magnetic dipole moment, which is an intrinsic dynamical quantity that for nonpolarizable FFs can be represented classically through the motion of the point charges on atomic sites.^[19,41]

Additional terms arise for a polarizable system, for which the induced electric dipole moment can also contribute to the magnetic dipole moment. The methodology proposed here extends the classical treatment of magnetic dipole moments^[19,41,42] to include several contributions arising from the induced electric dipole moment and its time derivative, as inspired from earlier work based on the nuclear velocity perturbation theory approach to VCD spectroscopy.^[31] As the scheme we develop is meant to be accurate, it is important to benchmark the electric and magnetic dipole moments against calculations performed at a higher level of theory. Here these ingredients entering the VCD calculation were thus compared to their corresponding values recalculated using DFT.

As a testing ground of our method, we have chosen to apply it to the alanine amino acid, in various phases covering the isolated molecule (in both neutral and zwitterionic states), but also condensed as a crystal, hydrated, or embedded in the much more inert nitrogen solvent. These chemically and physically diverse environments provide a wealth of situations under which we can assess not only the performance of the FF, but also the conditions for statistical convergence of the calculated spectra, thereby getting useful insight into reproducibility conditions that are appropriate to each phase. The AMOEBA FF^[43] was chosen as the main engine of our simulations, but the computational methodology is general to arbitrary force fields equipped with induced electric dipole moments. NVPT was chosen as our reference ab initio methods, but other approaches for computing VCD spectra from MD with an explicit account of electronic structure are also available.^[18,44]

The remainder of the article is organized as follows. The methodology is presented in the next section, including the general approach to VCD spectroscopy from time correlation functions and the determination of the magnetic dipole moment in MD simulations. Then, the results obtained in the gas and condensed phases are presented and discussed in comparison with experimental data. Finally, the last section concludes the article by suggesting some further extensions.

Methodology

This section lays out the main principles underlying the calculation of VCD spectroscopy from MD simulations, and the specific implementation of polarizable force fields. It also gives the numerical details of our simulations based on the AMOEBA FF.

IR and VCD absorption spectroscopies from time correlation functions

While IR spectroscopy measures the response of the electric dipole moment to an external excitation, the VCD signal originates from simultaneous changes in electric and magnetic dipole moments of the molecule due to vibrational transition.^[45] The latter corresponds to the differential absorption of right- and left-circularly polarized light. In the present work, both the IR and VCD spectral intensities are obtained from the Fourier transform of time correlation functions $C_{AB}(\tau) = \langle A(t)B(t+\tau) \rangle$ of appropriate dynamical variables $A(t)$ and $B(t+\tau)$, through the Wiener-Khinchine theorem.^[46] Here the ensemble average $\langle \cdot \rangle$ is taken over a sample of independent initial conditions at time zero or, given the ergodic hypothesis, over successive configurations of a single but very long trajectory.

The IR intensity is given by:^[32,36]

$$A_{\mu}(\omega) = \frac{2\pi\beta\omega^2}{3Vcn(\omega)} \int_{-\infty}^{+\infty} \langle \vec{\mu}(0) \cdot \vec{\mu}(\tau) \rangle e^{-i\omega\tau} d\tau, \quad (1)$$

with ω being the temporal frequency, V the volume of the sample, $n(\omega)$ the refraction index of the medium at temperature $T = 1/k_B\beta$ and c the velocity of light. Integration by parts straightforwardly shows that the IR spectral line shape can also be obtained from the time autocorrelation function of the time derivative of $\vec{\mu}$, also called the velocity form of the electric dipole moment: $\dot{\vec{\mu}}(t) = d\vec{\mu}(t)/dt$, giving equivalently:

$$A_{\dot{\mu}}(\omega) = \frac{2\pi\beta}{3Vcn(\omega)} \int_{-\infty}^{+\infty} \langle \dot{\vec{\mu}}(0) \cdot \dot{\vec{\mu}}(\tau) \rangle e^{-i\omega\tau} d\tau. \quad (2)$$

This form, based on the hypervirial theorem, avoids the origin dependence that arises through the positions alone.^[1]

In addition to the electric dipole moment $A = \vec{\mu}$ or its derivative $\dot{\vec{\mu}}$, the VCD signal requires, for C_{AB} , the magnetic dipole moment $B = \vec{m}$ that allows the rotation of the electric current around the axes of propagation, characteristic of circularly polarized light, to be considered. The resulting expression for the VCD absorption intensity reads:^[14,19,41,42]

$$\Delta A_{\dot{\mu}}(\omega) = \frac{8\pi\beta\omega}{3Vcn(\omega)} \int_{-\infty}^{+\infty} \langle \dot{\vec{\mu}}(0) \cdot \vec{m}(\tau) \rangle e^{-i\omega\tau} d\tau. \quad (3)$$

For the VCD spectra as well, the velocity form of the electric dipole moment will be employed since time correlation functions with $\dot{\vec{\mu}}$ decay faster than those with $\vec{\mu}$. This is especially useful for finite time simulations like in FPMD used in previous articles for the calculation of VCD spectra.^[32]

Electric and magnetic dipole moments from a polarizable force field

The dynamics of our molecular system is described by a polarizable atomistic force field in which the potential energy and gradient assume a particular form that explicitly accounts for multipolar electrostatics, including many-body polarization through interacting induced dipoles.^[47]

For any configuration of the system, the total electric dipole moment $\vec{\mu}$ contains a contribution at lowest order from the partial charges $\{q_i\}$ distributed at the respective

positions \vec{r}_i , and a contribution of the induced dipole moment $\vec{\mu}_i^{\text{ind}}$ on each polarizable site i :

$$\vec{\mu}(t) = \sum_{i=1}^N \left(q_i \vec{r}_i(t) + \vec{\mu}_i^{\text{ind}}(t) \right), \quad (4)$$

where N denotes the total number of atoms. It is this induced dipole moment that permits the description of polarization effects and is added compared to pure, fixed point charge models. The first term of electric dipole moment $\vec{\mu}$ is expressed here as a function of the position, thus this electric dipole moment corresponds to the position form. If the electric dipole moment is differentiated with respect to time, $\dot{\vec{\mu}} = \partial \vec{\mu} / \partial t$, the first term of the rhs of Eq. (4) becomes a function of the velocity, leading to the so-called velocity form of the electric dipole moment. However, there is no simple form for the time derivative of the induced dipole moment and numerical derivatives are used instead to calculate the time autocorrelation function needed to determine the IR absorption spectrum from Eq. (2).

The magnetic dipole moment $\vec{m}(t)$ contains contributions originating from the moving charges,^[48] but also from the moving induced electric dipole moments.^[49] Following the derivation of the NVPT approach,^[32] we also include a third contribution that describes the rotational motion of time varying dipole moments, resulting in the proposed expression as follows:^[48]

$$\begin{aligned} \vec{m}(t) = \sum_i^N \left[\frac{1}{2c} q_i \vec{r}_i(t) \times \vec{v}_i(t) \right. \\ \left. + \frac{1}{c} \vec{\mu}_i^{\text{ind}}(t) \times \vec{v}_i(t) \right. \\ \left. + \frac{1}{2c} \vec{r}_i(t) \times \dot{\vec{\mu}}_i^{\text{ind}}(t) \right] \end{aligned} \quad (5)$$

where \vec{v}_i denotes the velocity vector of atom i .

The calculation of the magnetic dipole moment vector thus requires prior knowledge of the time derivative of the induced electric dipole moment. In MD simulations, such a quantity is determined numerically at time t using a two-point finite differentiation scheme between values at times $t + \delta t$ and $t - \delta t$, δt being the time step. This entails some minor additional book keeping in the implementation of the trajectory integration, but without contributing significantly to the computational cost, which in AMOEBA essentially amounts to the self-consistent determination of electric dipole moments.

Computational details

The AMOEBA polarizable atomic multipole FF^[43] was used to describe the dynamics of the molecular systems of interest. AMOEBA was originally developed for organic molecules in condensed and gas phases,^[50,51] and was shown to perform very satisfactorily for IR frequencies and intensities alike.^[36,37,52–54] In practice, AMOEBA combines the traditional ingredients of nonpolarizable models, including bonded and non-bonded terms, with a multipolar description of electrostatics on each atom up to the quadrupole,^[55] as well as induced dipole moments $\vec{\mu}_i$ reacting self-consistently to the presence of local electric fields on each polarizable site. One of the most valuable features of AMOEBA is its high quality performance in reproducing quantum chemical data for the relative energies

of conformers,^[51] as the result of the additional multipolar contributions, all derived from dedicated high-level gas phase quantum chemical calculations.^[56]

In practice, the induced dipole moment $\vec{\mu}_i^{\text{ind}}$ on atomic site i is related to the local electric field $\vec{\mathcal{E}}_i$ acting on this site as

$$\vec{\mu}_i^{\text{ind}} = \alpha_i \vec{\mathcal{E}}_i \quad (6)$$

with α_i the atomic polarizability assumed to be isotropic. The electric field $\vec{\mathcal{E}}_i$ has a contribution from the induced dipole moments on other sites, and the entire set of induced dipole moments must be solved self-consistently, which in AMOEBA is achieved iteratively.

For the systems presently studied, AMOEBA parameters are readily available from the literature,^[56] except for multipoles of neutral isolated alanine. To determine these missing parameters, a set of multipoles was generated from the MP2/6-311G(d,p) electron density using the distributed multipole analysis developed by Stone,^[55,57] followed by the electrostatic potential fitting at the MP2/aug-cc-pVTZ level, according to the standard AMOEBA procedure.^[58] In addition, the correlation of the relative energies between DFT and AMOEBA was improved by varying manually the charge on the oxygen of the OH group, the other charges being adjusted so as to maintain global charge neutrality. The AMOEBA parameters for the alanine in its neutral form are provided as Supporting Information.

Molecular dynamics simulations with AMOEBA were performed with the Tinker software package.^[59] For all systems, an equilibration trajectory was first performed in the NVT canonical ensemble at the desired temperature (usually 300 K but occasionally lower, using the Nosé-Hoover thermostat) to generate appropriate initial conditions for the subsequent MD trajectories in the NVE microcanonical ensemble, from which the time series of electric and magnetic dipole moments were accumulated.

The canonical trajectories employed a time step of 1 fs and were propagated for 1 ns. Under constant total energy, the time step was reduced to 0.5 fs for a higher accuracy^[60] and the trajectories were propagated for successive series of either 200 ps or 40 ps in the gas and condensed phases, respectively. The time step of 0.5 fs was found to be sufficiently short and compatible with the spectroscopic range of relevance in the present article, namely 1000–2000 cm^{-1} . The electric and magnetic dipole moments accumulated over the course of the trajectories were subsequently processed using the ChirPy program^[61], from which the Fourier transform of the time correlation functions and the resulting IR and VCD spectra were determined, including nonlocal coupling terms.

To assess the present methodology, calculations employing an explicit description of electronic structure were performed to benchmark the electric and magnetic dipole moments whose time variations convey the vibrational information. Configurations from a reference trajectory were then also analyzed using DFT, where the Troullier-Martins pseudopotentials^[62] are used and the plane wave basis set has a cutoff of 100 Ry. These calculations used the BLYP functional and were carried out with the CPMD software.^[63] The magnetic dipole moments were evaluated using the NVPT approach.

Results in the gas phase

In the gas phase, the alanine amino acid is more stable in its neutral form than in the zwitterionic form, while the reverse holds once alanine is hydrated or as a crystal.^[13,64–66] For comparison across phases, both isolated forms were independently investigated in the present work.

Fig. 1 shows the IR and VCD spectra of zwitterionic alanine at 300 K, as obtained from microcanonical MD trajectories after averaging over 40 initial conditions in the canonical ensemble. This system provides a first testing ground for the method of incorporating polarization forces in the determination of dipole moments, and in particular the various contributions of Eq. (5) to the magnetic dipole moment. More precisely, three IR and VCD spectra were determined from the appropriate time correlation functions, assuming (i) a complete neglect of induced polarization forces, both in the electric dipole moment and the magnetic dipole moment, in which only the effects of fixed charges are included; (ii) the account of induced electric dipole moment in both dipole moments, but neglecting the contribution $\vec{r}_i \times \vec{\mu}_i^{\text{ind}}$ in Eq. (5); (iii) the full treatment of induced polarization and all terms in Eq. (5). It should be stressed here that the exact same underlying MD trajectories are used when comparing the spectra (with all induced dipole moments) when evaluating the energy and forces, hence the time variations are strictly identical, and the peak positions in the vibrational spectra are expected to be the same as well, only differences in peak intensities being expected across cases (i)–(iii).

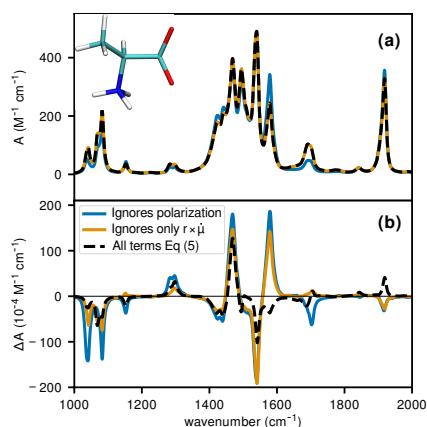


Figure 1. (a) IR and (b) VCD spectra of zwitterionic isolated alanine averaged over 40 MD trajectories at $T = 300$ K, as obtained using the complete treatment of induced polarization forces, ignoring them altogether, or ignoring only the contribution in $\vec{r}_i \times \vec{\mu}_i^{\text{ind}}$ in Eq. (5). Alanine in zwitterionic form is depicted as an inset in panel (a).

The spectra in Fig. 1(a) confirm that IR intensities are slightly affected by neglecting induced dipole moments, usually being lower in magnitude with respect to the full dipole moments. As expected, altering the magnetic dipole moment does not affect the IR spectrum. In contrast, all contributions from induced polarization to the dipole moments in Eqs. (4) and (5) are found to influence the VCD spectrum quite significantly. Ignoring induced polarization forces leads to excessively negative peaks below 1100 cm^{-1} and multiple peak inversions at 1590 , 1700 , and 1925 cm^{-1} , as well as another minor inversion near 1850 cm^{-1} . Ignoring only the term $\vec{r}_i \times \vec{\mu}_i^{\text{ind}}$ in the expression of the magnetic dipole moment, two peaks are now reversed at 1150 and

1700 cm^{-1} . Finally, once this term is restored, the peak at 1150 cm^{-1} is reversed again, the same occurring for the high peak at 1590 cm^{-1} , and the VCD intensities are generally found to be much lower when all polarizability contributions to the magnetic dipole moment are accounted for.

These results show that, even for a moderately polarizable system such as the alanine amino acid, induced polarization forces can all contribute to the VCD intensities in a larger extent as they do for the IR intensities. In the remainder of this article, the electric and magnetic dipole moments are thus obtained from all contributions from induced polarization forces.

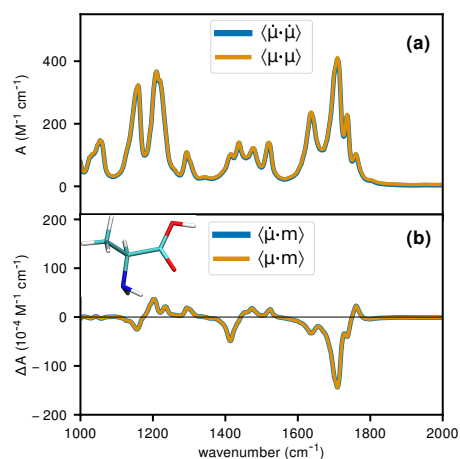


Figure 2. (a) IR and (b) VCD spectra of neutral isolated alanine averaged over 40 MD trajectories at $T = 300$ K, obtained from appropriate time correlation functions employing the position or the velocity forms of the electric dipole moment. Alanine in its neutral form is represented as an inset in panel (b).

Fig. 2 shows the resulting IR and VCD spectra obtained for isolated alanine in neutral form, using the same simulation settings as for the zwitterionic form. Here we examine the influence of choosing the position or the velocity form in the time correlation functions employed to determine the IR and VCD spectra, i.e. Eq. (1) versus Eq. (2). For both IR and VCD signals, the position and velocity forms of the electric dipole moment produce equivalent spectra, which is the expected result and confirms our correct implementation of numerical time derivatives. Comparing now the calculated spectra to results available in the literature for the zwitterionic and neutral forms^[66–68] the peaks above 1600 cm^{-1} correspond to the C=O stretch of the carboxylate group or carboxylic acid and to the bend of the amino or amine groups. The bendings of the CH_3 and CH groups are expected around 1450 cm^{-1} and 1300 – 1200 cm^{-1} , respectively. Below 1200 cm^{-1} the peaks are caused by the C-N-H bend.

The differences between the IR and VCD spectra obtained for the zwitterionic and neutral forms are significant and convey the importance of the protonation site on such a small molecule. The peaks can be assigned using the recent effective anharmonic mode analysis procedure.^[68] The IR spectrum is somewhat richer for the neutral form, with multiple peaks or broad bands near 1000 – 1250 cm^{-1} that correspond to OH and NH_2 bending modes, and near 1600 – 1700 cm^{-1} for NH_2 bending modes. In contrast, zwitterionic alanine shows one main series of peaks near 1400 – 1600 cm^{-1} that correspond to NH_3 bending modes combined to CH_3 bending modes, in addition to more localized peaks in the

1000–1100 cm^{-1} range that are mainly due to NH_3 rocking modes. The corresponding VCD spectra also differ with fewer positive peaks in the neutral case.

The two gas phase molecules provide convenient systems to address the issue of statistical convergence of the calculated spectra, in particular for the more sensitive case of vibrational circular dichroism which is a differentiated signal. We show in Fig. 3 the IR and VCD spectra obtained from time correlation functions based on reduced samples and a limited number $N_T \leq 40$ of MD trajectories with canonical initial conditions. The converged spectra, by definition, are those already shown in Figs. 1 and 2 and for which $N_T = 40$ MD trajectories were employed.

Global convergence of the spectra is here appreciated on a purely visual basis, by comparing the spectra obtained with $N_T < 40$ trajectories to the spectra assumed to be converged for $N_T = 40$, convergence being achieved when the intensities of the peaks stabilize after a certain sampling time. Fig. 3 shows that convergence is rather fast for the zwitterionic molecule, even for the VCD spectrum, positions and intensities being stabilized already with 4 trajectories of 20 ps each. In contrast, convergence of the spectra in the neutral molecule is much slower, positions being stabilized with at least 12 such trajectories, IR and VCD intensities requiring more than 20 and 30 trajectories, respectively. The slower convergence in the neutral case directly reflects its greater flexibility than the zwitterion, owing to fewer and smaller local charges. The greater difficulty of achieving converged conformational sampling in the neutral system is also confirmed by considering the spectra obtained on a limited number of MD trajectories randomly taken from the available set (see Fig. S1). Slower convergence of the VCD spectrum was previously noted by Hornivcek and coworkers^[60] who found that 100 MD simulations of 20 ps each were needed at 300 K to reach convergence for isolated α -pinene despite this molecule being very rigid. Similar conclusions are found here for neutral alanine with the convergence starting at 10 trajectories of 200 ps in total. We interpret the faster convergence of spectral features in the zwitterionic case as the possible consequence of stronger electrostatic forces arising from the larger charges, making this form more rigid relative to the neutral system.

Beyond the qualitative picture provided by Fig. 3, the global convergence of the IR and VCD spectra can be quantified by defining appropriate errors when using a subset of the available data accumulating the time correlation functions. The idea here is to represent quantitatively the convergence with respect to the increasing number of trajectories. For the IR spectra, for which the intensity is always positive, a simple normalized χ^2 expression is used to measure the difference between the signals obtained at convergence ($N_T = 40$) and for a subset of the data with $N_T < 40$ trajectories. The VCD signal, with positive and negative peaks, requires another measure, and here we have used the SimVCD method^[69] to determine the quantitative similarity between two VCD spectra. The normalized errors E_{IR} and E_{VCD} associated to the IR and VCD spectra are thus defined with the following equations:

$$E_{\text{IR}} = \frac{Y_{12}}{(Y_{11} + Y_{22})} \quad (7)$$

$$E_{\text{VCD}} = 1 - \frac{Z_{12}}{(Z_{11} + Z_{22} - |Z_{12}|)} = 1 - \text{SimVCD} \quad (8)$$

where Y_{ij} are the self and overlap integrals of the IR spectra

of i and j and Z_{ij} are the self and overlap integrals of the VCD spectra of i and j , both accumulated over the relevant frequency range of 1000–2000 cm^{-1} :

$$Y_{ij} = \int |A_i(\omega) - A_j(\omega)|^2 d\omega, \quad (9)$$

$$Y_{ii} = \int A_i(\omega)^2 d\omega, \quad (10)$$

$$Z_{ij} = \int \Delta A_i(\omega) \times \Delta A_j(\omega) d\omega, \quad (11)$$

$$Z_{ii} = \int \Delta A_i(\omega)^2 d\omega, \quad (12)$$

where $A_i(\omega)$ is the IR spectrum intensity $\Delta A_i(\omega)$ the VCD spectrum intensity. Note that we do not use the SimVCD measure itself, as it gives a number between -1 if the spectra are mirror of each other and 1 if they are identical, but $1 - \text{SimVCD}$ instead, so the error vanishes for identical spectra.

In the present case, we wish to compare the spectra accumulated over a subset of N_T trajectories to the spectra assumed to be converged at $N_T = 40$. In practice, for a fixed value of $N_T < 40$, a random sample of N_T dataset among the total number available is selected, the corresponding IR and VCD spectra are calculated from these data, and the resulting errors E_{IR} and E_{VCD} are evaluated using the above formulas. Repeating the random process over 10^4 independent realizations, average errors and the associated fluctuations can finally be estimated.

The results of this Monte Carlo procedure are shown in Fig. 4 for the present spectra.

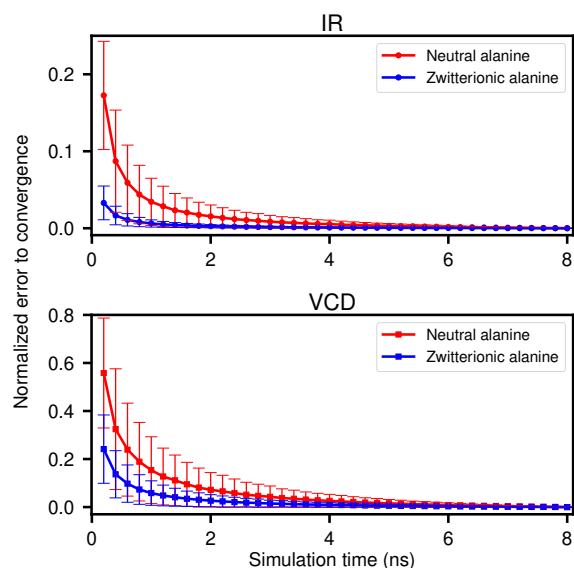


Figure 4. Normalized error in the IR and VCD spectra of isolated alanine at $T = 300$ K, as obtained using partial averages in the appropriate time correlation functions from subsets of N_T MD trajectories, as a function of the sampling time, and assuming the spectra are fully converged for a sampling time of 8 ns.

The variations of the normalized errors shown in this figure confirm the visual impression from the spectra themselves in Fig. 3, in that convergence is significantly faster for the zwitterionic system (both on the IR and VCD spectra), and that IR spectra converge generally faster than VCD spectra. For the present systems, convergence can be concluded to be acceptable with about 4 nanoseconds of sampling.

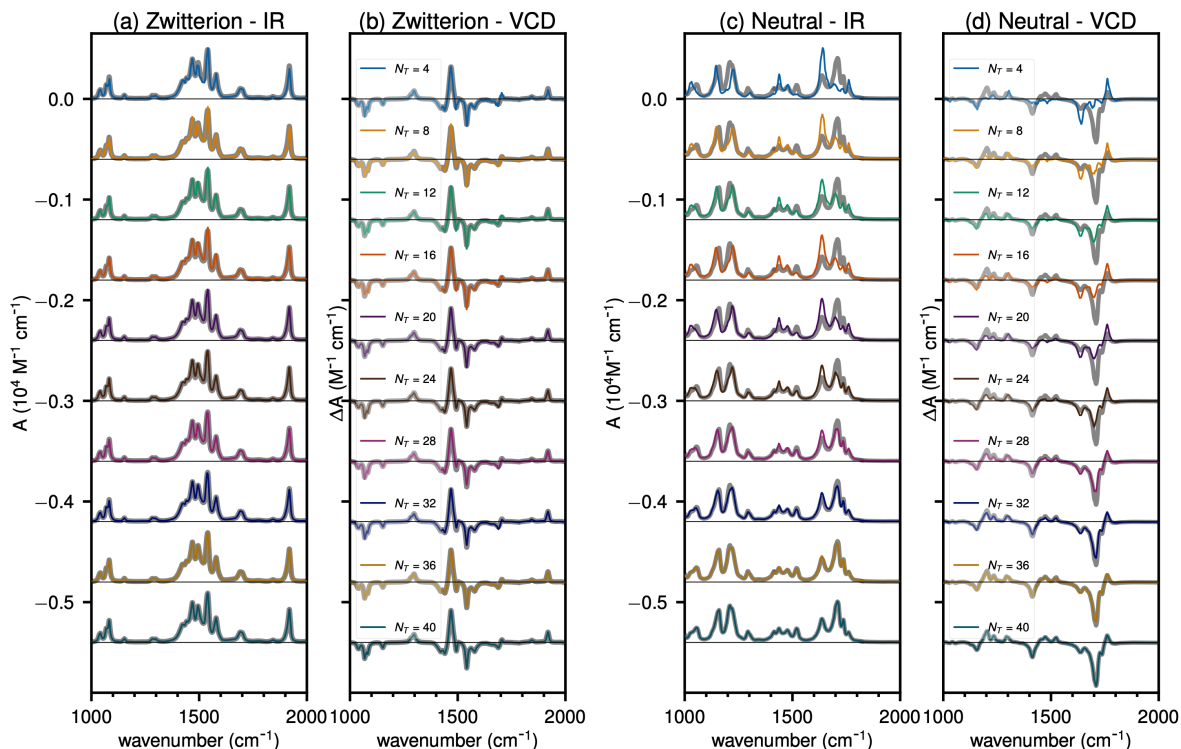


Figure 3. (a) and (c) IR spectra and (b) and (d) VCD spectra obtained from increasingly large samples of 2 to 40 MD trajectories at $T = 300$ K for (a,b) the zwitterionic and (c,d) the neutral forms of isolated alanine. For each spectrum, the converged spectrum obtained with 40 MD trajectories is superimposed in grey.

We end this section by discussing the accuracy of the spectra themselves by comparison with calculations employing an explicit description of electronic structure. More precisely, one of the 200 ps MD trajectories generated with AMOEBA was used to benchmark the electric and magnetic dipole moments predicted by the FF, and for the very same configurations the same quantities were determined at the DFT/BLYP level.

From these dipole moments, the IR and VCD spectra can be again determined using the Fourier transform of the corresponding time correlation functions. They are represented in Fig. 5 for both neutral and zwitterionic forms of isolated alanine. Since the underlying trajectories are equivalent, only differences in intensities are expected with the peak positions being identical between the FF and the DFT spectra. Overall, the spectra exhibit similar intensities both in the IR and VCD cases, the agreement being somewhat even better for the neutral form. For the zwitterionic form, the main discrepancies are found in the stronger activity predicted in the 1410–1450 cm^{-1} spectral range by the FF, and in the significantly lower intensity near 1925 cm^{-1} that points out at a possible inaccuracy in the description of the COO^- group. In addition, the peaks near 1050 cm^{-1} change sign in the VCD spectrum. For the neutral form, differences essentially reside in the generally higher intensities found with the FF, except at the peak located at 1740 cm^{-1} whose IR intensity is underestimated. The possible shortcomings of AMOEBA in describing the carboxylate group originate from the limited accuracy of distributed multipoles for anions, which cannot fully account for delocalized electronic clouds.

To interpret these spectra further, we show in Fig. 6 the statistical correlations between the three Cartesian compo-

nents of the dipole moments obtained at the FF and DFT levels. In anticipation of the results for the crystalline phase, and in order to avoid a poorly correlated electric dipole moment due to different molecular orientations in the crystalline phase, the vector $\vec{\mu}$ was shifted to the origin. No shift was applied to the components of $\dot{\vec{\mu}}$ or \vec{m} .

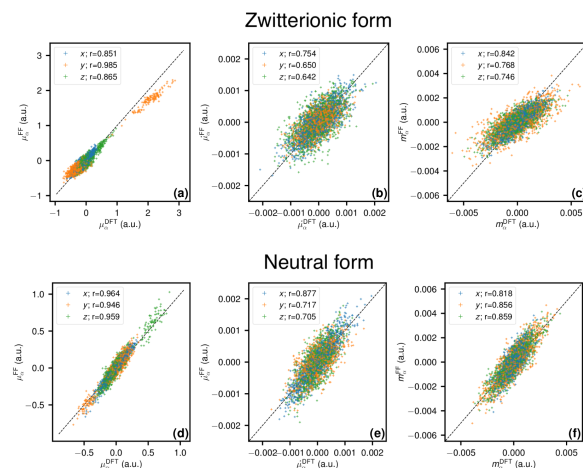


Figure 6. Cartesian components of the dipole moments predicted by AMOEBA as a function of the corresponding DFT values, for configurations along the MD trajectory used to generate the spectra of Fig. 5, for the (a–c) zwitterionic and (d–f) neutral forms of isolated alanine. (a) and (d) electric dipole moment $\vec{\mu}$; (b) and (e) time derivative $\dot{\vec{\mu}}$; (c) and (f) magnetic dipole moment \vec{m} . All quantities are given in atomic units, and for each correlation plot the Pearson correlation coefficient r is given.

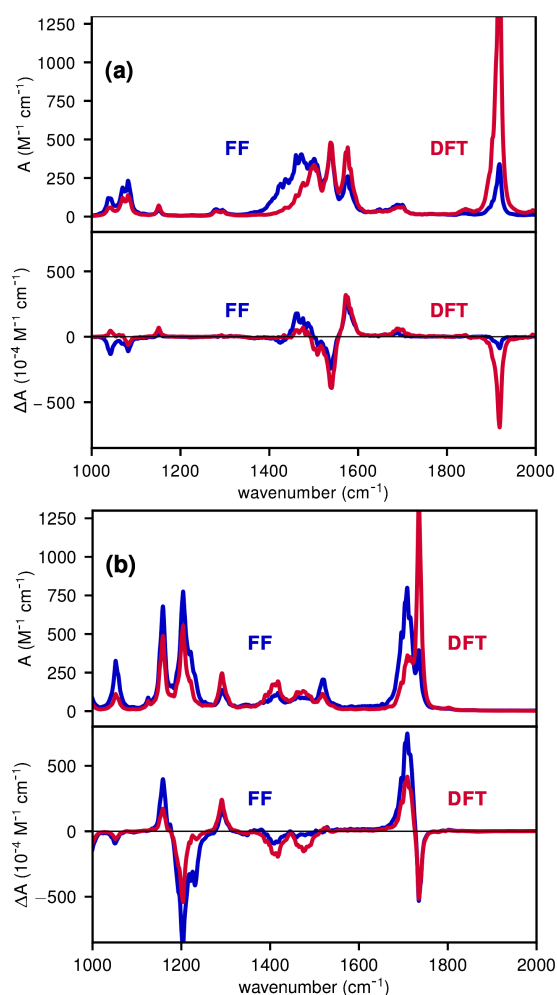


Figure 5. IR and VCD spectra obtained from a single 20 ps MD trajectory, using the dipole moments of the FF or after being recalculated at the DFT/BLYP level, for the (a) zwitterionic and (b) neutral forms of isolated alanine.

As a first comment, the components of the electric dipole moment are found to be rather disperse, but inspection of the MD trajectory for the zwitterion reveals that some functional groups occasionally undergo minor internal rotations, explaining the disconnected distributions of Fig. 6(a).

Even with this peculiarity, the FF and DFT values of the dipole moment components are all highly correlated to one another, the main deviation being the slightly more widespread distribution of the magnetic dipole moment found at the DFT level, relative to the AMOEBA value, which is interpreted as causing the stronger VCD signal associated to the carboxylate group in Fig. 5(a). The higher Pearson coefficients found for the neutral form are also consistent with the more similar FF and DFT spectra in Fig. 5(b).

Results in condensed phases

In the previous section we showed that for isolated molecules, AMOEBA can produce anharmonic vibrational spectra, including VCD signals, that are of good quality relative to electronic structure calculations obtained independently. This section focuses on condensed phases

of alanine, for which MD simulations were performed under periodic boundary conditions (PBC). For VCD spectroscopy, the traditional minimum image convention can cause a dependence of the magnetic dipole moment on the origin, and to address this gauge issue we have followed the recent suggestion by Jähnigen and coworkers^[14] by considering the nearest periodic image directly between molecular units. More precisely, denoting by $\delta\vec{r}_j$ the position of particle j and by $\vec{\Delta}_{ij}^{\text{PBC}}$ the lattice translation vectors needed to place the interparticle vector \vec{r}_{ij} between particles i and j in their minimum image, the following expression is used for the cross correlation function entering the VCD spectrum, between the time derivative of the electric dipole moment and the magnetic dipole moment:

$$C_{\text{PBC}}(t) = \left\langle \dot{\vec{\mu}}(0) \cdot \vec{m}(t) \right\rangle + \frac{1}{2c} \sum_i \left\langle \dot{\vec{\mu}}_i(0) \cdot \sum_j \left[\left(\vec{r}_j(t) + \frac{1}{2} \vec{\Delta}_{ij}^{\text{PBC}}(0) \right) \times \dot{\vec{\mu}}_j(t) \right] \right\rangle \quad (13)$$

This expression is used for the three condensed phases in which alanine is placed next, namely as a crystal or embedded in water or nitrogen solvents.

Crystalline alanine

Alanine in the crystal phase was simulated using the known structure of the $P2_12_12_1$ chiral space group.^[70] A periodic box of $12.073 \times 12.342 \times 11.57$ Å was constructed with ChirPy containing 16 alanine molecules in zwitterionic form. The simulation cell is depicted as an inset in Fig. 7. Because of the zwitterionic nature, the cohesion in the alanine crystal is relatively strong as the molecules are closely held together by electrostatic interactions between the ammonium and carboxylate groups, highlighted on the picture in Fig. 7 by black dotted lines for one amino group.

Following a similar protocol as in the gas phase simulations, a MD trajectory of 1 ns in the canonical ensemble at $T = 300$ K was first performed to generate initial configurations, every 10 ps, for the subsequent microcanonical trajectories from which time correlation functions were accumulated. The long-range electrostatic interactions in this periodic system were computed employing Ewald summation with a cut-off of 5 Å. Van der Waals interactions were also truncated at distances below a cut-off of 7 Å.

For the crystalline phase, 30 microcanonical trajectories of 40 ps each, or a total simulation time of 1200 ps, were thus simulated to provide the IR and VCD spectra considered to be the converged reference. These spectra are shown in Fig. 7 and are compared to the corresponding spectra obtained for the isolated molecule in zwitterionic form.

For this crystalline phase, the spectra appear to be well resolved, with peaks or groups of peaks that are mainly found around the same positions as in the gas phase. However, some residual shifts are clearly seen as manifestations of the influence of the condensed environment, some of them being highlighted in Fig. 7 by red arrows. For instance, the NH_3 bending mode shifts from 1580 to 1622 cm^{-1} , while the $\text{C}=\text{O}$ stretching mode decreases from 1919 to 1853 cm^{-1} . Both effects are direct consequences of the electrostatic interactions between the amino and carboxylate groups of neighboring molecules in the crystal.^[68]

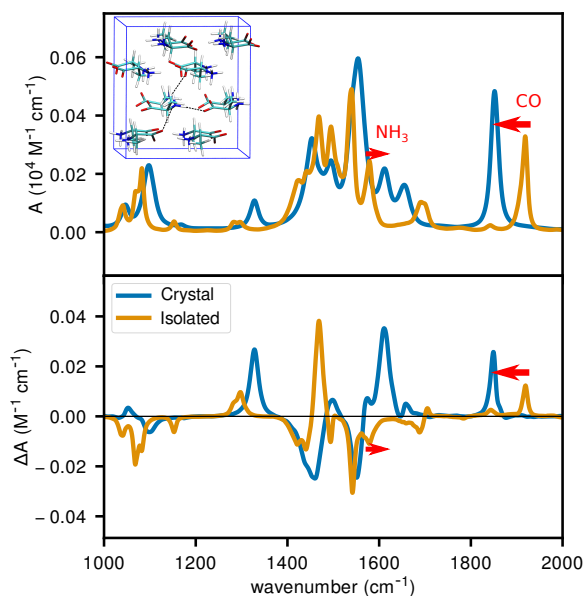


Figure 7. IR and VCD spectra obtained from MD simulations of crystalline alanine (blue lines). The corresponding spectra obtained for isolated alanine in zwitterionic form are superimposed as orange lines, the VCD spectrum being multiplied by a factor 3 for a better visibility. The red arrows highlight the shifts of peaks corresponding to the NH_3 bending and the $\text{C}=\text{O}$ stretching modes between the two phases.

Previous simulations at the FPMD level^[13] have shown additional features that are absent from the present calculations with AMOEBA. The carbonyl stretch was found as a doublet near 1600 cm^{-1} in Ref. 13 as the result of collective motion in the crystal and some interferences between intramolecular and phonon modes, intermolecular couplings acting over up to $4\text{--}5\text{ \AA}$. Instead, this mode appears here as a single peak at 1833 cm^{-1} , which points out at differences in the modeling and in particular some inaccuracies of the force field for the crystalline phase. For a fair comparison, the same periodic cell was employed for AMOEBA and DFT calculations, although both spectra may be affected by additional collective modes in larger periodic cells. This latter issue could be addressed by repeating the simulations over a longer spatial range using a larger periodic box.

Following again the previous discussion in the gas phase, the global convergence properties of the IR and VCD spectra were characterized by quantifying the similarity of spectra obtained from reduced sampling time, to the converged spectra resulting from the entire 1.2 ns integration. The variations of the IR and VCD spectra with increasing amount of simulation time are shown in Fig. S2.

On a more quantitative footing, the statistical errors associated to spectra obtained from a random subset of N_T trajectories among the 30 available were determined following the same procedure as defined for the gas phase, and the normalized errors are shown in Fig. 8 as a function of the sampling time. These errors highlight the much faster convergence for the crystal comparatively to the gas phase, which reflects the more rigid character of the crystalline system owing to rather strong intermolecular electrostatic interactions.

As for the gas phase systems, we also compared the accuracy of the dipole moments predicted by AMOEBA with the corresponding dipole moments obtained at the DFT/BLYP level of theory, for the very same configurations from a sin-

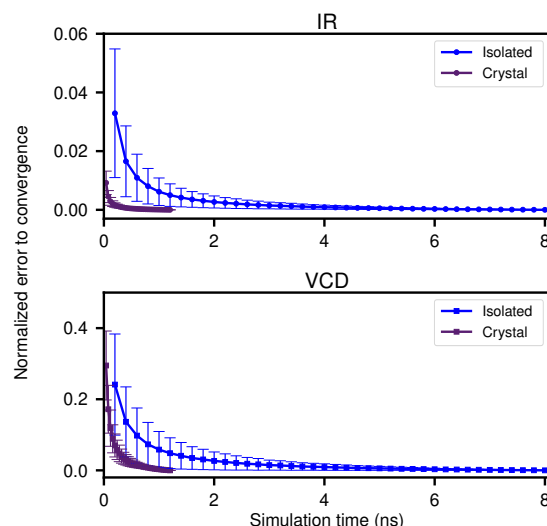


Figure 8. Normalized error of the average in the IR and VCD spectra of crystalline alanine obtained using N_T MD trajectories of 40 ps at 300 K , compared to the average from 30 trajectories, as a function of the sampling time. The corresponding data for the isolated alanine in zwitterionic form are superimposed.

gle MD trajectory generated with the FF. The IR and VCD spectra themselves, as well as the statistical correlation between the Cartesian components of the electric dipole moments $\vec{\mu}$, its time derivative $\dot{\vec{\mu}}$ and the magnetic dipole moment \vec{m} are shown as Supporting Information in Figs. S3 and S4, respectively. Once again the level of agreement for the IR and VCD spectrum is remarkable, the main difference lying in the underestimation of the IR intensity of the carboxylate bending peak near 1850 cm^{-1} and its VCD signature. For this condensed system, the dipole moments correlate better with the DFT data with respect to gas phase zwitterionic alanine.

The present computational results can also be compared to available experimental measurements from Rodríguez-Lazcan and coworkers who assigned the modes using periodic DFT calculations.^[71] Such a comparison can be made much easier by assigning the peaks of the present simulated spectra using effective anharmonic modes that can be extracted from an optimal decomposition of the molecular motion in terms of internal modes.^[68,72] The vibrational frequencies for the corresponding modes obtained with the FF are compared to the experimental values in Table 1. Here we used the same notation as in the experimental reference, namely ρ for rocking, δ for bending, ν for stretching modes.

The agreement between experimental and calculated frequencies is generally very satisfactory, especially for the vibrations corresponding to the CH_3 and NH_3 rocking modes, but also the CH , CH_3 and NH_3 bending modes. The rocking modes as well as the CH and CH_3 bending modes are generally coupled with one another, which is consistent with the attributions provided previously.^[71] The most significant discrepancy is found for $\text{C}=\text{O}$ stretching mode, which as noted earlier reflects some inaccuracies of the force field for the carboxylate group, related with the limited performance of the distributed multipolar description.

Table 1. Vibrational frequencies for crystalline alanine collected in the 1000–2000 cm^{-1} region from the present simulations with AMOEBA in comparison with experimental frequencies.^[71] All frequencies are given in wavenumbers.

Assigned mode ^[68]	Calculated frequency	Experimental frequency ^[71]
ρCH_3 ρNH_3 δCCH	1046	1014
ρCH_3 νCN	1104	1114
ρNH_3 δCH	1169	1152, 1237
δCCH	1326	1307
δCH_3 δCH	1451, 1483, 1479, 1503	1363, 1412, 1455
δNH_3 , νCO_2	1555, 1566, 1623, 1657	1506, 1519, 1594, 1621
νCO_2	1854	1648

Hydrated alanine

As the natural environment of organic biomolecules, we have considered the case of hydrated alanine as our next testing ground of the present methodology. Classical MD simulations were performed for a zwitterionic alanine molecule soaked in a previously equilibrated cubic periodic box of water with length $L = 18.6$ Å at the regular density of liquid water at room temperature, thereby yielding a box containing 211 water molecules in addition to the single alanine molecule. Such a sample is visually depicted in the inset of Fig. 9(a). These simulations employed a common cutoff of 7 Å for both the van der Waals potential and Ewald summation.

In the IR range of interest here (1000–2000 cm^{-1}), the solvent has a strong spectroscopic response due to the large electric dipole moment of the water molecule, likely to mask the contribution of the diluted alanine molecule. This issue also arises in experiments, where the spectrum of the pure solvent is just removed from the spectrum with the solute. Such a heuristic approach would be tricky in simulation because the pure solvent and the hydrated molecule do not share exactly the same periodic box. One way to circumvent this difficulty is to consider only a subset of solvent molecules around the solute when evaluating the dipole moments and processing them to determine the time correlation functions.

Instead of imposing a rigid cutoff that could cause discontinuities in the physical observables associated to a liquid system, we follow the suggestion by Scherrer and coworkers^[32] who introduced scaled molecular dipole moments that smoothly incorporate the contribution of distant molecules around the solute based on their distance through an exponentially decaying function. From a general perspective, and denoting by K any molecule of vibrational interest, its electric and magnetic dipole moments are thus scaled accordingly with

$$\vec{\mu}_K^{\text{scale}}(t) = \vec{\mu}_K(t) + \sum_{J \neq K} P_{KJ}(t) \vec{\mu}_J(t) \quad (14)$$

$$\vec{m}_K^{\text{scale}}(t) = \vec{m}_K(t) + \sum_{J \neq K} P_{KJ}(t) \vec{m}_J(t), \quad (15)$$

where the sum over J acts on all molecules surrounding K . In the above equations, the damping function P_{KJ} is chosen as $P_{KJ}(t) = \left(1 + e^{(|\vec{R}_{KJ}(t)| - R_c)/D}\right)^{-1}$, with R_c a cutoff parameter distance and D a sharpness parameter taken as 0.25 Å, following Heyden *et al.*^[73] It should be emphasized here that this scaling procedure is also needed for the VCD spectrum, despite pure water being an achiral solvent with

no VCD response, in order to account for possible chirality transfer effects.^[74] This functionality of scaling molecular dipole moments has been implemented in ChirPy.^[61]

As for crystalline alanine simulations, a 1 ns MD trajectory in the canonical ensemble at $T = 300$ K was first performed to generate starting configurations for the microcanonical trajectories from which time correlation functions were accumulated. Rather long simulations were needed to converge the IR and VCD spectra of hydrated alanine, and the following results were obtained using 50 MD trajectories of 40 ps each.

We first discuss the effects of the cutoff R_c on the spectra, using the very same underlying MD trajectories but different sets of dipole moments scaled accordingly with the value of R_c . These spectra are shown in Fig. 9 for $R_c = 0$, $R_c = 3.85$ Å and $R_c = 4.60$ Å. While $R_c = 0$ corresponds to considering the sole contribution of alanine on the spectra, the lowest, nonzero value accounts for water molecules in the first hydration shell.

As anticipated, the spectra obtained without scaled dipole moments do not show any characteristic signal due to the alanine solute, as the contribution of the solvent is totally excessive in this case. In contrast, only considering the dipole moments of the alanine solute produces well resolved IR and VCD spectra, but from which it is obviously difficult to assess the role of the solvent. Increasing the cut-off radius to 3.85, then 4.6 Å, the main contribution of the hydration shell is found as the emerging peak in the 1600–1800 cm^{-1} range, as well as an overall, shifted IR response.

Like the IR spectrum, the VCD spectrum is found to depend only weakly on the amount of water solvent included in the determination of the dipole moments, except in the region dominated by the solvent response. The most remarkable feature is probably the rather significant response of the pure solvent in the 1600–1800 cm^{-1} range in Fig. 9(a). The fluctuations in VCD intensity probably indicate still insufficient convergence, much longer trajectories being needed for the water signal to effectively average to zero. Yet we cannot exclude some possible chirality transfer effects, as several of the significant fluctuations in the 1100–1600 cm^{-1} range take place at the same frequencies as those where zwitterionic alanine is VCD active. One way to disentangle both effects could be, besides running even longer trajectories, to compare to the results obtained on larger boxes where chirality transfer should be attenuated.

Keeping such possible limitations in mind, the comparison between the spectra obtained for hydrated alanine and the isolated molecule in zwitterionic form provides useful insight into the effects of solvation on this amino acid. While

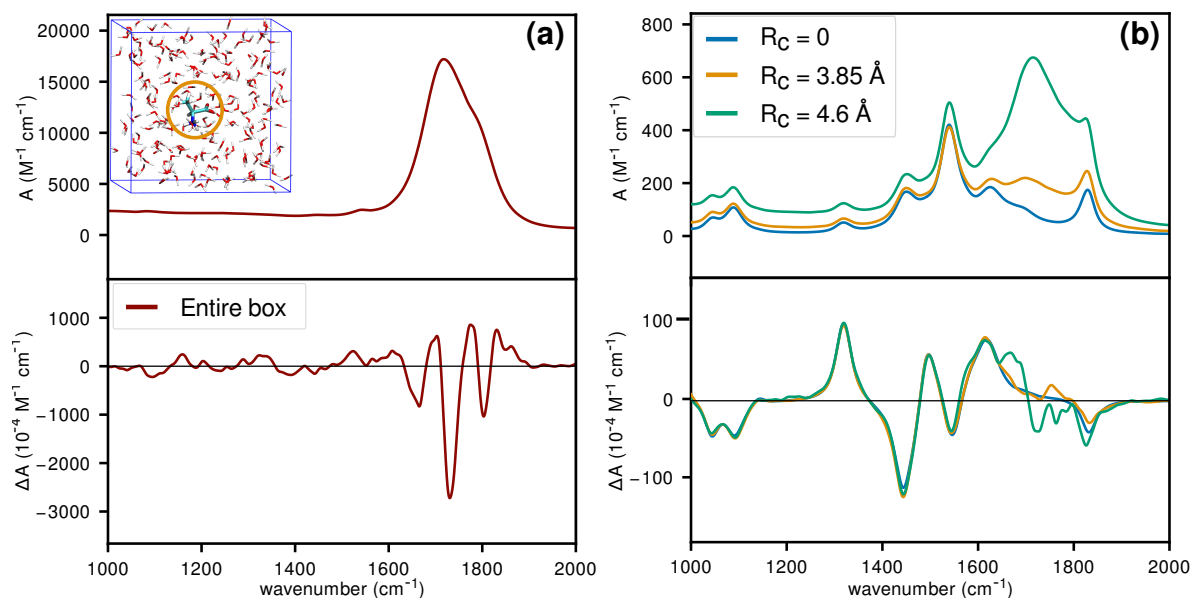


Figure 9. IR and VCD spectra of hydrated alanine at 300 K obtained using dipole moments scaled exponentially according to the distance to the solute, and for the same MD trajectories. (a) Entire box (no scaling) and (b) scaled dipole moments above cut-off values of $R_c = 0$, 3.85 and 4.6 Å.

the generic spectral features such as peak positions and even intensities are only mildly affected in the case of IR intensities, the VCD signals are reversed in the 1400–1600 cm^{-1} range.

The IR and VCD spectra obtained for hydrated alanine with increasing numbers of MD trajectories are shown in Fig. S5 of Supporting Information. For this system, the IR spectrum converges again rather fast with only a few MD trajectories being required for the signal to be indistinguishable from the reference, converged spectrum. In contrast, the VCD spectrum exhibits variable degrees of convergence, being rather fast below 2000 cm^{-1} where the signal is dominated by the contribution of the chiral solute.

The relative errors in the IR and VCD spectra were quantified by the same Monte Carlo procedure described above, and the results for hydrated alanine are shown in Fig. 10 for different values of the cutoff, including in the absence of cutoff ($R_c = L$, the size of the periodic box).

The error associated with the VCD spectrum decays particularly slowly with the increasing sampling time, especially as more and more solvent molecules are accounted for when evaluating the dipole moments. This behavior conveys on a more quantitative footing the difficulty of converging the VCD signal in a spectral region dominated by contributions from the achiral solvent. Furthermore, the non-local VCD contributions converge much slower than the local ones.^[15] In contrast, when the spectra ignore the explicit contributions from the solvent ($R_c = 0$) the corresponding error decays particularly fast owing to the absence of any VCD signal in the water bending region in that case. Such a slower convergence again points at the need to consider larger simulation cells or even longer trajectories to reduce the VCD signal more efficiently.

The spectra and dipole moments predicted by AMOEBA were also assessed against electronic structure calculations by comparing the results obtained on a same single MD trajectory. Here the same cutoff $R_c = 3.85$ Å was used also in the DFT analysis. The resulting IR and VCD spectra (see Fig. S6) show remarkable similarities, with no peak in-

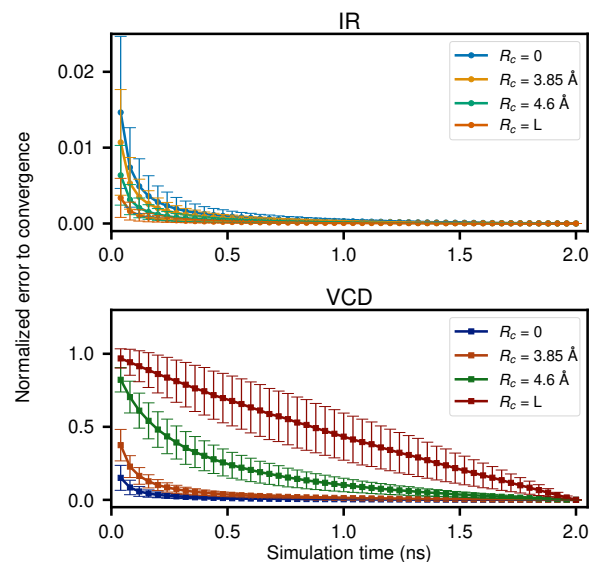


Figure 10. Normalized error of the average in the IR and VCD spectra of hydrated alanine at 300 K, as obtained from reduced numbers N_T of MD trajectories of 40 ps each, compared to the converged spectra for $N_T = 50$, as a function of the sampling time.

version in the VCD signal but the main discrepancy being, once again, the underestimation of the intensity associated with the carboxylate bending mode above 1800 cm^{-1} when described with the FF. The correlation plots between the dipole moments or their time derivatives were also considered (see Fig. S7, also distinguishing the specific contributions of the alanine solute alone in Fig. S8 or the water solvent molecules in Fig. S9). Consistently with the level of agreement between the FF and DFT spectra, the degree of correlation is particularly high with Pearson coefficients above 0.99 for the electric dipole moments. The most significant deviation between the FF and DFT properties relate to the magnetic dipole moment, whose distribution is more spread out at the DFT level, in agreement with the stronger

VCD activity at the DFT level found for the carboxylate mode.

Finally, and as in the crystalline case, experimental comparison can be attempted for hydrated alanine, using again the mode assignment carried with the effective modes procedure.^[68] Experimental modes and their associated frequencies were borrowed from Jalkanen and coworkers,^[75] and they are compared to the AMOEBA results in Table 2.

Overall, the agreement between the theoretical and experimental frequencies is satisfactory for the bending modes of CH, CH₃ and NH₃, but the C=O stretching frequency is overestimated with AMOEBA, which also lacks accuracy for the lower frequency bands below 1100 cm⁻¹. The bisignate signal experimentally found at 1300–1350 cm⁻¹ is reproduced by the simulations, although the associated frequencies appear as excessively separated. This feature, and especially the sensitivity of the higher frequency mode, are particularly prone to solvent effects.

Alanine solvated in dinitrogen

The water solvent was found to have a significant effect on the VCD spectrum of alanine, which reflects the relatively strong hydrogen bonding interactions between the solvent and the solute. As a complementary perspective on solvation effects on VCD spectra, and albeit on a more academic ground, we have considered the chemically inert case of nitrogen, in which alanine can be safely considered to remain in neutral form. The nitrogen solvent imposes cryogenic conditions to be in a liquid state, which further provides an opportunity to discuss temperature effects on the vibrational spectra.

The MD simulations were performed at 77.35 K, for a neutral alanine molecule solvated in a periodic box of 18.27 Å containing 104 N₂ molecules, hence a density of 0.8171 g/cm³. Once again the Ewald and van der Waals cutoffs were both set to 7 Å. At such low temperatures, the molecule is much less fluxional than at room temperature, and inspection of the trajectories reveals that it remains in the same conformer as chosen for the initial configuration. Excessively long trajectories would be required for the system to actually sample the various conformers in an ergodic fashion, which cannot be afforded by regular molecular dynamics. To cope with this time scale issue, we follow a hybrid approach in which the initial conditions are borrowed from two different conformations that correspond to minima with the FF, each trajectory giving specific IR and VCD spectra. Average spectra are then produced by appropriate reweighting over these individual anharmonic spectra, structure *i* contributing through a probability $p_i = Z_i / \sum_j Z_j$ that is proportional to its thermal weight Z_i taken for the present purpose in the harmonic approximation as:

$$Z_i = \exp(-\beta E_i) \prod_k \frac{\exp(-\beta \hbar \omega_{i,k}/2)}{1 - \exp(-\beta \hbar \omega_{i,k})} \quad (16)$$

where E_i is the equilibrium energy of structure *i*, $\beta = 1/k_B T$ with k_B the Boltzmann constant, and $\omega_{i,k}$ is the harmonic frequency of mode *k* of structure *i*, obtained here from static calculations performed with the FF. The resulting thermal weights are found to be 41.1% and 58.9% for these two conformations.

For each conformation, a MD trajectory in the canonical ensemble at 77.35 K was first performed for 5 ps, after which

the thermostat was removed and the simulation was pursued for 40 ps, with the dipole moments being recorded. The final configuration of this microcanonical trajectory was used as initial configuration for a subsequent 40 ps trajectory until ten successive trajectories are completed. For comparison, additional MD trajectories for the isolated neutral alanine were repeated at 77.35 K for both conformations as well, the IR and VCD spectra being also reweighted using the same probabilities p_i . The IR and VCD spectra obtained using this procedure are given in Fig. 11. A cutoff radius of 4.5 Å was used to partially include the solvent molecules (full inclusion of the solvent box leads to band broadening due to the rotation of the N₂ molecules, see Fig. S10).

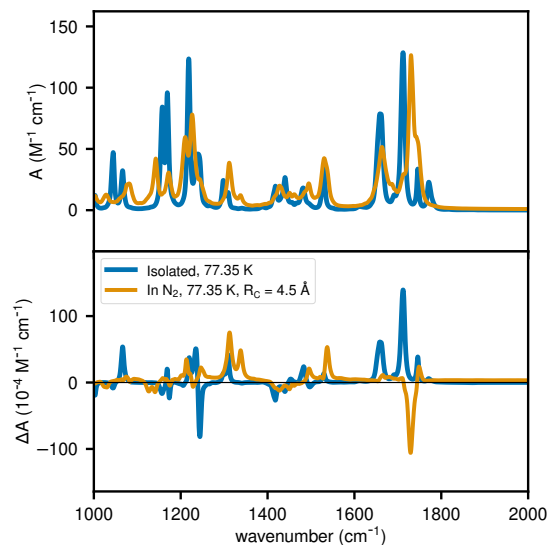


Figure 11. IR and VCD spectra of neutral alanine solvated in N₂ at 77.35 K ($R_c = 4.5$ Å), compared to the corresponding spectra in the gas phase, using the same initial conditions that sample the canonical ensemble of the isolated phase at the desired temperature, based on harmonic thermal weights.

Regarding the IR spectrum, the active peaks are weakly affected by the solvent, except near 1200 cm⁻¹, and the main effect of solvation is to broaden the bands. Solvation has a much stronger influence on the VCD spectrum, with a dramatic change with various peak inversions in the entire spectral range considered here. Inspection of the MD trajectories reveals a conformational change in the presence of the nitrogen solvent, the NH₂ group rotating by 180° and the hydrogen of the hydroxyl group turning towards the amino group. No such conformational change occurs in the gas phase.

These results confirm the sensitivity of the VCD spectrum to conformational changes, even driven by relatively inert solvents, and especially compared to the IR spectrum where the change is minimal.

Conclusion

In this article, a method to obtain vibrational spectra based on fully classical MD simulations that uses a chemically accurate polarizable atomic multipole force field was presented. In addition to the rather conventional IR intensities, the method was extended to the determination of vibrational circular dichroism intensities through the calcu-

Table 2. Frequencies for hydrated alanine collected in the 1000–2000 cm^{-1} region from anharmonic MD simulations with AMOEBA and experimental values from Ref. 75, modes being assigned following the same notations as in Table 1. All frequencies are given in wavenumbers.

Assigned mode ^[68]	Calculated frequency	Experimental frequency ^[75]
ωCH_3 ωNH_3	1044	–
νCCN	1134	1110
ξNH_3	1081	1220, 1145
δCH	1313	1351, 1301
δCH_3	–	1375
νCO_2	1444	1410
δCH_3	1486, 1471, 1456	1459
δNH_3	1584, 1530	1498
νCO_2	1835	1607
δNH_3	1621	1625
δNH_3	1655	1645

lation of magnetic dipole moments associated to the motion of the charges, induced dipole moments, and the rotational motion of these dipole moments.

The method was applied to the alanine amino acid in various gas and condensed phases. Isolated alanine in neutral and zwitterionic forms were used to assess the electric and magnetic dipole moments entering the determination of IR and VCD spectra, and the two contributions involving the induced dipole moments were found to be both important. The spectra for the crystalline phase show significant differences in the NH_3 bending and the $\text{C}=\text{O}$ stretching modes that are caused by the strong electrostatic supramolecular forces, some spectral features of collective nature being possibly missed due to the limited extent of the periodic cell. In the hydrated phase, the contribution of the solvent to the spectra had to be attenuated by scaling the dipole moments around the solute, which reveals the specific influence of the solvent on the VCD signal. Even embedded in the rather inert nitrogen fluid under cryogenic conditions, the solvent was found to have a qualitative impact on the vibrational spectra of neutral alanine due to the slight fluxional motion it induces, with no counterpart in the gas phase.

The statistical convergence of both IR and VCD spectra was carefully examined for all systems and found to be generally significantly slower for the VCD signal, as expected for a differential spectrum. Unsurprisingly, convergence was also found to be related to the degree of fluxionality in the system, being particularly fast for the crystal (a few tens of picoseconds) but much slower in the hydrated phase (several nanoseconds).

The quality of the force field itself, especially in its description of the dipole moments needed for the calculation of the vibrational spectra, was assessed against electronic structure calculations on the very same set of molecular configurations and found to be essentially of very high quality in all situations considered, thus further validating the proposed approach. However, the accuracy was found to be limited in the zwitterionic case owing to the poor performance of the distributed multipolar description for the carboxylate group.

Because AMOEBA derives from ab initio-based calculations, the present work could be straightforwardly extended to incorporate nuclear delocalization, e.g. through path-integral or quantum thermal bath methods in line with the recent development of the Q-AMOEBA water force field.^[76]

Other natural extensions include applications to other polarization models such as Drude oscillators^[77] or fluctuating charges.^[78]

As far as systems are concerned, it would be interesting to apply the present methodology to situations that are difficult to address using FPMD owing to the sampling bottleneck such as amyloid fibrils.^[79,80] The current version of the AMOEBA force field has been optimized for biomolecules,^[43] including peptides or proteins^[56] as well as nucleic acids,^[81] providing accurate structural and thermodynamic properties in solution as compared with experimental data. Furthermore, this force field performs well for molecular crystals and it has been used for peptide and protein structural refinement in X-ray crystallography.^[82] Calculating vibrational spectra still requires rather long convergence times (in excess of a few hundreds of ns), but the first results presented in this work are encouraging. Another potential application of the VCD methodology concerns viscous fluids with slow relaxation such as ionic liquids^[83] or deep eutectic solvents,^[84] which could be natural targets for the near future. The enhancement of VCD signals reported in supramolecular compounds,^[85–87] already hinted at here on the example of crystalline alanine, could also be investigated further using our approach.

Acknowledgements

This work was supported by the ANR Dichroprobe, Grant No. ANR-18-CE29-0001 of the French Agence Nationale de la Recherche. This work was performed using HPC resources from the GENCI (CINES/IDRIS, Grant No. 2021-A0120806830).

Conflict of Interest

References

- [1] L. A. Nafie, *Vibrational optical activity: principles and applications*, John Wiley & Sons **2011**.
- [2] J. Kessler, T. A. Keiderling, P. Bouř, *J. Phys. Chem. B* **2014**, *118*, 6937.
- [3] H. Chi, F. Tobias, J. Kessler, J. Kubelka, P. Bouř, T. A. Keiderling, *Biophysical Journal* **2015**, *108*, 523a.

- [4] J. M. Batista Jr., E. W. Blanch, V. d. S. Bolzani, *Nat. Prod. Rep.* **2015**, *32*, 1280.
- [5] T. A. Keiderling, *Chem. Rev.* **2020**, *120*, 3381.
- [6] T. Vermeyen, C. Merten, *Phys. Chem. Chem. Phys.* **2020**, *22*, 15640.
- [7] J. Dupont, R. Guillot, V. Lepère, A. Zehnacker, *J. Mol. Struct.* **2022**, *1262*, 133059.
- [8] K. Scholten, C. Merten, *Phys. Chem. Chem. Phys.* **2022**, *24*, 3611.
- [9] C. Müller, C. Merten, *Phys. Chem. Chem. Phys.* **2023**, *25*, 19462.
- [10] K. Le Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, A. Zehnacker, *Phys. Chem. Chem. Phys.* **2020**, *22*, 26047.
- [11] V. Declerck, A. Pérez-Mellor, R. Guillot, D. J. Aitken, M. Mons, A. Zehnacker, *Chirality* **2019**, *31*, 547.
- [12] A. Pérez-Mellor, K. Le Barbu-Debus, A. Zehnacker, *Chirality* **2020**, *32*, 693.
- [13] S. Jähnigen, A. Scherrer, R. Vuilleumier, D. Sebastiani, *Angew. Chem. Int. Ed.* **2018**, *57*, 13344.
- [14] S. Jähnigen, A. Zehnacker, R. Vuilleumier, *J. Phys. Chem. Lett.* **2021**, *12*, 7213.
- [15] S. Jähnigen, K. Le Barbu-Debus, R. Guillot, R. Vuilleumier, A. Zehnacker, *Angew. Chem. Int. Ed.* **2023**, *62*, e202215599.
- [16] S. Jähnigen, *Angew. Chem. Int. Ed.* **2023**, *62*, e202303595.
- [17] P. J. Stephens, F. J. Devlin, J. R. Cheeseman, *VCD spectroscopy for organic chemists*, CRC Press, Boca Raton **2012**.
- [18] J. Cheeseman, M. Frisch, F. Devlin, P. Stephens, *Chem. Phys. Lett.* **1996**, *252*, 211.
- [19] S. Yang, M. Cho, *J. Chem. Phys.* **2009**, *131*, 10B605.
- [20] K. Kwac, K.-K. Lee, J. B. Han, K.-I. Oh, M. Cho, *J. Chem. Phys.* **2008**, *128*, 03B606.
- [21] T. Giovannini, M. Olszówka, C. Cappelli, *J. Chem. Theory Comput.* **2016**, *12*, 5483.
- [22] J.-H. Choi, M. Cho, *J. Chem. Theory Comput.* **2011**, *7*, 4097.
- [23] J. Bloino, V. Barone, *J. Chem. Phys.* **2012**, *136*, 124108.
- [24] V. Barone, M. Biczysko, J. Bloino, M. Borkowska-Panek, I. Carnimeo, P. Panek, *Int. J. Quant. Chem.* **2012**, *112*, 2185.
- [25] C. Merten, R. McDonald, Y. Xu, *Inorg. Chem.* **2014**, *53*, 3177.
- [26] A. S. Perera, J. Thomas, M. R. Poopari, Y. Xu, *Front. Chem.* **2016**, *4*, 9.
- [27] R. Iftimie, P. Minary, M. E. Tuckerman, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6654.
- [28] D. Marx, J. Hutter, *Ab initio molecular dynamics: basic theory and advanced methods*, Cambridge University Press, Cambridge **2009**.
- [29] L. A. Nafie, T. B. Freedman, *J. Chem. Phys.* **1983**, *78*, 7108.
- [30] L. A. Nafie, *Annu. Rev. Phys. Chem.* **1997**, *48*, 357.
- [31] A. Scherrer, R. Vuilleumier, D. Sebastiani, *J. Chem. Theory Comput.* **2013**, *9*, 5305.
- [32] A. Scherrer, R. Vuilleumier, D. Sebastiani, *J. Chem. Phys.* **2016**, *145*, 084101.
- [33] S. Jähnigen, D. Sebastiani, R. Vuilleumier, *Phys. Chem. Chem. Phys.* **2021**, *23*, 17232.
- [34] V. Schultheis, R. Reichold, B. Schropp, P. Tavan, *J. Phys. Chem. B* **2008**, *112*, 12217.
- [35] G. S. Fanourgakis, S. S. Xantheas, *J. Chem. Phys.* **2008**, *128*, 074506.
- [36] D. Semrouni, A. Sharma, J.-P. Dognon, G. Ohanessian, C. Clavaguéra, *J. Chem. Theory Comput.* **2014**, *10*, 3190.
- [37] F. Thauunay, J.-P. Dognon, G. Ohanessian, C. Clavaguéra, *Phys. Chem. Chem. Phys.* **2015**, *17*, 25968.
- [38] J. Liu, W. H. Miller, G. S. Fanourgakis, S. S. Xantheas, S. Imoto, S. Saito, *J. Chem. Phys.* **2011**, *135*, 244503.
- [39] T. L. C. Jansen, *J. Phys. Chem. B* **2014**, *118*, 8162.
- [40] D. Koner, S. M. Salehi, P. Mondal, M. Mewly, *J. Chem. Phys.* **2020**, *153*, 010901.
- [41] S. Abbate, G. Longhi, K. Kwon, A. Moscovitz, *J. Chem. Phys.* **1998**, *108*, 50.
- [42] F. Calvo, *Chirality* **2015**, *27*, 253.
- [43] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. J. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, T. Head-Gordon, *J. Phys. Chem. B* **2010**, *114*, 2549.
- [44] M. Thomas, B. Kirchner, *J. Phys. Chem. Lett.* **2016**, *7*, 509.
- [45] N. Berova, K. Nakanishi, R. W. Woody, *Circular dichroism: principles and applications*, John Wiley & Sons, New York **2000**.
- [46] D. A. McQuarrie, *Statistical mechanics*, Harper's chemistry series, Harper and Row, New York **1976**.
- [47] F. H. Stillinger, *J. Chem. Phys.* **1979**, *71*, 1647.
- [48] J. D. Jackson, *Classical electrodynamics 2nd Ed.*, Wiley, New York **1977**.
- [49] A. Kholmetskii, O. Missevitch, T. Yarman, *Eur. Phys. J. Plus* **2014**, *129*, 1.
- [50] J. Kaminský, F. Jensen, *J. Chem. Theory Comput.* **2007**, *3*, 1774.
- [51] T. D. Rasmussen, P. Ren, J. W. Ponder, F. Jensen, *Int. J. Quant. Chem.* **2007**, *107*, 1390.
- [52] F. Thauunay, C. Clavaguéra, G. Ohanessian, *Phys. Chem. Chem. Phys.* **2015**, *17*, 25935.
- [53] F. Thauunay, A. A. Hassan, R. J. Cooper, E. R. Williams, C. Clavaguéra, G. Ohanessian, *Int. J. Mass spectrom.* **2017**, *418*, 15.
- [54] F. Thauunay, C. Jana, C. Clavaguéra, G. Ohanessian, *J. Phys. Chem. A* **2018**, *122*, 832.
- [55] A. J. Stone, *Chem. Phys. Lett.* **1981**, *83*, 233.
- [56] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, P. Ren, *J. Chem. Theory Comput.* **2013**, *9*, 4046.
- [57] A. Stone, M. Alderton, *Mol. Phys.* **1985**, *56*, 1047.
- [58] B. Walker, C. Liu, E. Wait, P. Ren, *J. Comput. Chem.* **2022**, *43*, 1530.
- [59] J. W. Ponder, *TINKER - Software Tools for Molecular Design (version 8)*, <http://dasher.wustl.edu/tinker> (accessed July 22th, 2020).
- [60] J. Horníček, P. Kaprálová, P. Bouř, *J. Chem. Phys.* **2007**, *127*, 084502.
- [61] S. Jähnigen, ChirPy - A Python Package for Chirality, Dynamics, and Molecular Vibrations (0.23.2) **2022**, <https://doi.org/10.5281/zenodo.4775330>.
- [62] N. Troullier, J. L. Martins, *Phys. Rev. B* **1991**, *43*, 1993.
- [63] CPMD, *Copyright IBM Corp 1990-2019, Copyright MPI für Festkörperforschung Stuttgart 1997-2001*. **2022**, <http://www.cpmd.org/> (accessed April 27th, 2022).

- [64] S. Blanco, A. Lesarri, J. C. López, J. L. Alonso, *J. Am. Chem. Soc.* **2004**, *126*, 11675.
- [65] R. M. Balabin, *Phys. Chem. Chem. Phys.* **2010**, *12*, 5980.
- [66] M. Rahmani, M. E. A. Benmalti, *J. Biomol. Struct.* **2022**, *40*, 8560.
- [67] Merck, *IR Spectrum Table & Chart 2022*, <https://www.sigmaaldrich.com/FR/fr/technical-documents/technical-article/analytical-chemistry/photometry-and-reflectometry/ir-spectrum-table> (accessed May 10st, 2022).
- [68] J. Bowles, S. Jähnigen, R. Vuilleumier, F. Calvo, C. Clavaguéra, F. Agostini, *J. Chem. Phys.* **2023**, *158*, 094305.
- [69] J. Shen, C. Zhu, S. Reiling, R. Vaz, *Spectrochim. Acta A* **2010**, *76*, 418.
- [70] C. C. Wilson, D. Myles, M. Ghosh, L. N. Johnson, W. Wang, *New J. Chem.* **2005**, *29*, 1318.
- [71] Y. Rodríguez-Lazcano, B. Maté, O. Gálvez, V. J. Herrero, I. Tanarro, R. Escribano, *J. Quant. Spectrosc. Radiat. Transf.* **2012**, *113*, 1266.
- [72] M. Martinez, M.-P. Gaigeot, D. Borgis, R. Vuilleumier, *J. Chem. Phys.* **2006**, *125*, 144106.
- [73] M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith, D. Marx, *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 12068.
- [74] J. Sadlej, J. C. Dobrowolski, J. E. Rode, *Chem. Soc. Rev.* **2010**, *39*, 1478.
- [75] K. J. Jalkanen, I. M. Degtyarenko, R. M. Nieminen, X. Cao, L. A. Nafie, F. Zhu, L. D. Barron, *Theor. Chem. Acc.* **2008**, *119*, 191.
- [76] T. Plé, N. Mauger, O. Adjoua, T. J. Inizan, L. Lagardère, S. Huppert, J.-P. Piquemal, *J. Chem. Theory Comput.* **2023**, *19*, 1432.
- [77] J. A. Lemkul, J. Huang, B. Roux, A. D. J. MacKerell, *Chem. Rev.* **2016**, *116*, 4983.
- [78] S. W. Rick, B. J. Berne, *J. Am. Chem. Soc.* **1996**, *118*, 672.
- [79] S. Ma, X. Cao, M. Mak, A. Sadik, C. Walkner, T. B. Freedman, I. K. Lednev, R. K. Dukor, L. A. Nafie, *J. Am. Chem. Soc.* **2007**, *129*, 12364.
- [80] M. Pazderková, T. Pazderka, M. Shanmugasundaram, R. K. Dukor, I. K. Lednev, L. A. Nafie, *Chirality* **2017**, *29*, 469.
- [81] C. Zhang, C. Lu, Z. Jing, C. Wu, J.-P. Piquemal, J. W. Ponder, P. Ren, *J. Chem. Theory Comput.* **2018**, *14*, 2084.
- [82] M. J. Schnieders, T. D. Fenn, V. S. Pande, *J. Chem. Theory Comput.* **2011**, *7*, 1141.
- [83] J. Blasius, R. Elfggen, O. Hollóczki, B. Kirchner, *Phys. Chem. Chem. Phys.* **2020**, *22*, 10726.
- [84] M. A. Ali, M. S. Rahman, R. Roy, P. Gambill, D. E. Raynie, M. A. Halim, *J. Phys. Chem. A* **2021**, *125*, 2402.
- [85] V. Setnička, M. Urbanová, K. Volka, S. Nampally, J.-M. Lehn, *Chem. Eur. J.* **2006**, *12*, 8735.
- [86] A. Brizard, C. Aimé, T. Labrot, I. Huc, D. Berthier, F. Artzner, B. Desbat, R. Oda, *J. Am. Chem. Soc.* **2007**, *129*, 3754.
- [87] H. Sato, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7671.

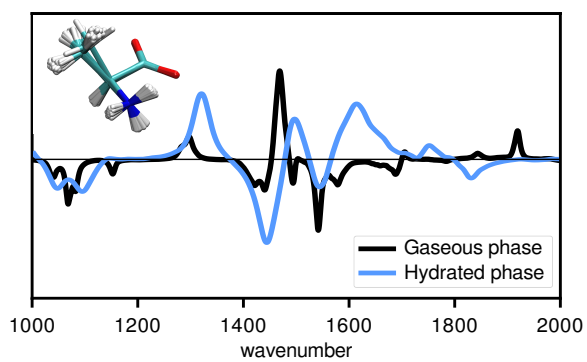


Table of Contents entry: Vibrational circular dichroism spectra in the gas and condensed phases are obtained from the Fourier transform of time correlation functions of electric and magnetic dipole moments generated along molecular dynamics trajectories with a polarizable force field