Supplementary material for: 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

Supplementary Information

Molecular mechanics simulations

We performed classical molecular dynamics (MD) simulations of the Ala-Leu peptide in a cubic simulation box of explicit water (844 molecules modelled as TIP3P water[1]) employing the AMBER ff99SB-ILDN[2,3] force field. We used a minimum distance of 1 nm between the solute and the periodic boundaries of the box, resulting in a total number of atoms of 2564. Water hydrogen atoms and polar hydrogen atoms of the peptide (ND3 and ND) were modelled with the mass of deuterium, mimicking the experimentally measured system. For Lennard-Jones interactions and electrostatic interactions (Particle-Mesh Ewald[4,5] with grid spacing of 0.16 and an interpolation order of 4) we used cut-off value of 1 nm.

The system was minimised and equilibrated for 500 ps. Then three MD simulations of 400 ns each were launched which yields a total simulation time of 1.2 μs. A V-rescale thermostat[6] was applied to control the temperature at 300 K (NVT ensemble). The positions of the solute atoms were saved to file every 0.25 ps. No constraints were applied and the leap-frog integrator with the time step of 1 fs was employed using the GROMACS simulation package[7].

Markov state model

A transition matrix has been computed for transitions between the micro-states in the classical molecular dynamics trajectory with varying the lag time τ up to 500ps. The spectrum of the matrix, calculated for these lag times, indicates a convergence of the implied time scales (see equation 3) of the slowest process at about 50 ps.

Using thus a lag time of $\tau = 50$ ps a transition matrix was set up by counting the transitions between the micro-states.

Employing then a robust Perron Cluster Analysis (PCCA+)[8] a coarse-grained tran-

sition matrix was constructed by merging the micro-states into meta-stable sets.

The implied time time-scales in Figure S1 suggest three slow processes, corresponding to transitions between four meta-stable sets. The spectral gap of the transition matrix is, however, after the third eigenvalue (see Table in Figure S1 b)). Therefore, we tried to perform PCCA+ to identify three and four meta-stable sets, respectively. The grouping of micro-states into the meta-stable sets is listed in Table S1. The fourth partitioning results in a regrouping of formerly separated micro-states, denoting that there are several transitions on this time-scale, as already suggested by the eigenvalues. We therefore worked with a MSM of three meta-stable sets as presented in Figure 3. The first three eigenvectors are given in (see Figure S1 b)). The first eigenvector corresponds to the stationary distribution and the other two correspond to the slowest processes which can be understood as transitions between meta-stable sets of clusters.

For the projection of the MD trajectory onto the micro-states defined by the torsion angles and the PCCA+ analysis we used PyEMMA[9].

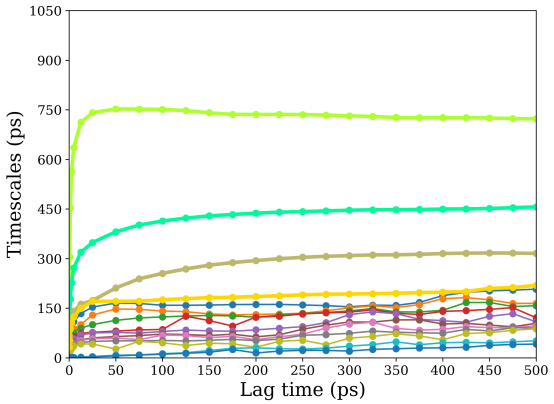
Table S1: Meta-stable sets of micro-states according to PCCA+

I	II		
[3, 5, 7, 9, 11, 13]	[0, 1, 2, 4, 6, 8, 10, 12, 14, 15]		
I	II	III	
[3, 5, 7, 9, 11, 13]	[0, 1, 4, 8, 12]	[2, 6, 10, 14, 15]	
I	II	III	IV
[5, 7, 13]	[0, 1, 8]	[4, 6, 12, 14]	[2, 3, 9, 10, 11, 15]

Table S2: Torsion angles of the representative conformations of the most probable micro-states from the classical MD simulations, shown in Figure 3 and subjected to further first-principles simulations (except for cluster 5).

Micro-state	ψ_{Ala}(rad)	χ_{Leu} (rad)	ϕ_{Leu} (rad)	ψ_{Leu}(rad)
0	2.48634	-3.00789	-1.26109	2.07314
4	2.39830	-1.24703	-1.54770	2.06933
2	2.31969	-3.09427	-1.35842	-1.17242
6	2.59943	-1.25551	-1.53055	-1.17207
5	2.21018	-1.12151	1.03061	1.40457

a)



b)

Eigenvalues
1.000
0.936
0.877
0.789
0.746
0.739

c)

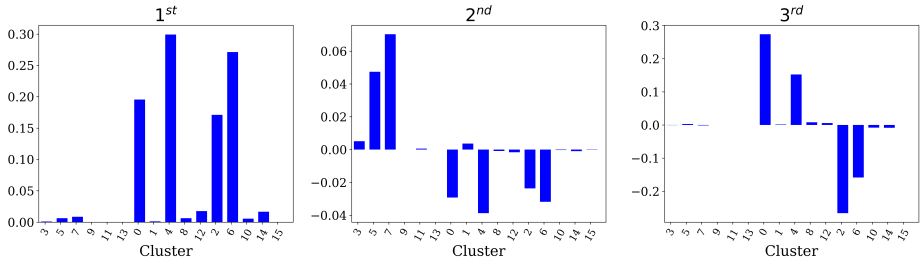


Figure S1: a) Implied time scales, b) First eigenvalues of the transitions matrix sampled with lag time τ=50ps. c) Corresponding first three eigenvectors expressed as contributions of the 16 micro-states.

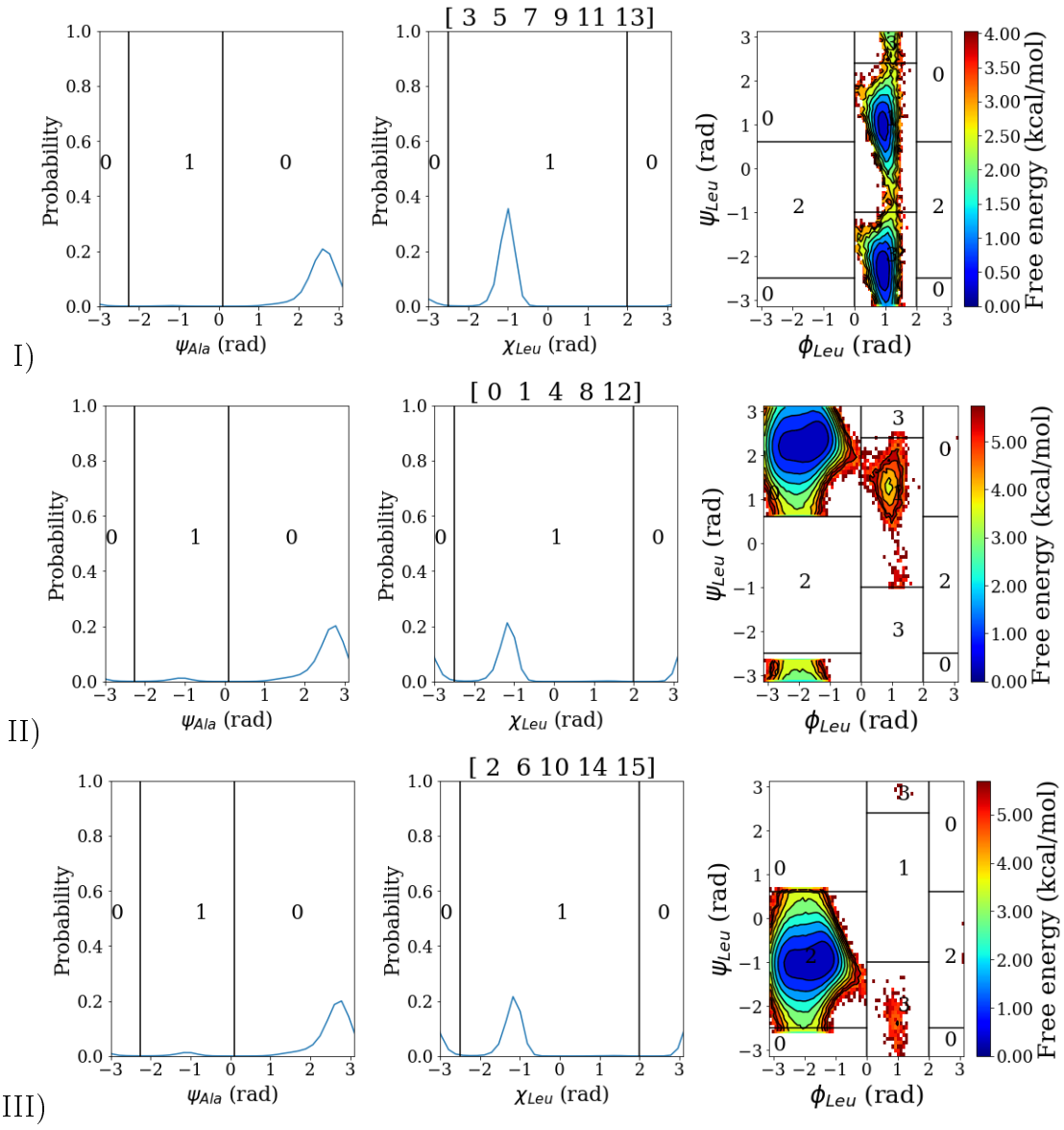


Figure S2: Distribution of torsion angles ψ_{Ala}, χ_{Leu} , ϕ_{Leu}, and ψ_{Leu} in the conformational clusters of the meta-stable sets I,II, and III, as obtained from the classical MD simulation of Ala-Leu in water.

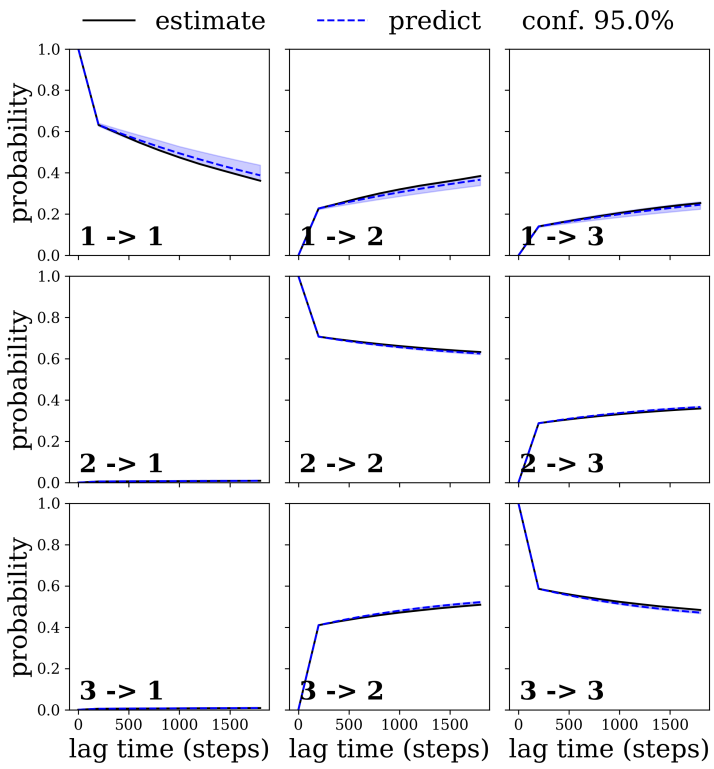


Figure S3: Chapman-Kolmogorov test of the Markov state model (MSM) with three states 1,2,3, corresponding to meta-stable sets I,II,III, respectively. Using the MSM estimated at lag time τ, transition probabilities are predicted for $n \cdot \tau$, and compared to an estimate of the model at lag time $n\tau$.

Normal mode calculations

For each cluster of conformations with a significantly high probability, we performed a geometry optimisation (convergence criterion 3.00E-04 $E_H/$Å) and subsequent normal mode calculation in implicit water (modelled by a polarisable continuum model, PCM, with a dielectric of ε=80) at the DFT level of theory. The BLYP exchange-correlation density functional and a $6-31G(d)$ basis set was employed for all these static calculations using the Gaussian programme package[10]. The N-terminus and the amino group of the peptide was set deuterated (ND3 and ND). All normal modes were scaled by 0.992 in frequency. This is the same scaling factor as used for the spectra computed from first-principles simulations but without an extra shift. The spectra obtained from normal mode calculations lack thermal and anharmonic effects. It is therefore reasonable that the exact frequencies in the vibrational spectra computed with the different approaches (normal modes and first-principles MD simulations) do not match exactly. Still, the spectra computed from normal modes facilitate the assignment of the spectra based on the first-principles simulations.

The IR-spectra computed from the normal-modes in implicit water are shown in Figure S4, top (dashed lines), together with the optimised geometries. With regard to the backbone conformation, all cluster snapshots converge to the same state which allows the charged COO and the polar ND group, and the CO and the terminal ND_3 group to come close to each other and interact optimally. The optimised structures of conformations 0 and 2 and 4,6 differ, however, in their leucine side chain conformations.

Normal modes computed for "microsolvated" Ala-Leu, that is with a total of seven water molecules added, one of each at each hydrogen bond donor or acceptor in the polar groups, ND_3 , CO, ND, and COO, respectively are shown as dotted grey lines in Figure S4, top.

The effect of adding one water molecule at a time at each polar group, ND_3, CO, ND, and COO, respectively, is shown in the bottom traces of Figure S4. Addition of a water

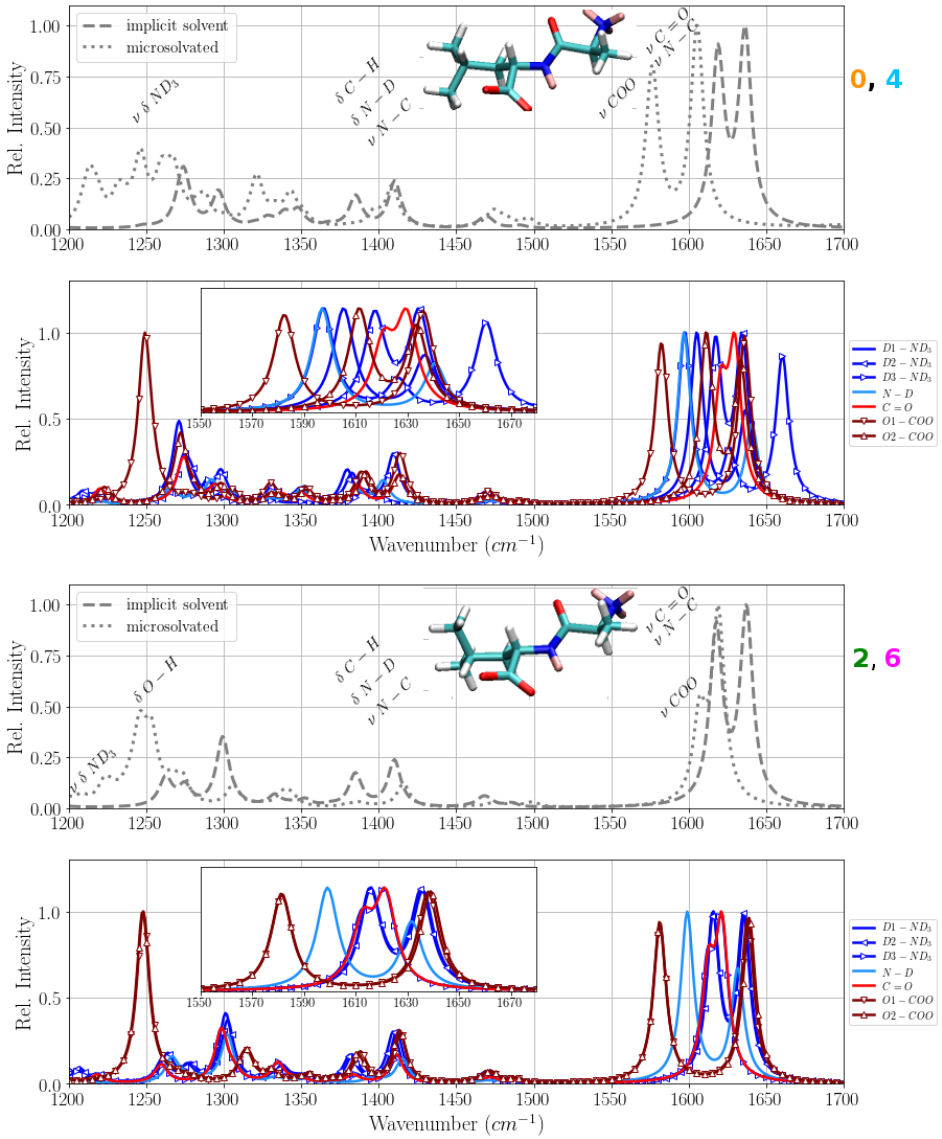


Figure S4: IR spectra computed in harmonic approximation for optimised geometries of the four conformers, corresponding to the most probable micro-states 0, 2, 4, and 6, respectively. Since conformations 0 and 2 and 4 and 6, respectively, are chemically equivalent only one of them is shown, and labelled as 0, 2 and 4, 6,respectively. Top: normal modes of Ala-Leu in implicit solvent (dashed) and additionally hydrogen-bonded to seven water molecules bound to the polar groups, ND3, CO, ND, and COO (dotted). Bottom: Normal modes of Ala-Leu in implicit solvent with one water molecule at a time hydrogen-bond to the polar groups, ND3 (dark blue), CO (red), ND (light blue), and COO (maroon), respectively.

molecule to the C=O group of conformers 0,4, leads to a red-shift of the corresponding stretch vibration. Addition of a water molecule to the C=O group in conformations 2,6 results in a slightly more pronounced red-shift of the CO band. Addition of a water molecule to the carboxyl group has also a similar effect on all conformers, i.e. the appearance of a rather strong COO band at ~ 1370 cm^{-1}. For the frequencies of the N-D stretch vibrations in all conformations upon addition of a water molecule to the respective group, another large red-shift, bringing this band to the amide I region, is observed .

First-principles Molecular dynamics simulations

For one representative conformation of each micro-state, we performed first-principles molecular dynamics simulations in explicit water. 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. The cubic simulation boxes had a minimum distance of 0.4 nm minimum between the solute and boundaries of the box and periodic boundary conditions were applied in all three dimensions. To keep the computational cost of the first principles simulations moderate, this box size, and therefore the number of water molecules, is smaller than in the classical simulations (see section "Molecular mechanics simulations") but is still just large enough to avoid interactions between periodic images.

Each Ala-Leu representative conformation was further energy minimised using a conjugate gradient algorithm[11] by keeping the atom positions of the solute fixed and let the water molecules relax around the solute.

The first principle simulations were performed using the CP2K package [12]. We employed the default Gaussian and plane waves (GPW) electronic structure method[13] as implemented in the Quickstep module[14] . We used double zeta valence plus (DZVP) basis set and interaction between valence and core electrons is described by Geodecker-

Teter-Hutter (GTH) norm conserving pseudopotentials. A plane wave expansion for the charge density is employed using energy cutoff of 500Ry. BLYP functional is used as exchange correlation functional.

Simulations were performed in an NVT ensemble. The temperature was controlled to be at 300 K using a chain of three Nose-Hoover thermostats[15] with a time constant for the thermostat chain of 100fs. CP2K default values of other thermostat parameters are used. The time-step for the numerical integration was 0.5 fs.

Atom positions were saved every step and every 5^{th} step Wannier localisation was performed so as to monitor the changes in dipole moment[16–21]. 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).

Table S4: Details of the first principles simulations

Microstate	Length of cubic simulation box (nm)	Number of water molecules	Total number of atoms
0	1.794	160	512
2	1.809	162	518
4	1.657	123	401
6	1.727	151	485

Spectra were calculated using the TRAVIS[22] programme.All calculated spectra are scaled by 0.992 in frequency. Such a scaling has been found to correct for the overestimation of high vibrational frequencies such as those in the amide region by BLYP functional[23] and is commonly applied although different approaches to obtain optimal scaling factors exist[24–26]. Since such scaling factors are dervied for harmonic mode calculations, i.e. to correct as well for anharmonic effects, different scaling procedures may be necessary for vibrational frequencies calculated from MD simulations. We have introduced an additional shift of 85 cm^{-1} so as to match best the most intense band in the experimental spectrum and thus simplify assignment.

Conformation and hydrogen-bonds analyses were carried out using MDtraj[27] and our own python and java scripts.

Conformational analyses

Time series of torsion angles and micro-states in the first-principles MD simulations

Table S4: Parts of first-principles trajectories of conformational clusters 0, 2, 4, and 6, respectively, used for the computation of spectra. The labels of the runs refer to the conformation in which the simulation was initiated.

0	2	4	6
0-run2-(0-20ps)	2-run1-(3.5-19ps)	0-run1-(34-56ps)	2-run3-(2-36ps)
0-run4-(0-20ps)	2-run1-(21-30ps)	4-run1-(30-50ps)	4-run1-(60-82ps)
2-run1-(10-50ps)	2-run2-(4-30ps)	6-run4-(0-24ps)	4-run3-(26-52ps)
2-run3-(0-20ps)			4-run4-(26-53ps)
2-run4-(0-20ps)			
6-run1-(10-50ps)			

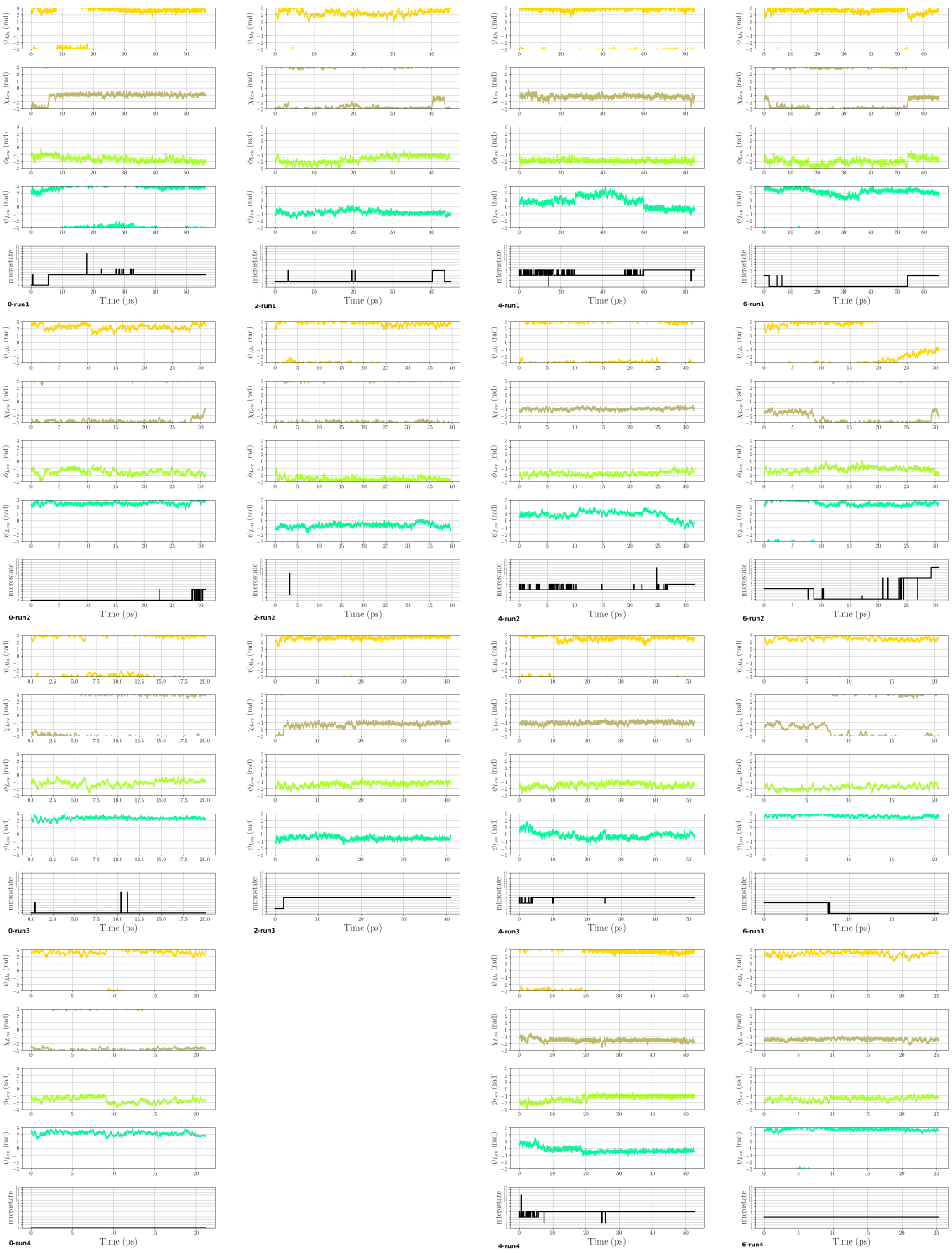


Figure S5: Time series of torsion angles ψ_{Ala}, χ_{Leu} , ϕ_{Leu} and ψ_{Leu} , and conformational states defined by them along the first-principles simulations. Note that a jump between conformational states is defined by the micro-states they belong to based on the initial discretisation. Such a jump may in fact be only a small change in one torsion angle.

Hydrogen bond analysis

The radial distribution functions of water molecules around the polar groups (Figures S6 and S7) show a significant probability for a first solvation shell and thus a water molecule to be within hydrogen bond distance for the terminal groups, ND3, and COO-, in all micro-states, albeit with significant variance between individual simulations of micro-states 0 and 6. The central peptide groups, CO and ND, in contrast, exhibit a lower density of water molecules within hydrogen-bond distance for micro-states 0, 2 and 4, indicating a less ordered water structure around these groups and a decreased likelihood of hydroge-bond formation between these groups and water molecules. Micro-state 2 shows a more proounced peak of water density around the CO group, yet smaller than those observed for the terminal carboxyl group. The reduced probability to find water in a first solvation shell around the central groups CO and ND, may be explained by a competition for binding water molecules with the charged termini which are in close vicinity in this short peptide.

This observation also explains the relatively low number of hydrogen bonds (1 to 1.5) between water and the CO group, compared to the two acceptor possibilities by the two lone-pairs at the oxygen atom. Other first-principles MD studies on small capped peptides[28,29], report an average number of hydrogen bonds of 2 to 2.5 per CO group. In those systems, however, there is no nearby charged terminal group, to which the surrounding water molecules are attracted instead. Indeed, a combined spectroscopic and computational study on uncapped di-alanine, find 1.5 to 2 hydrogen bonds to the amide carbonyl group, depending on the peptide conformation[30].

For each of the polar groups in the Ala-Leu peptide, i.e. the terminal ND3 group, CO, ND, and the two oxygen atoms of the terminal carboxyl group separately, we have analysed the hydrogen bonds with water. A hydrogen bond is defined based on geometrical criteria: the donor-acceptor distance is below 3.5Å and the donor-hydrogen$\cdots$acceptor angle is larger than 135°.

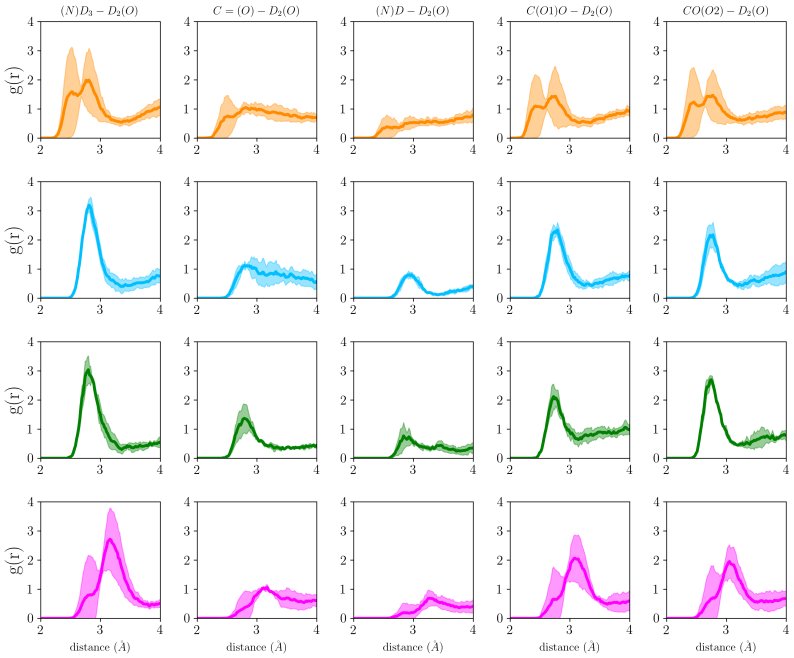


Figure S6: Radial distribution function, $g(r)$, of water oxygen atoms around the polar groups, analysed separately for ND3, CO, ND, and the two carboxyl oxygen atoms (COO1 and COO2, respectively) . The four micro-states 0, 2, 4, and 6 are shown in orange, green, blue, and magenta, respectively, averaged over the inidivdual simulation runs.

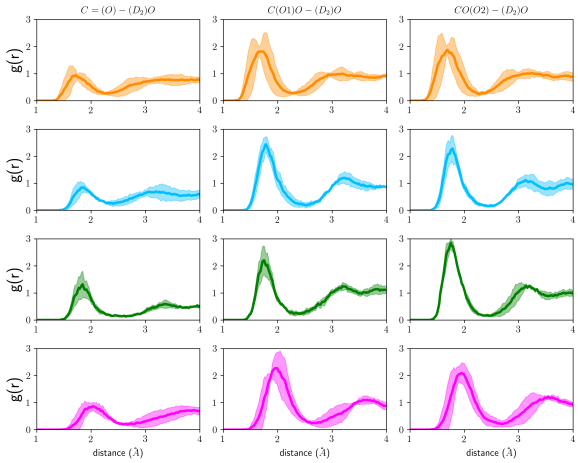


Figure S7: Radial distribution function, $g(r)$, of (a) water deuterium atoms around the oxygen atoms of the C=O group and the two carboxyl group (oxygen atoms COO1 and COO2, respectively). The four micro-states 0, 2, 4, and 6 are shown in orange, green, blue, and magenta, respectively, averaged over the inidivdual simulation runs.

In order to test whether the hydrogen bond probabilities are affected by the number of water molecules in the simulation setup, we have computed the average number of hydrogen bonds between the polar groups and water molecules from the three classical MD simulations (v.s.) with 844 water molecules and from another set of classical simulations (3 runs for each of the 4 microstates) with the system setup used in the first-principles simulations, i.e. 123–160 water molecules. As can be seen from Figure S8, the average number of hydrogen bonds is not affected by the number of water molecules and the box size.

Figure S9 lists the distribution of number of water molecules that are hydrogen-bounded to the polar groups as obtained from the individual first-principles simulations of the different micro-states.

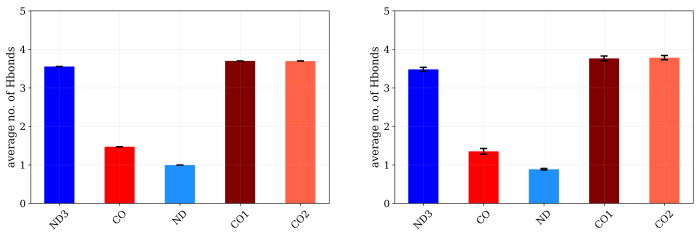


Figure S8: Number of hydrogen bonds between the polar groups of Ala-Leu and water from classical dynamics simulations (a) in a box with 844 water molecules and (b) in a box with 123–162 water molecules.

Figure S10 shows the distribution of distances between the donor/acceptor atom of the polar group and the oxygen atom of the water molecules that are hydrogen-bonded to that group.

Figure S11 shows the angular distribution of donor-hydrogen···acceptor angles for all water molecules that are within hydrogen-bond distance. Note that due to the angle criterion these water molecules are not necessarily forming a hydrogen bond to the polar groups of Ala-Leu.

The distribution of donor-acceptor distances within a hydrogen bond, ($\leq$3.5 Å) shows small differences between the individual simulations of some micro-states for

hydrogen bonds with the CO and the ND group. Whereas in most simulations the majority of hydrogen bonds is towards shorter distances to the CO group, simulations 0-run2, 0-run3 2-run3, 4-run2, 6-run1 6-run3 also show a significant probability for longer hydrogen-bond distances with CO. Hydrogen bonds with ND show distributions shifted towards larger distances mainly for simulations 0-run1, 2-run3, and 6-run4.

Analysis of the donor-hydrogen-acceptor angle distribution for water molecules that are within a hydrogen-bond distance, reveals the majority of donor-acceptor pairs to fulfill both hydrogen bond criteria, i.e. distance and angle. Still, a significant number of donor-acceptor pairs within hydrogen-bond distances exhibits donor-hydrogen-acceptor angles that are far from linear (up to 90 degree). This effect is most pronounced for hydrogen bonds with CO (in particular in simulations 4-run2, 2-run2, and all runs of micro-state 0). For hydrogen bonds with the ND group, simulations 0-run1 0-run5, 2-run1, 2-run2, 6-run3, 6-run4, exhibit smaller donor-hydrogen-acceptor angles, but no such effect is observed for any simulation of micro-state 4. These donor-acceptor pairs are not counted as hydrogen bonds, but likely still have a significant interaction, somewhat contributing to a weakening of the CO or ND bond, respectively.

The joint distributions of donor-hydrogen-acceptor angles and hydrogen–peptide oxygen distances (Figure S12) further support this idea. All micro-states exhibit a significant, rather localised, population of angles and distances that correspond to hydrogen bonds between water and the peptide oxygen atoms. In addition, a broader distribution at angles lower than the threshold we used for defining hydrogen bonds and below 4 Å with probabilities that vary between individual simulations can be observed in all micro-states. These interactions likely also have an effect on the vibration frequencies of νC=O and νCOO, broadening the corresponding bands in the infrared spectrum.

The hydrogen-bonds with longer distances are supposedly weaker than those with shorter distances and thus contribute less to a shift in the CO and ND frequencies,

respectively. But no such effect can be observed for micro-states 0, whereas for micro-state 2, number of hydrogen bonds and distribution of their distances agrees with the observation of a blue shifted νC=O band. For micro-state 4 in run2 the νC=O band is blue-shifted, seemingly contrasting the generally longer hydrogen-bonds. However, for the simulation 4-run2 the donor-hydrogen-acceptor angles significantly deviate from linearity, explaining the comparably low number and thus weak impact of hydrogen-bonded water molecules on the CO group. As for micro-state 6, only simulation 6-run4 shows a relation between the donor-acceptor orientation and CO frequency. In this simulation hydrogen-bond angles are far from linear, and thus hydrogen-bond interactions are weaker than in the other simulations of micro-state 6. Consequently, 6-run4, exhibits the νC=O band with the highest frequency among the simulations of micro-state 6. The precise hydrogen-bond distance, in contrast, appears not to have a large effect on the νC=O vibration frequencies.

Figure S9: Probability distribution of number of water molecules hydrogen-bounded to the peptide, analysed separately for ND3, CO, ND, and the two carboxyl oxygen atoms (COO1 and COO2, respectively). The four micro-states 0, 2, 4, and 6 are shown in orange, green, blue, and magenta, respectively. Each row of a sub-figure represents an individual simulation of that micro-state.

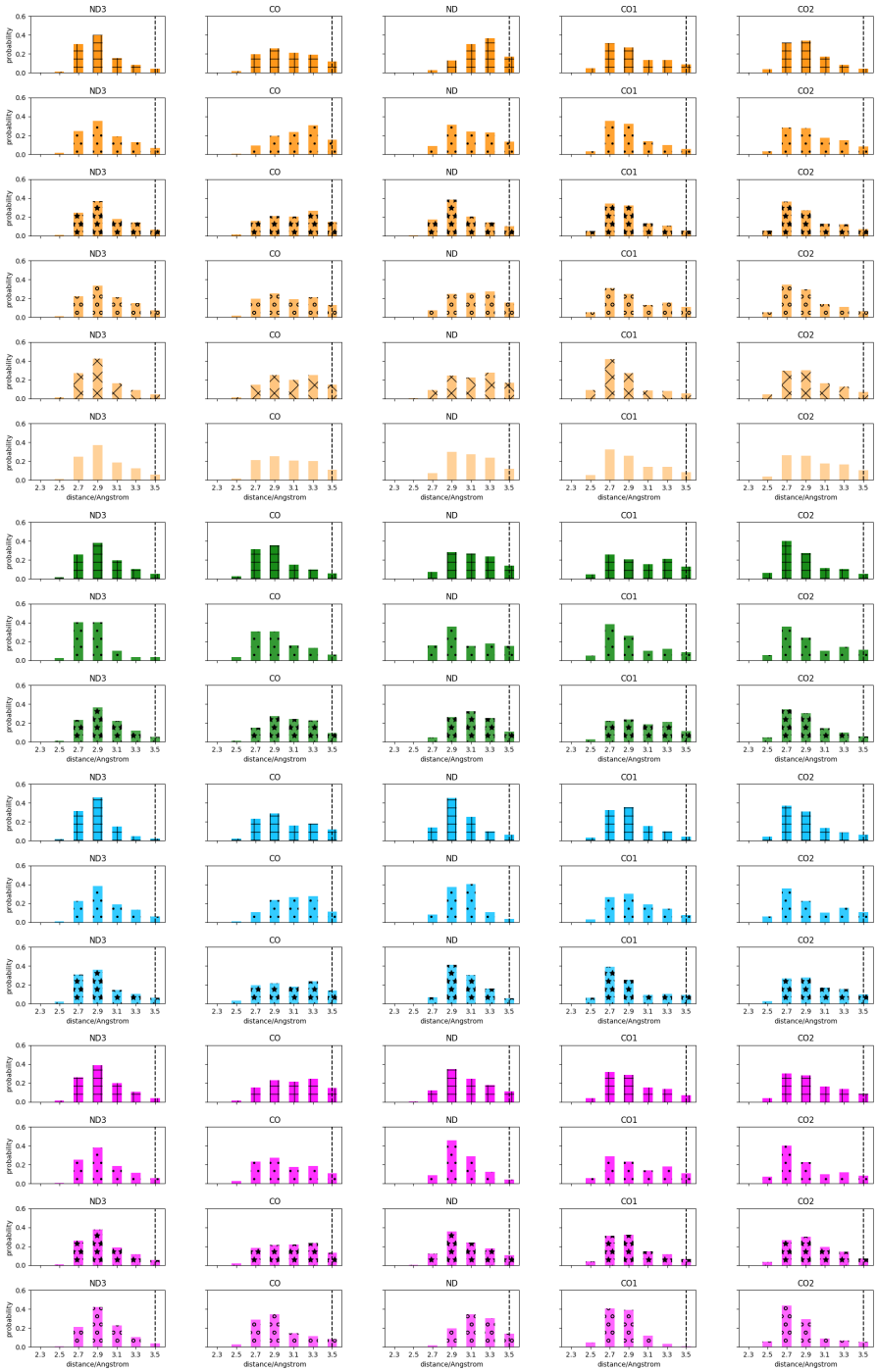


Figure S10: Probability distribution of hydrogen-bond distances between the peptide and water molecules, analysed separately for ND3, CO, ND, and the two carboxyl oxygen atoms (COO1 and COO2, respectively). The dashed line shows the distance threshold used as hydrogen bond criterion. The four micro-states 0, 2, 4, and 6 are shown in orange, green, blue, and magenta, respectively. Each row of a sub-figure represents an individual simulation of that micro-state.

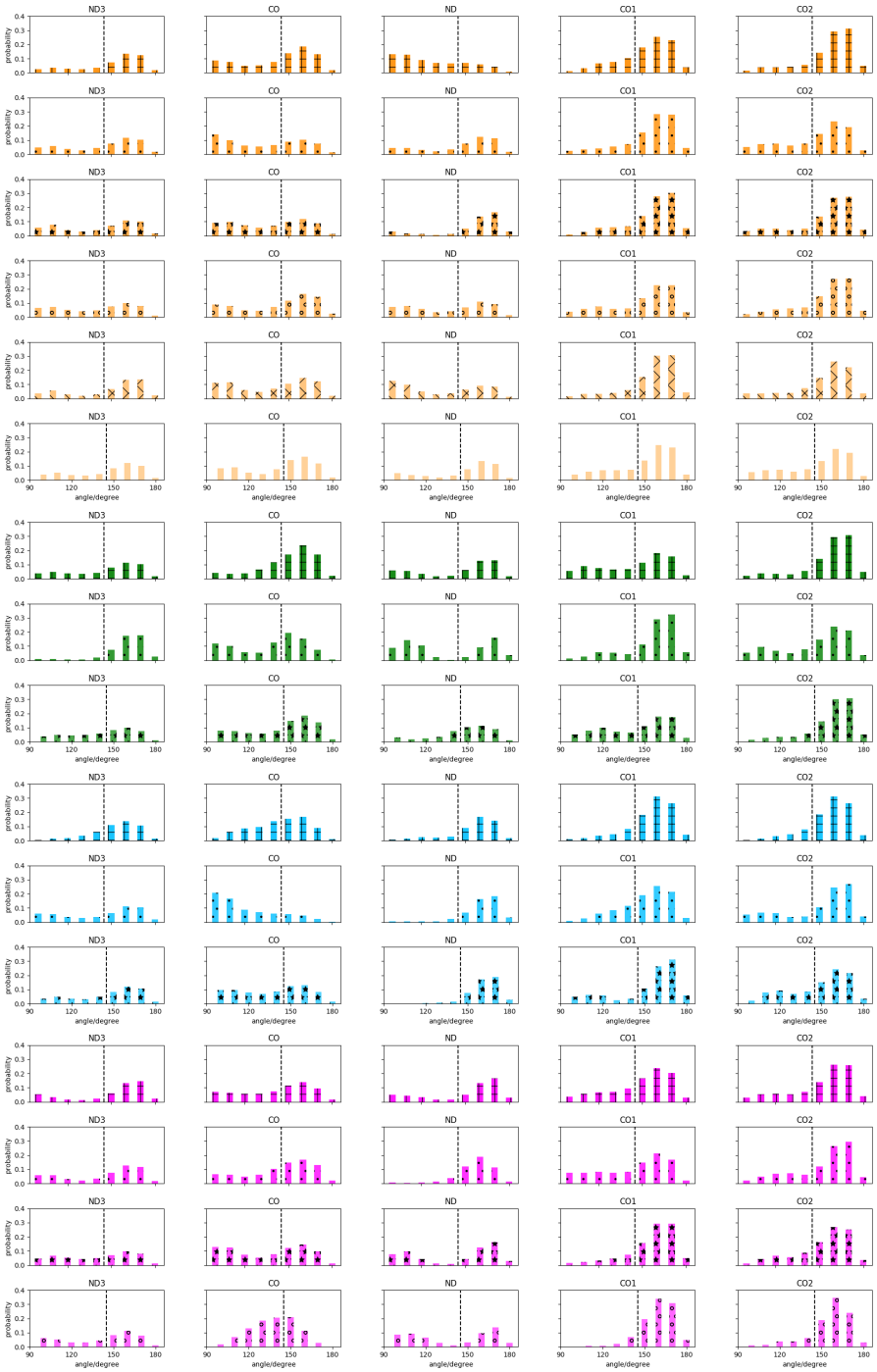


Figure S11: Probability distribution of donor-hydrogen-acceptor angles between the peptide and water molecules within hydrogen-bond distance, analysed separately for ND3, CO, ND, and the two carboxyl oxygen atoms (COO1 and COO2, respectively). The dashed line shows the angle threshold used as hydrogen bond criterion. The four micro-states 0, 2, 4, and 6 are shown in orange, green, blue, and magenta, respectively. Each row of a sub-figure represents an individual simulation of that micro-state.

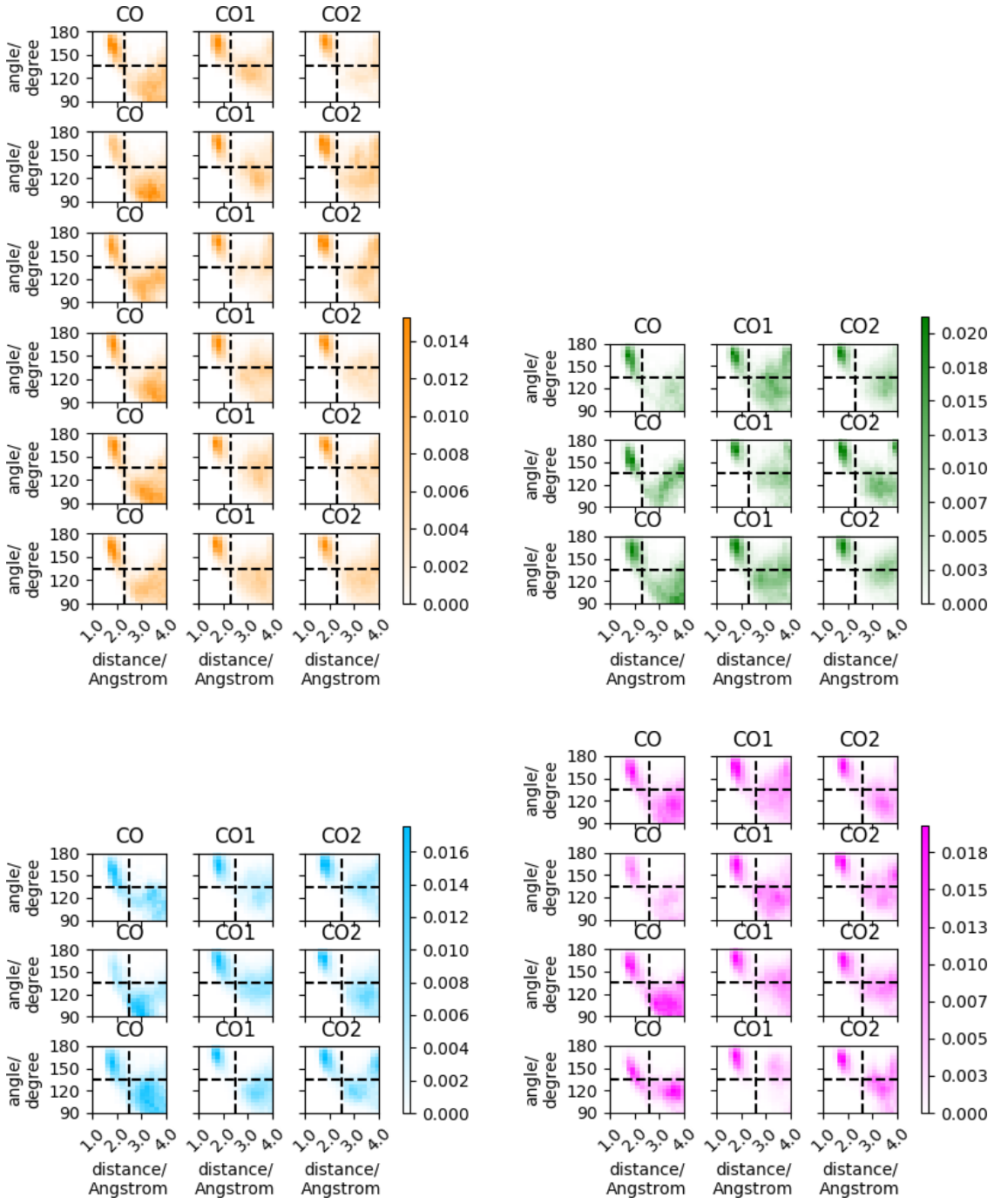
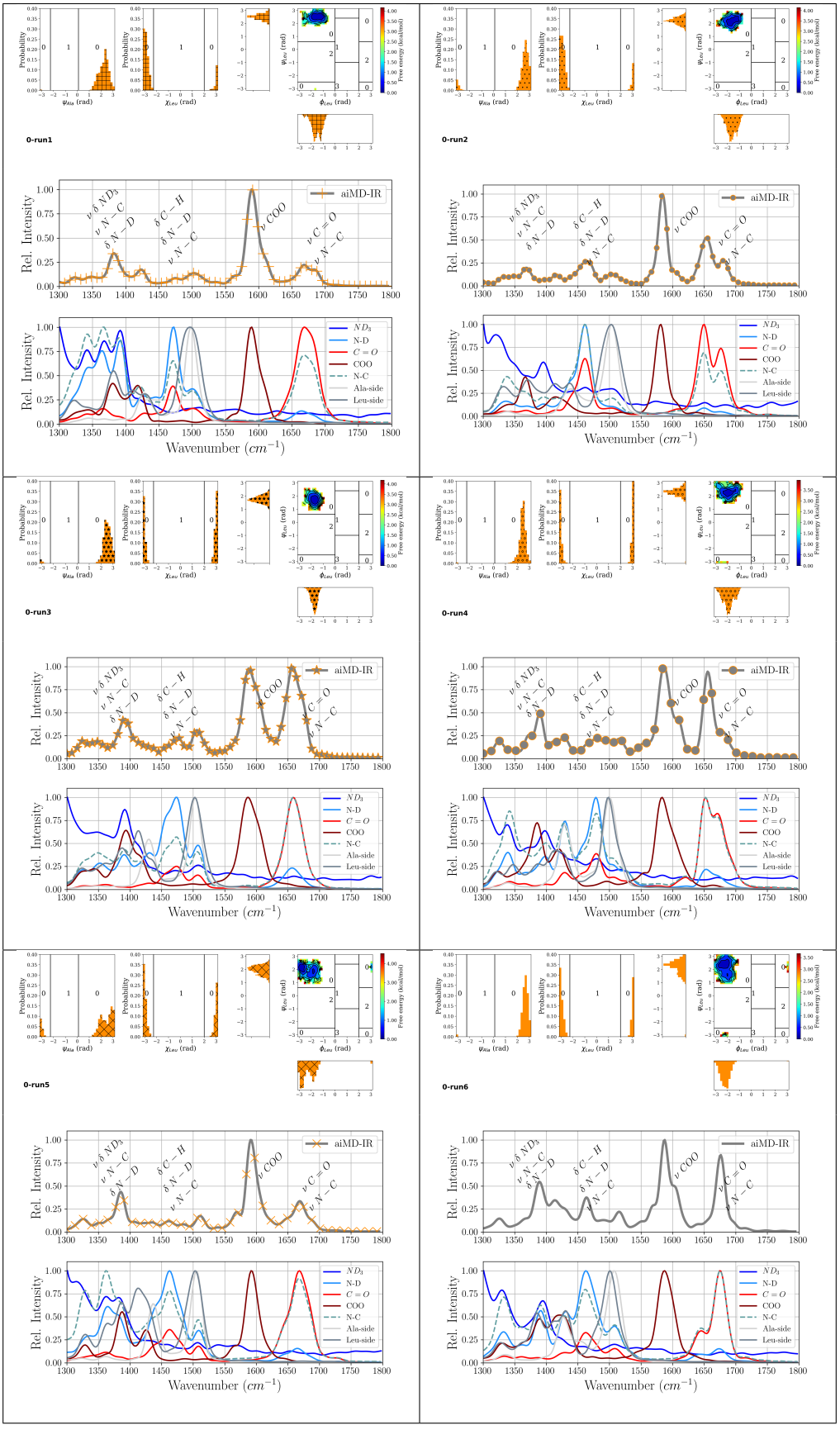
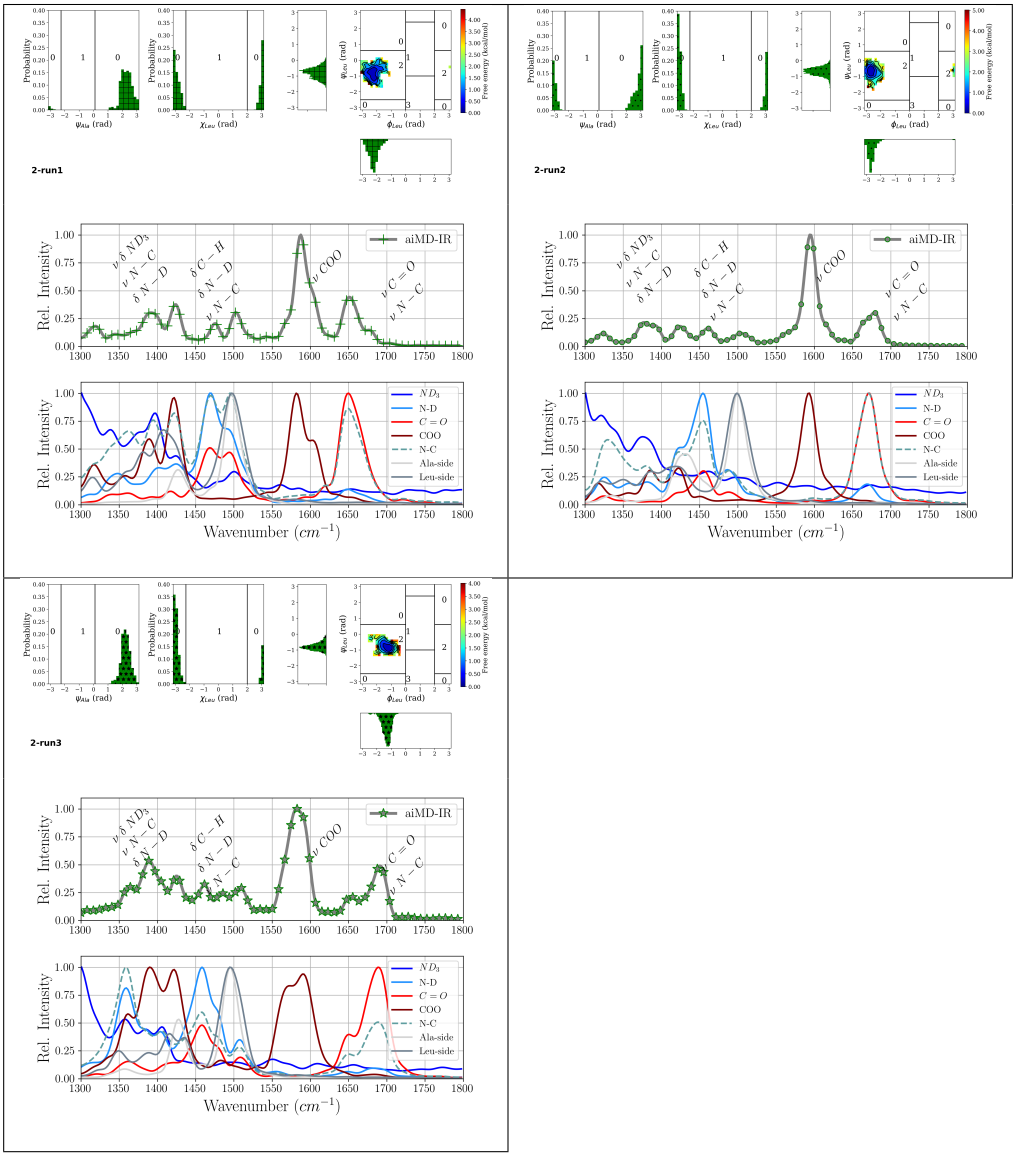
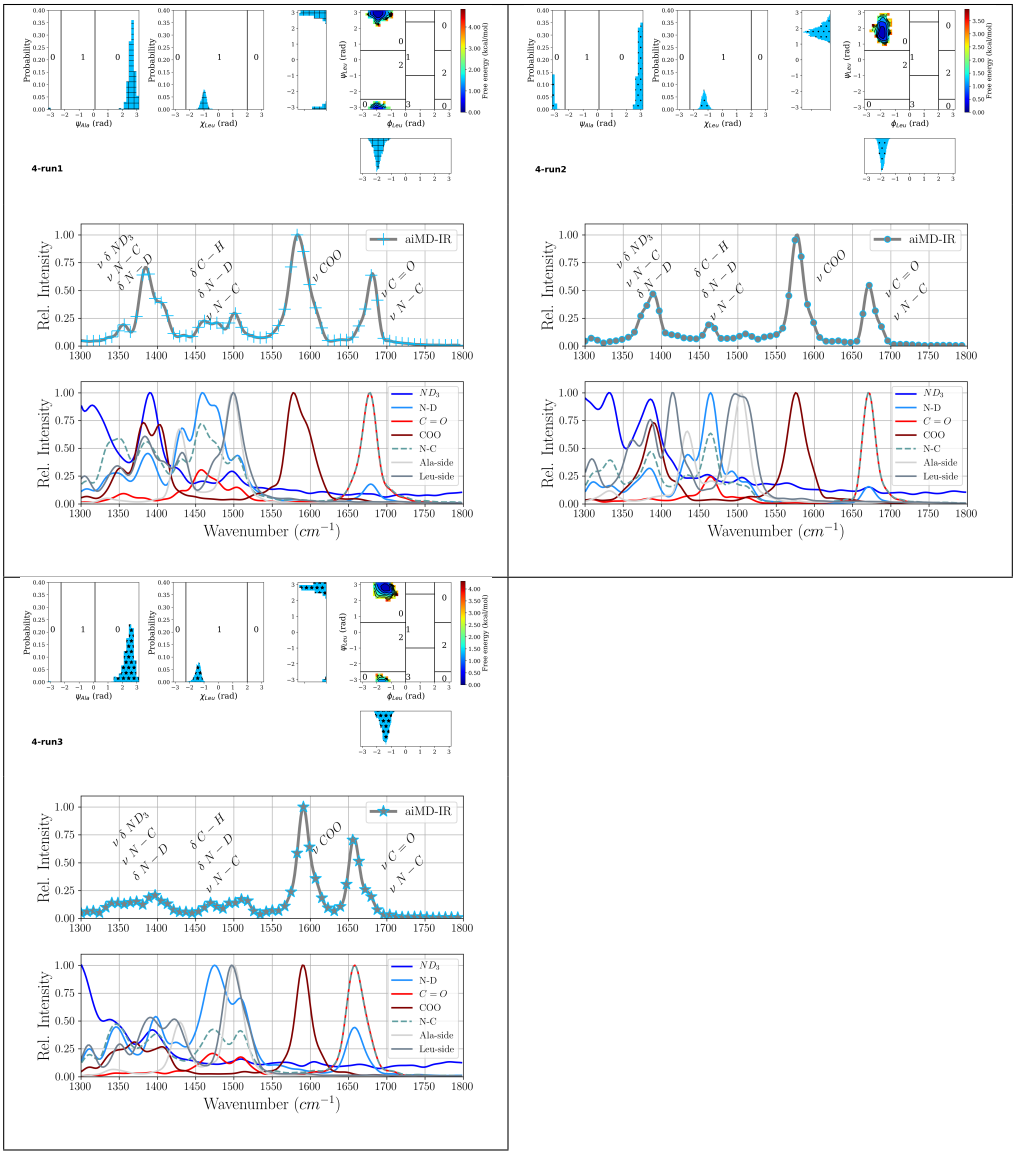


Figure S12: Joint distribution of angles between peptide oxygen atoms (C=O, COO1, and COO2, respectively), water deuterium, and water oxygen atoms with distances between peptide oxygen and water deuterium atoms. The four micro-states 0, 2, 4, and 6 are shown in orange, green, blue, and magenta, respectively. Each row of a sub-figure represents an individual simulation of that micro-state. The horizontal dashed line shows the angle threshold used as hydrogen bond criterion whereas the vertical dashed line indicates the first minimum of the radial distribution function shown in Figure S7 that can be understood as the limiting distance to find a water deuterium atom hydrogen-bonded to a peptide oxygen atom.

Torsion angle distributions, IR and power spectra from first principles MD simulations


0-run1
0-run2
0-run3
0-run4
0-run5
0-run6
Probability
Free energy (kcal/mol)
Rel. Intensity
aiMD-IR
$\nu\,\delta\,ND_3$
$\nu\,N-C$
$\delta\,N-D$
$\delta\,C-H$
$\delta\,N-C$
$\nu\,COO$
$\nu\,C=O$
$\nu\,N-C$
ND_3
N-D
$C=O$
COO
N-C
Ala-side
Leu-side
Wavenumber (cm^{-1})


2-run1
2-run2
2-run3
aiMD-IR
Rel. Intensity
Wavenumber (cm^{-1})
$\nu\,\delta\,ND_3$
$\nu\,N-C$
$\delta\,N-D$
$\delta\,C-H$
$\delta\,N-C$
$\nu\,N-D$
$\nu\,COO$
$\nu\,C=O$
$\nu\,N-C$
ND_3
N-D
$C=O$
COO
N-C
Ala-side
Leu-side


4-run1
4-run2
4-run3
aiMD-IR
Rel. Intensity
Wavenumber (cm^{-1})
$\nu\,\delta\,ND_3$
$\nu\,N-C$
$\delta\,N-D$
$\delta\,C-H$
$\delta\,N-D$
$\nu\,N-C$
$\nu\,COO$
$\nu\,C=O$
$\nu\,N-C$
ND_3
N-D
$C=O$
COO
N-C
Ala-side
Leu-side

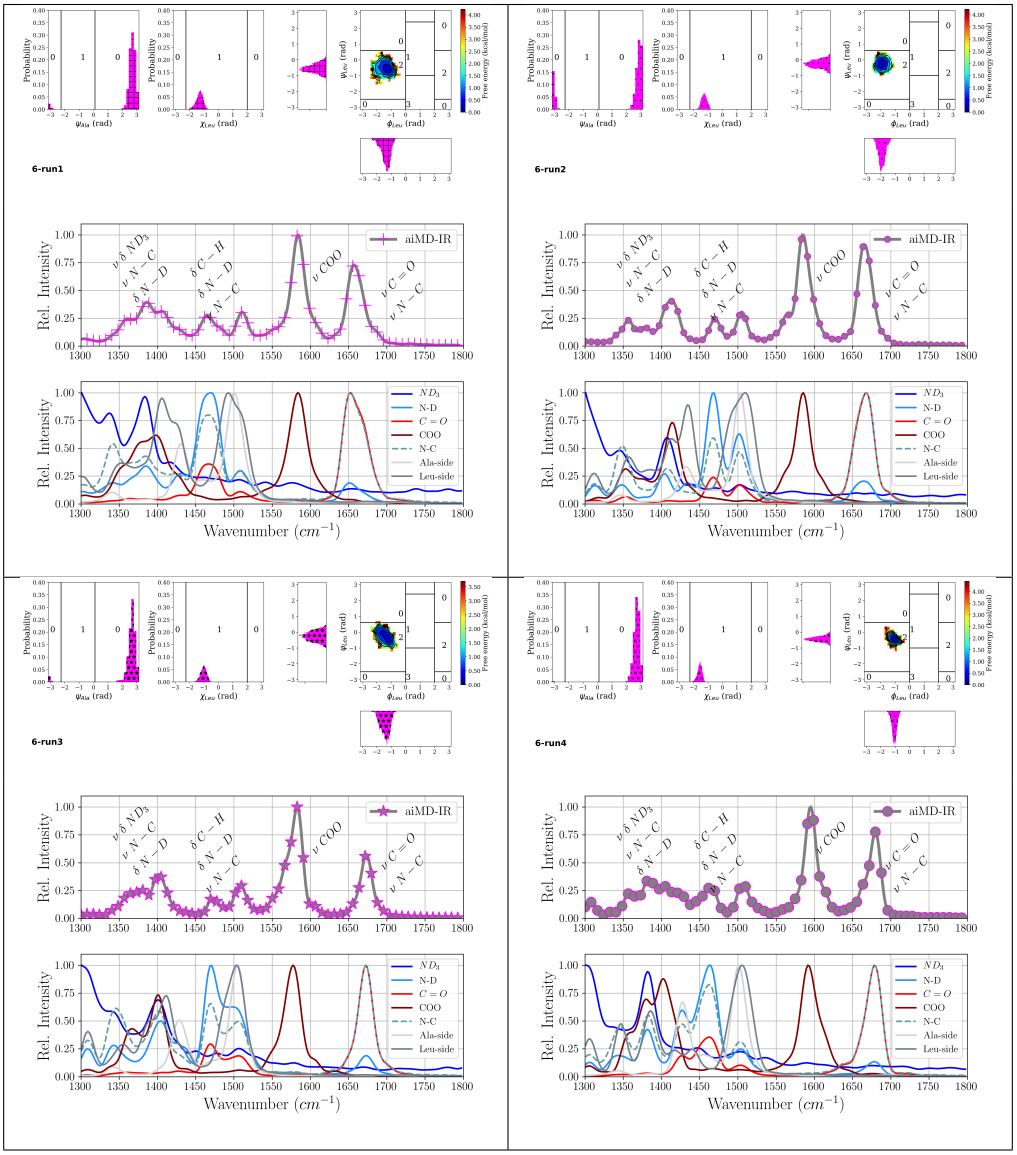


Figure S13: Torsion angle distributions (top), IR spectra (middle) and power spectra (bottom), obtained from the individual first-principles MD simulations of representative Ala-Leu conformations in water. The colours in distribtions and the IR spectra correspond to the different clusters as 0 (orange), 2 (green), 4 (purple), and 6 (magenta), respectively, with all runs of one cluster shown next to each other (run1 ... run3, run4, or run6, respectively, from left to right). The colours in the power spectra correspond to the different groups of atoms. The leucine backbone torsion angles, as shown in the ramachandran plot, would correspond to β-sheet for cluster 0 and 4, and to α-helix for clusters 2 and 6, respectively, but are chemically equivalent conformations due to the second oxygen atom in the carboxyl terminus.

References

(1) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(2) Hornal, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* **2006**, *65*, 712–725.

(3) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics* **2010**, *78*, 1950–1958.

(4) Darden, T.; York, D.; Pedersen, L. G. Particle mesh Ewald: an Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(5) Essman, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8503.

(6) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007**, *126*, 014101+.

(7) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, btt055.

(8) Deuflhard, P.; Weber, M. Robust Perron cluster analysis in conformation dynamics. *Linear algebra and its applications* **2005**, *398*, 161–184.

(9) Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Perez-Hernandez, G.; Hoffmann, M.; Plattne, N.; Wehmeyer, C.; Prinz, J.-H.; Noé, F. PyEMMA 2: A software package for estimation, validation, and analysis of Markov models. *J. Chem. Theory Comput.* **2015**, *11*, 5525–5542.

(10) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09 Revision E.01. Gaussian Inc. Wallingford CT 2009.

(11) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T.; Joannopoulos, J. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Rev. Modern Physics* **1992**, *64*, 1045.

(12) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: atomistic simulations of condensed matter systems. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014**, *4*, 15–25.

(13) Lippert, G.; Hutter, J.; Parrinello, M. The Gaussian and augmented-plane-wave density functional method for ab initio molecular dynamics simulations. *Theoretical Chemistry Accounts* **1999**, *103*, 124–140.

(14) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Computer Physics Communications* **2005**, *167*, 103–128.

(15) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268.

(16) King-Smith, R.; Vanderbilt, D. Theory of polarization of crystalline solids. *Physical Review B* **1993**, *47*, 1651.

(17) Resta, R. Macroscopic polarization in crystalline dielectrics: the geometric phase approach. *Reviews of modern physics* **1994**, *66*, 899.

(18) Marzari, N.; Vanderbilt, D. Maximally localized generalized Wannier functions for composite energy bands. *Physical review B* **1997**, *56*, 12847.

(19) Silvestrelli, P. L.; Parrinello, M. Water molecule dipole in the gas and in the liquid phase. *Physical Review Letters* **1999**, *82*, 3308.

(20) Silvestrelli, P. L.; Parrinello, M. Structural, electronic, and bonding properties of liquid water from first principles. *The Journal of chemical physics* **1999**, *111*, 3572–3580.

(21) Kirchner, B.; Hutter, J. Solvent effects on electronic properties from Wannier functions in a dimethyl sulfoxide/water mixture. *The Journal of chemical physics* **2004**, *121*, 5133–5142.

(22) Brehm, M.; Kirchner, B. TRAVIS-a free analyzer and visualizer for Monte Carlo and molecular dynamics trajectories. *J. Chem. Inf. Model.* **2011**, *51*, 2007–2023.

(23) Laury, M. L.; Carlson, M. J.; Wilson, A. K. Vibrational frequency scale factors for density functional theory and the polarization consistent basis sets. *Journal of Computational Chemistry* **2012**, *33*, 2380–2387.

(24) Halls, M. D.; Velkovski, J.; Schlegel, H. B. Harmonic frequency scaling factors for Hartree-Fock, S-VWN, B-LYP, B3-LYP, B3-PW91 and MP2 with the Sadlej pVTZ electric property basis set. *Theoretical Chemistry Accounts* **2001**, *105*, 413–421.

(25) Merrick, J. P.; Moran, D.; Radom, L. An Evaluation of Harmonic Vibrational Frequency Scale Factors. *The Journal of Physical Chemistry A* **2007**, *111*, 11683–11700.

(26) Sinha, P.; Boesch, S. E.; Gu, C.; Wheeler, R. A.; Wilson, A. K. Harmonic Vibrational Frequencies: Scaling Factors for HF, B3LYP, and MP2 Methods in Combination with Correlation Consistent Basis Sets. *J. Phys. Chem. A* **2004**, *108*, 9213–9217.

(27) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical J.* **2015**, *109*, 1528 – 1532.

(28) Gaigeot, M.-P. Unravelling the Conformational Dynamics of the Aqueous Alanine Dipeptide with First-Principles Molecular Dynamics. *J. Phys. Chem. B.* **2009**, *113*, 10059–10062.

(29) Kumar, V.; Chandra, A. First-Principles Simulation Study of Vibrational Spectral Diffusion and Hydrogen Bond Fluctuations in Aqueous Solution of N-Methylacetamide. *J. Phys. Chem. B* **2015**, *119*, 9858–0967.

(30) Feng, C.-J.; Tokmakoff, A. The dynamics of peptide-water interactions in dialanine: An ultrafast amide I 2D IR and computational spectroscopy study. *J. Chem. Phys.* **2017**, *147*, 085101.