Monitoring the Folding Kinetics of a β-Hairpin by Time-Resolved IR Spectroscopy in Silico

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · MARCH 2015

Impact Factor: 3.3 · DOI: 10.1021/acs.jpcb.5b01477 · Source: PubMed

4 AUTHORS, INCLUDING:

Isabella Daidone

Università degli Studi dell'Aquila

53 PUBLICATIONS **862** CITATIONS

SEE PROFILE

Lipi Thukral

Institute of Genomics and Integrative Biology

11 PUBLICATIONS **49** CITATIONS

SEE PROFILE

Andrea Amadei

University of Rome Tor Vergata

163 PUBLICATIONS **4,990** CITATIONS

SEE PROFILE

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Monitoring the Folding Kinetics of a β-Hairpin by Time-Resolved IR Spectroscopy in Silico

Isabella Daidone,*[,†] Lipi Thukral,[‡] Jeremy C. Smith,[¶,⊗] and Andrea Amadei*[,§]

[†]Department of Physical and Chemical Sciences, University of L'Aquila, via Vetoio (Coppito 1), 67010 L'Aquila, Italy

[‡]CSIR-Institute of Genomics and Integrative Biology, South Campus, Mathura Road, New Delhi 110020, India

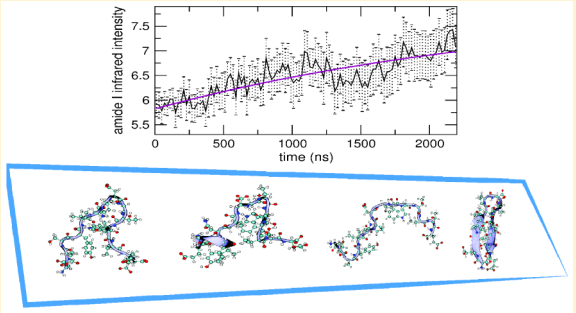
[¶]University of Tennessee/Oak Ridge National Laboratory, Center for Molecular Biophysics, P.O. Box 2008, Oak Ridge, Tennessee 37831-6309, United States

[⊗]Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, M407 Walters Life Sciences, 1414 Cumberland Avenue, Knoxville, Tennessee 37996, United States

[§]Department of Chemical Sciences and Technologies, University of Rome "Tor Vergata", via della Ricerca Scientifica 1, 00133 Rome, Italy

Supporting Information

ABSTRACT: Protein folding is one of the most fundamental problems in modern biochemistry. Time-resolved infrared (IR) spectroscopy in the amide I region is commonly used to monitor folding kinetics. However, associated atomic detail information on the folding mechanism requires simulations. In atomistic simulations structural order parameters are typically used to follow the folding process along the simulated trajectories. However, a rigorous test of the reliability of the mechanisms found in the simulations requires calculation of the time-dependent experimental observable, i.e., in the present case the IR signal in the amide I region. Here, we combine molecular dynamics simulation with a mixed quantum mechanics/molecular mechanics theoretical methodology, the Perturbed Matrix Method, in order to characterize the folding of a β-hairpin peptide, through modeling the time-dependence of the amide I IR signal. The kinetic and thermodynamic data (folding and unfolding rate constants, and equilibrium folded- and unfolded-state probabilities) obtained from the fit of the calculated signal are in good agreement with the available experimental data [Xu et al. *J. Am. Chem. Soc.* **2003**, *125*, 15388−15394]. To the best of our knowledge, this is the first report of the simulation of the time-resolved IR signal of a complex process occurring on a long (microsecond) time scale.



INTRODUCTION

Theoretical modeling of infrared (IR) spectra of polyatomic molecules in condensed phases is a challenging problem in physical chemistry. Beyond its intrinsically interesting aspects, the correct modeling of IR spectroscopic signals is of considerable practical importance, as this experimental technique is among the most utilized for addressing several relevant problems, in particular, in the biochemical field. For example, by means of time-resolved[1] and bidimensional[2] IR spectroscopy it is possible to characterize ligand migration in proteins[3,4] and folding/unfolding kinetics in peptides.[5−11] However, the complexity of the spectroscopic signal, arising from the structural fluctuations of the polypeptide chain and its interaction with the solvent, makes the interpretation of the spectra difficult in terms of structural features and transitions of the solvated peptide or protein. Theoretical−computational methods are needed to provide essential information on the complex relationship between spectroscopic absorption and dynamics, permitting the effects of structural transitions and of solvent interactions on the IR signal to be unraveled. In particular, the possibility of quantitatively modeling time-resolved signals could open the way to closer cooperation between theoreticians and experimentalists, with the former providing a detailed interpretation of the time evolution of the spectroscopic signal, permitting a full characterization of the kinetics of complex chemical−biochemical processes.

In the past decade several theoretical−computational approaches, typically based on mixed quantum-classical models, have been proposed[12−22] to reconstruct in detail the IR spectra of solvated peptides and hence to clarify the main atomistic determinants of peptide/protein IR signals. However, most of these methods suffer from poor sampling of the phase space. This is particularly problematic when the IR signal is to be monitored as a function of time in the study of complex processes occurring on long time scales (e.g., nanoseconds−microseconds), such as peptide/protein folding. In this context,

Received: February 12, 2015
Revised: March 13, 2015
Published: March 17, 2015

DOI: 10.1021/acs.jpcb.5b01477
J. Phys. Chem. B 2015, 119, 4849−4856

we have proposed, and applied, a mixed quantum mechanics/molecular dynamics (QM/MD) theoretical computational methodology based on the Perturbed Matrix Method (PMM)[23−27] to calculate amide I IR spectra of peptides and proteins. In PMM, as in other QM/MM procedures,[28] a portion of the system (the quantum center) is treated at an electronic level, while the rest of the system is described at a classical atomistic level, exerting an electrostatic perturbation on the quantum-center electronic states. As the method makes use of classical MD to provide phase space sampling, statistically relevant sampling of the quantum-center/environment configurations can be achieved, which is required for an accurate calculation of the spectra of complex systems such as amyloids[29] or peptide/protein unfolded states.[25,30−32] However, since the PMM methodology does not make use of empirical/adjustable parameters, it can sometimes be less accurate than other methodologies[33−35] in reproducing some of the details of the spectroscopic signals.

Here, we apply the PMM-MD procedure to study the folding kinetics of a β-hairpin peptide through the modeling of the time-dependent IR signal in the amide I region. The chosen peptide is a 15-residue model system, i.e., Peptide 1 (SESYINPDGTWTVTE), that has been studied experimentally using a T-jump IR method.[8] By monitoring the conformation-sensitive amide I absorption the time-resolved IR technique used allows the backbone conformation to be specifically followed during the course of folding/unfolding of the peptide. By using multiple, independent folding/unfolding MD simulations of Peptide 1 on the microsecond time scale, we were able to reconstruct the time-dependent amide I IR signal. The computational spectroscopic data were then utilized, by means of a fitting procedure, to obtain the relevant thermodynamic and kinetic properties of the folding/unfolding reaction, according to a simple two-state model equivalent to the one used to interpret the experimental data.[8]

METHODS

The Perturbed Matrix Method. Details of the PMM are given in previous articles[25,27] and in the SI. In this section we give an overview to provide a general picture of how the method works, paying particular attention to the calculation of IR spectroscopic properties.

In PMM a part (or parts) of the system is (are) chosen as the active quantum center (centers), hereafter termed QC (QCs). Quantum mechanical calculations are performed only on the QCs, while the rest of the system is described at a classical atomistic level. The quantum states of each QC are considered as perturbed by the electrostatic field exerted by the atomic charges of the environment, with the atomic motions provided by MD simulation. In the present case, the selected QCs are the 14 peptide backbone units of the system.

As concerns the specific task of modeling the vibrational excitations of each QC, the basic idea underlying the PMM-MD approach is to reconstruct, according to basic quantum mechanical principles, the dependence of the perturbed electronic ground state energy of each QC on the vibrational mode coordinate of interest (in the present case, the amide I mode) at each frame of the MD simulation. Trans-*N*-methylacetamide (trans-NMA) was chosen as the QC model for each peptide backbone unit. The mass-weighted Hessian eigenvectors of the isolated trans-NMA molecule, calculated quantum chemically, provide the unperturbed vibrational modes of each peptide backbone unit from which the amide I mode is selected (these calculations are performed only once). Then, at each configuration generated by the MD simulation and for each of the 14 peptide backbone units the following steps are performed. The perturbing environment, defined by the side chain of the residue considered, the atoms of the other *N*-1 residues, and the solvent atoms, is calculated and is used to obtain the perturbed vibrational mode, and corresponding frequency, of any individual peptide group. The perturbed frequencies of the 14 peptide units are then used to construct the excitonic coupling matrix (i.e., the Hamiltonian matrix for the interacting chromophores) which accounts for coupling between the vibrational modes of the chromophores. Diagonalization of the excitonic matrix provides the delocalized vibrational modes representing the vibrational (excitonic) states of the system.

To do the above, the following procedure is used:

1. A set of atomic configurations of the trans-NMA model (in the absence of the rest of the peptide), the "mode reference structures", is generated along the chosen mode coordinate and for each of these structures the corresponding unperturbed electronic eigenstates and properties (e.g., energies and dipoles) are evaluated. Note that this step needs to be performed only once, as it provides the unperturbed mode properties.

2. For each MD frame and for each IR chromophore (i.e., each backbone peptide unit) we calculate, using PMM, the perturbed electronic ground state and corresponding energy of each mode reference structure. Thus, for each MD frame and for each IR chromophore we obtain the electronic ground state energy as a function of the mode coordinate.

3. Each perturbed energy curve obtained above in step 2 at each MD frame and for each IR chromophore is then modeled by a Morse potential, thus providing the vibrational frequency of the perturbed mode.

4. At each MD frame the evaluated perturbed frequencies are used to construct the excitonic coupling matrix describing the excitonic coupling among the QCs' modes by means of the transition dipole coupling (TDC) approximation, i.e., the expansion of the chromophore−chromophore interaction operator up to the dipolar terms. Diagonalization of the excitonic coupling matrix provides the instantaneous vibrational eigenstates and eigenvalues (now including vibrational mode coupling), yielding the perturbed vibrational frequencies and corresponding transition dipoles of the whole peptide.

5. Finally, the obtained perturbed excitation frequencies and corresponding transition dipoles are used to reconstruct the complete vibrational spectrum.

The main approximations of the above approach are the following: (a) the invariant mode approximation,[27] which is based on the assumption that the perturbations acting on a single IR chromophore do not significantly modify the forms of the vibrational modes (i.e., the eigenvectors of the mass-weighted Hessian) of interest, but rather alter only the corresponding frequencies. (b) Similarly to other TDC calculations,[33−35] the PMM approach utilizes the dipolar approximation. However, in most of the available TDC methodologies the excitonic coupling matrix is commonly constructed by using unperturbed single-residue vibrational states (i.e., in the absence of the environment perturbation), and the inclusion of the perturbation effects then typically involves different levels of phenomenological approximation, trying to optimize the computed−experimental matching.[33−35] In contrast, in the present approach the excitonic coupling matrix is constructed from the basis set of the perturbed (and

DOI: 10.1021/acs.jpcb.5b01477
J. Phys. Chem. B 2015, 119, 4849−4856

not the unperturbed) vibrational states, thus explicitly including the perturbation of the atomic-molecular environment in the definition of the basis set used to provide the excitonic states. (c) The vibrational coupling between chromophore and solvent modes is neglected, and therefore, the method does not properly treat chromophore vibrational modes involving atomic coordinates of the first solvation shell.

The procedure briefly outlined above not only can provide the stationary spectral signal, but also can be used to obtain the time-dependent spectral observables, as in the present paper. In fact, by considering the MD frames provided by the set of independent trajectories within a short time interval (in our case we have chosen 20 ns as a good compromise between time resolution and sampling), we were able to obtain the corresponding ensemble time-dependent spectra. In particular, the time-dependent spectral intensity at a reference frequency was used as a *mechanical* time-dependent observable directly describing the kinetics of the folding/unfolding transition.

Unperturbed Quantum Chemical Calculations. Quantum chemical calculations were carried out on the isolated trans-NMA molecule with time dependent density functional theory (TD-DFT) with the 6-31+G(d) basis set, providing the unperturbed electronic eigenstates to be used in the PMM procedure. This level of theory was selected because it represents a good compromise between computational cost and accuracy. The mass-weighted Hessian matrix of the isolated trans-NMA was calculated on the optimized geometry at the B3LYP/6-31+G(d) level of theory and subsequently diagonalized to obtain the unperturbed eigenvectors and related eigenvalues. The coordinates of the optimized geometry and the corresponding ground state energy are reported in the SI. The eigenvector corresponding in vacuo to the amide I mode was, then, used to generate a number of points (i.e., configurations) as follows: a step of 0.05 au was adopted and the number of points was set to span at least an energy range of 20−30 kJ/mol (in the present case 31 points). The use of an energy interval of at least 20−30 kJ/mol is motivated by the fact that the energies of at least the ground and first excited vibrational states of the (unperturbed) amide I mode are within this range. For each point, six unperturbed electronic states were then evaluated at the same level of theory providing the basis set for the PMM calculations. Within the PMM procedure, the unperturbed electronic ground state and first few excited states are used to obtain the perturbed electronic ground state. This is because the perturbed electronic ground state is often a combination of the unperturbed electronic ground and first excited states.

Molecular Dynamics Simulations. A series of 17 2.5-μs-long atomistic MD simulations of Peptide 1 in explicit solvent was performed. Fifteen trajectories were started from unfolded structures that were extracted randomly from a simulation of 50 ns that itself started from a fully extended conformation of the peptide. The remaining two trajectories were started from a folded configuration with coordinates taken from the NMR structure,[36] using different initial velocities. The MD simulations were performed with the program GROMACS[37] and the OPLS-AA all atom force field[38] was used for the peptide. The water was modeled using the TIP4P water model.[39] Each of the 17 starting peptide structures was placed in a dodecahedral water box (volume = 48 nm^3) large enough to contain the peptide and at least 1.0 nm of solvent on all sides. Each simulation box contained 6647 atoms. Periodic boundary conditions were used and the long-range electrostatic interactions were treated with the Particle Mesh Ewald method.[40] The bond lengths were fixed[41] and a time step of 2 fs for numerical integration of the equations of motion was used. Coordinates were saved at every 0.2 ps. Simulations were performed in the NVT ensemble with the isokinetic temperature coupling[42] to keeping the temperature constant at 300 K. Three positive counterions (Na^+) were added by replacing three water molecules so as to produce a neutral simulation box.

It should be noted that the simulations were carried out using H_2O, while the experiments were performed in D_2O. The use of H_2O or D_2O for evaluating stationary (equilibrium) spectra are fully equivalent within the assumption that the solvent vibrational modes are not coupled to the amide I mode, as in the present case. However, when dealing with time-resolved spectroscopic observables the use of H_2O or D_2O as solvent could in principle provide weak kinetic variations due to a possible change in the solvent dynamical relaxations. We have disregarded these possible weak effects, as is also done by the experimentalists who assume that the use of D_2O does not significantly alter the chromophore vibrational properties utilized to obtain the reaction kinetics.

The folding of Peptide 1 was monitored using the R parameter[43] calculated as follows:

$$R = \sum_{i=1}^{5} \frac{R_i^N}{R_i}$$

where R_i^N is the ith interstrand $C_\alpha - C_\alpha$ distance in the β-hairpin structure (either the native NMR structure or the hairpin with the shift in the interstrand registry) and R_i is the same distance in the MD. The five interstrand $C_\alpha - C_\alpha$ pairs in the NMR hairpin are the following: N6-T10, I5-W11, Y4-T12, S3-V13, and E2-T14. The five interstrand $C_\alpha - C_\alpha$ pairs in the other hairpin are the following: N6-W11, I5-T12,Y4-V13, S3-T14, and E2-E15. A value of $R \approx 5$ indicates formation of a β-hairpin conformation. In particular, a cutoff value of $R \geqslant 4.8$ was used to define folded conformations.

The standard error, σ, on different properties evaluated from the 17 trajectories was estimated by calculating their standard deviations over n subsets of the trajectories and then extrapolating for the complete statistical sample (i.e., by considering the observable obtained from the complete sample equivalent to the corresponding average over the n subsets):

$$\sigma = \left(\frac{\sum_{i=1}^{n} (a_i - \bar{a})^2}{(n-1)n} \right)^{1/2} \quad (1)$$

$$\bar{a} = \frac{\sum_{i=1}^{n} a_i}{n} \quad (2)$$

where a_i is the generic parameter/observable evaluated in the ith subset. In the present case we used three independent subsets ($n = 3$) of the trajectories which was found to be a good compromise between the statistics within each subset and the sample size (given by n). The error on the observable a was, then, taken as the size of the 95% confidence interval, i.e., $\sim 2\sigma$.

RESULTS AND DISCUSSION

In a previous paper[44] we determined the folding mechanism of Peptide 1 by performing 15 independent MD simulations in explicit aqueous solvent, each 2.5 μs long, starting from different unfolded structures. The R parameter[43] was used, as

DOI: 10.1021/acs.jpcb.5b01477
J. Phys. Chem. B 2015, 119, 4849−4856

conformational order parameter, to follow the time evolution of the conformational transitions (see the Methods section for the definition of R). When $R \geqslant 4.8$ the peptide folds into β-hairpin conformations; otherwise, the peptide is unfolded. The mean folding time, calculated from fitting to the distribution of the residence times in the unfolded state, was $\tau_f \approx 1600$ ns.

The peptide forms the native NMR β-hairpin[36] in trajectories 1−6, 10, and 12, while it forms a β-hairpin with a shift in the interstrand registry in trajectories 7−9. In trajectories 11, 13−15, transitions into folded hairpins are not observed (see Figure 1). Folding into stable hairpin configurations occurs only in trajectories 1−8, while the other hairpins are unstable (see Figure 1B). Further inspection of these structures reveals that, although the R parameter is $\geqslant$4.8, the unstable hairpins are partly "distorted", e.g., not all backbone hydrogen bonds are formed or the hairpin is not fully extended. Representative examples of correctly folded hairpins (either the NMR β-hairpin or the hairpin with a shift in the interstrand registry) and distorted (misfolded) hairpins are given in Figure 1A and C, respectively. In the present work two additional simulations were performed starting from the native NMR β-hairpin configuration, and these remain basically folded over the entire simulation time of 2.5 μs (see trajectories 16 and 17 in Figure 1B).

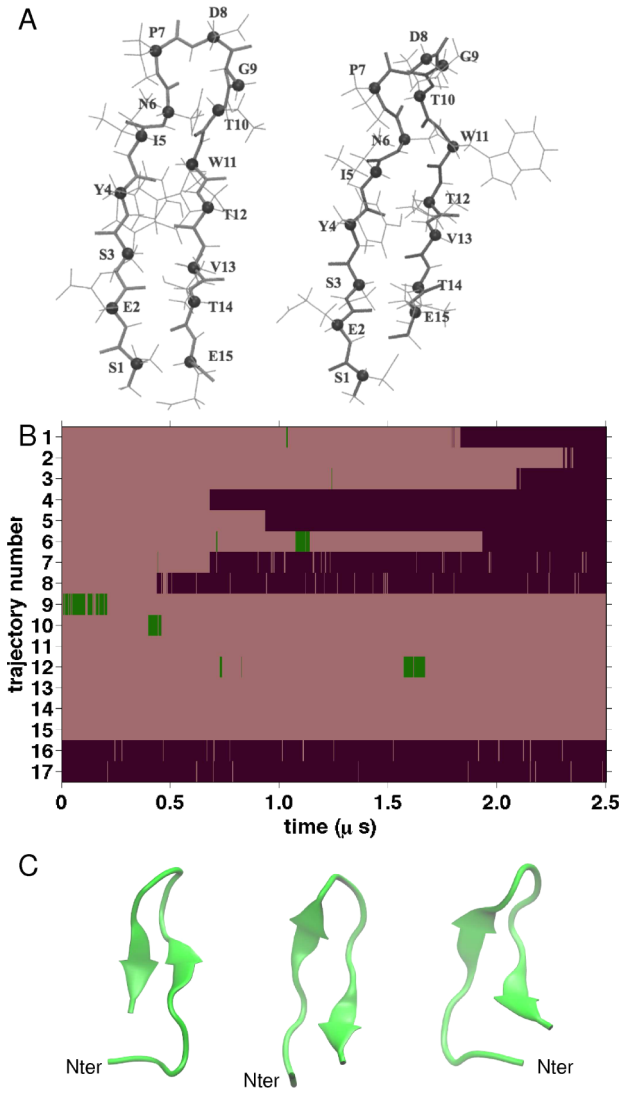


Figure 1. (A) Representative structures of the NMR β-hairpin (left) and the hairpin with the shift in the interstrand registry (right). (B) Time evolution of the secondary structure of Peptide 1 along the 17 trajectories. Trajectories 1−15 were started from folded-hairpin configurations, while trajectories 16 and 17 were started from unfolded structures. At time t = 0 the peptide is fully unfolded in all trajectories (R < 4.8, pink). When $R \geqslant 4.8$ the peptide forms either stable β-hairpin conformations, the NMR hairpin[36] or the hairpin with a shift in the interstrand registry (violet), or unstable, distorted β-hairpins (green). (C) Representative structures of distorted hairpins.

The main aim of the present work is to model the time-dependent IR signal of Peptide 1 in the amide I region to yield the kinetics of folding, in a similar manner as is done experimentally. In the IR experiment[8] Peptide 1 exhibits first-order folding/unfolding kinetics, as well as two-state thermodynamics. Based on these data, and on further pieces of evidence provided below, we decided to use a two-state model of the folding/unfolding process of the peptide under investigation. It should be noted that in the experimental work the authors make use of a T-jump procedure providing essentially an increase of the unfolded fraction. In comparison, in our simulated experiment we use a reactive initial ensemble largely constituted by unfolded structures (15 out of 17), and hence the kinetics obtained corresponds essentially to a folding relaxation. However, the two state model assumed here, which is in agreement with the experimental results, necessarily implies that the kinetic properties (e.g., the rate constants) must be independent of the initial reactive ensemble. Therefore, the kinetics obtained in our simulated folding/unfolding experiment should be equivalent to those experimentally obtained by means of the T-jump procedure.

We first calculate the stationary amide I IR spectrum of the unfolded and folded states of Peptide 1 (see Figure 2) using PMM-MD (see the Methods section). The folded state comprises all the fully folded β-hairpin structures (i.e., the NMR hairpin and the hairpin with the shift in the interstrand registry) sampled along the 17 trajectories, while the unfolded state comprises all the remaining configurations (including the distorted-hairpin structures). The two fully folded hairpin conformations are grouped together in a single folded state because they show nearly coincident amide I spectra. The distorted hairpins, instead, are included in the unfolded state (the reasons for this choice are given in the next paragraph). In the figure is also reported the unfolded−folded difference spectrum, showing the typical negative−positive feature. Comparison of the calculated difference spectrum with the experimental one[8] shows, indeed, that the main features of the difference spectrum are reasonably well reproduced (see Figure 2C). In particular, beyond the proper spectral shape, the calculations reproduce quantitatively the absolute intensity ratio of the two peaks (being ∼0.5−0.6, i.e., the intensity of the positive peak is ∼half that of the negative peak). The corresponding frequency shift is of the same order of magnitude but underestimated (∼15 cm^{-1} vs ∼30 cm^{-1}).

The frequency of the maximum peak of the folded-state spectrum (wavenumber = 1644 cm^{-1}) is used to monitor changes in the secondary structure of Peptide 1. The 17 MD trajectories were divided into segments of 20 ns and, for each segment, the amide I spectrum was calculated using the corresponding structures. The distribution of the intensities at 1644 cm^{-1} of each 20 ns spectrum is reported in Figure 3. The distribution is roughly bimodal, with the minimum at around 7.3. The peak corresponding to intensities in the range 3−7.3

DOI: 10.1021/acs.jpcb.5b01477
J. Phys. Chem. B 2015, 119, 4849−4856

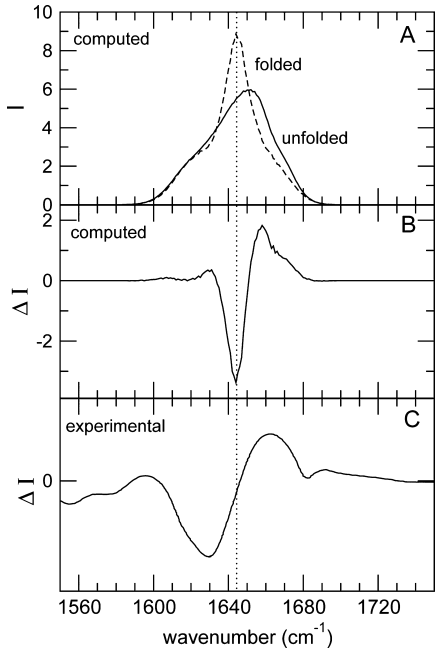


Figure 2. (A) Computed IR spectra in the amide I region of the folded and unfolded states of Peptide 1. (B) Computed difference spectrum generated by subtracting the spectrum of the folded state from that of the unfolded state. For the sake of comparison, the computed frequencies have been uniformly shifted to lower frequencies by 73 cm^{-1} in order to align the computed amide I peak of the folded state to the experimental maximum at 1644 cm^{-1} (the dotted line corresponds to wavenumber= 1644 cm^{-1}). (C) Experimental difference spectrum. It was generated by subtracting the spectrum collected at 3.0 °C from the one collected at 61.0 °C.[8] The intensities I are given in arbitrary units.

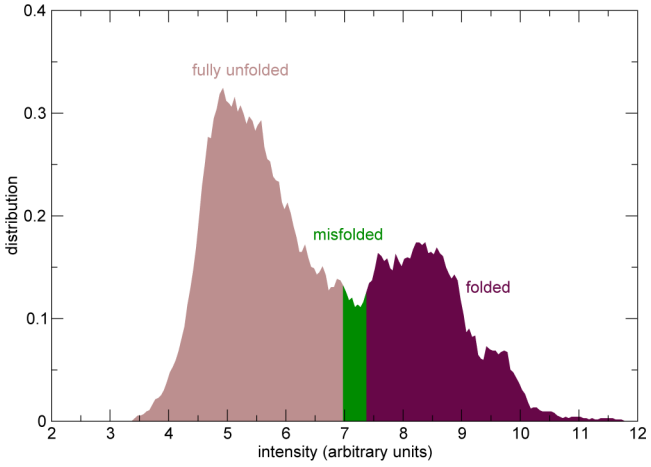


Figure 3. Distribution of IR intensities at wavenumber = 1644 cm^{-1}. The 17 MD trajectories were divided into segments of 20 ns and, for each segment, the amide I spectrum was calculated. The curve is the distribution of the intensities at 1644 cm^{-1} of each 20 ns spectrum.

mainly comprises unfolded configurations, while the other peak in the range 7.3−10.5 comprises folded configurations. The misfolded, distorted hairpin structures mentioned above have intensities in the range 6.9−7.4, i.e., they mainly belong to the unfolded-state peak, and their occurrence is less than 5−7% of all the unfolded configurations.

Based on Figure 3 (i.e., a bimodal distribution of the intensities and a low concentration of misfolded structures), we decided to group distorted hairpins together with fully unfolded configurations, rather than with the folded ones or considering them as a third state. Another reason for this choice is that misfolded configurations are kinetically connected to the unfolded state rather than to the folded state (i.e., transitions from/to the distorted-hairpin configurations occur only from/to fully unfolded structures).

To monitor the folding/unfolding relaxation kinetics, the time-dependence of the mean IR intensity at 1644 cm^{-1} along the 17 MD trajectories was calculated and is reported in Figure 4A. In order to ensure that the initial state is an equilibrium

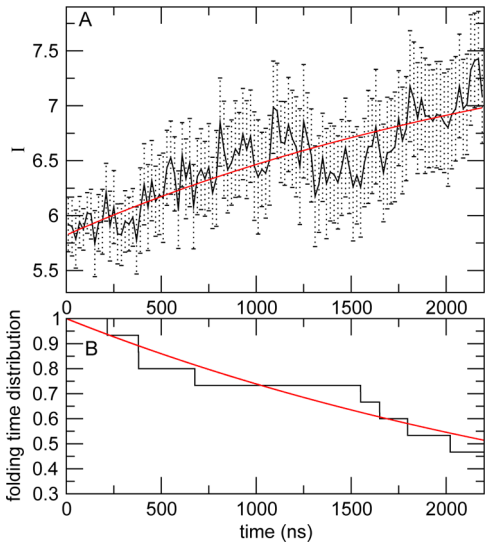


Figure 4. (A) Time-dependent amide I IR signal calculated over the 17 MD trajectories. The probing frequency was 1644 cm^{-1}. The 17 MD trajectories were divided into segments of 20 ns and, for each segment, the amide I spectrum was calculated using the corresponding structures. Each point in time of the kinetic trace corresponds to the intensity at wavenumber = 1644 cm^{-1} of the corresponding time-dependent spectrum averaged over the 17 trajectories. The red line is the fit to the function reported in eq 6. (B) Distribution of the residence times in the unfolded state. The peptide is considered folded if $R \geqslant 4.8$ and the hairpin is fully folded (i.e., the transitions to distorted hairpin configurations are excluded). The red line shows a monoexponential fit to the data. The time constant of the exponential fit is taken as the mean folding time.

state ensemble, the first 300 ns of each trajectory were removed from the kinetics evaluation. This time range was chosen because it corresponds to ∼3 times the time constant of the slowest relaxation process, τ_{slow}, taking place within the unfolded state, which corresponds to the "fully unfolded"/"misfolded" transition process ($1/\tau_{slow}$ = 1/105 ns^{-1} ≈ 1/2110 ns^{-1} + 1/110 ns^{-1}, with 2110 and 110 ns the time constants of the unfolded → misfolded and misfolded → unfolded transitions, respectively). Within the removed first 300 ns of the 17 trajectories, folding transitions did not occur.

By using a two-state model to describe the folding/unfolding equilibrium it is possible to derive a simple expression for the time-dependence of the IR spectroscopic signal as occurring during the folding process. The spectroscopic intensity, I, can then be expressed as a function of time t via

$$I(t) = I_f P_f(t) + I_u P_u(t) \quad (3)$$

where I_f and I_u are the spectroscopic IR intensities at the frequency of the folded-state maximum (1644 cm^{-1}) for the

DOI: 10.1021/acs.jpcb.5b01477
J. Phys. Chem. B 2015, 119, 4849−4856

folded and unfolded state, respectively (I_f = 8.62 and I_u = 5.45; see Figure 2), and $P_f(t)$ and $P_u(t)$ are the time dependent probabilities of the folded and unfolded state as estimated by the two-state model

$$P_f(t) = \frac{k_f}{k_u + k_f} + \left(\frac{k_u}{k_u + k_f} - P_u^0\right) e^{-(k_u+k_f)t} \quad (4)$$

$$P_u(t) = 1 - P_f(t) \quad (5)$$

In the above equations k_f and k_u represent the folding (u → f) and unfolding (f → u) rate constants, respectively, and P_u^0 the initial probability at time t = 0 for the unfolded state (in the present case P_u^0 = 15/17 ≈ 0.88). Equation 3 finally becomes

$$I(t) = I_u + (I_f - I_u)\left[\frac{k_f}{k_u + k_f} + \left(\frac{k_u}{k_u + k_f} - P_u^0\right) e^{-(k_u+k_f)t}\right] \quad (6)$$

The fit of the kinetic trace with eq 6 yields for k_f and k_u values of $1/k_f$ = 3500 ± 340 ns and $1/k_u$ = 13 000 ± 7400 ns, respectively. The correlation coefficient is 0.84, indicating that the two-state model used is rather accurate for describing the folding process of Peptide 1.

To further test the goodness of the fit, $1/k_f$ can be compared with the mean folding time calculated directly from the trajectories using structural order parameters. As mentioned above, the peptide is considered folded if $R \geqslant 4.8$, but this time the transitions to distorted hairpin configurations are excluded because, in the kinetic model, they occur within the unfolded state. The folding time distribution, and the corresponding monoexponential fit, is reported in Figure 4B. The mean folding time obtained from the fitting function is τ_f = 3304 ns, which is in good agreement with the value obtained from the kinetic model ($1/k_f$ = 3500 ± 340 ns).

From k_u and k_f the folded equilibrium probability can also be calculated, i.e., $P_f^{eq} = k_f/(k_u + k_f)$ = 0.78 ± 0.05. Comparison of the computed $1/k_f$ and P_f^{eq} with the corresponding experimental quantities (in the experiment $1/k_f \approx$ 800 ns and $P_f^{eq} \approx 0.5$)[8] shows that, although both are overestimated in the computational procedure, the equilibrium probability of the folded state is reasonably well reproduced and the mean folding time differs from the experimental value by less than an order of magnitude. The discrepancies are likely to result from the inaccuracies of the atomistic model used (e.g., the MD force field) and possibly from the uncertainties of the experimental data.

CONCLUSIONS

In atomistic simulations biological folding processes are typically monitored using structural order parameters. The folding mechanism emerging from a simulation may strongly depend on the choice of the parameter used, and therefore the reliability of the mechanism should be tested by calculation of the time-dependent experimental observable (as, for example, the amide I IR signal in the present case). The results of the present study show that the use of a common structural order parameter, such as R, might give an incomplete picture if not supported by the calculation of the time-dependent experimental observable. In fact, the mean folding time obtained using the order parameter (~1600 ns) is approximately half the value obtained from the IR signal (~3500 ns), the latter corresponding to transitions to fully folded hairpins only, while the former also includes misfolded hairpins.

The evaluation of an observable in complex processes often results from a huge number of events occurring within the observation time as, for example, in the measurement of kinetic rate constants. From a theoretical/computational point of view modeling such complex processes and the corresponding relevant observables is very challenging. The present PMM-MD approach was able to model the time-dependent spectroscopic signal as obtained by using an ensemble of molecules undergoing conformational transitions. To the best of our knowledge, this is the first time a "real" IR kinetic experiment has been performed in silico providing the complete kinetics and thermodynamics (k_f, k_u, and P_f^{eq}) of the folding process.

The present methodology, based on a fully quantum mechanical treatment of the chromophore vibrations and on extended phase space sampling to rigorously treat the chromophore-environment interaction, was shown, in the present and previous works, to be reliable in reproducing the main spectroscopic features of the amide I band in complex systems, such as amyloids,[29] extended β-sheets[31] and unfolded states of different peptides.[25,30,32] Moreover, the amide I signal originating from any residue of a peptide/protein chain may be isolated allowing the comparison with experimental isotope-labeled spectra.[29] Nevertheless, the PMM/MD methodology may not always reproduce the spectral fine structure. Many possible reasons for this exist, including inaccuracies in the basic quantum chemical calculations of the chromophore, approximations and errors of the classical force field used, the neglect of the vibrational coupling between the chromophore and solvent modes and the use of the dipolar approximation to model the chromophores interaction operator to be used in the excitonic coupling. However, the evaluation of the instantaneous perturbed chromophore quantum states within a rigorous statistical-mechanical treatment, based on extended MD sampling, allows the explicit and reliable reconstruction of the time-dependence of the observable. Moreover, the method, being based on first-principles, does not involve fitting to experiment in contrast to other computational approaches.[33−35]

The level of the theoretical computational approach described in this paper could be further improved by extending the definition of the QCs to include transient cluster structures, possibly involving solvent molecules and/or closely interacting peptide residues and modeling the excitonic coupling beyond the dipolar approximation. Such an extension of the method would allow the modeling of chromophore mode variations beyond the excitonic coupling utlilized in this paper, due to the peptide structural changes and/or solvent interactions.

ASSOCIATED CONTENT

Supporting Information

Theoretical details on the simulations performed. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: isabella.daidone@univaq.it.
*E-mail: andrea.amadei@uniroma2.it.

Notes

The authors declare no competing financial interest.

DOI: 10.1021/acs.jpcb.5b01477
J. Phys. Chem. B 2015, 119, 4849−4856

■ ACKNOWLEDGMENTS

We acknowledge the CINECA award IsC11$_{PHOSPHO}$ under the ISCRA initiative for the availability of high-performance computing resources and support. J.C.S. acknowledges support from the ORNL Adaptive Biosystems Imaging program of the United States Department of Energy.

■ REFERENCES

(1) Slayton, R. M.; Anfinrud, P. A. Time-Resolved Mid-Infrared Spectroscopy: Methods and Biological Applications. *Curr. Opin. Struct. Biol.* **1997**, *7*, 717−721.
(2) Zanni, M. T.; Hochstrasser, R. M. Two-Dimensional Infrared Spectroscopy: A Promising New Method for the Time Resolution of Structures. *Curr. Opin. Struct. Biol.* **2001**, *11*, 516−522.
(3) Kim, S.; Lim, M. Picosecond Dynamics of Ligand Interconversion in the Primary Docking Site of Heme Proteins. *J. Am. Chem. Soc.* **2005**, *127*, 5786−5787.
(4) Nienhaus, K.; Olson, J. S.; Franzen, S.; Nienhaus, G. U. The Origin of Stark Splitting in the Initial Photoproduct State of MbCO. *J. Am. Chem. Soc.* **2005**, *127*, 40−41.
(5) Huang, C.; Getahun, Z.; Zhu, Y.; Klemke, J.; DeGrado, W.; Gai, F. Helix Formation via Conformation Diffusion Search. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2788−2793.
(6) Bredenbeck, J.; Helbing, J.; Kumita, J. R.; Woolley, G. A.; Hamm, P. α-Helix Formation in a Photoswitchable Peptide Tracked from Picoseconds to Microseconds by Time-Resolved IR Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2379−2384.
(7) Ihalainen, J. A.; Bredenbeck, J.; Pfister, R.; Helbing, J.; Chi, L.; van Stokkum, I. H.; Woolley, G. A.; Hamm, P. Folding and Unfolding of a Photoswitchable Peptide from Picoseconds to Microseconds. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 5383−538.
(8) Xu, Y.; Oyola, R.; Gai, F. Infrared Study of the Stability and Folding Kinetics of a 15-Residue β-Hairpin. *J. Am. Chem. Soc.* **2003**, *125*, 15388−15394.
(9) Maness, S. J.; Franzena, S.; Gibbs, A. C.; Causgrove, T. P.; Dyer, R. B. Nanosecond Temperature Jump Relaxation Dynamics of Cyclic β-Hairpin Peptides. *Biophys. J.* **2003**, *84*, 3874−3882.
(10) Smith, A. W.; Tokmakoff, A. Probing Local Structural Events in Beta-Hairpin Unfolding With Transient Nonlinear Infrared Spectroscopy. *Angew. Chem., Int. Ed.* **2007**, *46*, 7984−7987.
(11) 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.
(12) Krimm, S.; Bandekar, J. Vibrational Spectroscopy and Conformation of Peptides, Polypeptides, and Proteins. *Adv. Protein Chem.* **1986**, *38*, 181−364.
(13) Torii, H.; Tasumi, M. Model Calculations on the Amide-I Infrared Bands of Globular Proteins. *J. Chem. Phys.* **1992**, *96*, 3379−3387.
(14) Lee, S. H.; Krimm, S. Ab Initio-Based Vibrational Analysis of α-poly(L-alanine). *Biopolymers* **1998**, *46*, 283−317.
(15) Moran, A.; Mukamel, S. The Origin of Vibrational Mode Couplings in Various Secondary Structural Motifs of Polypeptides. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 506−510.
(16) Brauner, J. W.; Flach, C. R.; Mendelsohn, R. A Quantitative Reconstruction of the Amide I Contour in the IR Spectra of Globular Proteins: From Structure to Spectrum. *J. Am. Chem. Soc.* **2005**, *127*, 100−109.
(17) Bour, P.; Keiderling, T. A. Vibrational Spectral Simulation for Peptides of Mixed Secondary Structure: Method Comparisons with the Trpzip Model Hairpin. *J. Phys. Chem. B* **2005**, *109*, 23687−23697.
(18) la Cour Jansen, T.; Dijkstra, A.; Watson, T.; Hirst, J.; Knoester, J. Modeling the Amide I Bands of Small Peptides. *J. Chem. Phys.* **2006**, *125*, 44312−44312−9.
(19) Smith, A. W.; Tokmakoff, A. Amide I Two-Dimensional Infrared Spectroscopy of β-Hairpin Peptides. *J. Chem. Phys.* **2007**, *126*, 045109−045109−11.
(20) Ham, S.; Cha, S.; Choi, J. H.; Cho, M. Amide I Modes of Tripeptides: Hessian Matrix Reconstruction and Isotope Effects. *J. Chem. Phys.* **2003**, *119*, 1451−1451−11.
(21) Hahn, S.; Ham, S.; Cho, M. Simulation Studies of Amide I IR Absorption and Two-Dimensional IR Spectra of β Hairpins in Liquid Water. *J. Phys. Chem. B* **2005**, *109*, 11789−11801.
(22) Choi, J. H.; Lee, H.; Lee, K. K.; Hahn, S.; Cho, M. Computational Spectroscopy of Ubiquitin: Comparison Between Theory and Experiments. *J. Chem. Phys.* **2007**, *126*, 045102−045102−14.
(23) Aschi, M.; Spezia, R.; Di Nola, A.; Amadei, A. A First Priciples Method to Model Perturbed Electronic Wavefunctions: The Effect of an External Electric Field. *Chem. Phys. Lett.* **2001**, *344*, 374−380.
(24) Amadei, A.; D'Abramo, M.; Daidone, I.; D'Alessandro, M.; Di Nola, A.; Aschi, M. Statistical Mechanical Modelling of Chemical Reactions in Complex Systems: The Kinetics of the Haem Carbon Monoxide Binding-Unbinding Reaction in Myoglobin. *Theor. Chem. Acc.* **2007**, *117*, 637−647.
(25) Daidone, I.; Aschi, M.; Zanetti-Polzi, L.; Nola, A. D.; Amadei, A. On the Origin of IR Spectral Changes Upon Protein Folding. *Chem. Phys. Lett.* **2010**, *488*, 213−218.
(26) Amadei, A.; Daidone, I.; Di Nola, A.; Aschi, M. Theoretical-Computational Modelling of Infrared Spectra in Peptides and Proteins: A New Frontier for Combined Theoretical-Experimental Investigations. *Curr. Opin. Struct. Biol.* **2010**, *20*, 155−161.
(27) Amadei, A.; Daidone, I.; Zanetti-Polzi, L.; Aschi, M. Modeling Quantum Vibrational Excitations in Condensed-Phase Molecular Systems. *Theor. Chem. Acc.* **2011**, *129*, 31−43.
(28) Gao, J.; Truhlar, D. G. Quantum Mechanical Methods for Enzyme Kinetics. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.
(29) Zanetti-Polzi, L.; Amadei, A.; Aschi, M.; Daidone, I. New Insight into the IR-Spectra/Structure Relationship in Amyloid Fibrils: A Theoretical Study on a Prion Peptide. *J. Am. Chem. Soc.* **2011**, *133*, 11414−11417.
(30) Zanetti-Polzi, L.; Daidone, I.; Anselmi, M.; Carchini, G.; Di Nola, A.; Amadei, A. Analysis of Infrared Spectra of β-Hairpin Peptides as Derived from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2011**, *115*, 11872−11878.
(31) Zanetti-Polzi, L.; Daidone, I.; Amadei, A. A Theoretical Reappraisal of Polylysine in the Investigation of Secondary Structure Sensitivity of Infrared Spectra. *J. Phys. Chem. B* **2012**, *116*, 3353−3360.
(32) Zanetti-Polzi, L.; Aschi, M.; Amadei, A.; Daidone, I. Simulation of the Amide I Infrared Spectrum in Photoinduced Peptide Folding/Unfolding Transitions. *J. Phys. Chem. B* **2013**, *117*, 12383−12390.
(33) Cheatum, C.; Tokmakoff, A.; Knoester, J. Signatures of β-Sheet Secondary Structures in Linear and Two-Dimensional Infrared Spectroscopy. *J. Chem. Phys.* **2004**, *120*, 8201−15.
(34) Lee, C.; Cho, M. Local Amide I Mode Frequencies and Coupling Constants in Multiple-Stranded Antiparallel β-Sheet Polypeptides. *J. Phys. Chem. B* **2004**, *108*, 20397−20407.
(35) Karjalainen, E.; Ravi, H.; Barth, A. Simulation of the Amide I Absorption of Stacked β-Sheets. *J. Phys. Chem. B* **2011**, *115*, 749−757.
(36) Santiveri, C. M.; Rico, M.; Jiménez, M. A. 13Cα and 13Cβ Chemical Shifts as a Tool to Delineate β-Hairpin Structures in Peptides. *J. Biomol. NMR* **2001**, *19*, 331−345.
(37) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *95*, 43−56.
(38) Jorgensen, W.; Tirado-Rives, J. The OPLS Potential Functions for Proteins: Energy Minimizations for Crystal of Cyclic Peptides and Cramin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(39) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Klein, R. I. M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 962.
(40) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N-log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

DOI: 10.1021/acs.jpcb.5b01477
J. Phys. Chem. B 2015, 119, 4849−4856

(41) 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.
(42) Brown, D.; Clarke, J. H. R. A Comparison of Constant Energy, Constant Temperature, and Constant Pressure Ensembles in Molecular Dynamics Dimulations of Atomic Liquids. *Mol. Phys.* **1984**, *51*, 1243−1252.
(43) Yang, S.; Onuchic, J. N.; García, A. E.; Levine, H. Folding Time Predictions from All-Atom Replica Exchange Simulations. *J. Mol. Biol.* **2007**, *372*, 756−763.
(44) Thukral, L.; Smith, J.; Daidone, I. Common Folding Mechanism of a β-Hairpin Peptide via Non-native Turn Formation Revealed by Unbiased Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2009**, *131*, 18147−18152.

DOI: 10.1021/acs.jpcb.5b01477
J. Phys. Chem. B 2015, 119, 4849−4856