The Vibrational Spectrum of the hydrated Alanine-Leucine Peptide in the Amide region from IR experiments and First Principles Calculations

Irtaza Hassan,[†] Luca Donati,[‡] Till Stensitzki,[†] Bettina G. Keller,[‡] Karsten Heyne,[†] and Petra Imhof[*,†]

[†]*Institute of Theoretical Physics, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany*

[‡]*Institute of Chemistry and Biochemistry, Freie Universität Berlin, Takustr. 3, 14195 Berlin, Germany*

E-mail: petra.imhof@fu-berlin.de

Abstract

We have combined infrared (IR) experiments with molecular dynamics (MD) simulations in solution at finite temperature to analyse the vibrational signature of the small floppy peptide Alanine-Leucine. IR spectra computed from first-principles MD simulations exhibit no distinct differences between conformational clusters of α-helix or β-sheet-like folds with different orientations of the bulky leucine side chain. All computed spectra show two prominent bands, in good agreement with the experiment, that are assigned to the stretch vibrations of the carbonyl and carboxyl group, respectively. Variations in band widths and exact maxima are likely due to small fluctuations in the backbone torsion angles.

Introduction

IR spectroscopy techniques are extensively used in the liquid and gas phase at finite temperature to probe molecular vibrations. This provides information about the chemical structure of a molecule and its interaction with the environment. Vibrational spectroscopy has furthermore been successful in revealing the secondary structure of peptides and proteins and allows to probe conformational dynamics via time-resolved measurements[1].

The dynamics of complex bio-molecules involves characteristics timescales ranging from 10^{-12} to 10^{0} seconds i.e from structural vibrations to complex conformational transitions. Peptides are frequently used model systems of proteins to study their conformational dynamics because of their reduced complexity and the shorter timescales (up to milli- or micro-seconds).

Molecular simulations can assist to evaluate the dynamics at atomic level and to find the conformations which are responsible for the vibrational fingerprints observed in experimental IR spectra. Often, the assignment of the measured IR spectra is aided by quantum chemical calculations (harmonic normal mode analysis) of one or few conformations of the respective molecule in vacuum or by using implicit solvent models. The spectra calculated for the different conformations are then matched with experimental spectra to identify which conformation is responsible for the measured vibrational signatures. This approach, albeit straightforward, lacks finite temperature, anharmonicites and usually does not account for the explicit interaction with the solvent. Recently, substantial progress has been made to improve the description of solute-solvent interactions by including specific solvent molecules explicitely into an otherwise implicit-solvent model[2,3] also for the computation of vibrational spectra[4,5]. For small molecules corrections for anharmonic effects in vibrational spectra can be obtained from vibrational self-consistend field approaches augmented by correlated methods[6–13].

Inclusion of explicit solvation, thermal and anharmonic effects altogether is possible

by employing molecular-dynamics based approaches which are enjoying increasing popularity and have been used successfully to calculate IR spectra of small peptides[14–16].

The computational expenses of such first-principles simulations, however, prevent the exploration of dynamical processes on time scales beyond $10's$ to $100's$ of picoseconds, even on density functional theory (DFT) level with generalised-gradient functionals with still limited accuracy. Correlated wave function methods with large basis sets, which would give spectroscopic accuracy, are computationally far too demanding for MD simulations..

On the other hand, classical, force-field based MD simulations are a means to exhaustively sample the conformational space of a molecule. Such MD simulations combined with Markov state models (MSMs) allow describing the conformational dynamics of poly-peptides[17–19]. The fixed point-charge model, typically used in force-field based simulations, does not allow to take any change in electronic density into account and therefore cannot accurately model the changes in dipole moment associated with the vibrational motions.

Approaches such as frequency maps[20–22], instantaneous normal mode calculations[23–26], quantum-classical descriptions[27,28], perturbed matrix methods[29], or polarisable force fields[26] all attempt at combining the strengths of quantum chemical calculations (accurate electron densities) with that of classical MD simulations (comprehensive sampling of the conformational space).

In this work, we follow a combined approach by sampling the conformational dynamics of a small floppy peptide in water by classic MD simulations. From these data, we constructed a MSM, identified the meta-stable sets and the time scales associated with the slowest processes. For representative conformations of the meta-stable sets, individual first-principles molecular dynamics simulations were performed. From these simulations the vibrational signatures were then computed in the frequency range that is most sensitive to the peptide conformation, so-called amide bands (1300-1800cm^{-1})[30,31].

The experimental IR spectrum of Alanine-Leucine (Ala-Leu) in water was then assigned by using the combined information from all the calculations.

We chose the peptide Ala-Leu because of its size: it is just large enough to exhibit conformational dynamics and small enough to allow for long enough first-principles simulations to compute IR spectra from. The uncapped zwitter-ionic form has the further advantage of one charged carboxyl group and a carbonyl group, absorbing at different frequencies, rather than the similar frequencies of two carbonyl groups in a capped peptide.. Furthermore, a transition that, in a longer peptide would correspond to a α-helix/β-sheet transition via a torsion around one of the backbone angles (Ψ_{Leu}), in the short peptide Ala-Leu rotates the carboxyl group. Because of the chemical equivalence of the two carboxyl atoms, this rotation yields two chemically equivalent conformations. The bulky side-chain of the Leucine amino acid furthermore gives rise to richer conformational dynamics.

Methods

Theoretical background

Markov state models

A Markov state model of the long-time conformational dynamics is constructed from discrete partitioning of the conformational space into micro-states[1] (Markov states)[18,32]. To this end, a transition matrix $\mathbf{T}$ is set up that estimates the underlying stochastic process, here transitions between conformational micro-states of the system. The entries of the matrix are

$$T_{ij}(\tau) = P(x_{t+\tau} = j | x_t = i) \tag{1}$$

[1]Please not that a micro-state in the context of this work refers to a set of conformations and not to a point in phase space.

The elements of the transition matrix represent the conditional probabilities of finding the molecule in state j at time $t+\tau$, given it was in state i at time t. This matrix defines a Markov process in which the propagation of the system is entirely determined by knowing its present state x_t and is independent of its past. The dynamic system furthermore fulfils detailed balance, that is in equilibrium all processes are reversible with the number of transitions $i \to j$ equal to the number of transitions $j \to i$.

The eigenvalues $\lambda_i(\tau)$ and eigenvectors $\mathbf{r}_i, \mathbf{l}_i^T$ (right and left) of the transition matrix are important ingredients to understand the prominent features of the conformational dynamics

$$\begin{aligned} \mathbf{T}(\tau)\,\mathbf{r}_i &= \mathbf{r}_i \lambda_i(\tau) \\ \mathbf{l}_i^T \mathbf{T}(\tau) &= \lambda_i(\tau)\,\mathbf{l}_i^T \end{aligned} \tag{2}$$

The transition matrix is row-stochastic

$$\sum_{j=1}^{N} T_{ij}(\tau) = 1 \ \forall \mathrm{i} \tag{3}$$

and for ergodic dynamics its largest eigenvalue $\lambda_1(\tau) = 1$. The corresponding left eigenvector is the stationary distribution.

The other eigenvalues $|\lambda_{i>1}(\tau)| < 1$ define the implied time-scales which can be understood as molecular relaxation timescales

$$t_i = -\frac{\tau}{\log|\lambda_i(\tau)|} \tag{4}$$

The corresponding eigenvectors represent the conformational transitions that occur on those timescales[18,32,33].

Theory of Infrared (IR) spectra

According to Fermi's Golden Rule, an infrared spectrum can be calculated through[34] :

$$I(\omega) = 3 \sum_i \sum_f \rho_i |\ \langle f \mid \vec{\mathcal{E}} \vec{\mu} \mid i \rangle\ |^2 \delta(\omega_{fi} - \omega) \tag{5}$$

where $\tilde{\mathcal{E}}$ is the applied external field vector, $\tilde{\mu}$ is the dipole vector of the molecular system. $I(\omega)$ is the intensity as a function of the reciprocal wavenumber of the radiation, ω (in cm $^{-1}$) and ω_{fi} is the reciprocal wavenumber associated with the transition between the initial and final vibrational states of the system $\mid i\rangle$ and $\langle f \mid$, respectively. ρ_i is the density of the molecules in the initial vibrational state $\mid i\rangle$. Within Linear Response Theory[35], the above equation can be rewritten as the Fourier transform of the dipole moment autocorrelation:

$$I(\omega) = \frac{2\pi k_B T \omega^2}{3cV} \int_{-\infty}^{\infty} \mathrm{d}t \langle \vec{\mu}(t) \cdot \vec{\mu}(0) \rangle \exp(i\omega t) \tag{6}$$

$$f(\omega) = \frac{2\pi k_B T \omega^2}{3cV} \int_{-\infty}^{\infty} \mathrm{d}t \langle \vec{r}(t) \cdot \vec{r}(0) \rangle \exp(i\omega t) \tag{7}$$

where T is the temperature, k_B the Boltzmann constant, c is the speed of light in vacuum, and V is the volume. The angular brackets represent the statistical average, as sampled by MD simulations, of the autocorrelation of the dipole moment $\vec{\mu}$ of the absorbing molecule. Equation 6 yields the complete IR spectrum of the molecular system. For the assignment to vibrations of molecular groups, power spectra are computed from the Fourier transform of the autocorrelation of the velocities (equation 7) of individual groups. See also review[16] on theoretical spectroscopy of floppy peptides at room temperature.

Molecular simulations and analysis

Markov state model

Using the trajectory of partially deuterated Ala-Leu (see Figure 1) in deuterated water obtained from classical MD simulations (see supplementary material for details), a Markov state model was constructed on the conformational space spanned by the torsion angles ψ_{Ala}, χ_{Leu} , ϕ_{Leu} and ψ_{Leu} (highlighted in Figure. 1). Such torsion coordinates have proven useful to capture the essential dynamics of small peptides such as Ala-Leu[32,36–39].

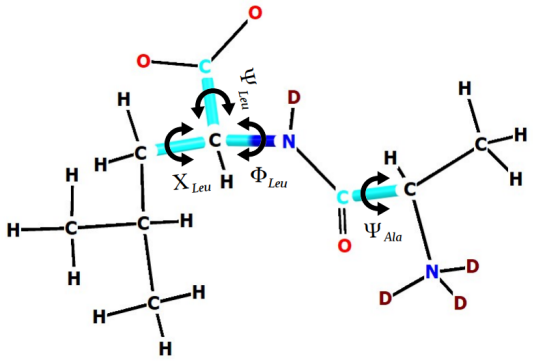


Figure 1: Scheme of the Ala-Leu peptide and the torsion angles ψ_{Ala}, χ_{Leu} , ϕ_{Leu} and ψ_{Leu} used for discretisation of the conformational space. Note that ψ_{Leu} is a pseudo-backbone dihedral angle since there is no nitrogen atom of a subsequent amino acid, but a second oxygen atom of the carboxyl terminus. "D" denotes a deuterium atom. Note that Ψ_{Leu} is a pseudo-backbone dihedral angle. That is, in the short peptide there is no nitrogen atom of the next amino acid but an oxygen atom of the C-teminus instead. A 180° torsion around Ψ_{Leu}, thus does not mean an actual conformational change, but rather an inter-conversion of two chemically equivalent structures.

The conformational space reduced to the four dimensions corresponding to torsion angles was then discretised into micro-states (corresponding to conformational clusters), based on the one-dimensional distribution of the torsion angles ψ_{Ala} and χ_{Leu} and the two-dimensional joint distribution (the ramachandran plot) of the torsion angles ϕ_{Leu} and ψ_{Leu} (see Figure 2). Two states for the torsion angles ψ_{Ala} and χ_{Leu}, respectively,

were found, while the ramachandran plot was divided into four conformational states. All the possible combinations define $2 \times 2 \times 4 = 16$ micro-states onto which the MD trajectory was projected. Based on the transition probabilities between them, micro-states were merged into three meta-stable sets and a coarse-grained transition matrix was constructed (see section Markov state models in the supplementary material).

First-principles Molecular dynamics simulations

For one representative conformation of each micro-state, we performed first-principles molecular dynamics simulations in explicit water at 300K. The same deuteration (ND_3 and ND in the peptide and D_2O water) as for the classical force field simulations (see Molecular mechanics simulations in supplementary material) was applied. For each of the four representative conformations of Ala-Leu three to four individual first-principles simulations were run for $20-80$ ps (see Figure S5 for individual runs) from which spectra were calculated. For further details please see the supplementary material.

Experimental setup

Infrared absorption spectra were taken with an Equinox 55 FTIR device (Bruker). Ala-Leu (Sigma-Aldrich, CAS 3303-34-2) was dissolved in D_2O and placed between two CaF_2 windows with a spacer thickness of 0.05 mm. Absorption of D_2O was subtracted in Figure 4 to stress the absorption signals of Ala-Leu. Note, in D_2O the exchangeable protons will exchange to deuterons. However, a residual amount of partially or undeuterated Ala-Leu remains to be present in the sample.

Results

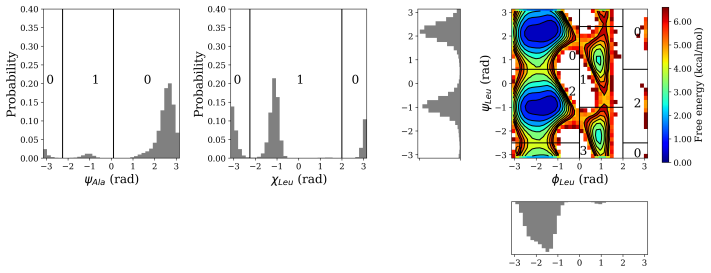


Figure 2: Distribution of the torsion angles ψ_{Ala}, χ_{Leu} , ϕ_{Leu} and ψ_{Leu} obtained from the classical MD simulation of Ala-Leu in water.

Markov state model

Figure 3 shows the coarse-grained model as a transition graph between meta-stable sets, together with representative conformations of the most probable micro-states in the set. Meta-stable set I consists of micro-states with a left-handed helix conformation and has the lowest probability. The transition into this set, corresponding to a torsion around Φ_{Leu}, is the slowest process. The two other meta-stable sets, II and III, have similar probability and are dominated by micro-states, labelled 0 and 4, and 2 and 6, respectively (see Table S1 in section Markov-state-model in the supplementary material for the complete list of micro-states in each meta-stable set and for further details).

The transition between the conformations in the two meta-stable sets II and III corresponds mainly to a torsion around Ψ_{Leu}. Transitions between micro-states within the same set, i.e. between 0 and 4, and between 2 and 6, respectively, both correspond to a torsion of the leucine side chain χ_{Leu}.

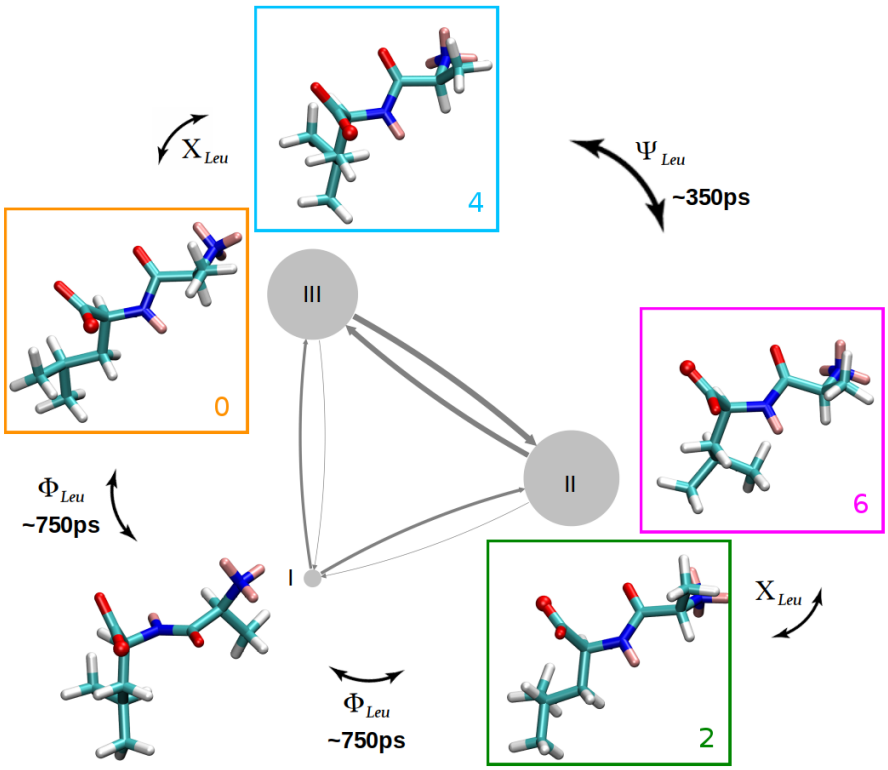


Figure 3: Coarse-grained model of the conformational dynamics of Ala-Leu in water. The three meta-stable sets, I-III, are represented as circles whose size corresponds to the probability of the respective set. The thickness of the arrows between the circles indicates the transition probability between a pair of meta-stable sets. The molecular structures in the boxes next to the circles are representative conformations of the most probable micro-states in the respective set. See Table S1 for the micro-states that constitute each set. Arrows between the coloured boxes indicate the coordinate of the conformational transition connecting two micro-states. The colour of the boxes correspond to micro-state 0 (orange), 2 (green), 4 (blue), and 6 (magenta), respectively. Carbon atoms are shown in cyan, oxygen atoms in red, nitrogen atoms in blue, hydrogen atoms in white and deuterium atoms in pink. One of the two carboxyl oxygen atoms is shown as sphere to illustrate the change in Ψ_{Leu} between meta-stable set II and III. Note that a Ψ_{Leu}-torsion of 180° inter-converts two chemically equivalent conformations.

Infrared spectrum

In the course of the individual first-principles simulations initiated from conformations representing micro-states 0, 2, 4, and 6, respectively, the system occasionally undergoes changes in the torsion angles ψ_{Ala}, χ_{Leu} , ϕ_{Leu} and ψ_{Leu} that correspond to transitions between micro-states. Hence, an individual simulation can be composed of parts that belong to different micro-states, e.g., 0 and 2. In most simulations, there are only a few jumps between micro-states. In order to analyse the vibrational fingerprint for individual micro-states, we have partitioned the first-principle trajectories by the micro-states 0, 2, 4, and 6 and computed spectra from the respective parts of the trajectories. For the time series of torsion angles and the assigned micro-state see supplementary Figure S5.

The experimental infrared spectrum of Ala-Leu in water is presented in Figure 4 together with the spectra computed from the first-principles MD simulations. The assignment (given as labels in Figure 4) is based on the computed power spectra with further aid from normal mode calculations (see supplementary material). The most prominent band is the stretch vibration of the carboxyl group (νCOO) at $\sim$1590 cm^{-1}. Note that the most intense band of the νCOO vibration has been used to normalise intensities and therefore shows a relative intensity of one for all spectra. The other band in the amide I region at $\sim$1660 cm^{-1} contains components of the carbonyl group (νC=O) and the peptide bond (νN–C). The intensity ratio of the two bands, νC=O and νCOO , is well reproduced by the computed spectra. The νC=O band is actually composed of two contributions with varying intensity ratios as can be seen from the spectra computed from the individual micro-states (Figure 4b)) and also indicated in the considerable error in the composed, weighted spectrum. (Figure 4 a) middle).

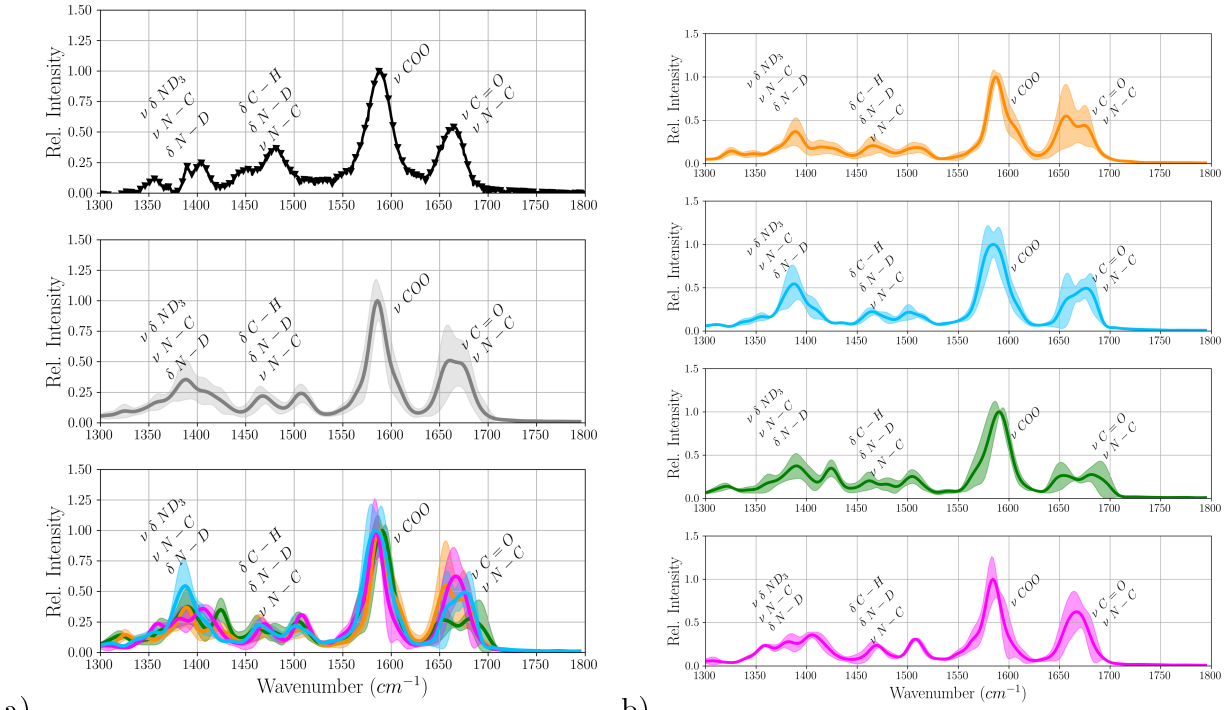
The amide II bands at $\sim$1450 cm^{-1} and $\sim$1480 cm^{-1}, assigned to bend (δN–D and δC–H) and stretch (νN–C), with some contribution from the CO group, are slightly less well reproduced; the higher frequency band is calculated at too high frequency

($\sim$1540 cm^{-1}) with too little intensity. The small shoulder at $\sim$1550 cm^{-1} in the experimental spectrum of Ala-Leu is likely due to remains of undeuterated Ala-Leu in the sample. Experiments on N-methylacetamide[40] report the δN-H bend vibration of the undeuterated species at this frequency ($\sim$1570 cm^{-1}) and the δN-D at $\sim$1450 cm^{-1} as in our spectrum of Ala-Leu).

The three smaller bands in the amid III region at $\sim$1360 cm^{-1} , $\sim$1390 cm^{-1}, and $\sim$1410 cm^{-1} in the experimental spectrum are computed as one broad band at $\sim$1380 cm^{-1} due to the averaging of several simulations, with also a significant variance in the intensities. The spectra computed individually for the micro-states (Figure 4b)) give rise to two shoulders, albeit with some error, which can be interpreted as corresponding to the lower and higher frequency bands resolved in the experimental spectrum. The main vibrational contribution stems from the terminal ND_3-group and δN-D and νN–C, but there are also COO contributions in varying amounts, depending on the individual simulation. The C=O group does not contribute to the bands in this frequency range. According to Krimm and Bandekar, both amide II and amide III bands are linear combinations of the same group movements, i.e. δN-D and νN–C[41]. In our simulations these bands show different intensity ratios for different simulation runs. As can be seen in the computed power spectra (Figure 1.6) not only do the δN-D and νN–C contributions fluctuate, there are additional contributions by the C=O and COO group to the bands in the amide II and amide III region, respectively, that vary considerably, explaining the different intensities computed for those bands.

Comparison of sampled conformations and resulting spectra

The distribution of the torsion angles within the partitions of trajectories belonging to micro-state 0, 4, 2, or 6, respectively, are shown in supplementary Figure 1.6. Within the range of the torsion angles ψ_{Ala}, χ_{Leu} , ϕ_{Leu} and ψ_{Leu}, assigned to the respective micro-states, there are different fluctuations within the individual runs, resulting in



a) b)

Figure 4: a) Experimental (top) and computed (middle) Infrared spectrum of Ala-Leu in water combined from the Boltzmann-weighted average of spectra computed for different micro-states (bottom). b) Individual computed spectra of micro-states 0 (orange), 2 (green), 4 (blue), and 6 (magenta). The shaded areas indicate the error as computed from the standard deviation from the mean.

wider, narrower, and occasionally almost bi-modal distributions.

The computed IR spectra show variations in the band assigned to the νC=O/νN–C stretch vibration, both in intensity and the exact location of the maximum. In many of the simulations, the computed νC=O/νN–C band has a shoulder, others show broadening, and in extreme cases (0-run4, 2-run2 in supplementary Figure 1.6), a second band can be observed. Comparison with the distributions of ϕ_{Leu} shows that these correspond to the shape of the νC=O/νN–C band in the IR spectrum in as much as narrow bands are only observed for narrow distributions, and broad distributions, with shoulders, correspond to a broadening or splitting in the νC=O/νN–C peak. This effect is generally more pronounced in the power spectra than in the IR spectra, indicating that the dipole moments are less affected by the variance in torsion angle than the velocities of the C=O and N–C groups.

A similar observation can be made for the distributions of ψ_{Leu} and the shapes of the νCOO bands. Narrow distributions correspond to narrow bands, shoulders in the ψ_{Leu} distribution correspond to shoulders or broadening in the νCOO band.

The two bands in the amide II region are present in almost all computed spectra, albeit with some fluctuation in their intensities. The δC-H contribution, mainly of the Leu side chain varies in exact frequency, not always matching the computed IR-band at $\sim$1540 cm^{-1}, suggesting only a minor contribution of δC–H to the IR intensity.

The amide III region exhibits considerable variation of the band intensities. Neither amide II nor amide III region show any relation to the torsion angle distributions, likely because the bands in these regions are composed of motions by several groups, N-C, ND, and ND_3 with contributions of the COO group.

Analysis of the distribution of number of hydrogen bonds from the same parts of the first-principles trajectories (Figure S8 in the supplementary material) shows that all polar groups are almost always engaged in at least one hydrogen-bond with a water molecule. In a few cases, (2-run1, 4-run3, and 6-run1) a second hydrogen bond

between C=O and water is counted with $> 30\%$ probability. These are the runs of the respective micro-state with the red-most νC=O band, indicating a slightly higher probability of further weakening of the C=O bond by an additional hydrogen-bond in these cases. Observed and computed broad or even split νC=O/νN-C bands can thus be explained by the CO group being in different hydrogen-bonded states, resulting in different amounts of red-shift. For the simulation of micro-state 0 with the highest probability of a second hydrogen-bond to C=O, the νC=O band is on the low frequency side, too. This is, however, also the case for other simulations of micro-state 0 with no particular relation to the hydrogen-bond probabilities. More detailed analysis of the hydrogen-bond interactions reveals differences in the hydrogen-bond distances and the distributions of donor-hydrogen-acceptor angles between the different simulations of the micro-states (see Figures S9 and S10). The impact of water molecules within hydrogen-bond distance on the frequency of the νC=O band, such as a red-shift, is thus further modulated by the orientation within the hydrogen-bond and thus the strength of that interaction.

Discussion

The probability distribution of peptide conformations obtained from the classical molecular dynamics simulations suggest a left-handed helix to be very improbable in the Ala-Leu peptide whereas conformations that correspond to a right-handed α-helix or β-sheet dominate. PolyProline (pPII) conformations, which have been suggested to coexist with β- conformers for the dialanine peptide[42], are observed in the classical simulations of Ala-Leu only transiently. Accordingly, such conformations have not been taken into account explicitly in the conformation-specifc first-principles simulations. In one first-principle simulation, backbone angles that correspond to pPII-conformations have been observed as transition states between two conformational states (4 and 6) and thus only for short life-times (see Figure S5 4- run1).

All transitions between meta-stable sets as well as those between the two most probable micro-states in the same meta-stable set ($0 \longleftrightarrow 4$ and $2 \longleftrightarrow 6$, respectively), equilibrate on time scales that are not accessible in the first-principles dynamics. Still, a few conformational transitions between different micro-states are observed in the course of some of the first-principles simulations. We have therefore dissected the first-principle trajectories into parts that sample only the same micro-state and used these parts for the computation of spectra.

The computed IR spectra show a considerable variance in the band intensities for simulations of the same micro-state. In contrast, there are no (further) differences between spectra computed for the different micro-states. This is to be anticipated for spectra of micro-states 0 and 2, and 4 and 6, respectively, since these correspond to chemically equivalent conformations. The conformational difference between micro-states 0 and 4, and between 2 and 6, is the orientation of the leucine side chain as defined by torsion angle χ_{Leu}. The effect on the amide region, if any, is smaller than or comparable to the variances between spectra from different runs initiated from the same conformation.

The experimental and the computed spectra are dominated by the bands assigned to the carbonyl (νC=O) and carboxyl (νCOO) stretch vibration, respectively. IR spectra of several other small peptides (N-acetyl-Gly-N-methylamide, N-acetyl-Ala-N-methylamide) with capped termini all show only a broad band assigned to the νC=O stretch vibration[43]. Computations of the spectra reveal that the νC=O of Alanine-dipeptide absorbs at almost the same frequency in either the α-helix or polyproline/β-sheet conformation but with a width that corresponds to the experimentally observed spectrum[42,44]. Slightly longer peptides, that can form intramolecular hydrogen bonds and thereby stabilise folds such as turns, Ac-Ph-Pro or trialanine, are reported with νC=O frequencies that differ by $\sim 20{-}30\,\mathrm{cm}^{-1}$ [45–47], giving rise to a shoulder or a double peak. Two-dimensional IR-experiments have furthermore revealed that the two peaks are due to coupled C=O dipoles rather than two conformations[48]. Spectra of $(\mathrm{Ala})_5$, unlabelled and isotope-labelled with ^{13}C=O and ^{13}C=^{18}O at individual C=O groups to shift their vibrational frequencies, show dual bands, separated by $\sim 20\ \mathrm{cm}^{-1}$, for both, the isotope-shifted and the unshifted groups[49]. Conformational analysis of classical MD simulations, combined with models for transition dipole coupling, reveals the coupling strength, and hence the detailed band shape, to depend on the conformation[49].

In the short peptide Ala-Leu studied in this work, there is only one C=O group. Any coupling would therefore have to be with the COO group. The νCOO band is observed at $1590\,\mathrm{cm}^{-1}$, at the same position reported for IR spectra of tripeptides ($(\mathrm{Ala})_3$,$(\mathrm{Ser})_3$, $(\mathrm{Val})_3$)[43]. In Ala-Leu, the two groups, C=O and COO, exhibit two well-separated ($\sim$50 cm^{-1}) vibrational signals of rather different intensity, suggesting no or only very weak coupling. The νC=O band maximum differs by$\sim 20-30\,\mathrm{cm}^{-1}$ between simulations, some of which indicate a dual peak also within the same simulation. The width of the frequency fluctuations for the computed νC=O band indicates a relation with the width of the distribution in torsion angle ϕ_{Leu} sampled in that particular simulation. The small differences in this torsion, and similarly of the ψ_{Leu} torsion, give

rise to fluctuations in the relative orientation of the C=O and the COO groups (and their corresponding dipoles). Likely, this also leads to fluctuations in the mutual impact of the two groups. Whether and how much the two groups are indeed coupled has to be revealed by future 2D-IR experiments.

Conclusions

The slow conformational dynamics of Ala-Leu in water are dominated by torsions around backbone angles ϕ_{Leu} and ψ_{Leu}. The slowest process can be attributed to changes in the ϕ_{Leu} torsion angles that lead to transitions to the least probable conformation, a left-handed helix. The most probable part of the conformational space can be formally assigned to the α-helix and β-sheet regions (as assigned by a discretisation of the relevant torsion angles). The inter-conversion of these two regions along ψ_{Leu} is the second slowest process. In the uncapped peptide, these two conformations are, however, actually chemically equivalent and correspondingly exhibit the same spectral signature. Another subdivision of conformations can be made by the orientation of the leucine side chain, corresponding to a torsion around χ_{Leu}.

The IR spectra computed from the first-principles MD simulations reproduce the experimental spectrum of Ala-Leu in reasonable agreement. In accordance with the chemical equivalence of the conformers with the same absolute ψ_{Leu} value, their spectra are very similar. The orientation of the leucine side chain is not reflected in the amide region of the vibrational spectrum of Ala-Leu as can be seen from comparison of the spectra computed for individual conformations.

The amide I region is very well reproduced by the simulations. The two prominent bands are assigned to the stretch vibrations of the carboxyl group, COO, and the carbonyl group, C=O, respectively. Fluctuations in the backbone torsion angle ψ_{Leu} result in a broadening of the νCOO band. The simulations furthermore reveal the

νC=O band to be composed of (at least) two frequency components. The variance in the exact frequency of this band can be attributed to mainly variations in the backbone torsion angle ϕ_{Leu} within the same area of the peptide fold. These small fluctuations occur on short time-scales and are therefore averaged out in the experimental spectrum, explaining the observation of only one broad νC=O band.

Acknowledgement

This research has been funded by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 "Scaling Cascades in Complex Systems", Project B05 "Origin of scaling cascades in protein dynamics". We thank the IT support of the Physics department at Freie Universität Berlin. Computational resources provided by the North-German Supercomputing Alliance (HLRN) are gratefully acknowledged.

References

(1) Barth, A. Infrared spectroscopy of proteins. *Bioch. Biophys. Acta (BBA)-Bioenergetics* **2007**, *1767*, 1073–1101.

(2) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003**, *24*, 669–681.

(3) Mennucci, B. Polarizable continuum model. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 386–404.

(4) Barone, V.; Biczysko, M.; Bloino, J.; Borkowska-Panek, M.; Carnimeo, I.; Panek, P. Toward anharmonic computations of vibrational spectra for large molecular systems. *International Journal of Quantum Chemistry* **2012**, *112*, 2185–2200.

(5) Barone, V.; Biczysko, M.; Bloino, J. Fully anharmonic IR and Raman spectra of medium-size molecular systems: accuracy and interpretation. *Physical Chemistry Chemical Physics* **2014**, *16*, 1759–1787.

(6) Chaban, G. M.; Jung, J. O.; Gerber, R. B. Anharmonic vibrational spectroscopy of glycine: Testing of ab initio and empirical potentials. *The Journal of Physical Chemistry A* **2000**, *104*, 10035–10044.

(7) Roy, T. K.; Gerber, R. B. Vibrational self-consistent field calculations for spectroscopy of biological molecules: new algorithmic developments and applications. *Phys. Chem. Chem. Phys.* **2013**, *15*, 9468–9492.

(8) Najbauer, E. E.; Bazso, G.; Apostolo, R.; Fausto, R.; Biczysko, M.; Barone, V.; Tarczay, G. Identification of Serine Conformers by Matrix-Isolation IR Spectroscopy Aided by Near-Infrared Laser-Induced Conformational Change, 2D Cor-

relation Analysis, and Quantum Mechanical Anharmonic Computations. *The Journal of Physical Chemistry B* **2015**, *119*, 10496–10510.

(9) Roy, T. K.; Sharma, R.; Gerber, R. B. First-principles anharmonic quantum calculations for peptide spectroscopy: VSCF calculations and comparison with experiments. *Physical Chemistry Chemical Physics* **2016**, *18*, 1607–1614.

(10) Panek, P. T.; Jacob, C. R. Anharmonic Theoretical Vibrational Spectroscopy of Polypeptides. *The journal of physical chemistry letters* **2016**, *7*, 3084–3090.

(11) Yatsyna, V.; Bakker, D. J.; Feifel, R.; Rijs, A. M.; Zhaunerchyk, V. Far-infrared amide IV-VI spectroscopy of isolated 2-and 4-Methylacetanilide. *The Journal of chemical physics* **2016**, *145*, 104309.

(12) Yu, Q.; Bowman, J. M. Communication: VSCF/VCI vibrational spectroscopy of $H7O3+$ and $H9O4+$ using high-level, many-body potential energy surface and dipole moment surfaces. *J. Chem. Phys.* **2017**, *146*, 121102.

(13) Faizan, M.; Alam, M. J.; Afroz, Z.; Bhat, S. A.; Ahmad, S. Anharmonic vibrational spectra and mode-mode couplings analysis of 2-aminopyridine. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2018**, *188*, 26 – 31.

(14) Gaigeot, M.-P.; Vuilleumier, R.; Sprik, M.; Borgis, D. Infrared spectroscopy of N-methylacetamide revisited by ab initio molecular dynamics simulations. *Journal of chemical theory and computation* **2005**, *1*, 772–789.

(15) Gaigeot, M.-P.; Martinez, M.; Vuilleumier, R. Infrared spectroscopy in the gas and liquid phase from first principle molecular dynamics simulations: application to small peptides. *Molecular Physics* **2007**, *105*, 2857–2878.

(16) Gaigeot, M.-P. Theoretical spectroscopy of floppy peptides at room temperature. A DFTMD perspective: gas and aqueous phase. *Physical Chemistry Chemical Physics* **2010**, *12*, 3336–3359.

(17) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *The Journal of chemical physics* **2007**, *126*, 155101.

(18) Prinz, J.-H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J. D.; Schütte, C.; Noé, F. Markov models of molecular kinetics: Generation and validation. *The Journal of chemical physics* **2011**, *134*, 174105.

(19) Chodera, J. D.; Noé, F. Markov state models of biomolecular conformational dynamics. *Current opinion in structural biology* **2014**, *25*, 135–144.

(20) la Cour Jansen, T.; Knoester, J. A transferable electrostatic map for solvation effects on amide I vibrations and its application to linear and two-dimensional spectroscopy. *J. Chem. Phys.* **2006**, *124*, 044502.

(21) Malolepsza, E.; Straub, J. E. Empirical Maps For The Calculation of Amide I Vibrational Spectra of Proteins From Classical Molecular Dynamics Simulations. *J. Phys. Chem. B* **2014**, *118*, 7848–7855.

(22) Cai, K.; Du, F.; Zheng, X.; Liu, J.; Zheng, R.; Zhao, J.; Wang, J. General Applicable Frequency Map for the Amide-I Mode in β-Peptides. *J Phys. Chem. B* **2016**, *120*, 1069–1079.

(23) Bastida, A.; Zúniga, J.; Requena, A.; Miguel, B.; Candela, M. E.; Soler, M. A. Conformational Changes of Trialanine in Water Induced by Vibrational Relaxation of the Amide I Mode. *The Journal of Physical Chemistry B* **2016**, *120*, 348–357.

(24) Faraq, M. H.; Bastida, A.; Ruiz-López, M. F.; Monard, G.; Ingrosso, F. Vibrational Energy Relaxation of the Amide I Mode of N-Methylacetamide in D2O Studied through Born–Oppenheimer Molecular Dynamics. *The Journal of Physical Chemistry B* **2014**, *118*, 6186–6197.

(25) Ingrosso, F.; Monard, G.; Farag, M. H.; Bastida, A.; Ruiz-López, M. F. Importance of Polarization and Charge Transfer Effects to Model the Infrared Spectra of Peptides in Solution. *J. Chem. Theory Comput.* **2011**, *7*, 1840–1849.

(26) Schwörer, M.; Wichmann, C.; Tavan, P. A polarizable QM/MM approach to the molecular dynamics of amide groups solvated in water. *The Journal of Chemical Physics* **2016**, *144*, 114504.

(27) Gorbunov, R. D.; Nguyen, P. H.; Kobus, M.; Stock, G. Quantum-classical description of the amide I vibrational spectrum of trialanine. *J. Chem. Phys.* **2007**, *126*, 054509.

(28) Kobus, M.; Nguyen, P. H.; Stock, G. Infrared signatures of the peptide dynamical transition: A molecular dynamics simulation study. *J. Chem. Phys.* **2010**, *133*, 034512.

(29) Zanetti Polzi, L.; Daidone, I.; Anselmi, M.; Carchini, G.; Di Nola, A.; Amadei, A. Analysis of Infrared Spectra of β-Hairpin Peptides As Derived from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2011**, *115*, 11872–11878.

(30) Venyaminov, S. Y.; Kalnin, N. N. Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. I. Spectral parameters of amino acid residue absorption bands. *Biopolymers* **1990**, *30*, 1243–1257.

(31) Venyaminov, S. Y.; Kalnin, N. N. Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. II. Amide absorption bands of polypeptides and fibrous proteins in α-, β-, and random coil conformations. *Biopolymers* **1990**, *30*, 1259–1271.

(32) Swope, W. C.; Pitera, J. W.; Suits, F. Describing protein folding kinetics by molecular dynamics simulations. 1. Theory. *The Journal of Physical Chemistry B* **2004**, *108*, 6571–6581.

(33) Buchete, N.-V.; Hummer, G. Coarse Master Equations for Peptide Folding Dynamics. *The Journal of Physical Chemistry B* **2008**, *112*, 6057–6069.

(34) McQuarrie, D. *Statistical Mechanics*; University Science Books, 2000.

(35) Kubo, R.; Toda, M.; Hashitsume, N. *Statistical physics II: nonequilibrium statistical mechanics*; Springer Science & Business Media, 2012; Vol. 31.

(36) Noé, F.; Horenko, I.; Schütte, C.; Smith, J. C. Hierarchical analysis of conformational dynamics in biomolecules: Transition networks of metastable states. *J. Chem. Phys.* **2007**, *126*, 155102.

(37) Swope, W. C.; Pitera, J. W.; Suits, F.; Pitman, M.; Eleftheriou, M.; Fitch, B. G.; Germain, R. S.; Rayshubski, A.; Ward, T. J. C.; Zhestkov, Y.; Zhou, R. Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 2. Example Applications to Alanine Dipeptide and a β-Hairpin Peptide. *J. Phys. Chem. B* **2004**, *108*, 6582–6594.

(38) Strodel, B.; Wales, D. J. Free energy surfaces from an extended harmonic superposition approach and kinetics for alanine dipeptide. *Chem. Phys. Lett.* **2008**, *466*, 105 – 115.

(39) Altis, A.; Otten, M.; Nguyen, P. H.; Hegger, R.; Stock, G. Construction of the free energy landscape of biomolecules via dihedral angle principal component analysis. *J. Chem. Phys.* **2008**, *128*, 245102.

(40) Myazawa, T.; Shimanouchi, T.; Mizushima, S.-I. Normal Vibrations of N-Methylacetamide. *J. Chem. Phys.* **1958**, *29*, 611–616.

(41) Krimm, S.; Bandekar, J. Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. *Adv. Prot. Chem.* **1986**, *38*, 181–364.

(42) Gaigeot, M.-P. Infrared spectroscopy of the alanine dipeptide analog in liquid water with DFT-MD. Direct evidence for PII/β conformations. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10198–10209.

(43) Eker, F.; Cao, X.; Nafie, L.; Schweitzer-Stenner, R. Tripeptides Adopt Stable Structures in Water. A Combined Polarized Visible Raman, FTIR, and VCD Spectroscopy Study. *J. Am. Chem. Soc.* **2002**, *124*, 14330–14341.

(44) Vitale, V.; Dziedzic, J.; Dubois, S. M.-M.; Fangohr, H.; Skylaris, C.-K. Anharmonic Infrared Spectroscopy through the Fourier Transform of Time Correlation Function Formalism in ONETEP. *Journal of Chemical Theory and Computation* **2015**, *11*, 3321–3332.

(45) Mahe, J.; Jaeqx, S.; Rijs, A. M.; Gaigeot, M.-P. Can far-IR action spectroscopy combined with BOMD simulations be conformation selective? *Phys. Chem. Chem. Phys.* **2015**, *17*, 25905–25914.

(46) Mukherjee, S.; Chowdhury, P.; Gai, F. Infrared Study of the Effect of Hydration on the Amide I Band and Aggregation Properties of Helical Peptides. *The Journal of Physical Chemistry B* **2007**, *111*, 4596–4602.

(47) Zanni, M. T.; Asplund, M. C.; Hochstrasser, R. M. Two-dimensional heterodyned and stimulated infrared photon echoes of N-methylacetamide-D. *J. Chem. Phys.* **2001**, *114*, 4579–4590.

(48) Woutersen, S.; Hamm, P. Structure Determination of Trialanine in Water Using Polarization Sensitive Two-Dimensional Vibrational Spectroscopy. *J Phys. Chem. B* **2000**, *104*, 11316–11320.

(49) Feng, Y.; Huang, J.; Kim, S.; Shim, J. H.; MacKerell, A. D.; Ge, N.-H. Structure of Penta-Alanine Investigated by Two-Dimensional Infrared Spectroscopy and

Molecular Dynamics Simulation. *The Journal of Physical Chemistry B* **2016**, *120*, 5325–5339.