THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

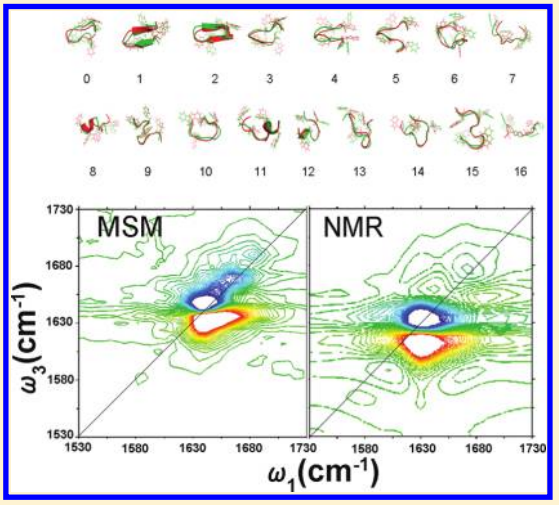
Simulating the T-Jump-Triggered Unfolding Dynamics of trpzip2 Peptide and Its Time-Resolved IR and Two-Dimensional IR Signals Using the Markov State Model Approach

Wei Zhuang,*[,†] Raymond Z. Cui,[‡] Daniel-Adriano Silva,[‡] and Xuhui Huang*[,‡]

[†]State Key Lab of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Dalian, Liaoning, China, 116023

[‡]Department of Chemistry, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

ABSTRACT: We proposed a computational protocol of simulating the T-jump peptide unfolding experiments and the related transient IR and two-dimensional IR (2DIR) spectra based on the Markov state model (MSM) and nonlinear exciton propagation (NEP) methods. MSMs partition the conformation space into a set of nonoverlapping metastable states, and we can calculate spectra signal for each of these states using the NEP method. Thus the overall spectroscopic observable for a given system is simply the sum of spectra of different metastable states weighted by their populations. We show that results from MSMs constructed from a large number of simulations have a much better agreement with the equilibrium experimental 2DIR spectra compared to that generated from straightforward MD simulations starting from the folded state. This indicates that a sufficient sampling of important relevant conformational states is critical for calculating the accurate spectroscopic observables. MSMs are also capable of simulating the unfolding relaxation dynamics upon the temperature jump. The agreement of the simulation using MSMs and NEP with the experiment not only provides a justification for our protocol, but also provides the physical insight of the underlying spectroscopic observables. The protocol we developed has the potential to be extended to simulate a wide range of fast triggering plus optical detection experiments for biomolecules.



1. INTRODUCTION

Protein folding is an important problem that is attracting scientists from a wide range of disciplines. However, after decades of studies by both experimental and theoretical methods,[1−13] we are still on the way of achieving a clear understanding of the folding mechanism even for the small, single domain peptides. One of the major challenges comes from the gap between the experimental and the theoretical studies. There are two main issues contributing to this gap. First, a time scale gap exists between the folding times of the peptides commonly chosen by the experiments (above microseconds[14−18]) and computing capability of the MD simulation method. Even with the recent encouraging progress in bridging this gap through large-scale simulations using graphic processors[19] or through the hardware and software specifically designed for MD simulations,[20] it is not an easy task for the MD simulations to achieve good statistics of trajectories at microseconds or longer. Second, even with sufficient sampling, to carry out a direct comparison between the MD simulations and the experimental observables is a difficult task. In order to bridge the gap between the theoretical and experimental investigations and achieve a better understanding of protein folding, an algorithm capable of simulating long time scale dynamics as well as calculating time-resolved spectroscopic signals based on the simulation result needs to be developed.

The introduction of pulsed laser techniques to trigger unfolding processes in nanoseconds has recently been widely used in protein-folding research.[21] Among this family of methods, T-jump,[15,17,18,21−25] which uses an intense laser pulse to produce a rapid temperature jump, is the most generally used since it can be used to trigger any process that is driven by a significant enthalpy change. In a T-jump experiment, the solution is heated by the pulsed laser, which emits in the near-infrared, and the temperature of the solution is caused to rise by a small amount in a very short time. This prepares the system with an unstable distribution of structures, and this allows the study of the relaxation to the equilibrium at the higher temperature after the T-jump at a typical time scale of microseconds or longer.

The fast dynamics triggered by T-jump are most commonly detected using optical spectroscopic tools.[15,17,18,21−25] Among

Special Issue: Shaul Mukamel Festschrift

Received: October 6, 2010
Revised: January 30, 2011
Published: March 09, 2011

which infrared (IR) spectroscopy has been shown to be a powerful and convenient option. The localized nature of the vibrational motion makes them a useful probe for secondary structures of proteins.[26,27] Various structural motifs such as α helices and β sheets can often be clearly distinguished by changes in the linear infrared absorption.[26,27] The amide-I infrared absorption band, which originates from the stretching motion of the C═O peptide bond coupled to in-phase N−H bending and C−H stretching,[26,28−30] provides a useful indicator of secondary structure changes because of its sensitivity to hydrogen bonding, dipole−dipole interactions, and geometry of the peptide backbone.

Over the past decade, two-dimensional infrared (2DIR) spectroscopy carried out with 20−100 fs laser pulses has emerged as a powerful tool in the investigation of ultrafast dynamics in the condense phase.[31−49] In 2D optical experiments, three ultrashort laser pulses interact with the molecular system to generate a coherent nonlinear signal, which depends parametrically on three consecutive time delays t_1, t_2, and t_3. 2D correlation plots are obtained by a double Fourier transform of the signal with respect to t_1 and t_3, holding t_2 fixed. By spreading the congested transitions of linear absorption in two dimensions, these techniques significantly enhance the spectral resolution. The cross-peaks (off diagonal peaks in 2D frequency correlation plots) carry signatures of intermode couplings, and their spectral line shapes probe solvent and conformational fluctuations. There have been considerable efforts in using 2DIR spectroscopy to study protein folding, misfolding, and aggregation.[37,41,49−54]

In this study, we demonstrate a protocol of simulating the T-jump (from T_1 to T_2) triggered peptide unfolding dynamics using Markov state models (MSMs) and calculating the related time-resolved Fourier transform IR (FTIR) and 2DIR signals using the direct nonlinear exciton propagation (NEP) method.

MSMs[55−59] partition conformational space into a number of nonoverlapping states, called macrostates or metastable states, such that intrastate transitions are fast, but interstate transitions are slow. Such separation of time scales ensures that the model is Markovian, in that the probability of being in a given state at time $t + \Delta t$ depends only on the state at time t. This property allows MSMs to be built from many short simulations (tens to hundreds of nanoseconds), which may only explore transitions between neighboring states, and then propagated to give global long time scale dynamics, such as processes occurring on microsecond or even longer time scales.

The method used for spectra calculation was direct NEP[60] developed by Mukamel et al. In recent years, several methods were proposed for the calculation of the optical response functions by the numerical integration of the Schrodinger equation (NISE).[61−65] NEP is similar to NISE, but carries out the integration on nonlinear exciton equations (NEEs) instead. Falvo and Mukamel demonstrated[60] that NEP can improve the computational efficiency by avoiding the calculation of several self-canceling Liouville pathways.

The complete protocol proposed in this manuscript can be summarized as follows. First, we construct MSMs to decompose conformation space into a set of metastable states at the temperature after the T-jump (T_2) from many molecular dynamics simulations. Next, we simulate the relaxation dynamics after the T-jump by calculating the evolution of populations of these metastable states using MSMs. Finally, we compute the FTIR or 2DIR spectra at different time points by the weighted sum of the spectra of individual metastable states. In order to demonstrate the power of our method, we use the trpzip2 hairpin peptide as an example.

trpzip2 (tryptophan zipper 2), which is designed by Cochran and co-workers[66] is one of the most stable β hairpins. trpzip2 has been investigated intensely as a benchmark for 2DIR's capability of revealing a peptide system's structure and dynamics. In 2006, Wang and Hochstrasser examined the unlabeled trpzip2 and two ^{13}C-isotopomers.[67] The results suggest a larger structure fluctuation in the terminal region than in the turn region as a result of the side chain effect and solvent−peptide interaction. A follow-up theoretical study in 2008 by Hochstrasser and Mukamel[68] suggested that the observed difference of the ^{13}C-shifted band, including its peak position and frequency distributions for different isotopomers, is likely to be due to the difference in the local environment of the solvated peptide. In 2007, Tokmakoff et al. carried out spectral simulations and isotope labeling to describe the 2DIR spectroscopy of trpzip2, together with several other hairpin peptides, in the amide I spectral region.[69] The authors compared several methods for calculating the site energies used in excitonic treatments of the amide I band. A theoretical study in 2008 by Jansen and Knoester[49] calculated the spectrum of the trpzip2 including vibrational population transfer induced by a fluctuating solvent. This study proposed the possibility of using vibrational population transfer to enhance the structural markers for protein motifs that occur in 2DIR spectroscopy. In 2010, a joint theoretical and experimental study by Tokmakoff, Jansen, and Knoester[54] used isotope-edited 2DIR to characterize the conformational heterogeneity of trpzip2 across its thermal unfolding transition. The experimental data were interpreted on the basis of structure-based spectroscopic modeling of conformers obtained from a MSM generated by Chodera et al.[70] based on simulation data sets from Pitera et al.[71]

This paper is organized as follows. In the Methodology section, we will discuss the protocol in more detail. The simulation result will be demonstrated in the Results section followed by further discussions.

2. METHODOLOGY

2.1. Simulation Data Sets. Our simulations were generated using the adaptive seeding method (ASM).[72] First, 4.5 ns replica exchange molecular dynamics (REMD)[73,74] simulations were performed with an exchange interval of 2 ps and all replicas starting from the folded protein structure. A 50-element, roughly exponentially distributed temperature list covering a range from 285 to 592 K was used for REMD. Second, conformations generated from REMD simulations were clustered into 20 states using the K-center clustering algorithm implemented in the MSMBuilder software package.[75] Next, the center of each cluster was used as the starting points for new constant temperature simulations (i.e., shooting simulations) at 300 and 350 K. The number of simulations launched from different clusters was proportional to their populations from REMD simulations around the shooting temperature. At 300 and 350 K, we performed 973 and 830 70-ns shooting simulations, respectively, with conformations saved every 100 ps. Finally, the last 50 ns of each of the shooting simulations were used to construct MSMs at 300 and 350 K, respectively. See the Constructing MSMs section below for the details on MSM construction.

The simulations were performed using the AMBER03 force field.[76,77] The folded protein structure was taken from the NMR

structure of Tryptophan Zipper (PDB ID: 1LE1).[78] The protein molecule was solvated in a water box with 2923 TIP3P[79] waters and two Cl^- ions. The simulation system was minimized using a steepest descent algorithm, followed by a 100 ps MD simulation applying a position restraint potential to the protein heavy atoms. All NVT simulations were coupled to a Nose−Hoover thermostat with a coupling constant of 0.05 ps^{-1}.[80] A cutoff of 9 Å was used for both van der Waals (VdW) and short-range electrostatic interactions. Long-range electrostatic interactions were treated with the particle-mesh Ewald (PME) method.[81] Nonbonded pair-lists were updated every 10 steps with an integration step size of 2 fs in all simulations. All bonds were constrained using the LINCS algorithm.[82] The REMD simulations were performed using Gromacs 4 on a local computer cluster,[83] while all the shooting simulations were performed using a version of the Gromacs modified for the Folding@home infrastructure.[84]

2.2. Constructing MSMs. In an MSM, the time evolution of a vector representing the population of each state may be calculated by repeatedly left-multiplying by the transition probability matrix as

$$P(n\Delta t) = [T(\Delta t)]^n P(0) \quad (1)$$

where $P(n\Delta t)$ is a vector of state populations at time $n\Delta t$, T is the column stochastic transition probability matrix, and Δt, the time interval to count the transitions, is the lag time, also called τ. The first left eigenvector of the transition matrix T corresponds to the equilibrium distribution.[56]

To construct MSMs, we first group each of the two sets of conformations (~486 000 at 300 K and ~415 000 at 350 K) into 2000 clusters or microstates using a K-centers clustering algorithm.[75] The structure similarity or distance between a pair of conformations is determined by their pairwise heavy atom root mean square deviation (rmsd). At this stage, we can construct 2000-state MSMs at 300 and 350 K. However, with a 2000-state MSM, it is difficult to illustrate the mechanisms for folding or unfolding. In addition, it also requires an enormous amount of effort to calculate the IR spectroscopic observables for all these 2000 states. Thus, we also lump microstates together by their kinetic connectivity to generate macrostate models.

We use the superlevel-set hierarchical clustering (SHC) algorithm[85] to lump microstates together to construct macrostate MSMs. The key insight of the SHC algorithm is to generate a set of super levels covering different density regions of phase space, then cluster each super level separately, and finally recombine this information into a single MSM. We constructed separated MSMs at 300 and 350 K. The details are as follows:

(1) We split 2000 microstates into n density levels ($L = \{L_1, ..., L_n\}$), each of which contains about the same number of the conformations. The K-centers algorithm gives clusters with approximately equal radius (the largest distance between any conformation in the cluster and the cluster center).[85] In this system, the microstates have an average radius of about 4 Å. As a result, there is a correlation between the population of each microstate and its density, allowing us to define density levels. Then, from this density level set, we construct the super density level set by defining $S_i = L_1 \cup L_2 ... \cup L_{i-1} \cup L_i$. At 300 K, L = {0.8, 0.85, 0.9, 0.95, 0.99}. At 350 K, L={0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 0.99}.

(2) We perform spectral clustering[86,87] at different super density levels (S_i) to group kinetically related microstates.

(3) Now, we can build a graph representing the connectivity of the states across super density levels. We can then generate gradient flows along the edges of the graph from low to high density levels. We assign each attraction node (or attractive basin) where the gradient flow ends to a new metastable state.

(4) Finally, we assign every microstate not belonging to an attraction node to the metastable state to which it has the largest transition probability. Thus we have a complete state decomposition for an MSM. We generate 17-state and 13-state MSMs at 300 and 350 K, respectively.

Implied time scales (τ_k) from MSMs are defined as below:

$$\tau_k = -\frac{\tau}{\ln \mu_k(\tau)} \quad (2)$$

where μ_k is an eigenvalue of the transition matrix with the lag time τ. Each implied time scale describes an aggregate transition between subsets of states in a MSM. If the model is Markovian and eq 1 holds, the exponentiation of T should be identical to an MSM constructed with a longer lag time, and the implied time scales will be independent of the lag time. Thus, the plateau of the implied time scale plots indicates that the model is Markovian.[58]

Implied time scale plots are shown in Figure 1 for four models: (a) 2000-state MSM at 300K, (b) 17-state MSM at 300K, (c) 2000-state MSM at 350K, and (d) 13-state MSM at 350K. All these implied time scales have leveled off at 10 ns, indicating that the models are Makovian. Thus we select 10 ns at the lag time to construct MSMs. The slowest implied time scale at 300 K is around a few hundred nanoseconds to 1 μs, which is consistent with the experimental observations.[88] At 300 K, the faster implied time scales of the 17-state MSM compared to the 2000-state model are due to the lumping effects and have been discussed before.[19,85]

2.3. Simulating Relaxation Dynamics after the Temperature Jump. We simulate the relaxation dynamics after the temperature jump by calculating the evolution of populations of metastable states using MSMs. If the temperature jump is very fast, the initial population of different metastable states will still be the equilibrium distribution at the temperature before the temperature jump (P_{T1}). In order to calculate these initial populations, all the conformations sampled at T_1 were assigned to the state decomposition from the MSM model constructed at T_2. The detailed procedure is as follows: First, a certain conformation at 300 K was first assigned to one of the 2000 microstates at 350 K, where the distance between this conformation and the microstate center was minimized. Second, this conformation was further assigned to a macrostate based on the mapping between microstates and macrostates defined by the 13-state MSM at 350 K.

This set of populations will then relax with time, and finally reach the equilibrium populations at the temperature after the T-jump (P_{T2}).

Our procedure to simulate the T-jump IR experiment is as follows. First, we simulate the state population relaxation kinetics by propagating the transition probability matrix by following eq 1 with $P(0) = P_{T1}$. Next, we calculate the one-dimensional (1D) and 2D IR spectra signals for each of the metastable states from the structure ensemble obtained from MSM. Finally, at a certain time point, we performed a weighted sum of these spectra signals of different states based on their populations to obtain an overall signal.

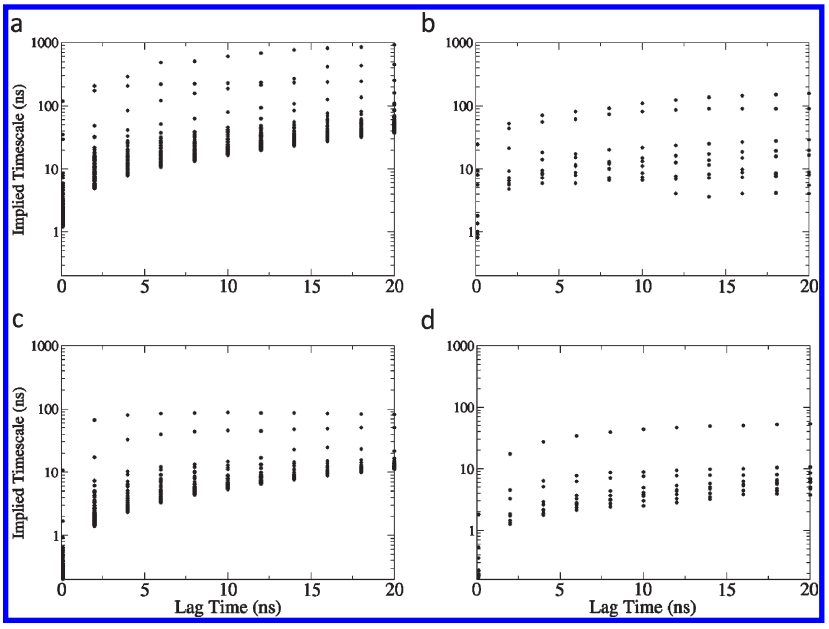


Figure 1. Implied time scales (ns) as a function of lag time (ns) are plotted for (a) a MSM with 2000 microstates at 300 K. Only the slowest 100 time scales are shown in the figure. (b) A macrostate MSM with 17 states at 300 K. (c) An MSM with 2000 microstates at 350 K. (d) A macrostate MSM with 17 states at 350 K.

2.4. Simulating the 1D and 2D IR Spectra of the Equilibrium State and the Unfolding Event. The FTIR and 2DIR spectra were simulated using the SPECTRON package[35] and NEP method.[60] For each metastable state, 100 randomly sampled configurations are generated as initial structures. Then, from each of these initial structures, a 10 ps short trajectory with 2 fs time step was created using the same parameters as described in section 2.1. The fluctuating vibrational Hamiltonian for the amide I band were then generated for each trajectory and used to integrate the NEE equation. A direct sum of the contributions from all the 100 configurations is then carried out to generate the final signals.

The construction of a fluctuating Hamiltonian for the primary vibrations and their coupling with a bath is the first step in the simulation of vibrational line shapes. We used the vibrational-exciton Hamiltonian

$$\hat{H} = \hat{H}_S + \hat{H}_F \tag{3}$$

where

$$\hat{H}_S = \sum_m \varepsilon_m \hat{B}_m^+ \hat{B}_m + \sum_{mn}^{m\neq n} J_{mn} \hat{B}_m^+ \hat{B}_m - \frac{1}{2}\sum_m \Delta_m \hat{B}_m^+ \hat{B}_m^+ \hat{B}_m \hat{B}_m \tag{4}$$

is the system Hamiltonian and $\hat{H}_F$ is the interaction with the optical field, $E(t)$

$$\hat{H}_F = -E(t)\sum_m \mu_m (\hat{B}_m^+ + \hat{B}_m) \tag{5}$$

$\hat{B}_m^+(\hat{B}_m)$ is the creation (annihilation) operator for the mth amide I mode, localized within the amide unit (O═C−N−H), with frequency ε_m, anharmonicity Δ_m, and transition dipole moment μ_m, These operators satisfy the Bose commutation relations $[\hat{B}_m, \hat{B}_n^+] = \delta_{mn}$, and J_{mn} are the harmonic intermode couplings. All parameters of $\hat{H}_S$ fluctuate due to conformational changes of the backbone, the solvent, and side-chain dynamics.

The fluctuating vibrational Hamiltonian was constructed for every snapshot along these trajectories using an approach based on the stark effect. Several models[35,47,89−91] have been used in the past to simulate the vibrational parameters and their corresponding fluctuations. We have used the CHO4[92] electrostatic map, which parametrized the amide-I vibration by identifying the electrostatic potential at four coordinates corresponding to the atoms C, O, N, and H of the amide bond, to simulate the fluctuations of the fundamental frequency ε_m of the amide-I modes of the trpzip2. The amide-I transition dipole μ_m and the intraband coupling J_{mn} are used as defined by the Torii and Tasumi transition-dipole coupling (TDC) model.[29] The coupling J_{mn} between nearest-neighbor amide-I modes is given by Cho's dihedral angles map.[92] The amide-I mode anharmonicity Δ_m is fixed to the measured value of -16 cm^{-1}.[32]

For comparison, we generated another set of folded configurations by an MD simulation from the NMR structure (1LEH.pdb), and the structure was equilibrated for 1 ns. Two hundred different configurations were generated by MD simulation, with 100 ps between each configuration, and used for the signal calculation.

3. RESULTS

3.1. Equilibrium State Distribution and the Spectra at 300 K. The equilibrium population distribution for the macrostates calculated from the 17-state MSM at 300 K is shown in Figure2a. The representative structures for all the states with population are shown in Figure2b. The macrostate 2, with a hairpin configuration, has the highest population (∼83%). Furthermore, distribution of rmsd to the folded structure (PDB ID: 1ELH) for conformations in this state have a narrow distribution center at 2.5 Å, indicating this state indeed contains structures close to the folded hairpin (see Figure 3). On the other hand, for the unfolded macrostate 8, the rmsd distribution centers at 7.5 Å (see Figure 3).

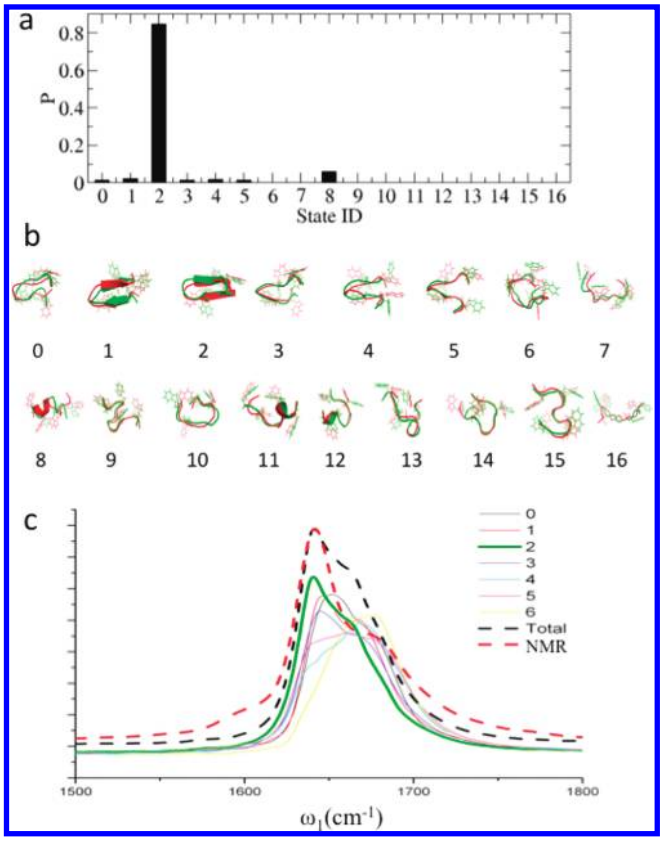


Figure 2. (a) Equilibrium populations for macrostates at 300 K computed from a 17-state MSM are displayed. (b) Representative structures from macrostates are shown. (c) The overall 300 K equilibrium linear absorption signal (black dash) and the contributions (solid lines) from all the macrostates with population >1%, simulated using MSM data. The overall signal is scaled by 1.2 for a better presentation. The green solid line gives the contribution from the most significant macro state (#2). The red dash line gives the simulated absorption signal based on the NMR structure, and it is scaled to have the same maximum as the MSM overall signal for a better comparison.

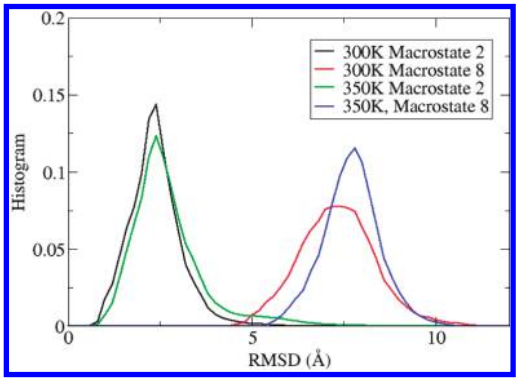


Figure 3. Histograms of the rmsd to the folded hairpin structures for conformations in four macrostates are plotted: macrostate 2 at 300 K in black, macrostate 8 at 300 K in red, macrostate 2 at 350 K in green, and macrostate 8 at 350 K in blue. The bin size is 0.2 Å.

The results for the simulated linear absorption signal (dashed black curve) at 300 K as well as the contributions from macrostates with population larger than 1% (solid colored curves) are plotted in Figure2c, the overall signal was rescaled to have a better demonstration. The overall signal shows two main transitions: the low-frequency, stronger transition peaked at ~1640 cm^{-1} and the high-frequency, weaker transition peaked at ~1670 cm^{-1}, which is consistent with the experimental result.[67,69,93] The overall signal resembles a typical feature of a folded hairpin due to the dominating population of folded state #2 at 300 K. On the other hand, the contribution from a typical unfolded state #8 has a single peak at ~1675 cm^{-1}. As a comparison, we show linear absorption signals calculated from the MD trajectories starting from the NMR structure only (dashed red curve), which shows a similar feature as in the MSM with a lower intensity for the high-frequency peak.

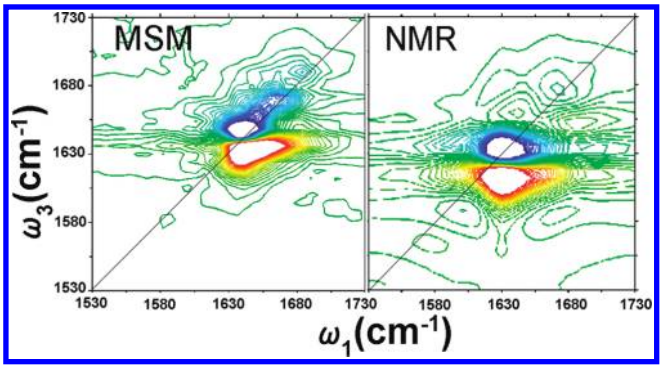


Figure 4. (left) The simulated absorptive 2DIR signal using MSM. (right) The same signal calculated on the basis of the NMR structure.

The simulated absorptive (KI+KII) 2DIR signals at 300 K using MSM (left) and NMR structure (right) are presented in Figure4. In the MSM result, the negative lobe is elongated along the diagonal line, the positive lobe is also elongated, with a small angle to the diagonal line. Both peaks have an asymmetric feature with the antidiagonal line width broadening at the red end. Negative features arise from the 0−1 vibrational transitions, whereas positive features arise from 1 to 2 transitons. These features agreed with what was observed in the experiments.[54,67] As the comparison, the NMR structure-based result (Figure 4, right panel) has the symmetric peaks elongated along the parallel direction.

3.2. Equilibrium States Distribution at 350 K and the Spectra. The macrostate equilibrium populations as well as representative structures from different states at 350K are shown in Figure5. The macro state 2 corresponds to the folded state with an average rmsd to the NMR structure at around 2.5 Å. This state is still the highest populated state at 350 K, but its population (~39%) is much smaller than that of the folded state at 300 K (~83%). This indicates that the temperature jump from 300K to 350K has induced significantly change of the folded state population from our simulations.

The simulated linear absorption signal of the system at 350 K is presented in Figure5c, together with the contributions from all the significant states. The linear IR shows one main transition centered at ~1660 cm^{-1}, which is consistent with the experimental results.[54] The contribution from state 2 has the maximum at 1640 cm^{-1} while almost all other contributions have the maxima at 1675 cm^{-1}.

The simulated 2DIR signal at 350 K is shown in Figure9. Compared with the 2DIR signal at 300 K, the spectral features broaden along the antidiagonal axis, and the spectral shifts along the diagonal axis resemble what was observed in the absorption. These features agreed with what was observed in the experiments[54]

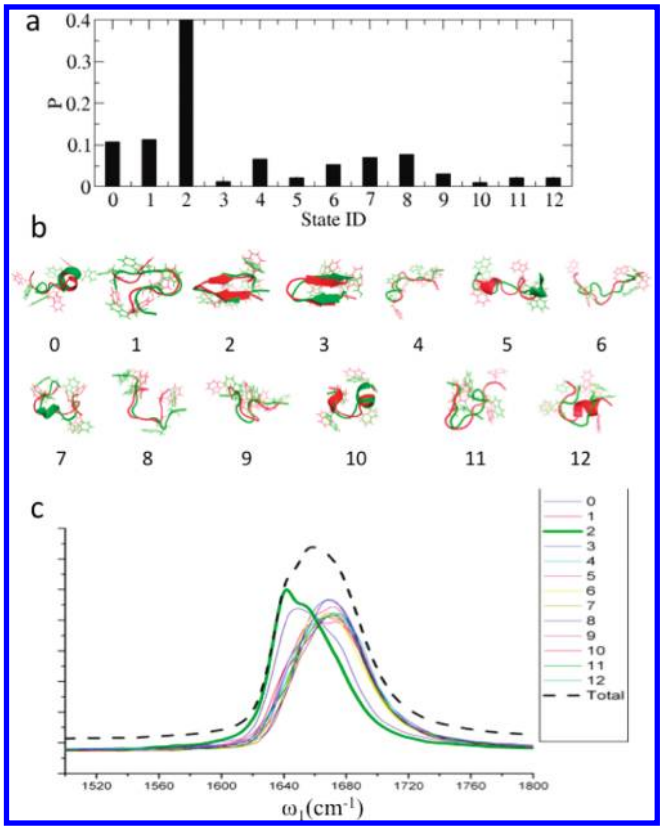


Figure 5. (a) Equilibrium populations for macrostates at 350 K computed from a 13-state MSM are displayed. (b) Representative structures from macrostates are shown. (c) The overall 350 K equilibrium linear absorption signal (black dash) and the contributions (solid lines) from all the macrostates, simulated using MSM data. The overall signal is scaled by 1.2 for a better presentation.

3.3. T-Jump Triggered Unfolding Dynamics and the Time-Resolved Spectra. We use MSMs to simulate the T-jump triggered unfolding dynamics. Upon the temperature jump, the system's temperature increases from T_1 to T_2. Thus the populations of metastable states will relax to the equilibrium distribution at T_2 starting from a set of initial distributions. In this study, we assume the time scale of the T-jump process is much faster than the time scale of the relaxation kinetics of peptide unfolding. Thus, we select the equilibrium populations at T_1 calculated from our simulations at T_1 as the initial populations. Figure 6 displays both the initial populations and final equilibrium populations for the relaxation process after the T-jump. We notice that the population of the folded state (macrostate 2) is much larger in the initial state than its equilibrium population at 350 K. With the selection of the initial populations, we can propagate eq 1 to compute the relaxation curves for populations of different macrostates. As shown in Figure 7, macrostate populations relax to the equilibrium populations within about 200 ns.

Next, we calculate the infrared spectra signals for each of the metastable states from the structure ensemble obtained from MSM. We then perform a weighted sum of these spectra signals of different states based on their populations at a certain time point to obtain an overall signal.

The simulated transient FTIR signals at different time points after the T-jump (upper) and their difference signals compared with the signal before the T-jump (lower) are shown in Figure 8. After the T-jump, the lower frequency transitions at 1640 cm^{-1} are getting weaker while the transitions at 1660 cm^{-1} are getting stronger, which indicates that the content of the hairpin is decreasing and the random coil is increasing, consistent with what we observed in the structure simulation. After 120 ns, the signal almost stops changing, which implies that the relaxation process finishes by this time.

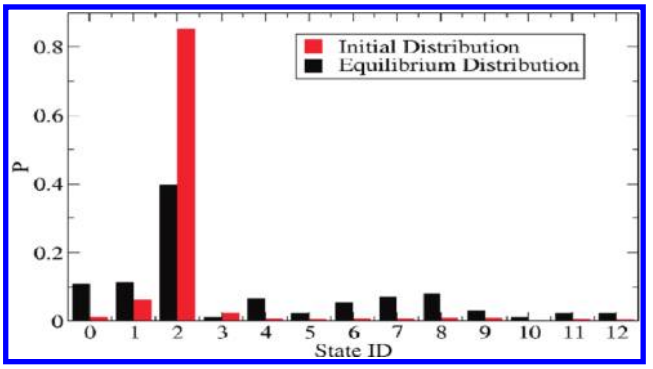


Figure 6. Initial and equilibrium populations for the relaxation process after the temperature jump from 300 to 350 K are shown in red and black, respectively.

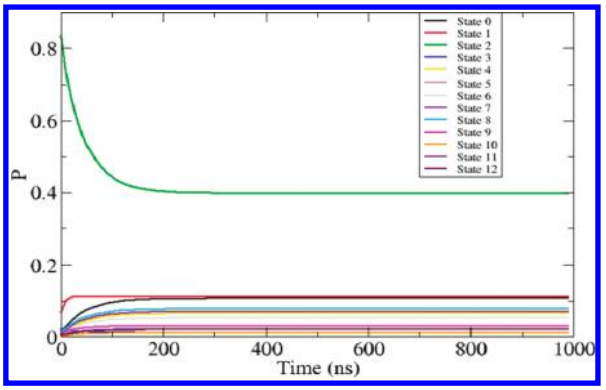


Figure 7. Macrostate populations after the temperature jump from 300 to 350 K as a function of time are shown. Population relaxations are calculated by propagating the MSM at 350 K.

The simulated transient 2DIR signals at different time points after the T-jump (upper) are shown in Figure 9. Along the diagonal line, the intensity is reduced at the lower frequency, while it increases at the higher frequency, which is consistent with what we observed in the FTIR. In order to have a clearer observation of how the signals change with time, we plotted the signals along the ω_3 with ω_1 = 1642 cm^{-1}, ω_1 = 1660 cm^{-1}, and ω_1 = 1684 cm^{-1} in Figure 10. After the T-jump, the intensities of the signals at ω_1 = 1642 cm^{-1} and ω_1 = 1660 cm^{-1} are decreasing, while the intensity of the signal at ω_1 = 1684 cm^{-1} is increasing. This is consistent with the fact that the population shifts from the folded state (macrostate #2) to other unfolded states.

4. DISCUSSION

Our simulation results show that the trpzip2 peptide mostly adopts a hairpin-like conformation at 300 K. The largest macrostate from our MSM with a population of more than 80% is a hairpin. However, other non-native states also exist in the structure ensemble. When the temperature was increased to 350 K, a substantial fraction of hairpin conformations unfolded.

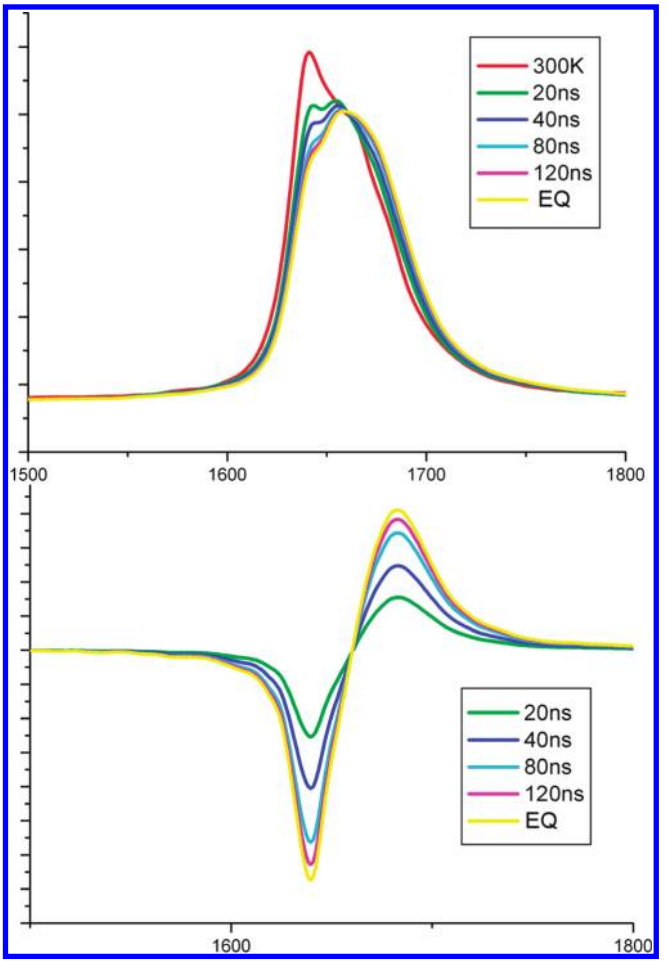


Figure 8. Simulated transient FTIR spectra at different moments after the T-jump (upper) and their difference spectra with the signal before the T-jump at 300 K (lower).

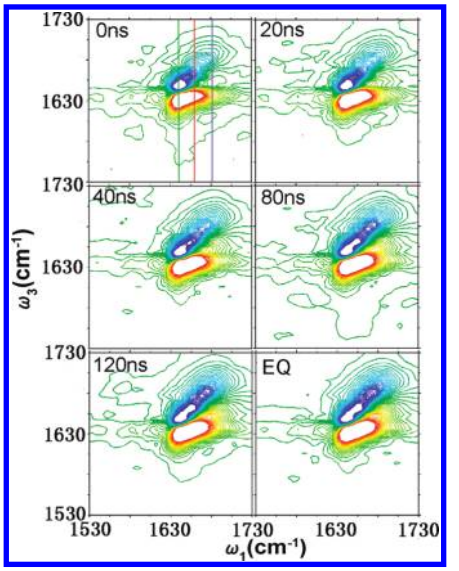


Figure 9. Simulated transient 2DIR at the different moments after the T-jump. The lines in the upper left panel present where we retrieved the slices of the signals (demonstrated in Figure10).

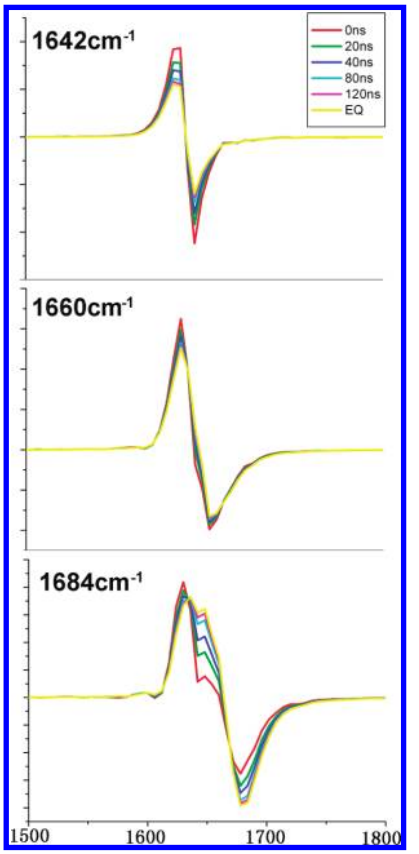


Figure 10. The slices of simulated Transient 2DIR at ω_1 = 1642 cm^{-1} (upper), ω_1 = 1660 cm^{-1}(middle) and ω_1 = 1684 cm^{-1}(lower).

For example, in our simulations the highest populated macrostate at 350 K is still a hairpin state; however, its population decreased from ~83% to ~39% compared to that at 300 K. Thus, the temperature jump from 300 to 350 K in simulations is sufficient to observe partial unfolding of the peptide.

The simulated 2DIR signals for the equilibrium at 300 K using MD and MSM demonstrate that MSM can provide a much better sampling compared to straightforward MD simulations from folded structures. For flexible polymers such as polypeptides, straightforward MD simulations are likely to be trapped in local free energy minimums, even if they reach the length of hundreds of nanoseconds. In the current study, we used the ASM[72] method to enhance our conformational sampling. For instance, we first used REMD simulations to explore the free energy landscape since high temperature will help in escaping the kinetics traps. Then, we shoot a large number of simulations at the temperature of interest, and construct MSMs based on these shooting simulations. The simulated linear absorption IR signals using these two methods both show the feature of a lower frequency strong transition and a higher frequency weak transition, consistent with the experiments. Although the intensities and the shape of the higher frequency peak are different, it is difficult to analyze which protocol gives a better interpretation due to the lack of the phenomenal differences. In the 2DIR, however, the MSM simulation has the asymmetrical feature elongated along the diagonal line, which agrees nicely with the experiment, while the MD generated signal has a symmetrical feature with no diagonal elongation. The straightforward MD's failure to reproduce the diagonal elongation is due to the underestimation of the structural inhomogeneity, which is known to be the cause of the

diagonal elongation (inhomogeneous broadening) in the 2D spectra. The observation above thus indicates that MSM creates a better sampling of the structure distribution.

In the transient FTIR signals shown in Figure9a, the intensity of the 1640 cm^{-1} peak decreases, while that of the 1675 cm^{-1} peak increases. This can be seen more clearly by making the difference signal with respect to the signal before the T-jump. An increasing pair with a negative lobe (1640 cm^{-1}) and a positive lobe (1675 cm^{-1}) is observed. By 120 ns, the feature has almost finished changing. According to Figure4b, the contribution to the FTIR signal can be divided into two categories: the contribution from state 2, which has a maximum at 1640 cm^{-1}, while almost all other contributions have the maximum at 1675 cm^{-1}. Thus the intensities of both the 1640 cm^{-1} and 1675 cm^{-1} peaks give a good probe for the unfolding dynamics. This is consistent with our simulated T-jump process using MSMs. Upon the temperature jump, the population of macrostate 2, which corresponds to the folded state, has a significant decrease. On the other hand, populations of all other non-native states increase as shown in Figure 6.

In a T-jump experiment, a temperature jump of 10−20 K is usually adopted. Here we carried out a 50 K jump in our simulation studies in order to induce more significant structural changes. In addition, it has been noted that temperature dependence of experimental and simulation observations may not be consistent in many cases such as melting temperature.[94] For the trpzip2 hairpin system, several experiments have been carried out previously to investigate its relaxation behavior. In 2004, Gruebele[88] reported a T-jump fluorescence experiment which revealed a two phase relaxation behavior for trpzip2. Within the first 100 ns, the fastest process leads to a rapid increase of trpzip2 fluorescence. Within 0.6−5 μs (depending on temperature and denaturant), the slower process completes the kinetics, leading to the equilibrium fluorescence signature. Also in 2004, Gai[93] reported a T-jump study of trpzip2 combined fluorescence and FTIR, the unfolding times reported are similar for both spectroscopies (~18 μs for fluorescence and ~24.7 μs for FTIR), the inconsistency between the relaxation times reported in the experiments and the simulation results in this manuscript can be due to various reasons such as the inconsistency of temperature dependence of the force field in the simulation and the experiment.[94] A thorough comparison of experimental and simulated T-jump results at various temperatures with different force fields should be carried out in the future to explore the current protocol's capability to reproduce the temperature dependence of the unfolding dynamics in the experiment.

In conclusion, we proposed in this manuscript a computational protocol of simulating T-jump peptide unfolding and the related transient IR and 2DIR experiments based on the MSM and NEP methods. Our results show that MSMs constructed from a large number of simulations initiated from different parts of the phase space can provide much better agreement with the equilibrium experimental FTIR and 2DIR spectra compared to that generated from straightforward MD simulations initiating from the folded state. We conclude that sufficient conformational sampling as well as including important relevant conformational states are crucial for obtaining accurate spectroscopic observables. MSMs are also capable of simulating the unfolding relaxation dynamics upon the temperature jump by calculating the time evolution of the 2DIR spectra. Finally, the protocol we introduced here provides a useful tool to validate the force fields. It can also be applied to study a wide range of biomolecular processes involving conformational changes.

AUTHOR INFORMATION

Corresponding Author

*E-mail: wzhuang@dicp.ac.cn (W.Z.); xuhuihuang@ust.hk (X. H.).

ACKNOWLEDGMENT

W.Z. gratefully acknowledges the support of the NSFC QingNian Grant 21003117 and NSFC Key Grant 21033008. X.H. acknowledges the support from the Hong Kong Research Grants Council HKUST2/CRF/10 and the University Grants Council RPC10SC03. Computing resources were provided by the Dalian Division of the CAS Computing Center in DICP, the CCB computer cluster in HKUST, and the Folding@home distributed computing environment which is supported by NIH Grant R01-GM062868 to Vijay Pande.

REFERENCES

(1) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. The protein folding problem. *Annu. Rev. Biophys.* **2008**, *37*, 289–316.

(2) Gruebele, M. The fast protein folding problem. *Annu. Rev. Phys. Chem.* **1999**, *50*, 485–516.

(3) Anfinsen, C. B.; Scheraga, H. A. Experimental and theoretical aspects of protein folding. *Adv. Protein Chem.* **1975**, *29*, 205–300.

(4) Kubelka, J.; Hofrichter, J.; Eaton, W. A. The protein folding "speed limit". *Curr. Opin. Struct. Biol.* **2004**, *14*, 76–88.

(5) Xu, Y.; Purkayastha, P.; Gai, F. Nanosecond folding dynamics of a three-stranded β-sheet. *J. Am. Chem. Soc.* **2006**, *128*, 15836–15842.

(6) Maity, H.; Maity, M.; Krishna, M. M.; Mayne, L.; Englander, S. W. Protein folding: tThe stepwise assembly of foldon units. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 4741–4746.

(7) Leopold, P. E.; Montal, M.; Onuchic, J. N. Protein folding funnels: A kinetic approach to the sequence−structure relationship. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8721–8725.

(8) Cecconi, C.; Shank, E. A.; Bustamante, C.; Marqusee, S. Direct observation of the three-state folding of a single protein molecule. *Science* **2005**, *309*, 2057–2060.

(9) Lei, H.; Wu, C.; Liu, H.; Duan, Y. Folding free-energy landscape of villin headpiece subdomain from molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4925–4930.

(10) Levy, Y.; Cho, S. S.; Onuchic, J. N.; Wolynes, P. G. A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. *J. Mol. Biol.* **2005**, *346*, 1121–1145.

(11) Matouschek, A.; Kellis, J. T., Jr.; Serrano, L.; Fersht, A. R. Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **1989**, *340*, 122–126.

(12) Sosnick, T. R.; Dothager, R. S.; Krantz, B. A. Differences in the folding transition state of ubiquitin indicated by φ and ψ analyses. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17377–17382.

(13) Zhou, R.; Berne, B. J.; Germain, R. The free energy landscape for β hairpin folding in explicit water. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931–14936.

(14) Bunagan, M. R.; Yang, X.; Saven, J. G.; Gai, F. Ultrafast folding of a computationally designed Trp-cage mutant: Trp(2)-cage. *J. Phys. Chem. B* **2006**, *110*, 3759–3763.

(15) Du, D. G.; Zhu, Y. J.; Huang, C. Y.; Gai, F. Understanding the key factors that control the rate of β-hairpin folding. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15915–15920.

(16) Zhu, Y.; Alonso, D. O. V.; Maki, K.; Huang, C.; Lahr, S. J.; Daggett, V.; Roder, H.; DeGrado, W. F.; Gai, F. Ultrafast folding of α3D:

A de novo designed three-helix bundle protein. *Proc. Natl. Acad. Sci. U.S. A.* **2003**, *100*, 15486–15491.
(17) Kubelka, J.; Henry, E. R.; Cellmer, T.; Hofrichter, J.; Eaton, W. A. Chemical, physical, and theoretical kinetics of an ultrafast folding protein. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18655–18662.
(18) Chung, H. S.; Tokmakoff, A Temperature-dependent downhill unfolding of ubiquitin. I. Nanosecond-to-millisecond resolved nonlinear infrared spectroscopy. *Proteins: Struct., Funct., Bioinf.* **2008**, *72*, 474–487.
(19) Voelz, V. A.; Bowman, G. R.; Beauchamp, K.; Pande, V. S. Molecular simulation of ab initio protein folding for a millisecond folder NTL9(1−39). *J. Am. Chem. Soc.* **2010**, *132*, 1526–1528.
(20) Klepeis, J. L.; Lindorff-Larsen, K.; Dror, R. O.; Shaw, D. E. Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol.* **2009**, *19*, 120–127.
(21) Eaton, W. A.; Munoz, V.; Hagen, S. J.; Jas, G. S.; Lapidus, L. J.; Henry, E. R; Hofrichter, J. Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 327–359.
(22) Ballew, R. M.; Sabelko, J.; Gruebele, M. Observation of distinct nanosecond and microsecond protein folding events. *Nat. Struct. Biol.* **1996**, *3*, 923–926.
(23) Ballew, R. M.; Sabelko, J.; Gruebele, M. Direct observation of fast protein folding: The initial collapse of apomyoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5759–5764.
(24) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. Folding dynamics and mechanism of β-hairpin formation. *Nature* **1997**, *390*, 196–199.
(25) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R.H.; Woodruff, W. H.; Dyer, R. B. Fast events in protein folding: Helix melting and formation in a small peptide. *Biochemistry* **1996**, *35*, 691–697.
(26) Mantsch, H. H., Chapman, D., Eds. *Infrared Spectroscopy of Biomolecules*; Wiley-Liss: New York, 1996; p 359.
(27) Spiro, E. T. *Biological Applications of Raman Spectroscopy*; Wiley-Interscience: New York, 1987; Vol. I.
(28) Creighton, T. *Proteins: Structures and Molecular Properties*; W. H. Freeman and Company: New York, 1993.
(29) Torii, H.; Tasumi, M. Ab initio molecular orbital study of the amide I vibrational interactions between the peptide groups in di- and tripeptides and considerations on the conformation of the extended helix. *J. Raman Spectrosc.* **1998**, *29*, 81–86.
(30) Krimm, S.; Bandekar, J. Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. *Adv. Protein Chem.* **1986**, *38*, 181.
(31) Asplund, M. C.; Zanni, M. T.; Hochstrasser, R. M. Two-dimensional infrared spectroscopy of peptides by phase-controlled femtosecond vibrational photon echoes. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8219–8224.
(32) Hamm, P.; Lim, M. H.; Hochstrasser, R. M. Structure of the amide I band of peptides measured by femtosecond nonlinear-infrared spectroscopy. *J. Phys. Chem. B* **1998**, *102*, 6123.
(33) Kim, Y. S.; Hochstrasser, R. M. Chemical exchange 2D IR of hydrogen-bond making and breaking. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11185–11190.
(34) Zheng, J.; Kwak, K.; Asbury, J.; Chen, X.; Piletic, I. R.; Fayer, M. D. Ultrafast dynamics of solute−solvent complexation observed at thermal equilibrium in real time. *Science* **2005**, *309*, 1338–1343.
(35) Zhuang, W.; Abramavicius, D.; Hayashi, T.; Mukamel, S. Simulation protocols for coherent femtosecond vibrational spectra of peptides. *J. Phys. Chem. B* **2006**, *110*, 3362–3374.
(36) Tanimura, Y.; Mukamel, S. Two-dimensional femtosecond vibrational spectroscopy of liquids. *J. Chem. Phys.* **1993**, *99*, 9496.
(37) Scheurer, C.; Piryatinski, A.; Mukamel, S. Signatures of β-peptide unfolding in two-dimensional vibrational echo spectroscopy: A simulation study. *J. Am. Chem. Soc.* **2001**, *123*, 3114.
(38) Fecko, C. J.; Eaves, J. D.; Loparo, J. J.; Tokmakoff, A.; Geissler, P. L. Ultrafast hydrogen-bond dynamics in the infrared spectroscopy of water. *Science* **2003**, *301*, 1698–1702.
(39) Demirdoven, N.; Cheatum, C. M.; Chung, H. S.; Khalil, M.; Knoester, J.; Tokmakoff, A. Two-dimensional infrared spectroscopy of antiparallel β-sheet secondary structure. *J. Am. Chem. Soc.* **2004**, *126*, 7981–7990.
(40) Shim, S. H.; Strasfeld, D. B.; Ling, Y. L.; Zanni, M. T. Automated 2D IR spectroscopy using a mid-IR pulse shaper and application of this technology to the human islet amyloid polypeptide. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14197–14202.
(41) Kolano, C.; Helbing, J.; Kozinski, M.; Sander, W.; Hamm, P. Watching hydrogen-bond dynamics in a β-turn by transient two-dimensional infrared spectroscopy. *Nature* **2006**, *444*, 469–472.
(42) Larsen, O. F.; Bodis, P.; Buma, W. J.; Hannam, J. S.; Leigh, D. A.; Woutersen, S. Probing the structure of a rotaxane with two-dimensional infrared spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13378–13382.
(43) Cowan, M. L.; Bruner, B. D.; Huse, N.; Dwyer, J. R.; Chugh, B.; Nibbering, E. T. J.; Elsaesser, T.; Miller, R. J. D. Ultrafast memory loss and energy redistribution in the hydrogen bond network of liquid H_2O. *Nature* **2005**, *434*, 199–202.
(44) Szyc, L.; Yang, M.; Nibbering, E. T. J.; Elsaesser, T. Ultrafast vibrational dynamics and local interactions of hydrated DNA. *Angew. Chem., Int. Ed.* **2010**, *49*, 3598–3610.
(45) Maekawa, H.; De Poli, M.; Toniolo, C.; Ge, N. H. Couplings between peptide linkages across a 3_{10}-helical hydrogen bond revealed by two-dimensional infrared spectroscopy. *J. Am. Chem. Soc.* **2009**, *131*, 2042–2043.
(46) Rubtsov, I. V. Relaxation-assisted two-dimensional infrared (RA 2DIR) method: Accessing distances over 10 Å and measuring bond connectivity patterns. *Acc. Chem. Res.* **2009**, *42*, 1385–1394.
(47) Kwac, K.; Cho, M. H. Molecular dynamics simulation study of *N*-methylacetamide in water. I. Amide I mode frequency fluctuation. *J. Chem. Phys.* **2003**, *119*, 2247–2255.
(48) Lawrence, C. P.; Skinner, J. L. Vibrational spectroscopy of HOD in liquid D_2O. I. Vibrational energy relaxation. *J. Chem. Phys.* **2002**, *117*, 5827.
(49) Jansen, T. L. C.; Knoester, J. Two-dimensional infrared population transfer spectroscopy for enhancing structural markers of proteins. *Biophys. J.* **2008**, *94*, 1818–1825.
(50) Chung, H. S.; Khalil, M.; Smith, A. W.; Ganim, Z.; Tokmakoff, A. Conformational changes during the nanosecond-to-millisecond unfolding of ubiquitin. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 612–617.
(51) Shim, S. H.; Gupta, R.; Ling, Y. L.; Strasfeld, D. B.; Raleigh, D. P.; Zanni, M. T. Two-dimensional IR spectroscopy and isotope labeling defines the pathway of amyloid formation with residue-specific resolution. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6614–6619.
(52) Bagchi, S.; Falvo, C.; Mukamel, S.; Hochstrasser, R. M. 2D-IR experiments and simulations of the coupling between amide-I and ionizable side chains in proteins: Application to the villin headpiece. *J. Phys. Chem. B* **2009**, *113*, 11260–11273.
(53) Roy, S.; Jansen, T. L. C.; Knoester, J. Structural classification of the amide I sites of a β-hairpin with isotope label 2DIR spectroscopy. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9347–9357.
(54) Smith, A. W.; Lessing, J.; Ganim, Z.; Peng, C. S.; Tokmakoff, A.; Roy, S.; Jansen, T. L.; Knoester, J. Melting of a β-hairpin peptide using isotope-edited 2D IR spectroscopy and simulations. *J. Phys. Chem. B* **2010**, *114*, 10913–10924.
(55) Noe, F.; Fischer, S. Transition networks for modeling the kinetics of conformational change in macromolecules. *Curr. Opin. Struct. Biol.* **2008**, *18*, 154–162.
(56) 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. *J. Chem. Phys.* **2007**, *126*, 155101.
(57) Buchete, N. V.; Hummer, G. Coarse master equations for peptide folding dynamics. *J. Phys. Chem. B* **2008**, *112*, 6057–6069.
(58) Swope, W. C.; Pitera, J. W.; Suits, F. Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 1. Theory. *J. Phys. Chem. B* **2004**, *108*, 6571–6581.

(59) Sriraman, S.; Kevrekidis, I. G.; Hummer, G. Coarse master equation from Bayesian analysis of replica molecular dynamics simulations. *J Phys Chem B* **2005**, *109*, 6479–6484.
(60) Falvo, C.; Palmieri, B.; Mukamel, S. Coherent infrared multidimensional spectra of the OH stretching band in liquid water simulated by direct nonlinear exciton propagation. *J. Chem. Phys.* **2009**, 130.
(61) Jansen, T. L.; Knoester, J. Nonadiabatic effects in the two-dimensional infrared spectra of peptides: Application to alanine dipeptide. *J. Phys. Chem. B* **2006**, *110*, 22910–22916.
(62) Jansen, T. L.; Zhuang, W; Mukamel, S. Stochastic Liouville equation simulation of multidimensional vibrational line shapes of trialanine. *J. Chem. Phys.* **2004**, *121*, 10577–10598.
(63) Torii, H. Effects of intermolecular vibrational coupling and liquid dynamics on the polarized Raman and two-dimensional infrared spectral profiles of liquid *N*,*N*-dimethylformamide analyzed with a time-domain computational method. *J. Phys. Chem. A* **2006**, *110*, 4822–4832.
(64) Auer, B. M.; Skinner, J. L. IR and Raman spectra of liquid water: Theory and interpretation. *J. Chem. Phys.* **2008**, 128.
(65) Jansen, T. L. C.; Ruszel, W. M. Motional narrowing in the time-averaging approximation for simulating two-dimensional nonlinear infrared spectra. *J. Chem. Phys.* **2008**, 128.
(66) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. Tryptophan zippers: Stable, monomeric β-hairpins. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5578–5583.
(67) Wang, J. P.; Chen, J. X.; Hochstrasser, R. M. Local structure of β-hairpin isotopomers by FTIR, 2D IR, and ab initio theory. *J. Phys. Chem. B* **2006**, *110*, 7545–7555.
(68) Wang, J. P.; Zhuang, W.; Mukamel, S.; Hochstrasser, R. Two-dimensional infrared spectroscopy as a probe of the solvent electrostatic field for a twelve residue peptide. *J. Phys. Chem. B* **2008**, *112*, 5930–5937.
(69) Smith, A. W.; Tokmakoff, A. Amide I two-dimensional infrared spectroscopy of β-hairpin peptides. *J. Chem. Phys.* **2007**, 126.
(70) 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. *J. Chem. Phys.* **2007**, 126.
(71) Pitera, J. W.; Haque, I.; Swope, W. C. Absence of reptation in the high-temperature folding of the trpzip2 β-hairpin peptide. *J. Chem. Phys.* **2006**, 124.
(72) Huang, X.; Bowman, G. R.; Bacallado, S.; Pande, V. S. Rapid equilibrium sampling initiated from nonequilibrium data. *Proc. Natl. Acad. Sci. U S A* **2009**, *106*, 19765–19769.
(73) Hansmann, U. H.; Okamoto, Y. New Monte Carlo algorithms for protein folding. *Curr. Opin. Struct. Biol.* **1999**, *9*, 177–183.
(74) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, 141–151.
(75) Bowman, G. R.; Huang, X.; Pande, V. S. Using generalized ensemble simulations and Markov state models to identify conformational states. *Methods* **2009**, *49*, 197–201.
(76) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
(77) Wang, J; Cieplak, P; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
(78) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. Tryptophan zippers: Stable, monomeric β-hairpins. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5578–5583.
(79) 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.
(80) Hoover, W. A Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev.* **1985**, *31*, 1695–1697.
(81) Darden, T.; York, D.; Pedersen., L. A smooth particle mesh Ewald potential. *J. Chem. Phys.* **1995**, *103*, 3014–3021.
(82) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije., J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
(83) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
(84) Shirts, M.; Pande, V. S. COMPUTING: Screen savers of the world unite!. *Science* **2000**, *290*, 1903–1904.
(85) Huang, X.; Yao, Y.; Bowman, G. R.; Sun, J.; Guibas, L. J.; Carlsson, G.; Pande, V. S. Constructing multi-resolution Markov state models (MSMs) to elucidate RNA hairpin folding mechanisms. *Pac. Symp. Biocomput.* **2010**, 228–239.
(86) Deuflhard, P.; Weber, M. Robust Perron cluster analysis in conformation dynamics. *Linear Algebra Appl.* **2005**, *398*, 161–184.
(87) Deuflhard, P.; Huisinga, W.; Fischer, A.; Schütte, C. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Linear Algebra Appl.* **2000**, *315*, 39–59.
(88) Yang, W. Y.; Gruebele, M. Detection-dependent kinetics as a probe of folding landscape microstructure. *J. Am. Chem. Soc.* **2004**, *126*, 7758–7759.
(89) Bour, P.; Keiderling, T. A. Empirical modeling of the peptide amide I band IR intensity in water solution. *J. Chem. Phys.* **2003**, *119*, 11253.
(90) Jansen, T. L.; 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.
(91) Schmidt, J. R.; Corcelli, S. A.; Skinner, J. L. Ultrafast vibrational spectroscopy of water and aqueous *N*-methylacetamide: Comparison of different electronic structure/molecular dynamics approaches. *J. Chem. Phys.* **2004**, *121*, 8887–8896.
(92) 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.
(93) Wang, T.; Xu, Y.; Du, D. G.; Gai, F. Determining β-sheet stability by Fourier transform infrared difference spectra. *Biopolymers* **2004**, *75*, 163–172.
(94) Huang, X; Bowman, G. R.; Pande, V. S. Convergence of folding free energy landscapes via application of enhanced sampling methods in a distributed computing environment. *J. Chem. Phys.* **2008**, *128*, 205106.