Influence of the environment on the infrared spectrum of alanine: an effective modes analysis

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

Influence of the environment on the infrared spectrum of alanine: an effective modes analysis

Jessica Bowles,[1] Sascha Jähnigen,[2,3] Rodolphe Vuilleumier,[3]
Florent Calvo,[4] Carine Clavaguéra,[1] and Federica Agostini*[1]

[1] *Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, 91405, Orsay, France*
[2] *Institut des Sciences Moléculaires d'Orsay (ISMO), CNRS, Université Paris-Saclay, 91405 Orsay, France*
[3] *PASTEUR Laboratory, Département de Chimie, Ecole Normale Supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France*
[4] *Université Grenoble Alpes, CNRS, LIPhy, F-38000 Grenoble, France*

The vibrational spectrum of the alanine amino acid was computationally determined in the infrared range of 1000–2000 cm^{-1}, under various environments encompassing the gas, hydrated, and crystalline phases, by means of classical molecular dynamics trajectories carried out with the AMOEBA polarizable force field. An effective mode analysis was performed in which the spectra are optimally decomposed into different absorption bands arising from well-defined internal modes. In the gas phase, this analysis allows to unravel the significant differences between the spectra obtained for the neutral and zwitterionic forms of the amino acid. In condensed phases, the method provides invaluable insight into the molecular origins of the vibrational bands and further shows that peaks with similar positions can be traced to rather different molecular motions.

*Email: federica.agostini@universite-paris-saclay.fr

I. INTRODUCTION

Infrared (IR) spectroscopy is a powerful tool to characterize the chemical physics of molecular systems in their environment [1]. In interpreting IR spectra, the synergy between experiments and computer simulations is particularly relevant for the assignment of absorption bands to well-defined molecular vibrations or to functional groups as well as for the identification of how dynamical changes in the environment affect the absorption bands. When combined with vibrational circular dichroism (VCD), i.e., the difference in the response of a molecular system to left- and right-circularly polarized IR light [2–5], IR spectroscopy represents an extremely sensitive tool for characterizing the structure and to efficiently identify absolute and relative configurations of chiral molecules [6].

In complex systems, the assignment of IR and VCD bands associated to the excitation of particular vibrational modes in a molecule, which can be interpreted as the vibrational signature of that molecule, often turns out to be a cumbersome task [7–15]. This is due, for instance, to anharmonic effects arising at increasing temperature, or to the presence of an environment, e.g., a solvent, that can mask the vibrational signature of the solute molecule of interest. In addition, while molecular dynamics (MD) simulations at finite temperature are able to realistically capture those effects in IR and VCD spectra, applying a standard normal modes analysis for their interpretation is clearly not sufficient, since a theory beyond the harmonic approximation is needed.

To overcome these limitations, several methods have been developed to interpret the IR vibrational spectrum obtained from a MD trajectory [11–13, 16]. The principal modes analysis has been proposed in which the average mass-weighted normal modes and associated frequencies are calculated using the Hessian matrix [12] or the covariances of Cartesian atomic velocities [13]. Martinez *et al.* have developed an effective modes analysis, with the aim to decompose the IR spectrum as an approximate sum of effective modes contributions [16]. These effective modes are defined as linear combinations of mass-weighted atomic displacements, similarly to the harmonic normal modes analysis, and in the zero temperature limit, the harmonic normal modes are recovered. Since IR and VCD peaks are located at the same frequencies, the identification of the effective modes from the IR absorption bands can be used as an interpretation tool for VCD as well. Another useful approach for computational vibrational spectroscopy, called driven molecular dynamics (DMD), was proposed by Bowman *et al.* [11]. In this method, a sinusoidal driving force is added to the Hamiltonian. The driving frequency can be varied through a frequency range and energy absorption should maximize locally at the normal mode frequencies. An examination of the molecular motion at a given resonance provides the corresponding normal mode. This DMD method was coupled with the determination of the IR spectra itself through the Fourier transform of the autocorrelation function by Thaunay *et al.* [17].

The analysis of IR spectra based on the identification of effective modes [16] has been previously applied to a variety of molecular systems [18–23] where diverse MD techniques were employed to generate the trajectories. In the present work, we focus on the IR spectroscopy of the alanine amino acid [24–28], a chiral molecule that some of the authors have recently studied based on ab initio and classical MD simulations in different phases [6]. The choice of alanine is motivated by the availability of experimental data [25], but this molecule also serves here as a convenient testbed for implementing and comparing

methodologies for vibrational spectroscopy. In particular, we aim at developing tools for simulating VCD spectra based on polarizable force fields, and in the present article we address the specific issue of modes assignment. The effective modes analysis is used here to (i) identify the spectral signatures of neutral and zwitterionic alanine in the gas phase in the region 1000–2000 cm^{-1} as well as to (ii) determine the effect of the environment on the modes and on the absorption bands in the same spectral region by studying the hydrated form of alanine and its crystal form.

The paper is organized as follows. In Section II, we present the theory and illustrate the numerical procedures employed to determine the effective modes in alanine (Section II A) and to generate the MD trajectory (Section II B). Section III is devoted to a detailed presentation of the results: in Section III A, the neutral and zwitterionic forms of alanine are compared in the gas phase; in Section III B, the hydrated and crystalline zwitterionic forms are studied. Our conclusions are presented in Section IV.

II. METHODS

The IR vibrational spectra of isolated alanine as well as its hydrated and crystalline forms were investigated based on classical MD simulations using the AMOEBA polarizable force field [29]. In the gas phase, we studied both zwitterionic and neutral forms, which are shown in Fig. 1 (upper panels), even though, when isolated, alanine is most stable in the neutral form. In contrast, once alanine is hydrated or in a crystal, as shown in Fig. 1 (lower panels), the zwitterionic form becomes more stable.

In the following, we will present the theoretical and technical background at the basis of our methodology, with the aim to characterize the vibrational modes of alanine in different environments in the spectral region 1000–2000 cm^{-1} (Section II A) based on the calculation of the IR spectra of the molecule in the various phases introduced above (Section II B).

A. Effective modes analysis including temperature and anharmonic effects

Our approach to characterize the vibrational modes of alanine in the region 1000–2000 cm^{-1}, and thus to assign absorption bands to well-defined internal molecular vibrations, is by determining the effective modes [16, 30]. The present effective modes analysis was introduced in Refs. [16, 30] and then extensively used and documented in Refs. [18–23]. Therefore, we recall here the background theory only briefly.

The IR spectrum $I_{\mathrm{IR}}(\omega)$ of a molecular system is computed from the Fourier transform of the autocorrelation

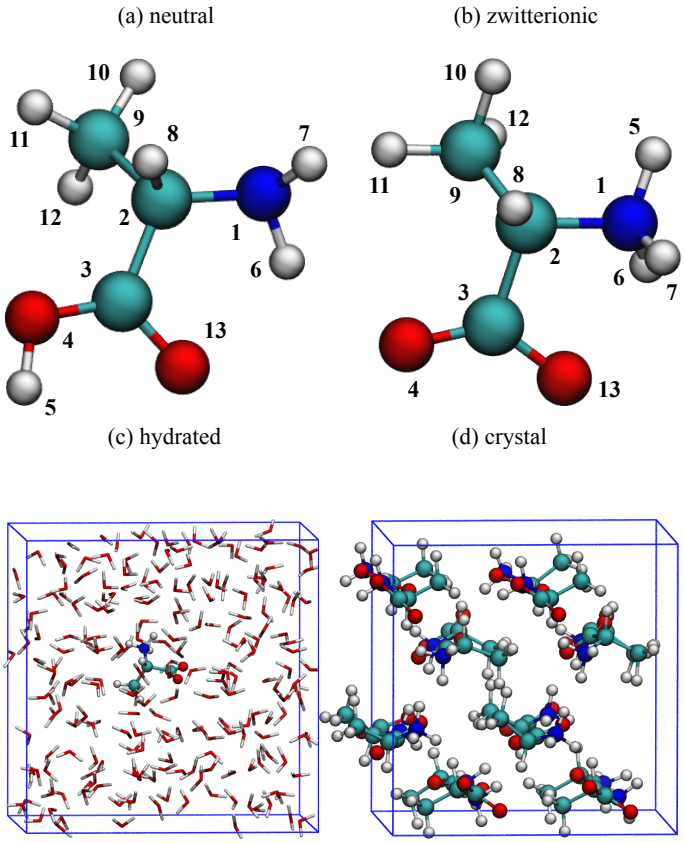

FIG. 1. (a) Neutral and (b) zwitterionic forms of isolated alanine. The atoms are labeled by numbers which are used below to define the set of internal coordinates needed to initialize the effective modes analysis; (c) hydrated and (d) crystal forms of zwitterionic alanine.

function of the time derivative of the electric dipole moment $\boldsymbol{\mu}$, namely

$$I_{\mathrm{IR}}(\omega) \propto \int_{-\infty}^{+\infty} \langle \dot{\boldsymbol{\mu}}(0) \cdot \dot{\boldsymbol{\mu}}(t) \rangle \, e^{-i\omega t} dt \qquad (1)$$

The averaging operation $\langle \cdot \rangle$ is performed in our simulations over the canonical ensemble at $T = 300$ K (see Section II B for details).

Introducing the effective modes, we approximate the IR spectrum $I_{\mathrm{IR}}(\omega)$ of a molecule of N atoms as a sum of contributions from individual effective modes, $I_{\mathrm{EM}}(\omega)$, namely as

$$I_{\mathrm{IR}}(\omega) \simeq I_{\mathrm{EM}}(\omega) \propto \sum_{i=1}^{3N-6} \left\langle \left| \frac{\partial \boldsymbol{\mu}}{\partial q_i}(0) \right|^2 \right\rangle P_{ii}^{(q)}(\omega) \qquad (2)$$

$$= \sum_{i=1}^{3N-6} I_{\mathrm{EM}}^{(i)}(\omega) \qquad (3)$$

In Eq. (3), $I_{\mathrm{EM}}^{(i)}(\omega)$ defines the contribution of effective mode i to the global effective spectrum $I_{\mathrm{EM}}(\omega)$. The sum over i runs over the $3N-6$ internal degrees of freedom of

the molecule; $\mathbf{q} = \{q_i\}_{i=1,\ldots,3N-6}$ indicates a set of effective coordinates, or effective modes, to be determined by the effective modes analysis; $P^{(q)}_{ii}(\omega)$ is the power spectrum, i.e., the Fourier transform of the velocity autocorrelation function, expressed using the effective modes $\mathbf{q}$

$$P^{(q)}_{ii}(\omega) = \int_{-\infty}^{+\infty} \langle \dot{q}_i(t)\dot{q}_i(0)\rangle \, e^{-i\omega t} dt \tag{4}$$

The purpose of the effective modes analysis is to determine $I_{\mathrm{EM}}(\omega)$ that best captures the absorption bands of $I_{\mathrm{IR}}(\omega)$ via a suitable construction of the effective modes $\mathbf{q}$ as linear combinations of internal coordinates $\boldsymbol{\zeta}$ of the molecule at hand, alanine in this study. Such a suitable construction is achieved by imposing that the power spectrum of each effective mode $P^{(q)}_{ii}(\omega)$ is maximally localized in frequency.

In order to derive the effective modes approximation (2) to the IR spectrum (1), the following hypotheses are made [18]:

1. positions and velocities in the effective modes coordinate system are uncorrelated variables, on average, along the trajectory;
2. rotations and vibrations are uncorrelated on average, and the rotational time correlation function decay is much slower than the vibrational time correlation function;
3. the time correlation function for the derivatives of the dipole moment with respect to the effective modes decays more slowly than the vibrational time correlation function, it is sufficient to consider its value at correlation time $t = 0$;
4. the off-diagonal elements of the q-power spectrum are negligible if compared to the diagonal elements (normal mode condition).

The linear combination of internal coordinates yielding the effective modes is

$$\mathbf{q} = \mathbf{Z}^{-1}\boldsymbol{\zeta} \quad \text{or} \quad q_i = \sum_{k=1}^{3N-6} Z^{-1}_{ik}\zeta_k \tag{5}$$

with $\mathbf{Z}^{-1}$ identifying the transformation matrix from internal $\boldsymbol{\zeta}$ to effective $\mathbf{q}$ modes. The elements of the transformation matrix, i.e., Z^{-1}_{ik}, are parameters to be determined by minimizing the functional $\Omega^{(m)}$

$$\Omega^{(m)} = \sum_{i=1}^{3N-6} \left[\frac{\beta}{2\pi} \int_{-\infty}^{+\infty} d\omega \, \omega^{2m} P^{(q)}_{ii}(\omega) - \left(\frac{\beta}{2\pi} \int_{-\infty}^{+\infty} d\omega \, \omega^{m} P^{(q)}_{ii}(\omega) \right)^2 \right] \tag{6}$$

The power spectrum $P^{(q)}_{ii}(\omega)$ is real and positive, which follows from the Wiener-Khintchine theorem. Therefore, the quantity $P^{(q)}_{ii}(\omega)$ can be interpreted as a probability distribution with normalization condition chosen as

$$\frac{\beta}{2\pi} \int_{-\infty}^{+\infty} d\omega \, P^{(q)}_{ij}(\omega) = \delta_{ij} \tag{7}$$

This condition allows a system in thermal equilibrium in the canonical ensemble to satisfy the relation $\langle \dot{x}_i \dot{x}_j \rangle = \delta_{ij}/(\beta m_i)$ where $\dot{x}_i$ is the Cartesian velocity along the i direction and m_i is the i-th mass. With such an interpretation of the power spectrum, the functional defined in Eq. (6) can be rewritten as

$$\Omega^{(m)} = \sum_{i=1}^{3N-6} \left(\langle \omega^{2m} \rangle_i - \langle \omega^m \rangle_i^2 \right) \tag{8}$$

where $\langle f(\omega) \rangle_i$ is the average of a function $f(\omega)$ over all positive frequencies weighted by the power spectrum of mode i. With $m = 1$, the quantity $\langle \omega^2 \rangle_i - \langle \omega \rangle_i^2$ is the variance of the power spectrum of mode i, whose minimization leads to maximizing the localization of the corresponding band $P^{(q)}_{ii}(\omega)$. With $m = 2$ this method is equivalent to a normal modes analysis performed with the thermally averaged Hessian matrix [16, 22], which differs from a standard (harmonic) normal modes analysis as it includes anharmonicities through their thermal contribution. Once a set of internal coordinates $\boldsymbol{\zeta}$ is defined to initialize the minimization procedure, calculating the stationary variations of $\Omega^{(m)}$ with respect to the Z^{-1}_{ik} yields the optimal linear combination of $\boldsymbol{\zeta}$ that produces a set of effective modes $\mathbf{q}$ with maximally localized absorption bands. Note that $P^{(q)}_{ii}(\omega)$ depends on the whole MD trajectory via the ensemble average operation, so does the search for the optimal linear combination. It is worth mentioning here that the idea of determining maximally localized effective modes in the harmonic limit has also been exploited by Jakob and Reiher [31] in the context of quantum chemical calculations.

The minimization procedure that defines the effective modes is as follows.

- First, we use a matrix notation for the power spectra, namely we use the symbols $\mathbf{P}^{(q)}(\omega)$ and $\mathbf{P}^{(\zeta)}(\omega)$ in effective and internal coordinates, respectively. Since the Fourier transform is a linear operator, the two quantities are related as

$$\mathbf{P}^{(q)}(\omega) = \mathbf{Z}^{-1}\mathbf{P}^{(\zeta)}(\omega)\mathbf{Z}^{-1^T} \tag{9}$$

where $\mathbf{Z}^{-1^T}$ is the transpose of the transformation matrix $\mathbf{Z}^{-1}$ from internal to effective coordinates.

- Second, we introduce the symbol

$$\mathbf{K}_q^{(m)} = \frac{\beta}{2\pi}\int_{-\infty}^{+\infty} d\omega\, \omega^m \mathbf{P}^{(q)}(\omega) = \mathbf{Z}^{-1}\mathbf{K}_\zeta^{(m)}\mathbf{Z}^{-1^T} \quad (10)$$

such that the functional in Eq. (6) can be rewritten as

$$\Omega^{(m)} = \sum_{i=1}^{3N-6}\left[\left(\mathbf{Z}^{-1}\mathbf{K}_\zeta^{(2m)}\mathbf{Z}^{-1^T}\right)_{ii} - \left(\mathbf{Z}^{-1}\mathbf{K}_\zeta^{(m)}\mathbf{Z}^{-1^T}\right)_{ii}^2\right] \quad (11)$$

- Third, the functional $\Omega^{(m)}$ is minimized with respect to the matrix $\mathbf{Z}^{-1}$, calculating $\delta\Omega^{(m)}/\delta\mathbf{Z}^{-1} = 0$ with normalization constraint $\mathbf{K}_q^{(0)} = \mathbf{Z}^{-1}\mathbf{K}_\zeta^{(0)}\mathbf{Z}^{-1^T} = \mathbf{1}$ following from Eq. (7).
- Fourth, the minimization procedure is translated into a generalized eigenvalue problem

$$\mathbf{K}_\zeta^{(m)}\mathbf{Z}^{-1^T} = \mathbf{K}_\zeta^{(0)}\mathbf{Z}^{-1^T}\mathbf{\Lambda} \quad (12)$$

where $\mathbf{\Lambda}$ is a diagonal eigenvalue matrix containing the frequencies of the effective modes. Such an eigenvalue problem is equivalent to that derived within the harmonic normal modes analysis, where the Hessian matrix is replaced by $\mathbf{K}_\zeta^{(m)}$ and the mass matrix is replaced by $\mathbf{K}_\zeta^{(0)}$.

The results presented in Section III were obtained by choosing $m = 2$. We performed the analysis of the zwitterionic form of alanine in the gas phase for $m = 4$ as well, and no major differences were observed.

The effective modes analysis itself is not very costly from the computational point of view, the main limiting factor being memory storage, and convergence of the MD trajectory: positions and velocities have to be read in input and stored before solving the generalized eigenvalue problem of Eq. (12).

Internal coordinates are naturally based on elementary geometrical parameters defining bond lengths, angles or/and dihedrals. For the present system in zwitterionic form, the chosen internal coordinates are listed in Table I. Other choices of internal coordinates are always possible. We have tested alternative possibilities, without, however, identifying major differences either in the decomposition of the spectrum or in the identity of the effective modes.

TABLE I. List of the 33 internal coordinates used for the effective modes analysis of alanine in the zwitterionic form (12 bond lengths, 16 valence angles, 5 dihedral angles). The labels of the atoms refer to Fig. 1(b). For the neutral form, the bond N1-H5 is replaced by the bond O4-H5, the valence angle H5-N1-H6 is suppressed, and the dihedral angle H5-O4-C3-O13 is added.

bond lengths	**valence angles**	**dihedral angles**
N1-C2	C3-C2-N1	O4-C3-C2-N1
C2-C3	O4-C3-C2	H5-N1-C2-C3
C3-O4	C2-N1-H5	H8-C2-N1-H7
N1-H5 (neutral: O4-H5)	C2-N1-H6	C9-C2-N1-H7
N1-H6	C2-N1-H7	H10-C9-C2-N1
N1-H7	N1-C2-H8	(neutral: H5-O4-C3-O13)
C2-H8	C9-C2-N1	
C2-C9	H10-C9-C2	
C9-H10	H11-C9-C2	
C9-H11	H12-C9-C2	
C9-H12	O13-C3-C2	
C3-O13	H5-N1-H6 (neutral: none)	
	H7-N1-H6	
	H12-C9-H11	
	H10-C9-H11	
	O13-C3-O4	

B. Molecular dynamics with the AMOEBA force fields

In addition to bonding and non-bonding interactions of conventional force fields, polarizable force fields allow to include the polarization of the atomic sites, which can be to some extent related to the deformation of the electronic density cloud in response to the induced electric field from the surrounding charges. Such polarization effects turned out to be the key to capture the essential non-Born-Oppenheimer effects in VCD spectroscopy of chiral molecules [2, 3] based on classical MD, since the force-field polarizability adds correlation between charges and geometries that is absent in a fixed charges model.

In our simulations, the 2013 AMOEBA protein version polarizable force field was used [29, 32]. AMOEBA was originally developed for biomolecules such as proteins and DNA, and some organic molecules [29, 33, 34]. The main interest of this force field lies in its ability to reproduce relative conformational energies. Conventional

force fields that do not include polarization effects only reproduce about 50% of the conformations found at the MP2/aug-cc-pVTZ level of theory whereas AMOEBA reproduces about 80% of them [35]. The good reproduction of the relative energy is due to the use of permanent electrostatic multipolar moments (charges, dipole and quadrupole moments) on each atom, and induced dipoles reacting self-consistently to the presence of local electric fields on each polarizable site. Polarization at very short distances was damped through the effective charge smearing method of Thole [36]. This leads to a good precision for thermodynamic properties such as free energies of solvation and complexation. The AMOEBA force field has also been extended to IR spectroscopy of peptides in the gas phase and hydrated ions very satisfactorily [17, 37, 38].

The AMOEBA force field is distributed with the Tinker software package [39]. To determine the IR spectrum, the electric dipole moments were collected for multiple trajectories, then processed through the Fourier transform of the time autocorrelation function of the electric dipole moments by the ChirPy program [40]. This gave an individual spectrum for each trajectory from which an average spectrum was determined. In the gas phase, 40 trajectories of 200 ps each were performed and in the condensed phases 30 trajectories of 40 ps each were assembled. A time step of 0.5 fs was employed as it is found to be optimal for IR spectroscopy in this frequency range [41], and the electric dipole moments were collected with a time step of 4 fs for the Fourier transform of the autocorrelation function. All systems were pre-equilibrated at 300 K with the Nosé-Hoover thermostat before the electric dipole moments were recorded.

MD simulations of the hydrated phase were carried out using periodic boundary conditions with a cubic box with an edge length of 18.6 Å that contained 211 water molecules. This box was pre-equilibrated before soaking the alanine solute.

In accordance with experimental crystallographic data, an alanine crystal was prepared for simulations. Given the chiral space group of alanine ($P2_12_12_1$), a periodic box of 12.073 × 12.342 × 11.57 Å was constructed containing 16 alanine molecules. The Ewald real-space cutoff was 7 Å, and the van der Waals cutoff was 7 Å for the hydrated phase and 5 Å for the crystalline phase.

III. EFFECTIVE MODES AND BANDS ASSIGNMENT OF ALANINE

In this section, we report our effective modes analysis performed on gas phase alanine in the neutral and zwitterionic forms in Section III A, and on hydrated and crystalline alanine in zwitterionic form in Section III B.

Note that, in our effective modes analysis, we only report the power spectra of the effective modes $P_{ii}^{(q)}(\omega)$ whose integral over the whole frequency range satisfies the normalization condition of Eq. (7). Therefore, the reported intensities cannot be used to compare with experiments. The reported power spectra were smoothened out by performing a locally estimated scatterplot smoothing (LOESS) [42], to attenuate the significant noise of the bare spectra. The bare spectra were all normalized to unity when integrated over the whole spectral domain before applying the LOESS procedure. Differences in intensities when the bands have different delocalizations are then justified by the normalization. All along the analysis and bands assignment procedure, the modes were labelled with their average frequencies, which are defined as the average values over the bare normalized spectra.

Before reporting on the full IR spectrum of alanine, let us briefly analyze the spectrum of its gas phase zwitterionic form in terms of fragments contributions. Since the spectral region 1000–2000 cm^{-1} of interest in this work is densely populated by overlapping and not well-defined absorption bands, this preliminary study aims to clarify the results that will be discussed in the following sections on the alanine molecule. From the trajectory of the alanine molecule, positions and velocities of three fragments, namely CH_3, NH_3, and CO_2 were extracted in order to identify their effective modes and the bands of the full IR spectrum arising from their vibrations. Note that, while the analysis is performed on the isolated fragments, their dynamics accounts for the presence of the remaining degrees of freedom of the molecule.

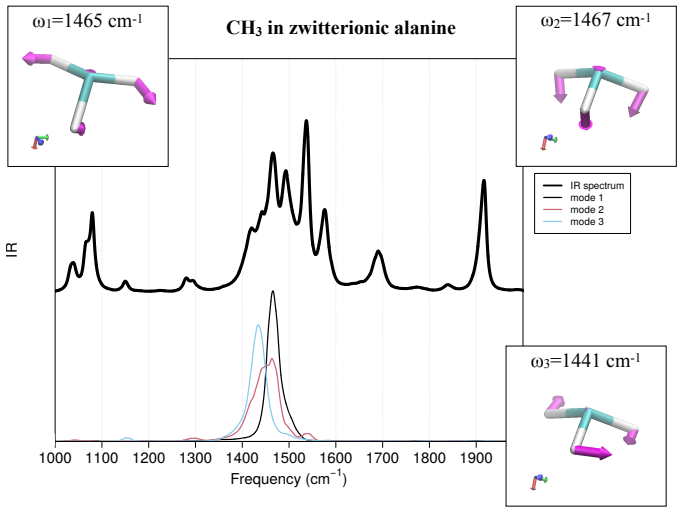

FIG. 2. IR spectrum of alanine in zwitterionic form in the gas phase. Thick black line: anharmonic IR spectrum from the MD trajectory. Colored lines: power spectra determined by the effective modes analysis, corresponding to the 3 modes of the fragment CH_3. Insets: stick representation of the fragment CH_3, along with the effective modes, depicted as arrows.

The IR absorption bands of the fragment CH_3 are mainly located in the region 1400–1500 cm^{-1}, as shown by the colored bands in the lower panel of Fig. 2, and thus contribute to the large absorption region of zwitterionic alanine in the gas phase in the region 1300–1600 cm^{-1} (black line in the upper panel of Fig. 2). The effective

modes identified by our analysis are, rather unsurprisingly, combinations of bending modes (mode 1 and mode 3, depicted as arrows in the insets of Fig. 2) and the umbrella mode (mode 2 in Fig. 2). These modes will be referred to as $\delta_{as}CH_3$ (modes 1 and 3) and $\delta_s CH_3$ (mode 2) in Section III A.

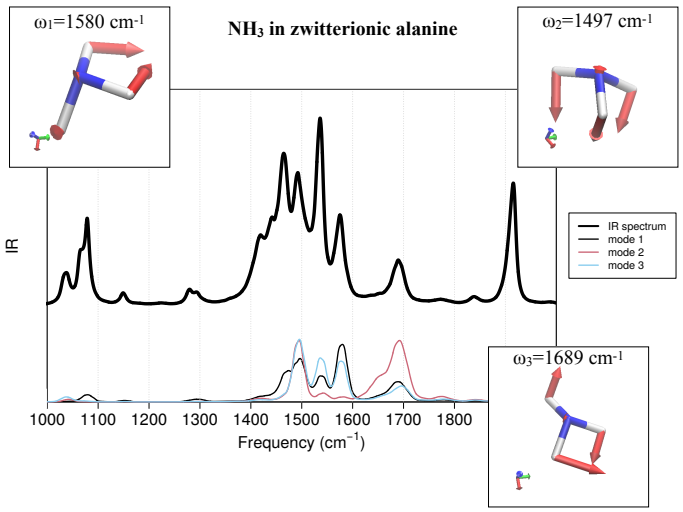

FIG. 3. IR spectrum of alanine in zwitterionic form in the gas phase. Thick black line: anharmonic IR spectrum from the MD trajectory. Colored lines: power spectra determined by the effective modes analysis, corresponding to the 3 modes of the fragment NH_3. Insets: stick representation of the fragment NH_3, along with the effective modes, depicted as arrows.

The IR absorption bands of the fragment NH_3 are slightly blue shifted with respect to those of CH_3, and are mainly distributed in the region 1450–1750 cm^{-1}, as shown by the colored bands in the lower panel of Fig. 3. Despite the blue shift, the vibrations within the NH3 fragments are expected to be coupled to those of CH_3 since the absorption bands partially overlap. Therefore, similarly to CH_3, NH_3 contributes to the large absorption region of zwitterionic alanine in the gas phase in the region 1300–1600 cm^{-1} (black line in the upper panel of Fig. 3) as well as to the isolated band at 1700 cm^{-1}. In contrast, we find that the bands are not well localized in frequency as it was the case for the CH_3 fragment, which might be a manifestation of strong anharmonicity and coupling of NH_3 with other modes that have not been considered in this fragment analysis [43]. The effective modes identified by our analysis are very similar to those in CH_3, namely they are a HNH bending mode (mode 1, depicted as arrows in the inset of Fig. 3), the umbrella mode (mode 2 in Fig. 3), and a combination of a HNH bending mode with an out-of-plane vibration of the remaining NH bond (mode 3 in Fig. 3). Similar to the CH_3 fragment, these modes will be denoted as $\delta_{as}NH_3$ and $\delta_s NH_3$ in Section III A.

The IR absorption band of the fragment CO_2 is composed of two main bands: one band appears essentially in the region 1500–1600 cm^{-1}, but is coupled to smaller

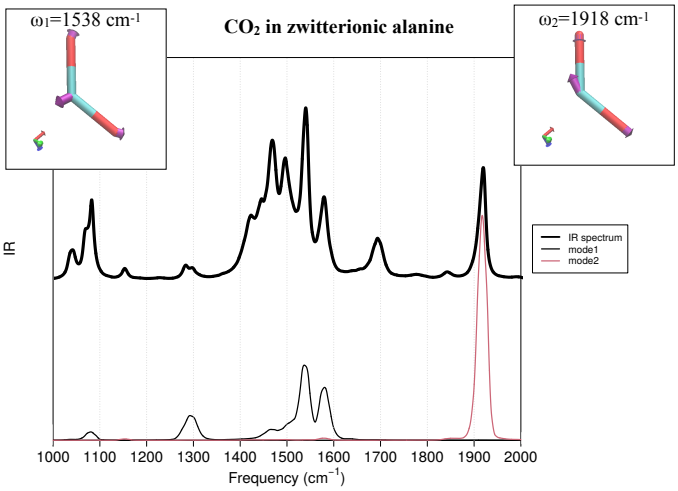

FIG. 4. IR spectrum of alanine in zwitterionic form in the gas phase. Thick black line: anharmonic IR spectrum from the MD trajectory. Colored lines: power spectra determined by the effective modes analysis, corresponding to the 2 modes of the fragment CO_2. Insets: stick representation of the fragment CO_2, along with the effective modes, depicted as arrows.

bands at (a bit less than) 1100 cm^{-1} and at 1300 cm^{-1}; one band is very well defined and localized at around 1900 cm^{-1}. Therefore, also for this fragment, we find the participation of the bending motion in CO_2 (mode 1, depicted as arrows in the inset of Fig. 4) to the large absorption region of zwitterionic alanine in the gas phase in the region 1300–1600 cm^{-1} (black line in the upper panel of Fig. 4), while the asymmetric stretching in CO_2 (mode 2, in the inset of Fig. 4) produces an absorption peak in the high frequency region of the analyzed spectrum. In Section III A, those modes will be referred to as δCO, indicating a bending motion where one or both CO bonds are involved, and as ν_{as}COO, respectively. It is worth mentioning here that, in general, the CO stretch would be expected at lower frequencies, i.e., at around 1650–1700 cm^{-1} [24, 25]. Finding this mode near 1900 cm^{-1} is a consequence of the parametrization of the force field.

A. Isolated zwitterionic and neutral forms of alanine

Figure 5 reports the full IR absorption spectrum of zwitterionic alanine in the gas phase (upper black line), along with the normal modes peaks (upper gray bars) and its decomposition in effective modes contributions (lower colored curves).

In the low frequency region of Fig. 5, four bands can be clearly identified: two overlapping bands in the region 1000–1100 cm^{-1}, one band in the region 1100–1200 cm^{-1}, and one band localized at 1300 cm^{-1}. The four bands see the participation of the three fragments of the molecule analyzed above, together with a contri-

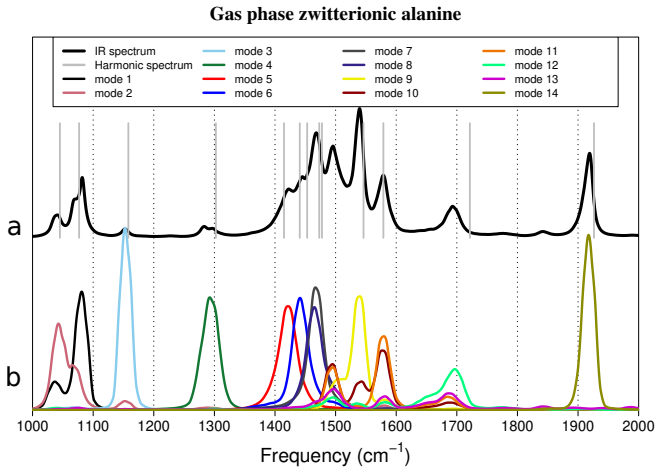

FIG. 5. IR spectrum of zwitterionic alanine in the gas phase [(a) upper black line] along with the normal modes peaks [(a) gray vertical bars] and its decomposition in effective modes bands [(b) lower colored curves].

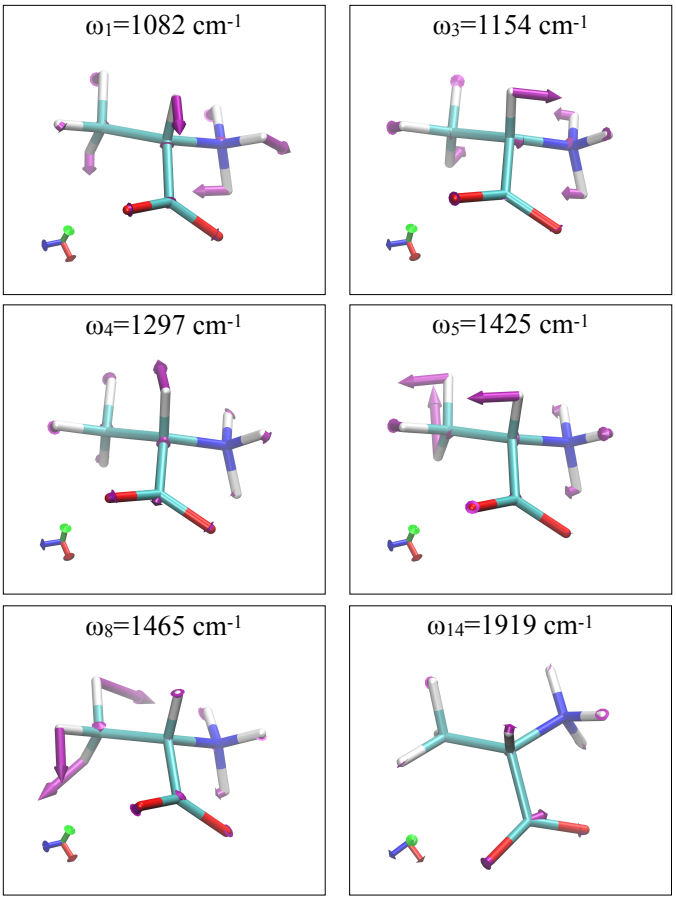

FIG. 6. Effective modes associated to the spectral regions: 1000–1350 cm^{-1} (modes 1, 3 and 4), 350–1500 cm^{-1} (modes 5 and 8), 1500–1900 cm^{-1} (mode 14) of the IR spectrum of zwitterionic alanine in the gas phase.

bution from the vibration of the CH bond that is not involved in the CH_3 fragment (see Fig. 6, modes 1, 3 and 4).

The region 1350–1500 cm^{-1} of the spectrum in Fig. 5 is populated by four main bands, associated to the four modes. Two of these modes, i.e., 5 and 6, are shown in Fig. 6. As expected from the fragment analysis reported above, this region of the spectrum sees the main participation from the bending and umbrella modes in the two fragments CH_3 and NH_3. Some distorted CO_2 bending motions contribute as well to the higher frequency modes (for instance, see mode 8 in Fig. 6), giving rise to the overlapping band slightly above 1450 cm^{-1} in Fig. 5.

In the remaining part of the spectrum of zwitterionic alanine in the gas phase, we identify five modes between 1500–1900 cm^{-1}. The high energy band at 1900 cm^{-1} is easily assigned to the asymmetric stretching motion in the fragment CO_2, as observed in Fig. 4 as well: this is mode 14 represented in Fig. 6. Centered at 1700 cm^{-1}, we mainly identify a band corresponding to mode 12 (mode not represented in Fig. 6), which is produced by the excitation of the bending and the umbrella motion in NH_3 (see Fig. 3 for reference), which is now coupled to the bending motion in the fragment CH_3 as well as to the vibration of the CH bond that is not involved in CH_3. We confirm here, as seen in Fig. 3, that vibrations in NH_3 cannot be completely decoupled from those in other fragments, probably due to strong anharmonicity. This is further supported by the lack of full localization of the corresponding absorption bands, which remain delocalized and present several peaks.

The three remaining bands are located in the region 1500–1600 cm^{-1} and are mainly produced by the absorption of the NH_3 fragment, with subsequent excitation of the bending and umbrella vibrations (modes not reported in Fig. 6).

The analysis of the effective modes and the bands assignment just discussed are summarized in Table II, where we report, in the right column, the average frequencies of the modes identified from their effective modes analysis. The left column of the table reports experimental IR data from Ref. [25]. The modes assignment is presented in the central column, where we adopt the notation for the modes used in Ref. [25]. Note that, in Ref. [25] the bands were assigned based on harmonic normal modes calculations with electronic energy calculated at the Hartree-Fock level. In general, we find that the effective modes analysis captures many features of the experimental IR spectrum, even though quantitative agreement is lacking in some regions of the analyzed domain. To quantify the deviation of the effective modes frequencies from the experimental values, we calculated the mean absolute error (MAE) for the modes whose assignment corresponds to the experimental one. For the zwitterionic form of alanine, Table II leads to a MAE of 77.5 cm^{-1}. It is plausible to assume that the deviations are mainly caused by either the force field used to generate the MD trajectory or by the strongly anharmonic

TABLE II. Left column: experimental frequencies collected from Ref. [25] in the spectral region 1000–2000 cm^{-1} for the zwitterionic form of alanine. Central column: bands assignment using the abbreviation from Ref. [25] (ρ for rocking, δ for bending, ν for stretching, ω for wagging; s for symmetric and as for asymmetric modes). Right column: calculated effective frequencies from this work.

Experimental frequency (cm^{-1})	Assignment	Calculated frequency (cm^{-1})
1014	ρCH_3	1043
1027	δCH	
1114	ρCH_3	1082
1149	ρNH_3	1154
1235	δCH	
1307	$\delta_s NH_3$	1425 1441
1354	νCOO	1297
1408	$\delta_s CH_3$	
1452	$\delta_{as} CH_3$	1467
1523	$\delta_{as} CH_3$	1465
	$\delta_{as} CH_3$, $\delta_s NH_3$, δCO	1538
1592	$\delta_{as} NH_3$	1580
1603	$\nu_{as} COO$	1919
	ωCH_3, ωNH_3	1702

dynamics of alanine. Another contribution to the discrepancies might originate from the level of theory and the harmonic approximation used in Ref. [25] to assign the modes listed in Table II.

As expected from their different structures, the IR absorption spectrum of gas phase neutral alanine is very different from that calculated for the zwitterionic form. By comparing Fig. 5 and Fig. 7 some major differences can be pointed out: (i) in the region 1000–1300 cm^{-1}, the low intensity band at 1150 cm^{-1} in zwitterionic alanine becomes more intense and splits into two contributions for the neutral form; (ii) in the region 1350–1500 cm^{-1}, the large absorption band of zwitterionic alanine becomes less intense in the neutral form, and four main peaks appear to be better resolved; (iii) the region 1500–1900 cm^{-1} becomes more IR active for the neutral form of alanine that for the zwitterionic form, with the complete disappearance of absorption above 1800 cm^{-1}.

The effective modes of neutral alanine in the gas phase were determined, in order to identify the modes causing the differences between the corresponding IR spectra of the two forms of the molecule. For the sake of clarity, we do not describe here all the modes determined in the spectral region 1000–2000 cm^{-1}, but we only comment on the effective modes mainly responsible for the differences (i), (ii) and (iii) discussed above.

In Fig. 8, we find that the selected effective modes responsible for the different spectral features of neutral alanine with respect to zwitterionic alanine clearly involve some vibrational motions of the fragments that mainly distinguish the two species: the OH bond in the COOH and the NH_2 fragments. Mode 2 of Fig. 8, centered around 1150 cm^{-1} is a seemingly wagging motion of the

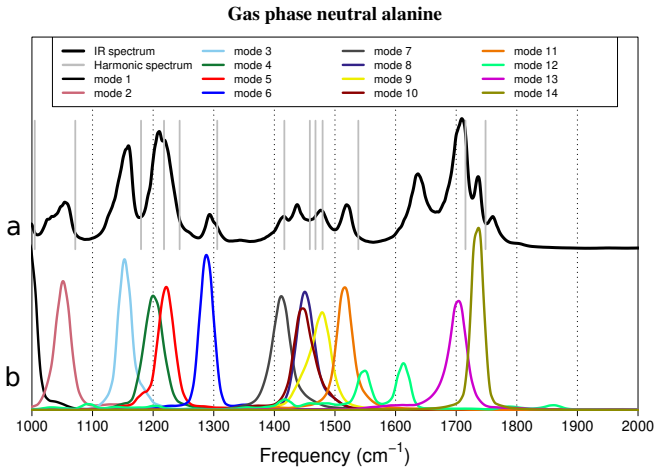

FIG. 7. IR spectrum of neutral alanine in the gas phase [(a) upper black line] along with the normal modes peaks [(a) gray vertical bars] and its decomposition in effective modes [(b) lower colored curves].

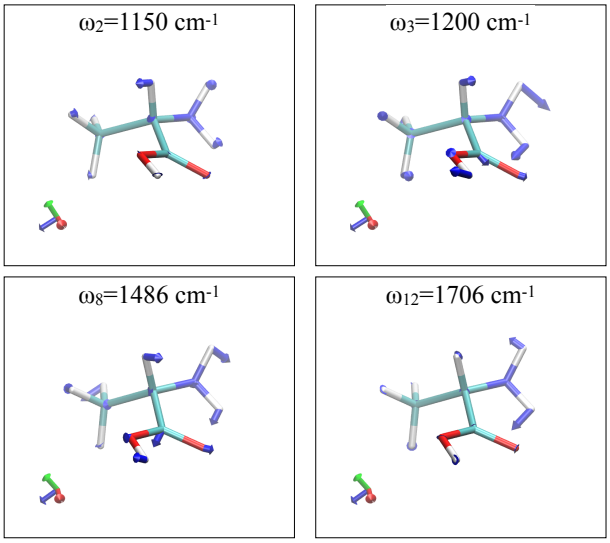

FIG. 8. Effective modes associated to the spectral region 1000–1800 cm^{-1} of the IR spectrum of neutral alanine in the gas phase, responsible for the major differences with the IR spectrum of zwitterionic alanine in the gas phase.

NH_2 moiety, which is absent in the zwitterionic form. Similarly, the wagging motion of NH_2 at 1200 cm^{-1}, which is coupled to the in-plane oscillations of the OH bond (in plane with COO), is totally absent in the absorption spectrum of zwitterionic alanine of Fig. 5. Mode 8, around 1486 cm^{-1}, is similar to mode 3 just described, but is coupled to a vibration in the CH_3 fragment that looks like an umbrella mode. Finally, we find that mode 12, slightly above 1700 cm^{-1}, involves solely the HNH bending motion in NH_2, and was not so well defined in the analysis reported in Fig. 6.

The overall decomposition of the IR spectrum of neu-

TABLE III. Left column: experimental frequencies collected from Ref. [25] in the spectral region 1000–2000 cm^{-1} for the neutral form of alanine. Central column: bands assignment using the abbreviation from Ref. [25] (ρ for rocking, δ for bending, ν for stretching, ω for wagging; s for symmetric and as for asymmetric). Right column: theoretical effective frequencies from this work.

Experimental frequency (cm^{-1})	Assignment	Calculated frequency (cm^{-1})
1037	ρCH_3	1048
1064	ρCH_3	
	δCH, ωNH_2	1150
1110	νCN	
1153	δCO	
1206	δCOH	1200 1219
	ρNH_2, ωCH_3, δCOH	1285
1335	δCH	
1376	$\delta_s CH_3$	1416
1408	δCH	
1454	$\delta_{as} CH_3$	1454
1460	$\delta_{as} CH_3$	1443
	ωNH_2, ωCOO	1517
	ρNH_2, δOH	1610
1622	δNH_2	1706
1774	$\nu C{=}O$	1732

tral alanine in the gas phase is summarized in Table III, where we report in the right column the average frequencies of the effective modes. As done previously in Table II, we compare our analysis to the results of Ref. [25], whose frequencies are listed in the left column. The central columns report the bands assignment, using again the notation introduced in Ref. [25]. With respect to the zwitterionic form of alanine, the agreement between our effective mode analysis and the experimental spectrum of Ref. [25] for the neutral form turns our to be even slightly better, with a MAE of 26.6 cm^{-1}.

Figures 5 and 7 (upper panels) show the comparison of the IR spectra calculated along the MD trajectories at 300 K with the normal modes peaks shown as gray vertical bars. Such a comparison emphasizes the added value of the present analysis method, in which anharmonicities are naturally included. In both cases, several anharmonic bands cannot be correctly interpreted based on the harmonic modes alone. In Table IV, we list the harmonic frequencies and their assignment. In both zwitterionic and neutral forms, we identified the highest frequency modes in the spectral range of interest from the effective modes analysis, i.e., the CO stretching for the zwitterionic form and the NH_2 bending for the neutral form, and we report in the table the 14 modes that we suppose being analogous to the anharmonic bands. Note that the lowest frequencies lie outside the $1000 - 2000$ cm^{-1} range, as the CO stretching of the zwitterionic form, and, thus, do not appear in the spectra in Figs. 5 and 7. As for the values of the harmonic frequencies, we find some disagreement in the assignments of the harmonic modes with respect to the effective modes analysis.

TABLE IV. Left column: harmonic frequencies in the spectral region of interest for the zwitterionic and neutral forms of alanine in the gas phase. Right column: peak assignment with the notation used in Tables II and III.

Harmonic frequency (cm^{-1})	Assignment
Zwitterionic alanine	
952	δCH, ρCH_3, ρNH_3
1045	δCH, ρCH_3, ρNH_3
1077	δCH, ρCH_3, ρNH_3
1158	δCH, ρCH_3, ρNH_3
1302	δCH
1415	$\delta_s CH_3$, $\delta_s NH_3$, δCH
1441	$\delta_{as} CH_3$, $\delta_{as} NH_3$
1453	$\delta_s CH_3$, $\delta_s NH_3$
1473	$\delta_s CH_3$, $\delta_s NH_3$, δCH
1478	$\delta_{as} CH_3$
1546	$\delta_s NH_3$, δCH, δCOO
1579	$\delta_s NH_3$, δCH, δCOO
1722	$\delta_{as} NH_3$
1926	$\nu_{as} COO$
Neutral alanine	
980	ωCH_3, δCH
1005	ωCH_3, δCH
1072	ωNH_2, δCH
1180	ωNH_2, δCH
1218	δOH, δCH, δNH_2
1244	δOH, δCH, δNH_2
1306	δCH
1416	$\delta_s CH_3$, δCH
1459	$\delta_{as} CH_3$
1468	$\delta_{as} CH_3$
1480	$\delta_s CH_3$, δCH
1539	δNH_2, δCH
1715	δNH_2
1748	δNH_2, νCO

B. Hydrated and crystalline zwitterionic forms of alanine

The IR absorption spectrum of hydrated alanine was determined by considering the electric dipole moments of the solute and the solvent molecules within the first solvation shell only, i.e. within a radius of 3.85 Å, in such a way as not to obscure the signal in the relevant spectroscopic region of 1000–2000 cm^{-1} by that of the water solvent. In the crystalline phase, the IR spectrum was computed considering the whole environment.

Fig. 9 depicts the resulting IR absorption spectra for alanine in both condensed phases. Both spectra were again decomposed into effective modes, shown as colored lines in the figure. While some similarities can be identified, in the full spectra as well as in their decompositions, some features appear to be different.

The main differences between the absorption spectra of hydrated and crystalline alanine are as follows:

1. in the region 1000–1200 cm^{-1}, the two main bands are well resolved in the crystal form, as well as the shoulder around 1180 cm^{-1}, by the identification of three distinct effective modes bands. In the hydrated form, the two main bands are predicted as

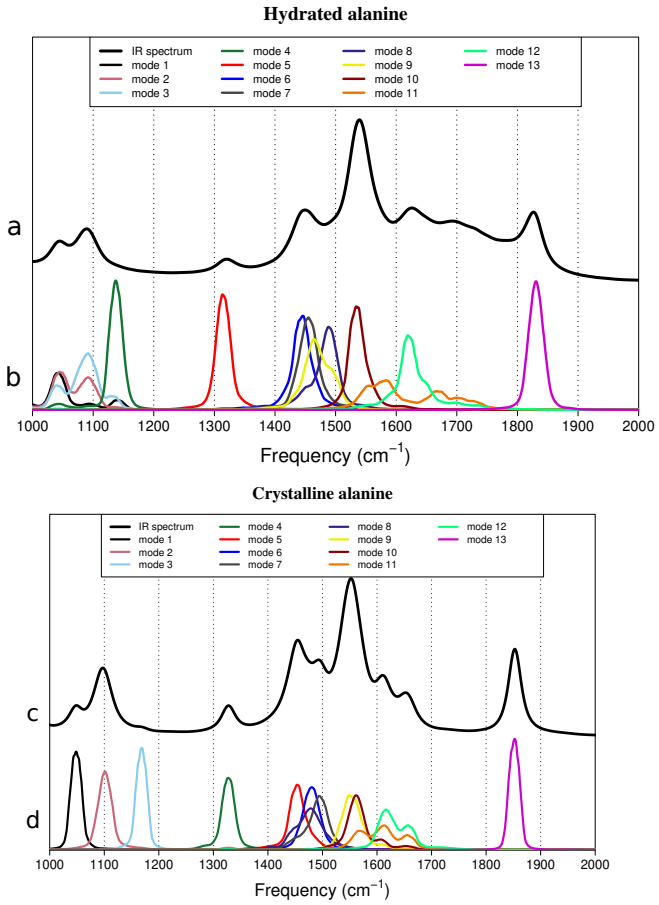

FIG. 9. Upper panel: (a) IR spectrum of zwitterionic alanine in water (black line) along with (b) its decomposition in effective modes bands (colored curves). Lower panel: (c) IR spectrum of zwitterionic alanine in crystal structure (black line) along with (d) its decomposition in effective modes bands (colored curves).

arising from the superposition of three modes. A fourth mode is responsible for the shoulder, which seems, however, suppressed in water;

2. in the region 1400–1650 cm^{-1}, the low intensity bands at the sides on the central most intense bands are both split into two contributions in the crystal form, while only one band appears to be present on each side of the intense band in the case of hydrated alanine;

3. in the region around 1700 cm^{-1} a large and poorly resolved band appears in the hydrated form of alanine, which is completely absent in the crystal form.

Additionally, we find that, compared to zwitterionic alanine in the gas phase, the band near 1300 cm^{-1} is slightly blue shifted in water and in the crystal, while, concomitantly, the band near 1900 cm^{-1} is red shifted when the molecule is not isolated.

In order to analyze the points just discussed, we start

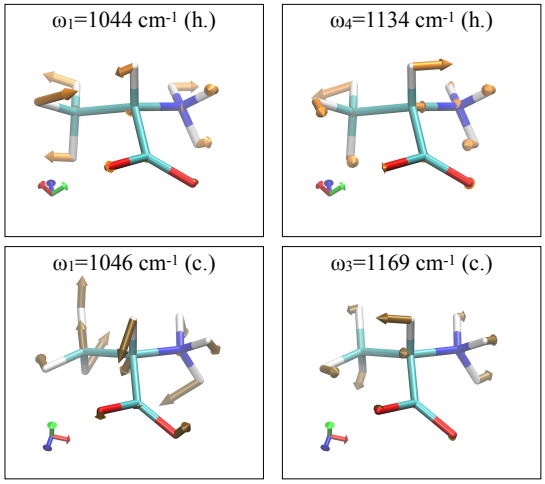

FIG. 10. Effectives modes in the region 1000–1200 cm^{-1} for the hydrated (h.) form of zwitterionic alanine (upper panels) and for the crystal (c.) form of zwitterionic alanine (lower panels).

by reporting in Fig. 10 some of the effective modes giving rise to the IR absorption bands in the region 1000–1200 cm^{-1}. Note that all modes are very similar to each other, involving some kinds of wagging and twisting of the fragments CH_3 and NH_3, mainly coupled to the vibrations of the CH bond that is not involved in CH_3. Only mode 3 in the crystal form appears slightly different, as it clearly involves the umbrella mode in NH_3.

Concerning the appearance of the double peaks in the region 1450–1650 cm^{-1} of the IR spectrum of crystalline alanine in comparison to the single peaks structure of hydrated alanine, we report the corresponding effective modes in Fig. 11. In the figure, we selected two modes, one around 1450 cm^{-1} and one slightly above 1600 cm^{-1}. The modes on the left panels and on the right panels of Fig. 11 are very similar: on the left, both modes involve the umbrella motion of NH_3 coupled to the (out-of-phase) umbrella motion in CH_3; on the right, both modes involve the HNH bending motion in NH_3, coupled to an out-of-plane vibration of the third NH bond in the fragment. Therefore, it appears reasonable to assume that despite the fact that the IR spectrum of the crystal form of alanine presents a better resolved structure than its hydrated form, the source of the absorption bands is to be found in modes that are fundamentally very similar. Presumably, the constrained vibrations in the crystal reduce the effects of anharmonicity, thus allowing the absorption bands to be better resolved in frequency.

Finally, in Fig. 12, we analyze the effective modes responsible for the appearance of the absorption bands at around 1350 cm^{-1} (left panels) and 1850 cm^{-1} (right panels). The effective modes at lower frequency appear to be slightly different, as in water no umbrella motion involving the CH_3 fragment could be discerned, but rather a combination of HCH bending and CH vibration. As previously noted, this slight difference can be ascribed to

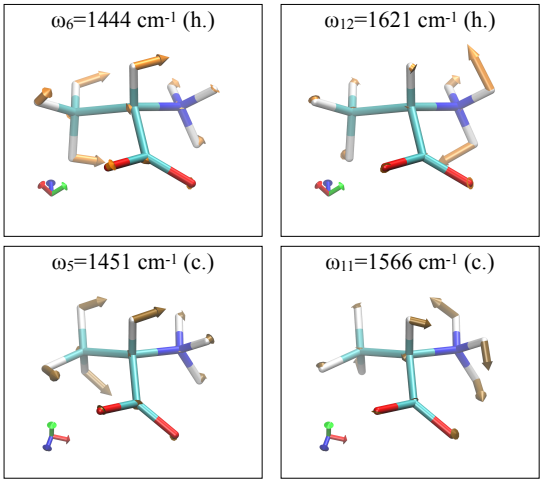

FIG. 11. Effectives modes in the region 1400–1650 cm^{-1} for the hydrated (h.) form of zwitterionic alanine (upper panels) and for the crystal (c.) form of zwitterionic alanine (lower panels).

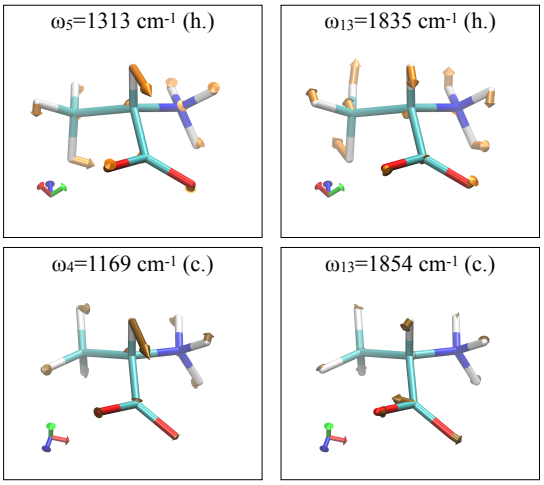

FIG. 12. Effectives modes around 1350 cm^{-1} (left) and 1850 cm^{-1} (right) for the hydrated (h.) form of zwitterionic alanine (upper panels) and for the crystal (c.) form of zwitterionic alanine (lower panels).

the different rigidities of the environment where alanine is embedded. The effective mode at higher frequency clearly points at the asymmetric stretching motion of CO_2 in the crystalline form, which becomes instead a wagging motion in the water solvent that is coupled to vibrations in the other fragments of the solute molecule. We can thus conclude that in water, in the region spectral region 1000–2000 cm^{-1}, the pure antisymmetric stretching motion of CO_2 is suppressed.

IV. CONCLUSIONS

We reported an in-depth theoretical analysis of the IR spectroscopy of alanine in various forms. We applied the effective modes analysis, aiming to decompose the IR absorption spectra of alanine in the gas phase, i.e., in zwitterionic and neutral forms, and of hydrated and crystalline zwitterionic alanine. The decomposition relies on the identification of well-defined, maximally localized in frequency, absorption bands calculated as the power spectra of the effective modes, whose sum yields with good approximation the full IR spectrum of the molecule at hand.

The effective modes analysis was briefly presented to demonstrate that it can be understood as a finite-temperature and anharmonic generalization of the standard normal modes analysis relevant in the static limit. Since alanine is quite flexible and its dynamics was simulated at room temperature, the harmonic normal modes analysis is not be sufficient to satisfactorily characterize the complex absorption spectra discussed in the present work.

The reported decomposition of the spectra of alanine demonstrates the capability of the effective modes analysis to unravel absorption spectra that can appear at first glance very broad and not well resolved. The comparison between the two gas phase forms of alanine allowed us to identify the modes responsible for the changes in the IR spectra: even though it is expected that the changes in the neutral form are produced by the fragments of the molecule that are not present in the zwitterionic form, it is interesting to find that those changes are confirmed based on a rigorously-derived numerical procedure. In addition, the comparison with the literature in the assignment of the absorption bands of the gas phase species allowed us to validate the new parametrization of the AMOEBA polarizable force field, which we used to predict the IR absorption spectroscopy of hydrated and crystal alanine. Future studies of alanine in combination with the effective modes analysis will potentially address issues such as the optimal identification of internal coordinates to initialize the minimization procedure and a more systematic comparison of the effective modes analysis to alternative techniques of analysis of IR spectra, as those mentioned in the introduction.

The field of the possible applications of this method is wide. It can complement any IR spectroscopy study, using polarizable molecular dynamics, of fluxional molecules such as peptides in hydrated or crystalline phases, or under less conventional environments such as deep eutectic solvents.

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).

[1] E. Ditler and S. Luber, WIREs Comput. Mol. Sci. **12**, e1605 (2022).

[2] A. Scherrer, F. Agostini, D. Sebastiani, E. K. U. Gross, and R. Vuilleumier, J. Chem. Phys. **143**, 074106 (2015).

[3] A. Scherrer, R. Vuilleumier, and D. Sebastiani, J. Chem. Phys. **145**, 084101 (2016).

[4] K. L. Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, and A. Zehnacker, Phys. Chem. Chem. Phys. **22**, 26047 (2020).

[5] S. Jähnigen, K. L. Barbu-Debus, R. Guillot, R. Vuilleumier, and A. Zehnacker, Angew. Chem. Int. Ed. , e202215599 (2022).

[6] S. Jähnigen, A. Scherrer, R. Vuilleumier, and D. Sebastiani, Angew. Chem. Int. Ed. **57**, 1 (2018).

[7] R. M. Levy, O. de la Luz Rojas, and R. A. Friesner, J. Phys. Chem. **88**, 4233 (1984).

[8] J. Kohanoff, Comput. Mater. Sci. **2**, 221 (1994).

[9] R. M. Stratt, Acc. Chem. Res. **28**, 201 (1995).

[10] B. R. Brooks, D. Janežič, and M. Karplus, J. Comput. Chem. **16**, 1522 (1995).

[11] J. M. Bowman, X. Zhang, and A. Brown, J. Chem. Phys. **119**, 646 (2003).

[12] R. A. Wheeler, H. Dong, and S. E. Boesch, ChemPhysChem **4**, 382 (2003).

[13] A. Strachan, J. Chem. Phys. **120**, 1 (2004).

[14] M. Schmitz and P. Tavan, J. Chem. Phys. **121**, 12233 (2004).

[15] M. C. G. Lebrero, L. L. Perissinotti, and D. A. Estrin, J. Phys. Chem. A **109**, 9598 (2005).

[16] M. Martinez, M.-P. Gaigeot, D. Borgis, and R. Vuilleumier, J. Chem. Phys. **125**, 144106 (2006).

[17] F. Thaunay, J.-P. Dognon, G. Ohanessian, and C. Clavaguéra, Phys. Chem. Chem. Phys. **17**, 25968 (2015).

[18] F. Agostini, R. Vuilleumier, and G. Ciccotti, J. Chem. Phys. **134** (2011).

[19] F. Agostini, R. Vuilleumier, and G. Ciccotti, J. Chem. Phys. **134** (2011).

[20] D. Bovi, A. Mezzetti, R. Vuilleumier, M.-P. Gaigeot, B. Chazallon, R. Spezia, and L. Guidoni, Phys. Chem. Chem. Phys. **13**, 20954 (2011).

[21] L. Tanzi, F. Ramondo, and L. Guidoni, J. Phys. Chem. A **116**, 10160 (2012).

[22] C. Zhang, R. Z. Khaliullin, D. Bovi, L. Guidoni, and T. D. Kühne, J. Phys. Chem. Lett. **4** (2013).

[23] M. Basire, F. Mouhat, G. Fraux, A. Bordage, J. Hazemann, M. Louvel, R. Spezia, S. Bonella, and R. Vuilleumier, J. Chem. Phys. **146**, 134102 (2017).

[24] M. Rahmani and M. E. A. Benmalti, J. Biomol. Struct. Dyn. **doi: 10.1080/07391102.2021.1918580**, 1 (2021).

[25] M. T. S. Rosado, M. L. R. Duarte, and R. Fausto, J. Mol. Struct. **410**, 343 (1997).

[26] M.-P. Gaigeot, Phys. Chem. Chem. Phys. **12**, 3336 (2010).

[27] D. R. Galimberti, A. Milani, M. Tommasini, C. Castiglioni, and M.-P. Gaigeot, J. Chem. Theory Comput. **13**, 3802 (2017).

[28] S. Bakels, M.-P. Gaigeot, and A. M. Rijs, Chem. Rev. **12**, 3233 (2020).

[29] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, J. Chem. Theory Comput. **9**, 4046 (2013).

[30] M.-P. Gaigeot, M. Martinez, and R. Vuilleumier, Mol. Phys. **105**, 2857 (2007).

[31] C. R. Jacob and M. Reiher, J. Chem. Phys. **130**, 084106 (2009).

[32] 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. Di Stasio Jr., M. Head-Gordon, G. N. I. Clark, M. E. Johnson, and T. Head-Gordon, J. Phys. Chem. B **114**, 2549 (2010).

[33] P. Ren, C. Wu, and J. W. Ponder, J. Chem. Theory Comput. **7**, 3143 (2011).

[34] C. Zhang, C. Lu, Z. Jing, C. Wu, J.-P. Piquemal, J. W. Ponder, and P. Ren, J. Chem. Theory Comput. **14**, 2084 (2018).

[35] T. D. Rasmussen, P. Ren, J. W. Ponder, and F. Jensen, Int. J. Quantum Chem. **107**, 1390 (2007).

[36] B. Thole, Chem. Phys. **59**, 341 (1981).

[37] D. Semrouni, A. Sharma, J.-P. Dognon, G. Ohanessian, and C. Clavaguéra, J. Chem. Theory Comput. **10**, 3190 (2014).

[38] F. Thaunay, C. Jana, C. Clavaguéra, and G. Ohanessian, J. Phys. Chem. A **122**, 832 (2018).

[39] J. W. Ponder, "TINKER - Software Tools for Molecular Design (version 8)," http://dasher.wustl.edu/tinker (2019), (accessed July 22th, 2020).

[40] S. Jähnigen, "ChirPy A python package for chirality, dynamics, and molecular vibrations (version 0.23.2)," (2022), https://github.com/sjaehnigen/chirpy.

[41] J. Horníček, P. Kaprálová, and P. Bouř, J. Chem. Phys **127**, 084502 (2007).

[42] A. Savitzky and M. J. E. Golay, Anal. Chem. **36**, 1627 (1964).

[43] To isolate the effect of these two problems separetly, one way to proceed further would be by including other parts of the molecule when analyzing NH_3 in order to check if the localization improves; in this case, the coupling to other modes would be the most reasonable cause for the lack of localization.