JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

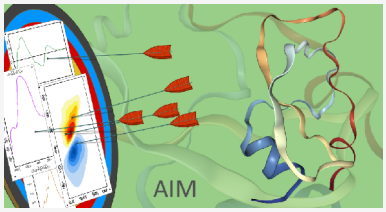
AIM: A Mapping Program for Infrared Spectroscopy of Proteins

Kim E. van Adrichem* and Thomas L. C. Jansen*

Cite This: J. Chem. Theory Comput. 2022, 18, 3089−3098 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Here, we present a new analysis program, AIM, that allows extracting the vibrational amide-I Hamiltonian using molecular dynamics trajectories for protein infrared spectroscopy modeling. The constructed Hamiltonians can be used as input for spectral calculations allowing the calculation of infrared absorption spectra, vibrational circular dichroism, and two-dimensional infrared spectra. These spectroscopies allow the study of the structure and dynamics of proteins. We will explain the essence of how AIM works and give examples of the information and spectra that can be obtained with the program using the Trypsin Inhibitor as an example. AIM is freely available from GitHub, and the package contains a demonstration allowing easy introduction to the use of the program.



1. INTRODUCTION

Knowledge about the structure and dynamics of proteins is of great importance to understand their function.[1−3] X-ray crystallography,[4] nuclear magnetic resonance (NMR) spectroscopy,[5] and cryogenic electron microscopy (cryo-EM)[6] are powerful techniques to address protein structure. However, these methods are not well suited for obtaining the structure of all proteins, and in particular, information about protein dynamics, therefore alternatives are needed. Infrared (IR) spectroscopy[7,8] often represents such an alternative due to its time resolution. However, this method does not have the desired atomic resolution. Therefore, modeling is of crucial importance for the interpretation of the IR spectra. In this paper, we will report on the development of a new program that improves the current schemes for connecting information on structure and dynamics obtained from molecular dynamics (MD) simulations with spectroscopic observables, which can be calculated with spectral simulation software. Our focus will be on the protein amide-I band, which is known to reveal information on the protein secondary structure.[9−12]

The amide-I vibration is dominated by the CO-stretch in a peptide bond. In Figure 1, typical amide-I vibrations are illustrated in a small peptide. The absorption cross-section of the amide-I vibration is large as the highly polar CO-bond results in a large transition-dipole moment of the vibration. This also results in strong couplings between different local amide-I vibrations and delocalization of the normal modes, which are the observable vibrations in the spectroscopic experiments. The resulting collective nature of the amide-I vibrations makes this spectral region sensitive to the secondary structure and therefore popular for studying protein structure and dynamics.[8]

The bend vibration of water absorbs at ∼1650 cm^{-1}, which overlaps with the amide-I spectral region. To ease interpretation of the spectra and limit spectral overlap, experiments are often performed in heavy water, which shifts the water bend vibration out of the spectral window to ∼1200 cm^{-1}.[13] This also results in

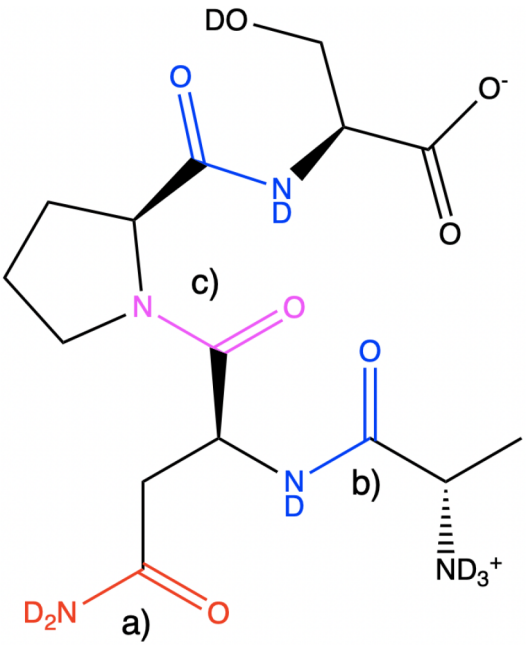
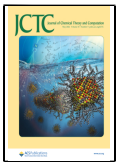


Figure 1. a) The primary amide found in the glutamine and asparagine side chains. b) Secondary amide found in the protein backbone. c) The tertiary amide group found at sites before proline in the backbone (also denoted preproline) units. The carbon, oxygen, nitrogen, and deuterium atoms involved in the amide-I vibration are highlighted with color.

Received: February 2, 2022
Published: April 7, 2022



https://doi.org/10.1021/acs.jctc.2c00113
J. Chem. Theory Comput. 2022, 18, 3089−3098

(partial) exchange of the acidic hydrogen atoms in the protein with deuterium. This includes exchange of the hydrogen atoms involved in the amide-I vibrations, which shifts these vibrations by about 10 cm^{-1}.[14,15] To distinguish the new vibrations, they are often denoted amide-I′. Here, we will use the term amide-I to describe both, but unless otherwise stated explicitly, we will assume that the amide-I′ vibration is actually considered.

Gas-phase infrared spectra of peptides can be calculated efficiently using standard quantum chemical methods. However, for solvated proteins, such calculations become prohibitively computationally costly. This stems from three basic reasons. First, real proteins contain hundreds of atoms. Second, the solvent affects the vibrations and must be included explicitly. Finally, the systems are dynamic, and therefore, calculating the vibrational frequencies once does not suffice. For absorption spectra, a few hundred frequency calculations might suffice for the full protein, while for two-dimensional infrared spectra,[7] thousands of frequency calculations are needed to account for the dynamic line shape. To overcome this challenge, the idea of frequency mappings emerged.[16] Essentially, the vibrational frequencies are approximated using maps connecting them with the local electrostatic environment. Such maps are based on first-principles calculations or empirical fitting. Combined with spectral simulation methods[17] this allows predicting the infrared spectra of entire proteins. This simulation protocol has successfully been applied in the study of protein structure and dynamics both using infrared absorption and two-dimensional infrared spectroscopy. The applications include the study of protein folding,[18−21] the dimerization of insulin,[22,23] the gating mechanism of the influenza M2 channel,[24] the function of potassium channels,[25−27] the formation of amyloid fibril structures,[28] and the electrochemistry of cytochrome *c*.[29] Computational spectroscopy has, thus, been demonstrated as a powerful tool in interpreting and predicting infrared spectra. However, general software, which is easy to use for nonexperts, is still not readily available.

The remainder of this paper is organized as follows. First, in the Methods section, the essential simulation protocol is summarized, and the role of the AIM program[30] in this protocol is outlined. Furthermore, the advantages of AIM compared to existing mapping programs are described. This is followed by a section demonstrating examples of the application of AIM to both GROMACS and NAMD molecular dynamics trajectories. AIM is freely available for download on GitHub,[30] and example files are included in the Supporting Information of this paper. Finally, the paper ends with the Conclusion, which includes an overview of our ideas for near future extensions of the program.

2. METHODS

The overall workflow for calculating amide-I spectra is illustrated in Figure 2. The AIM program takes a central place in this workflow. In essence, one will have to start by using an initial guess structure for the protein. This could, for example, come from the Protein Data Bank[31] or structure prediction software such as AlphaFold.[32] A molecular dynamics simulation is then performed to generate a representative structure and dynamics. This can involve sampling methods such as replica exchange[33] or coarse-grained molecular dynamics coupling with backmapping schemes[34] to produce atomistic trajectories. Atomistic molecular dynamics trajectories with dense sampling (10−20 fs between stored frames) are then used as input for the AIM program. AIM will generate a time-dependent Hamiltonian for the amide-I vibrations along with a trajectory of the transition

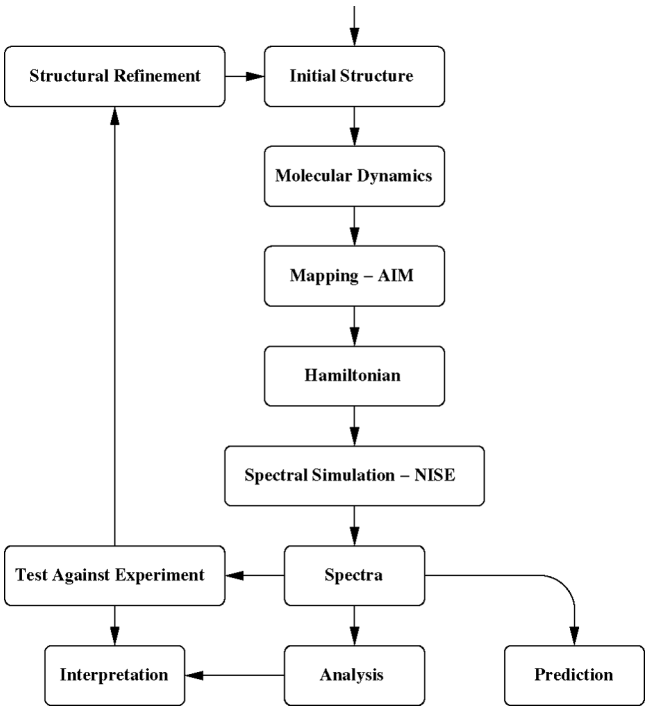


Figure 2. A schematic illustration of the workflow for calculating amide-I spectra.

dipoles and a trajectory of the center of the positions of the amide-I modes. This is the required input for the amide-I spectral calculations of linear absorption, linear dichroism, vibrational circular dichroism (VCD), isotope-label spectra, and two-dimensional infrared spectroscopic (2DIR) signals. AIM is designed to directly interface with the NISE[35] program package, but the generated trajectories can also be converted to be used in the Spectron[36,37] and g_spec[38] packages. These programs also allow for the analysis of the origin of spectral features. Comparing the calculated spectra with experimental observations then allows spectral interpretation and potentially refinement of the structural model of the investigated protein. The following description will focus on the general properties of the Hamiltonian trajectories and the main options available in AIM to extract these Hamiltonians from molecular dynamics trajectories. For proteins of which the spectra are not yet known, the procedure can further be used to predict the spectra and for proposing ways of distinguishing different potential structures.[39] A brief outline of the procedure used for spectral calculations in the NISE program will also be provided.

The amide-I vibrations are described with a vibrational Hamiltonian of the form[7]

$$H(t) = \sum_i \omega_i(t) B_i^\dagger B_i + \sum_{i\neq j} J_{ij}(t) B_i^\dagger B_j - \sum_i \frac{\Delta_i(t)}{2} B_i^\dagger B_i^\dagger B_i B_i \quad (1)$$

where i labels each amide-I site, and $B_i^\dagger$ and B_i are the Bosonic creation and annihilation operators, respectively. $\omega_i(t)$ is the vibrational frequency of the local amide-I mode i at time t, $J_{ij}(t)$ is the coupling between two local modes, and $\Delta_i(t)$ is the anharmonicity of the amide-I vibration on site i. The anharmonicity is needed for two-dimensional infrared spectra, where double excited states can be reached, as a result of two consecutive excitations. The time-dependence of the parameters in this Hamiltonian originates from the fluctuations of the protein structure and the local environment. Extensive research

https://doi.org/10.1021/acs.jctc.2c00113
J. Chem. Theory Comput. 2022, 18, 3089−3098

has been performed in the community to develop so-called frequency mappings to predict these parameters based on information about the local structure and environment.[16] Below a more detailed discussion of the different aspects of such mappings will be outlined. The AIM program allows using the most popular of these mappings to create vibrational amide-I Hamiltonians from molecular dynamics trajectories of proteins.

The local amide-I vibration is typically described with a mapping that depends on the local electrostatic potential, field, and/or gradient on the carbon, oxygen, nitrogen, and/or deuterium atom of the central amide group. Furthermore, a nearest neighbor frequency shift is sometimes added as described below. We have implemented the most common electrostatic maps in AIM.[40−44] The package also allows for user-defined mappings which should be straightforward to add when they depend only on the parameters discussed above.

The simplest approach for extracting couplings is using the transition-dipole coupling with the point-dipole approximation. The local amide-I modes are then considered vibrating point dipoles, and the coupling between pairs of vibrating point dipoles is given by the equation

$$J_{ij} = \frac{1}{4\pi\epsilon_0\epsilon_r}\left(\frac{\vec{\mu}_i\cdot\vec{\mu}_j}{|r_{ij}|^3} - \frac{3\vec{\mu}_i\cdot\vec{r}_{ij}\vec{r}_{ij}\cdot\vec{\mu}_j}{|r_{ij}|^5}\right) \quad (2)$$

Here, r_{ij} is the distance between the two dipoles, $\vec{\mu}_i$ and $\vec{\mu}_j$ are the transition dipoles, and ϵ_r is the relative dielectric constant. AIM also includes the more elaborate transition charge coupling scheme,[45−47] which includes higher order multipole corrections to the interaction. The relative dielectric constant can be provided as an input parameter, scaling the long-range couplings depending on the (optical) dielectric environment. Previously, a value of ϵ_r = 1 gave reasonable results compared with experiment.[47−49] A benchmark study[50] suggests that setting ϵ_r = 1.5 may also be an appropriate choice.

As the electrostatic approximation may not be sufficient for short-range interactions, DFT calculations have been used to parametrize vibrational frequency shifts caused by the nearest neighbor groups along the protein backbone.[42] These mappings also provide the couplings between the neighboring amide vibrations. In essence, the maps were created by calculating the amide-I modes for peptide dimers with DFT and varying the Ramachandran angles systematically in combination with a Hessian reconstruction procedure. The frequency shift of the two neighboring units and the coupling between the two can then be determined using bilinear interpolation on the grid of values determined with DFT.

The anharmonicity of the amide-I mode is typically set to 16 cm^{-1} as determined in early experiments.[7] This value was also found in one mapping study.[42] It was found that the anharmonicity does not fluctuate a great deal, and using a fixed value is often a reasonable approximation. The current version of AIM does not include a specific mapping of the anharmonicities, and a constant value is assumed.

A backbone amide unit located just before a proline unit will be a tertiary amide (see Figure 1c). This unit has a vibrational frequency that is typically about 27 cm^{-1} below that of the secondary amide units.[51] Only the Jansen map was developed for this unit, and other maps implemented in AIM at this point apply a simple 27 cm^{-1} redshift for preproline groups to correct the frequency. We note that this value can be changed by the user, and it is possible to combine the preproline Jansen map with the other maps for the backbone in the program.

The side chains of glutamine and asparagine contain a primary amide-I unit (see Figure 1). The vibrational frequency of a primary amide is typically 30 cm^{-1} above that of a secondary amide.[41] The Skinner map was developed with an explicit side-chain map, which is included in AIM. For all other map choices, the side-chain frequency is calculated with the secondary amide map and blue-shifted by 30 cm^{-1}. We note that this value can be changed by the user, and it is possible to combine the side-chain Skinner map with the other maps for the backbone in the program.

Isotope edited vibrational spectroscopy is a powerful tool for revealing local structural and dynamical information. The formal frequency shift is given by the harmonic oscillator formula $\omega = \sqrt{\frac{k}{\mu}}$, where k is the force constant, and μ is the reduced mass. Assuming that only the reduced mass changes upon isotope editing, this results in a frequency scaling. In practice, a systematic frequency shift of about −41 cm^{-1} is observed for ^{13}C labeling[21] and −66 cm^{-1} for combined $^{13}C^{18}O$ labeling.[50] Often vibrational spectroscopy for the amide-I band is performed in D_2O to avoid interference with the H_2O bend around 1650 cm^{-1}.[52] This leads to an exchange of hydrogen with deuterium of the acidic protons including that of the peptide bond.[53] This changes the gas-phase frequency by 10 cm^{-1}.[54] It was demonstrated that this is neither well described by a simple shift nor with a simple change of reduced mass.[15] Therefore, the spectra in D_2O can better be modeled explicitly with mappings taking the deuteration explicitly into account as the Skinner,[41] Jansen,[42] and Tokmakoff[40] ones implemented in AIM. Isotope edited spectroscopy with ^{13}C and ^{18}O labels can easily be performed as postprocessing of the Hamiltonian trajectory generated by AIM. Often the approximation that the isotope labeled units are uncoupled from the other units is further applied. This allows for efficient calculations by truncating the Hamiltonian to the labeled units only for the spectral calculations.[26] Such truncated Hamiltonian can also be obtained directly with AIM.

The main function of AIM is to extract the parameters for the Hamiltonian trajectory described in eq 1. This is typically done with 20 fs timesteps between the stored MD frames, as this is needed to sample the frequency fluctuations, which are important for motional narrowing observed in FTIR spectra and lineshapes observed in 2DIR spectra. AIM further allows storing a file with positions related to the vibrations. As a default, this is chosen to be the position of the carbon atom in the peptide bond. This, for example, allows calculation not only for the VCD spectra[12,55] but also for various analysis purposes, where the knowledge of the position of the different sites is important.

The predecessor of AIM, AmideImaps,[56] has many of the functionalities of this new program. However, that was a tool developed specifically for GROMACS[57] and only worked for that specific package with support for GROMACS versions 3 and 4. AIM utilizes the MDAnalysis[58,59] package and, in principle, allows the use of trajectories of all the MD-software packages that MDAnalysis can analyze. We have at this point implemented and tested the use of GROMACS,[57] CHARMM,[60] NAMD,[61] and AMBER[62] trajectories. This is also an advantage over other spectroscopic mapping programs such as g_amide.[63,64] Other advantages of AIM include that protein recognition has been implemented to automatically identify and

https://doi.org/10.1021/acs.jctc.2c00113
J. Chem. Theory Comput. 2022, 18, 3089−3098

treat diverse structures including cyclic proteins and multichain systems. For most systems, this means that the program automatically detects all relevant chromophores and assigns the proper connectivity for the application of nearest neighbor mappings. Preproline residues are identified, and in the case of consecutive proline residues, all preproline units are treated as such. The program allows users to add mappings for nonstandard residues. It has been optimized and is about twice as fast as the previous AmideImaps when the optional precompiled c-libraries are used. This speedup is partially thanks to the implementation of a neighbor searching algorithm.[65]

3. EXAMPLES FOR THE TRYPSIN INHIBITOR

In the following, we will demonstrate the capability of the AIM program by the application to the Trypsin Inhibitor. This protein was chosen because experimental spectra are available for the same species as the Protein Data Bank structure and because it contains both an α-helix, β-sheet, and a short 3^{10}-helix segment. The initial structure was taken from the Protein Data Bank[31] using the 4PTI entry.[66] The structure is illustrated in Figure 3. In the following, we will first describe the MD simulations performed with GROMACS and NAMD, then we will present the Hamiltonians obtained with AIM, and finally, we will present the resulting spectra obtained with NISE.

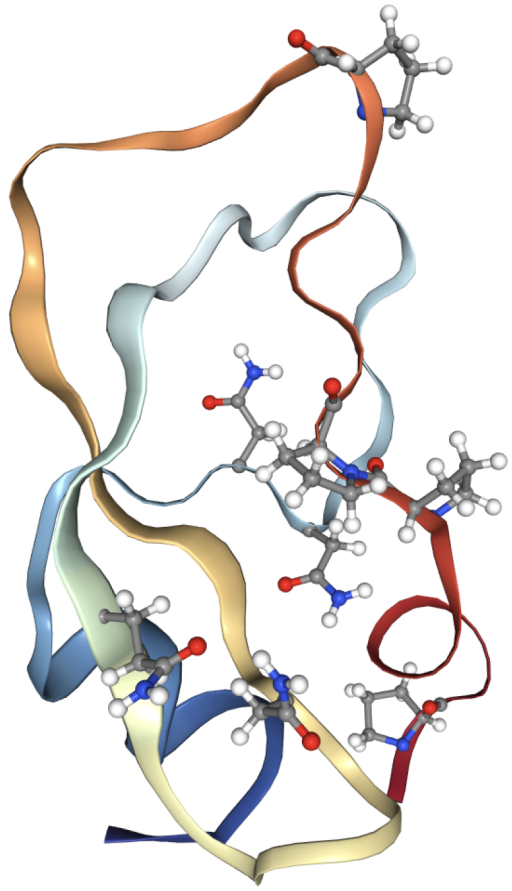


Figure 3. Structure of the Trypsin Inhibitor plotted using NGLview. The highlighted amino acids are the proline (Pro2, Pro8, Pro9, and Pro13), asparagine (Asn24, Asn43, and Ans44), and glutamine (Gln31) residues.[67] The N-terminal is near the 3^{10}-helix colored in red, and the C-terminal is near the α-helix colored in blue.

3.1. Molecular Dynamics Simulations. The GROMACS simulation structure was prepared by solvating the protein with SPC/E water[68] covering the protein with a layer of at least 1 nm of water using periodic boundary conditions. A total of 6898 water molecules were included. Six chlorine counterions were added to keep the simulation box neutral. The protein and chlorine were treated using the OPLS-AA force field.[69] The system preparation and simulations were performed with GROMACS version 4.6.3.[70] Following an initial equilibration, the system was simulated for 1 ns using 2 fs time steps and saving every tenth frame. A 1 nm cutoff was used for the short-range interactions, and Particle-Mesh-Ewald[71] was used for the long-range interactions. A simple 1 nm cutoff was used for the Lennard-Jones interactions. The temperature was kept constant at 300 K using the velocity rescaling scheme,[72] and the pressure was kept constant at 1 bar using the Parrinello−Rahman barostat.[73]

The MD trajectory generated with GROMACS was analyzed using AIM. For all calculations, a cutoff of 2 nm was used for the electrostatic mappings,[74] and the nearest neighbor coupling and frequency shifts[46] were applied. The neighbor searching algorithm was applied with updates for every 50^{th} stored frame using a cutoff of 2.5 nm. The dielectric constant used for the long-range coupling was set to one. These settings replicate typical choices of previous studies.[49] Two separate Hamiltonian trajectories were generated. For the first, we used the Jansen electrostatic map[42] for frequencies and dipoles and the transition-dipole coupling scheme of Tasumi.[75] For the other Hamiltonian trajectory, the Skinner electrostatic map was used[41] together with transition dipoles of Torii[75] and the transition-charge coupling scheme.[46] In both cases, the position of the peptide bond carbon atom was stored as the position for the vibrational modes. These mapping combinations were chosen, as they were found to perform the best in the previous benchmark studies.[49]

For the NAMD simulations, the structure was prepared by solvating the protein with TIP3P water[76] covering the protein with a layer of at least 1 nm of water using periodic boundary conditions. A total of 2547 water molecules were included. Six chlorine counterions were added to keep the simulation box neutral. The protein and chlorine were treated using the CHARMM36m force field.[77] The system preparation and simulations were performed with NAMD version 2.14.[61] Following an initial equilibration, the system was simulated for 1 ns using 2 fs time steps and saving every tenth frame. A 1 nm cutoff was used for the short-range interactions. and Particle-Mesh-Ewald[71] was used for the long-range interactions. A 1 nm cutoff with a shifting function and a 0.9 nm shifting distance was used for the Lennard-Jones interactions. The temperature was kept constant at 300 K, and the pressure was kept constant at 1 bar using the Langevin algorithms for temperature and pressure.

The MD trajectory generated with NAMD was analyzed using AIM. For all calculations, a cutoff of 1.8 nm was used for the electrostatic mappings, and the nearest neighbor coupling and frequency shifts[46] were applied. The slightly shorter cutoff was needed to avoid issues with the periodic boundary, as the simulation box was smaller than that of the GROMACS calculations. The neighbor searching algorithm was applied with updates for every 50^{th} stored frame using a cutoff of 2.5 nm. The dielectric constant used for the long-range coupling was set to one just as for the GROMACS trajectories. One Hamiltonian trajectory was generated. The Skinner electrostatic map was

https://doi.org/10.1021/acs.jctc.2c00113
J. Chem. Theory Comput. 2022, 18, 3089−3098

used[41] together with transition dipoles of Torii[75] and the transition-dipole coupling scheme of Tasumi.[75]

The reason for sampling the trajectories at 20 fs intervals is that the site energies are known to fluctuate on a sub-100 fs time scale[78] leading to motional narrowing of the spectra. Furthermore, the used time step defines the spectral width available in the NISE algorithm.[79] Previous benchmark studies found 20 fs to provide a good compromise between accuracy and computational cost.[48] The use of thermostat and/or barostat may affect the dynamics in the molecular dynamics simulations and therefore the resulting spectra, which depend on the dynamics. Using the microcanonical ensemble or comparing results with and without a thermostat and barostat is, thus, recommended. For the present GROMACS simulations, this was tested, and no significant difference was observed on the spectra.

3.2. Hamiltonians. The average of each calculated Hamiltonian trajectory is illustrated in Figure 4. The pattern of the couplings reflects the main secondary structures. The 3^{10}-helix from Asp3 to Leu6 is accompanied by strong negative neighbor couplings. The double-stranded antiparallel β-sheet from Ile18 to Tyr35 is visible through the antidiagonal line showing the cross-strand couplings where Lys26 is in the middle of the β-turn. The α-helix from Ala48 to Gly56 is connected with negative couplings between the involved units resulting from the largely aligned transition dipoles in the helix. The side-chain amides exhibit rather large couplings with each other as well as numerous backbone units.

3.3. Spectral Simulations. The spectral simulations were performed using the NISE_2017 spectral simulation package.[35] For 2DIR simulations, couplings with a magnitude smaller than 0.01 cm^{-1} were neglected to speed up the calculations. The anharmonicity was set to 16 cm^{-1}. The spectra were calculated in the spectral window from 1550 to 1750 cm^{-1} using coherence times from 0 to 2.56 ps in 20 fs increments. An exponential apodization function corresponding to an effective lifetime of 1.8 ps was used for spectral smoothening. The spectra were obtained by averaging over 1000 equidistant starting points along the trajectory. This follows commonly used simulation protocols.[12]

The linear absorption for all three simulations is shown in Figure 5 together with the experimental data from ref 80. A systematic peak shift was introduced by maximizing the spectral overlap with the experimental data as discussed in refs 48 and 49. The obtained shifts were −18.8, 7.2, and 21.0 cm^{-1} for the three trajectories which is comparable to the numbers reported in refs 48 and 49. The general line width and peak shape of the simulated spectra are in fairly good agreement with experiment. However, all the predicted spectra are slightly narrower than the experimental one. Subsystem spectra were calculated by projection on subsets of residues[29] for interpretation. The projection on the preproline residues shows that these units predominantly contribute to the red part of the spectrum in all cases. The projection on the side-chain units gives very different results, when using the Skinner map specifically developed for these units and when using the shifted Jansen map. The shifted map underestimates the blue shift of the side-chain amides significantly. It is, thus, a better idea to use the Skinner map for side chains directly in combination with the Jansen map than to simply shift the Jansen map frequencies. This is in contrast to the preproline maps, where the shifted backbone Skinner map gives a reasonable result. More elaborate comparisons of different mappings are found in the literature.[47−49]

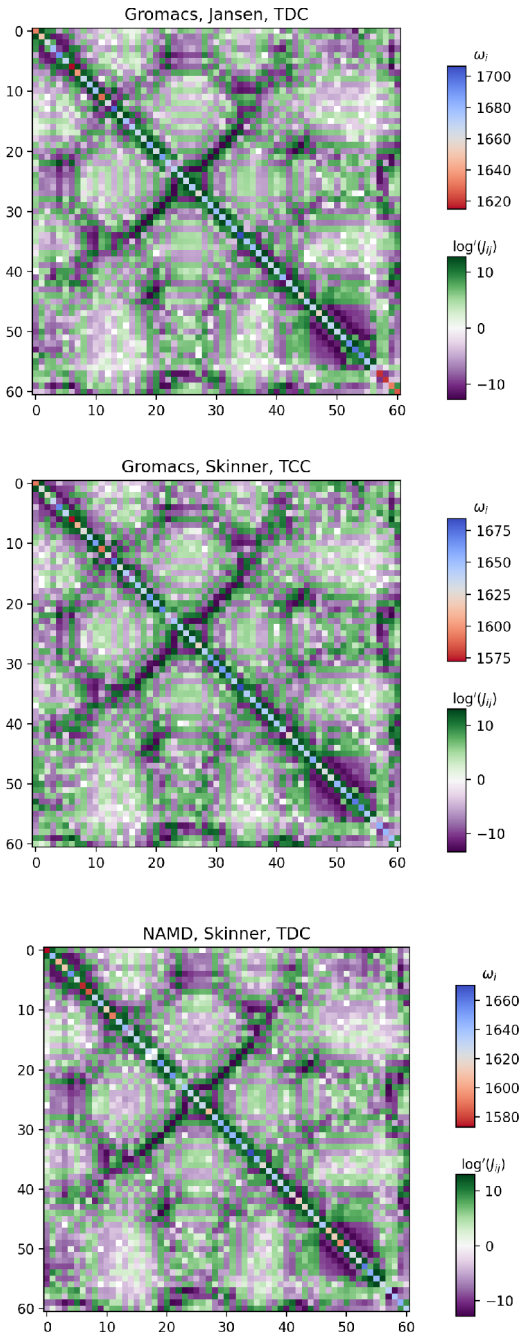


Figure 4. Comparison of the average of the two GROMACS based (top) and the NAMD based (bottom) Hamiltonian trajectories of the Trypsin Inhibitor. The indices 0 to 56 are the backbone residues. Indexes 57 through 60 are the side-chain units with Gln31 being index 58. The couplings are plotted using a logarithmic scale to enhance smaller values.

The simulated VCD spectra[12,55,81] are shown in Figure 6. For these simulations, the carbonyl carbon of the backbone was assumed to be the center of each site vibration, and intrinsic magnetic transition-dipole moments were neglected. The dependence on the simulation method is very clear, and the deviation from the experimentally reported spectrum[82] is sizable. The used algorithm provides conservative circular dichroism spectra, while the experimental spectrum is predominantly negative. This can potentially be explained by mixing with the amide-II vibration, which is neglected here.[81,83]

https://doi.org/10.1021/acs.jctc.2c00113
J. Chem. Theory Comput. 2022, 18, 3089−3098

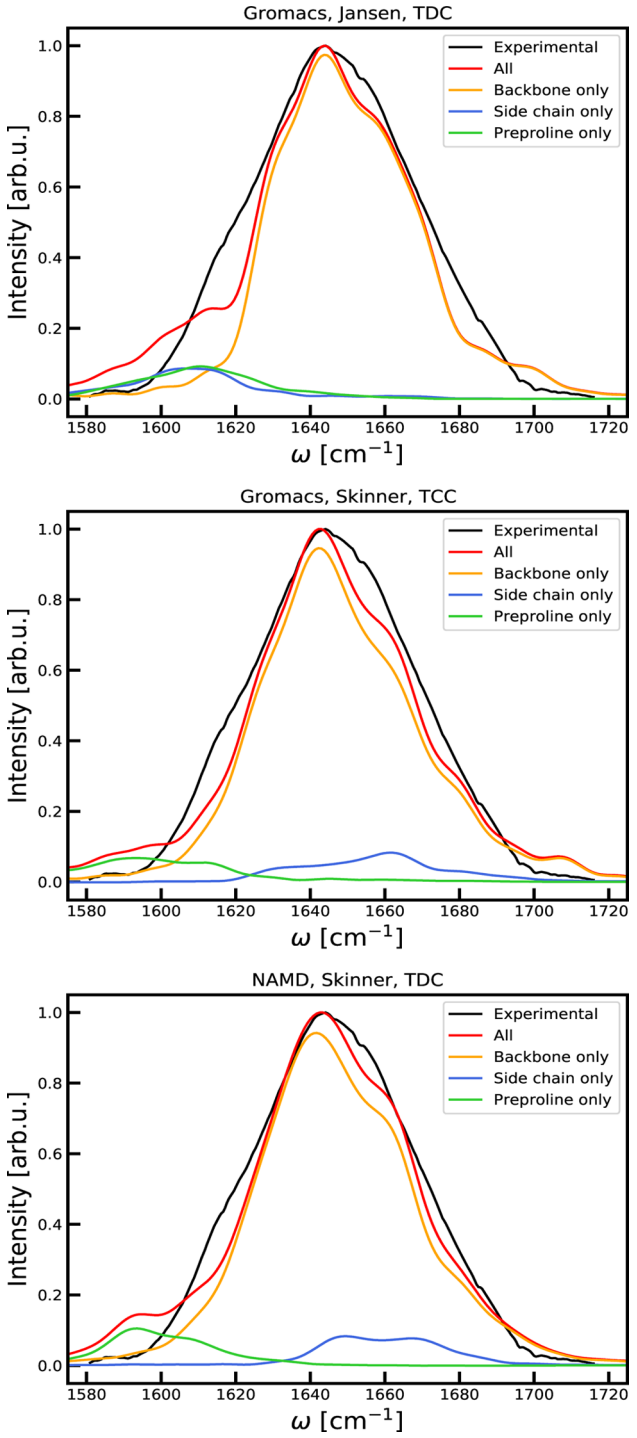


Figure 5. Predicted linear absorption spectra of the Trypsin Inhibitor. The experimental data are taken from ref 80. The frequencies of the simulated spectra were shifted to maximize the spectral overlap, as described in the text.

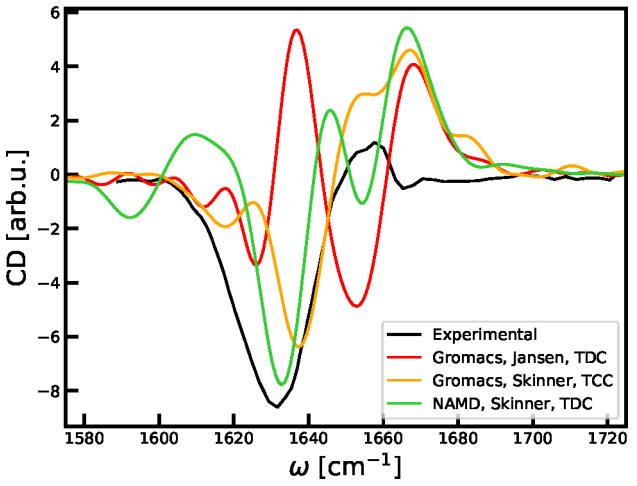


Figure 6. Predicted VCD spectra of the Trypsin Inhibitor. The experimental data was digitized from ref 82. The frequencies of the simulated spectra were shifted as for the absorption spectra.

It is generally known that circular dichroism is a very sensitive technique and, thus, challenging to predict. This, in principle, provides a handle for further improving the mappings in the future. Furthermore, the inclusion of coupling with other modes, the inclusion of magnetic dipole contributions, and the adjustment of the assigned positions of the transition dipoles may help improve the agreement in the future.

The two-dimensional infrared spectra obtained from the three trajectories described above are compared with experimental data[80] in Figure 7. The frequencies in these spectra were not shifted. The predicted spectra are all narrower than the experimental spectrum. This is expected as the absorption spectra were already narrower. This means that the separate peak features stand out more in the simulated spectra. These features are quite similar in the three predicted spectra. Diagonal slices and broad-band pump−probe spectra are shown in Figure 8 to further highlight the differences between the spectra. The pump−probe spectra were obtained by integrating the 2DIR spectra over ω_1. Overall, the NAMD trajectory seems slightly closer to the experimental result. The underestimation of the line width may be a result of either the amount of disorder in the system or an overestimation of the couplings. A potential source of undersampling of disorder may be the length of the trajectories, which were only 1 ns long here. Based on the calculations presented here and the previous benchmark studies already discussed, the mapping combination combining the Skinner backbone, the Skinner side chain, and the shifted Skinner map for the preproline units combined with the TCC coupling and the nearest neighbor corrections used in this study is made the default choice in the AIM program. Depending on, for example, the force field choice, the user may, of course, want to choose an alternative mapping combination.

The use of the NumPy[84] and Numba[85] libraries is important for the speed of the developed code. The time duration of the calculations described here will, of course, be highly system dependent. As an indication of the expected computational time requirement, the Skinner mapping calculations using the c-library took about 1 h on both a 2.3 GHz 8-Core Intel Core i9 under Mac OS and a 2.5 GHz (two Intel Xeon E5 2680v3 CPUs) with Linux. An absorption spectrum was calculated in about 12 s. A VCD spectrum took 10 min to calculate. Each 2DIR spectrum was obtained in 128 CPU h. The NAMD molecular dynamics simulations took about 2 CPU h. The AIM program is not running parallel, but it is easy to distribute subsets of a trajectory over different calculations. In the future, we intend to make a native parallel AIM version available via the GitHub site.[30]

4. CONCLUSION

We have described a protocol for simulating the amide-I band of proteins in solution. Our focus is on AIM, which is a new

https://doi.org/10.1021/acs.jctc.2c00113
J. Chem. Theory Comput. 2022, 18, 3089−3098

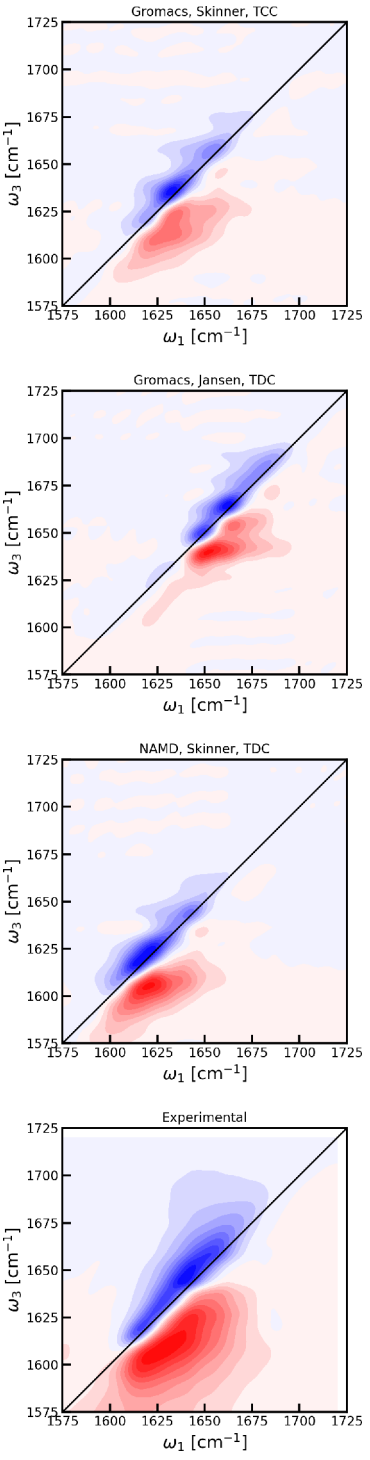


Figure 7. Predicted two-dimensional infrared spectra with zero waiting time and perpendicular laser polarization of the Trypsin Inhibitor. The contour lines are drawn for each 10% of the maximal signal. The red color indicates induced absorption, while the blue color indicates bleach. The experimental data are taken from ref 80.

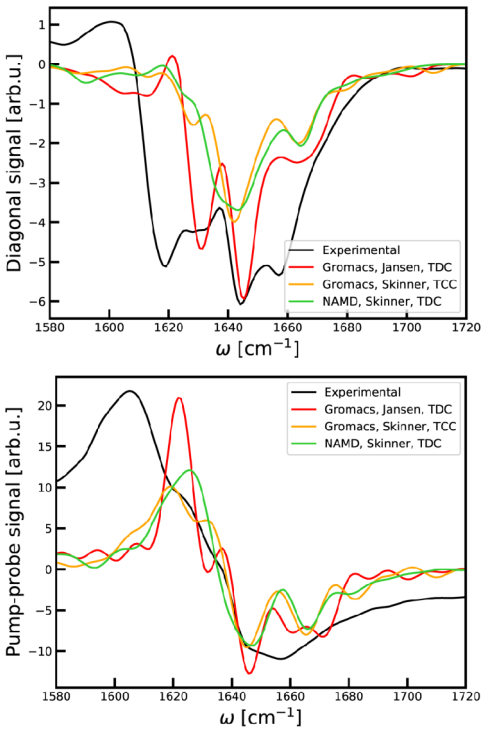


Figure 8. Top: Diagonal cuts through the 2DIR spectra of the Trypsin Inhibitor. Bottom: Broad-band pump−probe spectra of the Trypsin Inhibitor. The experimental data were obtained by analysis of data from ref 80. The frequencies of the simulated spectra were shifted as for the absorption spectra.

program developed to convert the structural and dynamic information obtained in molecular dynamics simulations to Hamiltonian trajectories that are used as input for spectral calculations. AIM is distributed freely as open-source software via GitHub.[30] We demonstrated the application of the program to trajectories generated with GROMACS and NAMD. We discussed a few points that one needs to keep in mind for such calculations.

For the demonstration of AIM, we found the general coupling pattern to be very similar between models, while some variation was found between absorption spectra in accordance with previous studies.[47−49] The variation between predicted VCD reflects the higher sensitivity to structural details of this technique. For the two-dimensional infrared spectra, we observed that they generally reproduce the same pattern, while the inhomogeneous line width was slightly underestimated in the simulations. The shift correction applied to preproline units for the Skinner map was seen to give sensible results, while a similar correction to the side-chain units for the Jansen map underestimates the blue shift predicted by the Skinner map. This suggests that the amide-I vibration of primary amides is quite different from that of secondary and tertiary amides.

The main advantages of AIM are that the program allows generating Hamiltonian trajectories from multiple molecular dynamics packages, it automatically identifies the different units in proteins, and it is about twice as fast as the predecessor. AIM in its current form is already a very powerful tool for the simulations of infrared spectra of proteins. However, we foresee a number of improvements in the future. One of these would be the inclusion of the calculation of transition polarizabilities needed for the simulation of Raman and sum-frequency generation signals. Verified interfacing with more MD packages will also be a useful addition. Furthermore, it should be easy to extend the code with more mappings, not only for the amide-I region but also for other vibrational modes. The code is already designed so that users can add these mappings independently.

https://doi.org/10.1021/acs.jctc.2c00113
J. Chem. Theory Comput. 2022, 18, 3089−3098

5. DATA AND SOFTWARE AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c00113.

Description of zip file (PDF)

AIM_Examples.zip containing input for GROMACS and NAMD production run calculations and input files for AIM and NISE calculations (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

Kim E. van Adrichem − *Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands;* orcid.org/0000-0002-2260-9269; Email: k.e.van.adrichem@rug.nl

Thomas L. C. Jansen − *Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands;* orcid.org/0000-0001-6066-6080; Email: t.l.c.jansen@rug.nl

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c00113

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This publication is part of the project "Nanoscale Regulators of Photosynthesis" (with project number OCENW.GROOT.2019.086) of the research programme NWO Groot which is (partly) financed by the Dutch Research Council (NWO). We are grateful to Markus Meuwly and Maryam Salehi for their support with testing the code for CHARMM trajectories and to Christopher M. Cheatum and Lauren Lambach for their help with testing the code for AMBER trajectories. We are grateful to Tsjerk Wassenaar, Siewert-Jan Marrink, and Ana Cunha for helpful discussions. We thank the Center for Information Technology of the University of Groningen for its support and for providing access to the Peregrine high performance computing cluster. Part of this work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative.

■ REFERENCES

(1) Gray, W.; Sandberg, L.; Foster, J. Molecular Model for Elastin Structure and Function. *Nature* **1973**, *246*, 461−466.

(2) Karplus, M.; Kuriyan, J. Molecular Dynamics and Protein Function. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6679−6685.

(3) Dunker, A. K.; Brown, C. J.; Lawson, J. D.; Iakoucheva, L. M.; Obradović, Z. Intrinsic Disorder and Protein Function. *Biochem.* **2002**, *41*, 6573−6582.

(4) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. A Three-Dimensional Model of the Myoglobin Molecule Obtained by x-Ray Analysis. *Nature* **1958**, *181*, 662−666.

(5) Wüthrich, K. The Way to NMR Structures of Proteins. *Nat. Struct. Biol.* **2001**, *8*, 923−925.

(6) de la Cruz, M. J.; Hattne, J.; Shi, D.; Seidler, P.; Rodriguez, J.; Reyes, F. E.; Sawaya, M. R.; Cascio, D.; Weiss, S. C.; Kim, S. K.; Hinck, C. S.; Hinck, A. P.; Calero, G.; Eisenberg, D.; Gonen, T. Atomic-Resolution Structures from Fragmented Protein Crystals with the cryoEM Method MicroED. *Nat. Methods* **2017**, *14*, 399−402.

(7) Hamm, P.; Lim, M.; Hochstrasser, R. M. Structure of the Amide I Band of Peptides Measured by Femtosecond Nonlinear-Infrared Spectroscopy. *J. Phys. Chem. B* **1998**, *102*, 6123−6138.

(8) Barth, A. Infrared Spectroscopy of Proteins. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2007**, *1767*, 1073−1101.

(9) Woutersen, S.; Hamm, P. Time-Resolved Two-Dimensional Vibrational Spectroscopy of a Short Alpha-Helix in Water. *J. Chem. Phys.* **2001**, *115*, 7737−7743.

(10) Cheatum, C. M.; Tokmakoff, A.; Knoester, J. Signatures of Beta-Sheet Secondary Structures in Linear and Two-Dimensional Infrared Spectroscopy. *J. Chem. Phys.* **2004**, *120*, 8201−8215.

(11) Maekawa, H.; Toniolo, C.; Moretto, A.; Broxterman, Q. B.; Ge, N.-H. Different Spectral Signatures of Octapeptide 3^{10}- and Alpha-Helices Revealed by Two-Dimensional Infrared Spectroscopy. *J. Phys. Chem. B* **2006**, *110*, 5834−5837.

(12) Jansen, T. L. C. Computational Spectroscopy of Complex Systems. *J. Chem. Phys.* **2021**, *155*, 170901.

(13) Bodis, P.; Larsen, O. F. A.; Woutersen, S. Vibrational Relaxation of the Bending Mode of HDO in Liquid D_2O. *J. Phys. Chem. A* **2005**, *109*, 5303−5306.

(14) DeFlores, L. P.; Ganim, Z.; Nicodemus, R. A.; Tokmakoff, A. Amide I'-II' 2D IR Spectroscopy Provides Enhanced Protein Secondary Structural Sensitivity. *J. Am. Chem. Soc.* **2009**, *131*, 3385−3391.

(15) Chelius, K.; Wat, J. H.; Phadkule, A.; Reppert, M. Distinct Electrostatic Frequency Tuning Rates for Amide I and Amide I' Vibrations. *J. Chem. Phys.* **2021**, *155*, 195101.

(16) Baiz, C. R.; Błasiak, B.; Bredenbeck, J.; Cho, M.; Choi, J.-H.; Corcelli, S. A.; Dijkstra, A. G.; Feng, C.-J.; Garrett-Roe, S.; Ge, N.-H.; Hanson-Heine, M. W. D.; Hirst, J. D.; Jansen, T. L. C.; Kwac, K.; Kubarych, K. J.; Londergan, C. H.; Maekawa, H.; Reppert, M.; Saito, S.; Roy, S.; Skinner, J. L.; Stock, G.; Straub, J. E.; Thielges, M. C.; Tominaga, K.; Tokmakoff, A.; Torii, H.; Wang, L.; Webb, L. J.; Zanni, M. T. Vibrational Spectroscopic Map, Vibrational Spectroscopy, and Intermolecular Interaction. *Chem. Rev.* **2020**, *120*, 7152−7218.

(17) Jansen, T. L. C.; Saito, S.; Jeon, J.; Cho, M. Theory of Coherent Two-Dimensional Vibrational Spectroscopy. *J. Chem. Phys.* **2019**, *150*, 100901.

(18) Hamm, P.; Helbing, J.; Bredenbeck, J. Stretched versus Compressed Exponential Kinetics in Alpha-Helix Folding. *Chem. Phys.* **2006**, *323*, 54.

(19) Meuzelaar, H.; Marino, K. A.; Huerta-Viga, A.; Panman, M. R.; Smeenk, L. E. J.; Kettelarij, A. J.; van Maarseveen, J. H.; Timmerman, P.; Bolhuis, P. G.; Woutersen, S. Folding Dynamics of the Trp-Cage Miniprotein: Evidence for a Native-like Intermediate from Combined Time-Resolved Vibrational Spectroscopy and Molecular Dynamics Simulations. *J. Phys. Chem. B* **2013**, *117*, 11490−11510.

(20) Chung, H. S.; Ganim, Z.; Jones, K. C.; Tokmakoff, A. Transient 2D IR Spectroscopy of Ubiquitin Unfolding Dynamics. *P. Natl. Acad. Sci.* **2007**, *104*, 14237−14242.

(21) Smith, A. W.; Lessing, J.; Ganim, Z.; Peng, C. S.; Tokmakoff, A.; Roy, S.; Jansen, T. L. C.; Knoester, J. Melting of a Beta-Hairpin Peptide Using Isotope-Edited 2D IR Spectroscopy and Simulations. *J. Phys. Chem. B* **2010**, *114*, 10913−10924.

(22) Ganim, Z.; Jones, K. C.; Tokmakoff, A. Insulin Dimer Dissociation and Unfolding Revealed by Amide I Two-Dimensional Infrared Spectroscopy. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3579−3588.

(23) Salehi, S. M.; Koner, D.; Meuwly, M. Dynamics and Infrared Spectrocopy of Monomeric and Dimeric Wild Type and Mutant Insulin. *J. Phys. Chem. B* **2020**, *124*, 11882−11894.

(24) Manor, J.; Mukherjee, P.; Lin, Y.-S.; Leonov, H.; Skinner, J. L.; Zanni, M. T.; Arkin, I. T. Gating Mechanism of the Influenza a M2 Channel Revealed by 1D and 2D IR Spectroscopies. *Structure* **2009**, *17*, 247−254.

(25) Stevenson, P.; Götz, C.; Baiz, C. R.; Akerboom, J.; Tokmakoff, A.; Vaziri, A. Visualizing KcsA Conformational Changes upon Ion Binding by Infrared Spectroscopy and Atomistic Modeling. *J. Phys. Chem. B* **2015**, *119*, 5824−5831.

(26) Kopec, W.; Köpfer, D. A.; Vickery, O. N.; Bondarenko, A. S.; Jansen, T. L. C.; de Groot, B. L.; Zachariae, U. Direct Knock-on of Desolvated Ions Governs Strict Ion Selectivity in K+ Channels. *Nat. Chem.* **2018**, *10*, 813−820.
(27) Strong, S. E.; Hestand, N. J.; Kananenka, A. A.; Zanni, M. T.; Skinner, J. IR Spectroscopy Can Reveal the Mechanism of k Transport in Ion Channels. *Biophys. J.* **2020**, *118*, 254−261.
(28) Wang, L.; Middleton, C. T.; Singh, S.; Reddy, A. S.; Woys, A. M.; Strasfeld, D. B.; Marek, P.; Raleigh, D. P.; de Pablo, J. J.; Zanni, M. T.; Skinner, J. L. 2DIR Spectroscopy of Human Amylin Fibrils Reflects Stable Beta-Sheet Structure. *J. Am. Chem. Soc.* **2011**, *133*, 16062−16071.
(29) El Khoury, Y.; Le Breton, G.; Cunha, A. V.; Jansen, T. L. C.; van Wilderen, L. J. G. W.; Bredenbeck, J. Lessons from Combined Experimental and Theoretical Examination of the FTIR and 2D-IR Spectroelectrochemistry of the Amide I Region of Cytochrome *c*. *J. Chem. Phys.* **2021**, *154*, 124201.
(30) For AIM, see: https://github.com/Kimvana/AIM (accessed 2022-03-06).
(31) Burley, S. K.; Berman, H. M.; Bhikadiya, C.; Bi, C.; Chen, L.; Costanzo, L. D.; Christie, C.; Dalenberg, K.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ghosh, S.; Goodsell, D. S.; Green, R. K.; Guranović, V.; Guzenko, D.; Hudson, B. P.; Kalro, T.; Liang, Y.; Lowe, R.; Namkoong, H.; Peisach, E.; Periskova, I.; Prlić, A.; Randle, C.; Rose, A.; Rose, P.; Sala, R.; Sekharan, M.; Shao, C.; Tan, L.; Tao, Y.-P.; Valasatava, Y.; Voigt, M.; Westbrook, J.; Woo, J.; Yang, H.; Young, J.; Zhuravleva, M.; Zardecki, C. RCSB Protein Data Bank: Biological Macromolecular Structures Enabling Research and Education in Fundamental Biology, Biomedicine, Biotechnology and Energy. *Nucleic Acids Res.* **2019**, *47*, D464−D474.
(32) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.
(33) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141.
(34) Rzepiela, A. J.; Schäfer, L. V.; Goga, N.; Risselada, H. J.; de Vries, A. H.; Marrink, S. J. Reconstruction of Atomistic Details from Coarse Grained Structures. *J. Comput. Chem.* **2010**, *31*, 1333−1343.
(35) For NISE_2017, see: https://github.com/GHlacour/NISE_2017 (accessed 2022-02-02).
(36) Zhuang, W.; Abramavicius, D.; Venkatramani, R.; Jansen, T. L. C.; Voronine, D.; Robinson, B.; Hayashi, T.; Mukamel, S. *SPECTRON*. https://mukamel.ps.uci.edu/software.html (accessed 2022-02-02).
(37) Segatta, F.; Nenov, A.; Nascimento, D. R.; Govind, N.; Mukamel, S.; Garavelli, M. iSPECTRON: A Simulation Interface for Linear and Nonlinear Spectra with Ab-Initio Quantum Chemistry Software. *J. Comput. Chem.* **2021**, *42*, 644−659.
(38) For G_spec, see: Reppert, M.; Feng, C.-J. https://github.com/mreppert/g_spec (accessed 2022-02-02).
(39) Saxena, V.; Steendam, R.; Jansen, T. L. C. Distinguishing Islet Amyloid Polypeptide Fibril Structures with Infrared Isotope-Label Spectroscopy. *J. Chem. Phys.* **2022**, *156*, No. 055101.
(40) Reppert, M.; Tokmakoff, A. Electrostatic Frequency Shifts in Amide I Vibrational Spectra: Direct Parameterization against Experiment. *J. Chem. Phys.* **2013**, *138*, 134116.
(41) Wang, L.; Middleton, C. T.; Zanni, M. T.; Skinner, J. L. Development and Validation of Transferable Amide I Vibrational Frequency Maps for Peptides. *J. Phys. Chem. B* **2011**, *115*, 3713−3724.
(42) 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*, No. 044502.
(43) Ham, S.; Cho, M. Amide I Modes in the N-methylacetamide Dimer and Glycine Dipeptide Analog: Diagonal Force Constants. *J. Chem. Phys.* **2003**, *118*, 6915−6922.
(44) Watson, T. M.; Hirst, J. D. Theoretical Studies of the Amide I Vibrational Frequencies of [Leu]-Enkephalin. *Mol. Phys.* **2005**, *103*, 1531.
(45) Hamm, P.; Woutersen, S. Coupling of the Amide I Modes of the Glycine Dipeptide. *Bull. Chem. Soc. Jpn.* **2002**, *75*, 985.
(46) la Cour Jansen, T.; Dijkstra, A. G.; Watson, T. M.; Hirst, J. D.; Knoester, J. Modeling the Amide I Bands of Small Peptides. *J. Chem. Phys.* **2006**, *125*, No. 044312.
(47) Ganim, Z.; Tokmakoff, A. Spectral Signatures of Heterogeneous Protein Ensembles Revealed by MD Simulations of 2DIR Spectra. *Biophys. J.* **2006**, *91*, 2636.
(48) Bondarenko, A.; Jansen, T. L. C. Application of Two-Dimensional Infrared Spectroscopy to Benchmark Models for the Amide I Band of Proteins. *J. Chem. Phys.* **2015**, *142*, 212437.
(49) Cunha, A. V.; Bondarenko, A. S.; Jansen, T. L. C. Assessing Spectral Simulation Protocols for the Amide I Band of Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 3982−3992.
(50) Woys, A. M.; Almeida, A. M.; Wang, L.; Chiu, C.-C.; McGovern, M.; de Pablo, J. J.; Skinner, J. L.; Gellman, S. H.; Zanni, M. T. Parallel Beta-Sheet Vibrational Couplings Revealed by 2D IR Spectroscopy of an Isotopically Labeled Macrocycle: Quantitative Benchmark for the Interpretation of Amyloid and Protein Infrared Spectra. *J. Am. Chem. Soc.* **2012**, *134*, 19118−19128.
(51) Roy, S.; Lessing, J.; Meisl, G.; Ganim, Z.; Tokmakoff, A.; Knoester, J.; Jansen, T. L. C. Solvent and Conformation Dependence of Amide I Vibrations in Peptides and Proteins Containing Proline. *J. Chem. Phys.* **2011**, *135*, 234507.
(52) Cunha, A. V.; Salamatova, E.; Bloem, R.; Roeters, S. J.; Woutersen, S.; Pshenichnikov, M. S.; Jansen, T. L. C. Interplay between Hydrogen Bonding and Vibrational Coupling in Liquid *N*-Methylacetamide. *J. Phys. Chem. Lett.* **2017**, *8*, 2438−2444.
(53) DeFlores, L. P.; Tokmakoff, A. Water Penetration into Protein Secondary Structure Revealed by Hydrogen-Deuterium Exchange Two-Dimensional Infrared Spectroscopy. *J. Am. Chem. Soc.* **2006**, *128*, 16520−16521.
(54) Mayne, L. C.; Hudson, B. Resonance Raman Spectroscopy of N-methylamide: Overtones and Combinations of the C-N Stretch (Amide II') and Effect of Solvation on the C = O Stretch (Amide I) Intensity. *J. Phys. Chem.* **1991**, *95*, 2962.
(55) Choi, J.-H.; Cho, M. Amide I Vibrational Circular Dichroism of Dipeptide: Conformation Dependence and Fragment Analysis. *J. Chem. Phys.* **2004**, *120*, 4383−4392.
(56) For AmideImaps, see: https://github.com/GHlacour/AmideImaps (accessed 2022-02-02).
(57) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(58) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32*, 2319−2327.
(59) Gowers, R.; Linke, M.; Barnoud, J.; Reddy, T.; Melo, M.; Seyler, S.; Domański, J.; Dotson, D.; Buchoux, S.; Kenney, I.; Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proceedings of the 15th Python in Science Conference*; 2016; DOI: 10.25080/Majora-629e541a-00e.
(60) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(61) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU

and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, No. 044130.
(62) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package: Amber Biomolecular Simulation Package. *WIREs Comput. Mol. Sci.* **2013**, *3*, 198−210.
(63) Reppert, M.; Tokmakoff, A. Computational Amide I 2D IR Spectroscopy as a Probe of Protein Structure and Dynamics. *Annu. Rev. Phys. Chem.* **2016**, *67*, 359−386.
(64) For G_amide, see: Reppert, M. https://github.com/mreppert/g_amide (accessed 2022-02-02).
(65) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, 1987; DOI: 10.1093/oso/9780198803195.001.0001.
(66) Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R. The Geometry of the Reactive Site and of the Peptide Groups in Trypsin, Trypsinogen and Its Complexes with Inhibitors. *Acta Crystallogr. B Struct. Sci.* **1983**, *39*, 480−490.
(67) Nguyen, H.; Case, D. A.; Rose, A. S. NGLview−Interactive Molecular Graphics for Jupyter Notebooks. *Bioinformatics* **2018**, *34*, 1241−1242.
(68) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(69) Jorgensen, W. L.; Tirado-Rives, J. The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(70) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(71) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(72) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.
(73) Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980**, *45*, 1196−1199.
(74) Roy, S.; Jansen, T. L. C.; Knoester, J. Structural Classification of the Amide I Sites of a Beta-Hairpin with Isotope Label 2DIR Spectroscopy. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9347.
(75) Torii, H.; Tasumi, M. Model Calculations on the amide-I Infrared Bands of Globular Proteins. *J. Chem. Phys.* **1992**, *96*, 3379.
(76) 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.
(77) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71−73.
(78) DeCamp, M. F.; DeFlores, L.; McCracken, J. M.; Tokmakoff, A.; Kwac, K.; Cho, M. Amide I Vibrational Dynamics of *N*-Methylacetamide in Polar Solvents: The Role of Electrostatic Interactions. *J. Phys. Chem. B* **2005**, *109*, 11016−11026.
(79) Jansen, T. L. C.; Knoester, J. Nonadiabatic Effects in the Two-Dimensional Infrared Spectra of Peptides: Alanine Dipeptide. *J. Phys. Chem. B* **2006**, *110*, 22910−22916.
(80) Baiz, C. R.; Peng, C. S.; Reppert, M. E.; Jones, K. C.; Tokmakoff, A. Coherent Two-Dimensional Infrared Spectroscopy: Quantitative Analysis of Protein Secondary Structure in Solution. *Analyst* **2012**, *137*, 1793.
(81) Hayashi, T.; Mukamel, S. Vibrational-Exciton Couplings for the Amide I, II, III, and A Modes of Peptides. *J. Phys. Chem. B* **2007**, *111*, 11032−11046.
(82) Pancoska, P.; Yasui, S. C.; Keiderling, T. A. Statistical Analyses of the Vibrational Circular Dichroism of Selected Proteins and Relationship to Secondary Structures. *Biochem.* **1991**, *30*, 5089−5103.
(83) Dijkstra, A. G.; Jansen, T. L. C.; Knoester, J. Modeling the Vibrational Dynamics and Nonlinear Infrared Spectra of Coupled Amide I and II Modes in Peptides. *J. Phys. Chem. B* **2011**, *115*, 5392−5401.
(84) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; del Río, J. F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E. Array Programming with NumPy. *Nature* **2020**, *585*, 357−362.
(85) Lam, S. K.; Pitrou, A.; Seibert, S. Numba: A LLVM-based Python JIT Compiler. *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC - LLVM '15*; Austin, Texas, 2015; pp 1−6, DOI: 10.1145/2833157.2833162.