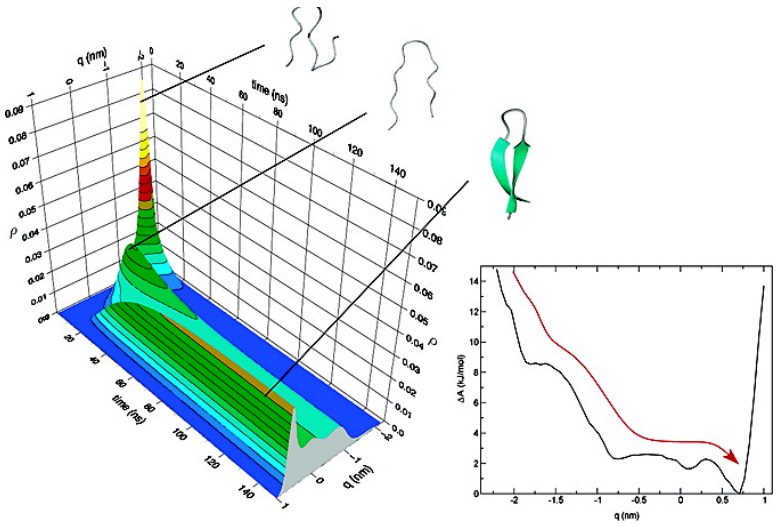
Article

Theoretical Characterization of α-Helix and β-Hairpin Folding Kinetics

Isabella Daidone, Marco D'Abramo, Alfredo Di Nola, and Andrea Amadei

J. Am. Chem. Soc., **2005**, 127 (42), 14825-14832• DOI: 10.1021/ja053383f • Publication Date (Web): 30 September 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 10 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 09/30/2005

Theoretical Characterization of α-Helix and β-Hairpin Folding Kinetics

Isabella Daidone,† Marco D'Abramo,† Alfredo Di Nola,† and Andrea Amadei*,‡

Contribution from the Department of Chemistry, University of Rome "La Sapienza", P.le Aldo Moro 5, 00185 Rome, Italy, and Dipartimento di Scienze e Tecnologie Chimiche, University of Rome "Tor Vergata", via della Ricerca Scientifica 1, I-00133 Rome, Italy

Received May 24, 2005; E-mail: andrea.amadei@uniroma2.it

Abstract: By means of the conformational free energy surface and corresponding diffusion coefficients, as obtained by long time scale atomistic molecular dynamics simulations (μs time scale), we model the folding kinetics of α-helix and β-hairpin peptides as a diffusive process over the free energy surface. The two model systems studied in this paper (the α-helical temporin L and the β-hairpin prion protein H1 peptide) exhibit a funnel-like almost barrierless free energy profile, leading to nonexponential folding kinetics matching rather well the available experimental data. Moreover, using the free energy profile provided by Muñoz et al. [Muñoz et al. *Nature* **1997**, *390*: 196−199], this model was also applied to reproduce the two-state folding kinetics of the C-terminal β-hairpin of protein GB1, yielding an exponential folding kinetics with a time constant (∼5 μs) in excellent agreement with the experimentally observed one (∼6 μs). Finally, the folding kinetics obtained by solving the diffusion equation, considering either a one-dimensional or a two-dimensional free energy surface, are also compared in order to understand the relevance of the possible kinetic coupling between conformational degrees of freedom in the folding process.

Introduction

According to the energy landscape theory,[1−5] protein folding kinetics can be viewed as a diffusive process over a free energy hypersurface defined by a few collective coordinates describing protein conformational transitions. The free energy gradient, determined by the competition between energy and entropy, determines the average drift up or down the funnel. In such a folding funnel the presence of free energy barriers, depending on the roughness of the surface, may provide a folding time of milliseconds or more. However, in many proteins a faster folding kinetics is observed[6,7] (microseconds time scale) that can be essentially described as flowing downhill to the folded state without crossing significant free energy barriers.[8] Small α-helices and β-hairpins, the basic structural elements of proteins, exhibit greatly simplified free energy landscapes compared to those of proteins[9] and their typical folding time scale ranges from hundreds of nanoseconds (α-helix) to a few microseconds (β-hairpin) (for a review, see ref 10). Most experimental and theoretical α-helix folding studies have been carried out on alanine-rich peptides, due to their high propensity to form α-helices.[11−13] The kinetics of formation of the more complex β-hairpin motif has also been addressed, although not to the same extent. The first β-hairpin forming peptide to be characterized in thermodynamic and kinetic studies was the 16-residue C-terminal hairpin from the GB1 protein.[14] This peptide exhibits a two-state behavior with a folding mean time of about 6 μs at room temperature. Only quite recently a very fast folding β-hairpin has been reported, the 15-residue peptide I, with a folding mean time of only 0.8 μs.[15]

Molecular dynamics (MD) simulations, one of the most popular computational tools, can access simulation times up to ∼1 μs for a solvated peptide, thus providing direct information on peptide folding processes.[16−18] However, the typical amount of folding/unfolding transitions in such simulations, although providing a reasonable thermodynamic characterization, is still not sufficient to obtain a reliable kinetic evaluation. Therefore direct folding kinetics studies, even of peptides, are still challenging and require the use of simplified models[19−23] or massive parallel computing to enable the generation of large numbers of independent, although still short, trajectories.[24,25]

† University of Rome "La Sapienza".
‡ University of Rome "Tor Vergata".

(1) Dill, K.; Chan, H. *Nat. Struct. Biol.* **1997**, *4*, 10−19.
(2) Onuchic, J.; Nymeyer, H.; Garcia, A.; Chahine, J.; Socci, N. D. *Adv. Protein Chem.* **2000**, *53*, 87−152.
(3) Bryngelson, J.; Wolynes, P. *J. Phys. Chem.* **1989**, *93*, 6902−6915.
(4) Plotkin, S.; Onuchic, J. *Q. Rev. Biophys.* **2002**, *35*, 111−167.
(5) Plotkin, S.; Onuchic, J. *Q. Rev. Biophys.* **2002**, *35*, 205−286.
(6) Jackson, S.; Fersht, A. *Biochemistry* **1991**, *30*, 10428−10435.
(7) Jackson, S. *Fold. Des.* **1998**, *3*, R81−R91.
(8) Garcia-Mira, M.; Sadqi, M.; Fischer, N.; Sanchez-Ruiz, J.; Muñoz, V. *Science* **2002**, *298*, 2191−2195.
(9) Brooks, C., III; Case, D. *Chem. Rev.* **1993**, *93*, 2487−2502.
(10) Kubelka, J.; Hofrichter, J.; Eaton, W. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76−88.
(11) Gilmanshin, R.; Williams, S.; Callender, R.; Woodruff, W.; Dyer, R. *Biochemistry* **1997**, *36*, 15006−15012.
(12) García, G. H. A.; Garde, S. *Phys. Rev. Lett.* **2000**, *85*, 2637−2640.
(13) Huang, C.; Getahun, Z.; Zhu, Y.; Klemke, J.; DeGrado, W.; Gai, F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2788−2793.
(14) Muñoz, V.; Thompson, P.; Hofrichter, J.; Eaton, W. *Nature* **1997**, *390*, 196−199.
(15) Xu, Y.; Oyola, R.; Gai, F. *J. Am. Chem. Soc.* **2003**, *125*, 15388−15394.
(16) Daidone, I.; Amadei, A.; Di Nola, A. *Proteins: Struct., Funct., Genet.* **2005**, *59*, 510−518.
(17) van Gunsteren, W.; Bürgi, R.; Peter, C.; Daura, X. *Angew. Chem., Int. Ed.* **2001**, *40*, 351−355.
(18) Gnanakaran, S.; Nymeyer, H.; Portman, J.; Sanbonmatsu, K.; García, A. *Curr. Opin. Struct. Biol.* **2003**, *13*, 168−174.

On the other hand, several theoretical models have been developed[12,26,27] in order to use short time scale trajectories, which can be even shorter than the time required to evolve from an unfolded state to a fully folded one, to extrapolate long time behaviors. In such a way information about folding rates and mechanisms can be provided and compared with experimental observations.

In the present study, inspired by the energy-landscape theory,[19,28] we model peptide folding as a diffusion over the free energy surface defined by a reduced number of degrees of freedom and we choose as reaction coordinates the essential (conformational) eigenvectors as provided by the essential dynamics (ED) analysis[29,30] performed on equilibrium MD simulations at room temperature. Such full atomistic long time scale trajectories of the peptide in explicit water are used to obtain the free energy surface and the corresponding diffusion coefficient, to be used in the diffusion equation utilized in our model. Hence, the complete kinetics and folding time constant can be evaluated.

This model is here applied to study the folding kinetics of an α-helix, the 13 residues temporin L peptide, and a β-hairpin, the 14 residues syrian hamster prion protein H1 peptide, performing long time scale (∼300 ns and ∼1.1 μs for the α-helix and β-hairpin, respectively) MD trajectories.[31] Given the lack of experimentally derived folding times for our peptides, to evaluate the reliability of our kinetic model, we test such a diffusive model on the GB1 β-hairpin using the experimentally based free energy profile provided by Muñoz et al.[14] The computed folding time results are in excellent agreement with the experimental one. Finally, the folding kinetics obtained by solving the diffusion equation considering either a one-dimensional or a two-dimensional free energy surface are also compared in order to understand the relevance of the possible kinetic coupling between conformational degrees of freedom in the folding process.

Theory

Protein folding is a collective self-organization process that occurs by a multiplicity of routes down a folding funnel, hence making it very hard to devise the proper reaction coordinates to describe such a high-dimensional complex system. To utilize a proper set of coordinates to describe the folding kinetics, we need to identify degrees of freedom such that all the orthogonal coordinates are well equilibrated during the folding relaxation. Hence, different coordinates could be used depending on the exact definition of the thermodynamic states characterizing the folding process (i.e., unfolded, folded, and intermediate states) and the initial conditions of the system (i.e., the coordinates equilibrated at the beginning of the process). In principle, if these reaction coordinates are properly defined, we may assume that their kinetic relaxation is described by a diffusive process under an external potential given by the free energy.[19,28] It has to be noted that a certain variation of the folding free energy profile, and hence of the kinetics, is also possible as a consequence of the choice of the folding coordinates and of the orthogonal planes used to obtain the corresponding free energy. The use of suboptimal reaction coordinates to describe the folding kinetics would result in an inaccurate model of the process, based on a set of thermodynamic conformational states not fully corresponding to the experimentally observed ones.

To construct a kinetic model for the folding process, we make the following simplifications.

(1) We consider the folding process as properly described by a single backbone collective degree of freedom q (the first essential eigenvector, ev 1, which is the slowest relaxing backbone collective coordinate describing the most relevant conformational transitions) and a set of other coordinates defining the secondary structure of the peptide (see Methods for the definition of secondary structure states and for the evaluation of essential coordinates).

(2) The rate constants for secondary structure interconversions provided by the analysis of the MD trajectories (see Methods) invariably show for each secondary structure interconversion and essential coordinate position a mean life up to a few picoseconds. Hence, we may safely consider such interconversions as instantaneously equilibrated during the diffusion along ev 1. In fact, given the diffusion coefficient evaluated for the first essential coordinate (∼10^{-5} nm^2/ps), the diffusion over the accessible range (few nanometers) requires several nanoseconds.

(3) We also assume the diffusion coefficient as independent of the essential coordinate position, and we neglect the initial kinetic behavior characterized by a time dependence of the apparent diffusion coefficient (vide infra).

We then obtain for each secondary structure state the corresponding probability in time as

$$P_{\mathcal{l}}(t) = \int \rho(q, t)\Omega_{\mathcal{l}}(q)\, dq \tag{1}$$

$$\Omega_{\mathcal{l}}(q) = \frac{e^{-\beta\Delta A_{\mathcal{l}}(q)}}{\sum_{\mathcal{l}'} e^{-\beta\Delta A_{\mathcal{l}'}(q)}} \tag{2}$$

where $\Omega_{\mathcal{l}}(q)$ is the equilibrium fraction of each secondary structure state, defined by the index vector $\mathcal{l}$, as a function of the position q along the essential eigenvector (associated to the free energy $\Delta A_{\mathcal{l}}(q)$), and the probability density $\rho(q, t)$ is obtained solving the diffusion equation (DE)

$$\frac{\partial\rho(q, t)}{\partial t} = \frac{\partial}{\partial q}\left[D\left(\frac{\partial\rho(q, t)}{\partial q} + \frac{\partial\Delta A(q)}{\partial q}\frac{\rho(q, t)}{RT}\right)\right] \tag{3}$$

$$\Delta A(q) = -RT \ln\left[\frac{\rho_{eq}(q)}{\rho_{eq}(q_{ref})}\right] \tag{4}$$

(19) Socci, N.; Onuchic, J.; Wolynes, P. *J. Chem. Phys.* **1996**, *104*, 5860−5868.
(20) Klimov, D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2544−2549.
(21) Favrin, G.; Irbäck, A.; Samuelsson, B.; Wallin, S. *Biophys. J.* **2003**, *85*, 1457−1465.
(22) Zhou, Y.; Zhang, Q.; Stell, G.; Wang, J. *J. Am. Chem. Soc.* **2003**, *125*, 6300−6305.
(23) Ulmschneider, J.; Jorgensen, W. *J. Am. Chem. Soc.* **2004**, *126*, 1849−1857.
(24) Snow, C.; Nguyen, H.; Pande, V.; Gruebele, M. *Nature* **2002**, *420*, 102−106.
(25) Snow, C.; Zagrovic, B.; Pande, V. S. *J. Am. Chem. Soc.* **2002**, *124*, 14548−14549.
(26) Bolhuis, P. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12129−12134.
(27) Swope, W.; Pitera, J.; Suits, F.; Pitman, M.; Eleftheriou, M.; Fitch, B.; Germain, R.; Rayshubskiy, A.; Ward, T.; Zhestkov, Y.; Zhou, R. *J. Phys. Chem. B* **2004**, *108*, 6582−6594.
(28) Bryngelson, J.; Onuchic, J.; Socci, N.; Wolynes, P. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167−195.
(29) García, A. *Phys. Rev. Lett.* **1992**, *66*, 2696−2699.
(30) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412−425.
(31) The structure of the amyloidogenic H1 peptide has not been resolved at high resolution, since in aqueous solution it aggregates very rapidly to form β-sheet rich fibrils. However, it is believed to adopt a β-hairpin conformation when in the aggregated form.[16,38,43] Peptides are widely distributed in nature, being a crucial component in the innate immunity of both invertebrates and vertebrates. The antimicrobial temporin L peptide is found to adopt an α-helical conformation when bound to liposomes; in water its helical propensity is still present, although highly reduced.[44]

where $\rho_{eq}(q)$ is the equilibrium probability density in q and q_{ref} is a reference position (typically the overall free energy minimum). Note that $\Delta A(q)$ and $\Omega_{/}(q)$ can be obtained by an extended MD simulation providing a reasonable convergence of the configurational sampling, while in order to obtain D a more complex procedure is required.

To evaluate the free diffusive kinetics along ev 1 we use the extended model described in a previous paper,[16] based on modeling the velocity autocorrelation function relaxation via a multiexponential decay. According to this model the diffusion along the essential degrees of freedom is described by a dual regime: a fast diffusion (typically up to a few picosecond) switching multiexponentially to a slower one. In that paper[16] we considered the β-hairpin free diffusion kinetics in different regions (with an almost constant free energy) of the plane defined by the first two essential eigenvectors. We investigated the free diffusive behavior up to 100 ps, which required a biexponential decay to be properly modeled in each essential space region. In the present paper we consider only the first essential eigenvector, and since the corresponding one-dimensional free energy landscape is less corrugated than the bidimensional surface, we are able to characterize the free diffusion process within a larger region. Thus, a better statistics is afforded, allowing us to extend our investigation over a much longer time (up to 1 ns), showing that a third much slower relaxation mode is necessary to accurately describe the free diffusion over such an extended time interval.

Hence, for the two peptides studied in this paper we utilize such a model based on a triexponential switching from the fast to the slow diffusion regime, providing for the mean square displacement

$$\langle \Delta q^2(t) \rangle \cong 2D_\infty t + 2[D_0 - A_1]\tau_1[1 - e^{-t/\tau_1}] + 2[D_0 - A_2]\tau_2[1 - e^{-t/\tau_2}] + 2[D_0 - A_3]\tau_3[1 - e^{-t/\tau_3}] \quad (5)$$

where τ_1, τ_2, and τ_3 are the "relaxation times" of the three switching modes, A_1, A_2, and A_3 are three parameters defined by the integrals of the velocity autocorrelation function (see Appendix of Daidone et al. 2005[16]), D_0 is the fast diffusion constant providing the initial diffusion process, and D_∞ is the long time diffusion constant characterizing the diffusion process beyond the modes relaxation. The obtained D_∞ for both peptides are used as diffusion constants into the corresponding DE to evaluate the folding kinetics, which therefore can be considered accurate only for $t > \tau_3$, i.e., beyond the slowest relaxation time. Note that in the whole time interval considered (corresponding for both peptides to about 10 times τ_3) only a single additional relaxation mode is detected, thus suggesting that three relaxation modes are probably accurately describing the complete velocity autocorrelation function relaxation; i.e., no slower relaxation modes are present.

Finally, to investigate the possible effects of the conformational coordinate coupling in the folding kinetics, we apply, for the β-hairpin peptide, the same procedure also in the plane defined by the first two essential eigenvectors. This was accomplished by solving the corresponding two-dimensional DE

$$\frac{\partial \rho(q_1, q_2, t)}{\partial t} = \frac{\partial}{\partial q_1}\left[D_1\left(\frac{\partial \rho(q_1, q_2, t)}{\partial q_1} + \frac{\partial \Delta A(q_1, q_2)}{\partial q_1}\frac{\rho(q_1, q_2, t)}{RT}\right)\right] + \frac{\partial}{\partial q_2}\left[D_2\left(\frac{\partial \rho(q_1, q_2, t)}{\partial q_2} + \frac{\partial \Delta A(q_1, q_2)}{\partial q_2}\frac{\rho(q_1, q_2, t)}{RT}\right)\right] \quad (6)$$

where we assume a diagonal diffusion coefficient matrix and the bidimensional free energy surface, $A(q_1, q_2)$, is modeled using a combination of bidimensional Gaussian functions to reproduce the free energy as provided by the MD simulations. We use $D_1 = D_2 = D$, since the diffusion coefficients along the two eigenvectors resulted in being almost identical (data not shown).

Methods

MD Simulations Protocol. MD simulations, in the NVT ensemble, with fixed bond lengths[32] and a time step of 2 fs for numerical integration were performed with the GROMACS software package[33] and with the GROMOS96 force field.[34] Water was modeled by the simple point charge (SPC) model.[35] A nonbond pairlist cutoff of 9.0 Å was used, and the pairlist was updated every four time steps. The long-range electrostatic interactions were treated with the particle mesh Ewald method.[36] The isokinetic temperature coupling[37] was used to keep the temperature constant at 300 K. The peptides, in their different starting conformations, were solvated with water and placed in a periodic truncated octahedron large enough to contain a peptide molecule and ~1.0 nm of solvent on all sides. One and three negative counterions (Cl^-) for the H1 peptide and the temporin L, respectively, were added by replacing the corresponding number of water molecules to achieve a neutral condition. The side chains were protonated as to reproduce a pH of about 7. The N-terminal and C-terminal of the H1 peptide were amidated and acetilated, respectively, to reproduce the experimental conditions.[38]

Two all atom MD simulations in explicit water at 300 K of the H1 peptide (MKHMAGAAAAGAVV), for a total of ~1.1 μs of simulation time, were carried out: 240 ns starting from the α-helix conformation obtained from the simulation in a 30% (v/v) TFE/water mixture of a previous work[39] and 850 ns starting from the β-hairpin conformation observed during the 240 ns simulation, using a new set of initial velocities. The temporin L peptide (FVQWFSKFLGRIL) was simulated at 300 K for 300 ns starting from an ideal α-helix conformation.

Essential Dynamics Analysis. The principles of the ED analysis are described in detail elsewhere.[30,40] Briefly, from all the structures generated by equilibrium MD simulations, a covariance matrix of positional fluctuations (C_α only) is built and diagonalized. Eigenvectors

(32) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(33) van der Spoel, D.; van Drunen, R.; Berendsen, H. J. C. *GROningen MAchine for Chemical Simulation;* Department of Biophysical Chemistry, BIOSON Research Institute: Nijenborgh 4 NL-9717 AG Groningen, 1994.
(34) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide;* Hochschulverlag AG an der ETH Zürich: Zürich, 1996.
(35) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(36) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(37) Brown, D.; Clarke, J. H. R. *Mol. Phys.* **1984**, *51*, 1243−1252.
(38) Nguyen, J.; Baldwin, M. A.; Cohen, F. E.; Prusiner, S. B. *Biochemistry* **1995**, *34*, 4186−4192.
(39) Daidone, I.; Simona, F.; Roccatano, D.; Broglia, R. A.; Tiana, G.; Colombo, G.; Di Nola, A. *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 198−204.
(40) de Groot, B. L.; Amadei, A.; Scheek, R. M.; van Nuland, N. A.; Berendsen, H. J. C. *Proteins: Struct., Funct., Genet.* **1996**, *26*, 314−322.

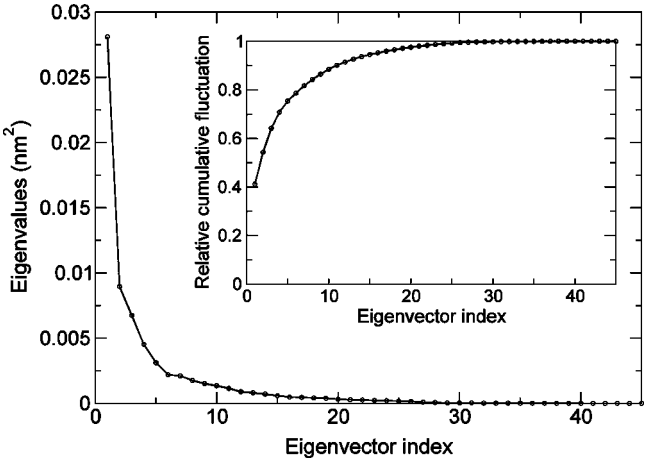


Figure 1. Eigenvalues, in decreasing order of magnitude, obtained from the H1 peptide Cα coordinates covariance matrix as a function of eigenvectors index. The corresponding relative cumulative positional fluctuation is given in the inset.

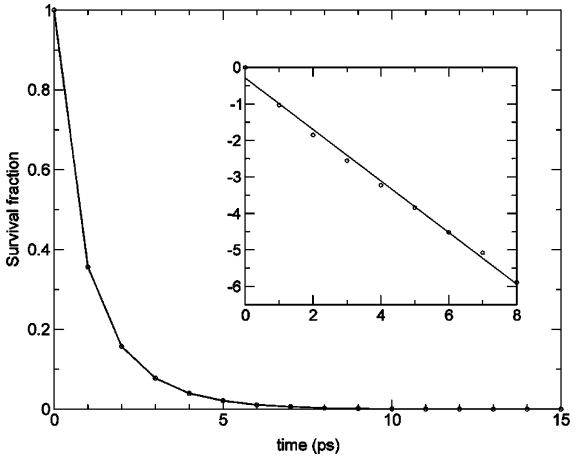


Figure 2. Example of a $\mathcal{I} \rightarrow \mathcal{I}'$ secondary structure interconversion. The survival fraction of state $\mathcal{I}$ is reported as a function of time. In the inset we show the logarithm of the curve and the corresponding linear fit.

are directions in configurational space, and the corresponding eigenvalues indicate the mean square fluctuations along these axes. The procedure corresponds to a linear multidimensional least-squares fitting of a trajectory in configurational space. Sorting the eigenvectors by the size of the eigenvalues shows that the configurational space can be divided in a low dimensional (essential) subspace in which most of the positional fluctuations are confined and a high dimensional (near-constraints) subspace in which merely small vibrations occur.

In Figure 1 the eigenvalues obtained from the Cα coordinate covariance matrix for the β-hairpin peptide are reported as a function of the eigenvector index and are ordered in descending order of magnitude. The corresponding relative cumulative positional fluctuation (with respect to the total positional fluctuation) is given in the inset. The first and second essential eigenvectors, i.e., the ones with the two largest eigenvalues, account for 45% and 15% of the overall positional fluctuation, respectively (inset of Figure 1). Taken together they define 60% of the overall positional fluctuation, hence describing the most relevant conformational transitions of the peptide backbone. The ED analysis performed on temporin L provides similar results; in particular the contribution to the overall positional fluctuation of the first two essential eigenvectors is almost identical to the H1 peptide one (data not shown).

Secondary Structure Definition. To predict β-hairpin and α-helix folding times, we first defined secondary structure states. For the β-hairpin we defined eight secondary structure states corresponding to eight different hydrogen bond (HB) patterns. The status of the three pairs of HB, characteristic of the fully folded hairpin, is indicated by a tridimensional index vector, $\mathcal{I}$, with components given by either 1 if the HB pair is formed or 0 if it is absent. The first component of this vector represents the status of the HB pair closest to the termini of the strands, and the last component represents the status of the pair closest to the turn of the folded hairpin. Thus 000 represents the condition with no HB pairs formed, 001, the condition with only the turn HB pair formed, 111, the condition with all the three HB pairs formed, and so on. This characterization results in $2^3 = 8$ possible HB states and 56 possible transitions. For temporin L we defined three secondary structure states according to the number of helical, either α or helix 5, turns indicated by the one-dimensional index l: the unfolded state (no helical turns, $l = 0$), the partially folded state (more than 1 and less than 2 helical turns, $l = 1$), and the completely folded state (2 or more helical turns, $l = 2$). Note that we did not explicitly consider intermediate states for the partially folded condition as only a limited population with 1.5 turns is present (see Results).

To evaluate secondary structure interconversion rate constants we considered each $\mathcal{I} \rightarrow \mathcal{I}'$ transition kinetics as obtained by all the subtrajectories (subparts) starting in state $\mathcal{I}$ and directly converting in state $\mathcal{I}$.

For each subpart the time interval required for the $\mathcal{I} \rightarrow \mathcal{I}'$ transition (transition time) was monitored. Hence, the secondary structure interconversion was obtained by plotting the fraction of subtrajectories that has not yet converted (survival fraction) as a function of time. An example, for a β-hairpin peptide interconversion, is given in Figure 2. In the inset, the logarithm of the curve and the corresponding linear fit used to evaluate the interconversion time constant are shown.

Free Energy and Diffusion Coefficient Evaluation. To obtain the free energy along ev 1, the MD structures sampled every 1 ps were projected onto 50 grid cells used to divide the overall accessible range. For every cell the number of points were counted, and the relative probability density, $\rho_{eq}(q)$, was calculated as well as the equilibrium fraction, $\Omega_{\mathcal{I}}(q)$, of each secondary structure state. Finally, we chose as the reference state the grid cell with the highest probability density $\rho_{eq}(q_{ref})$, i.e., the cell corresponding to the overall free energy minimum, and $\Delta A(q) = -RT \ln \rho_{eq}(q)/\rho_{eq}(q_{ref})$ was fitted by a polynomial function to be used into the DE (eq 3). To check the effect of different grid spacings on the thermodynamic properties, the same types of free energy landscapes were constructed using different numbers of cells: 40, 50, 60, and 100 (data not shown). Interestingly, all the different grids provided similar free energy landscapes with the same free energy maximum variation (~14 kJ/mol for the β-hairpin and ~12.5 kJ/mol for the α-helix), the surface being slightly more corrugated on going from the grid with a lower cell density (40) to the denser one (100). A similar procedure was employed to obtain the free energy surface in the plane of the first two essential eigenvectors of the β-hairpin. In this case a 20×20 grid was used to divide this plane in 400 cells and a combination of bidimensional Gaussian functions, fitting the free energy surface, was used in the bidimensional DE (eq 6).

To obtain the diffusion coefficient along ev 1, different regions of the eigenvector range, where no relevant free energy gradient is encountered, were analyzed separately. To generate an ensemble of independent trajectories we used all the trajectory fragments starting within each selected region providing the ensemble mean square displacement, from the corresponding initial point, as a function of time. To increase the statistics, we tested the procedure using increasing size of the chosen region, from 0.1 nm to almost all the accessible range (2.0 and 1.5 nm for the β-hairpin and α-helix, respectively). We found rather similar values of the diffusion coefficients and relaxation times. The values reported in the results section are obtained considering an interval of 2.0 and 1.5 nm for the β-hairpin and α-helix, respectively, both centered at −0.5 nm of the corresponding ev 1.

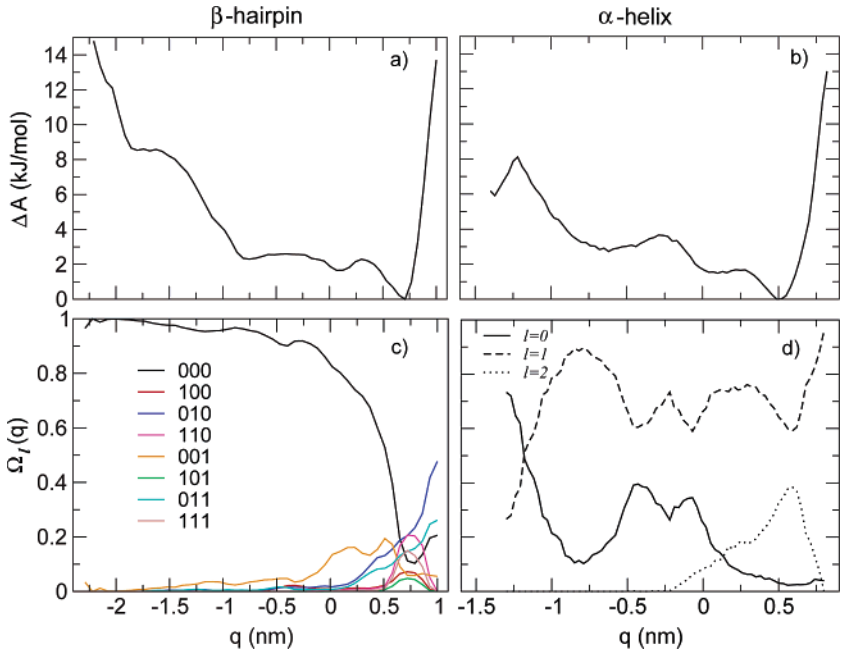


Figure 3. Upper panels: free energy profile, $\Delta A(q)$, as a function of the position q along the first essential eigenvector for the β-hairpin H1 peptide (a) and for the helical temporin L peptide (b). Lower panels: equilibrium fraction, $\Omega_{\mathcal{l}}(q)$, of each secondary structure state as a function of the position q along the first essential eigenvector. The eight secondary structure states of the β-hairpin, defined by the index vector $\mathcal{l}$, are reported in (c). The three secondary structure states of the helix, indicated by the one-dimensional index l, are reported in (d).

All the curve fittings were obtained using the graphing tool Xmgrace (http://plasma-gate.weizmann.ac.il/Grace/doc/UsersGuide.html), which makes use of the Levenberg−Marquardt algorithm and provides χ^2 and correlation coefficient evaluations. Moreover, we also evaluated the noise for the model parameters, calculating their standard deviations, σ, over a sample of n subtrajectories and then extrapolating for the complete statistical sample:

$$\sigma = \left(\frac{\sum_{i=1}^{n} (a_i - \bar{a})^2}{(n-1)n} \right)^{1/2} \tag{7}$$

$$\bar{a} = \left(\frac{\sum_{i=1}^{n} a_i}{n} \right) \tag{8}$$

where a_i is the generic parameter evaluated in the ith subtrajectory. Note that such a noise evaluation is based on the approximation that the parameters obtained using the whole trajectory are equivalent to the ones obtained averaging over the a_i values. In the present case we used three independent subtrajectories which resulted in being a good compromise between the statistics within each subtrajectory and the sample size used in the last equations, i.e., n.

Finally, the DE were solved numerically using a finite-difference scheme.

Results

Structural and Thermodynamic Analyses. A structural analysis and an accurate thermodynamic characterization of the folding of the H1 peptide at room temperature were provided in the previous articles.[16,39] From long time scale (1.1 μs) equilibrium MD simulations in aqueous solvent, the H1 peptide, initially modeled as an α-helix, preferentially adopts a rather stable β-hairpin structure and several unfolding/refolding events are observed. Such a β-hairpin, either partially or completely folded, is populated for ∼30% of the total simulation time; α-helical conformations are sampled only a few times, with an equilibrium probability less than 1%. For the rest of the time the peptide is essentially unfolded. For temporin L a structural analysis of the 300 ns simulation reveals that the most populated conformation (∼63% of time) is the one with just one helical turn, whereas longer helical stretches are sampled for ∼21% (∼7% and ∼14% with 1.5 and more than 2 turns, respectively). The unfolded state is populated for ∼16%.

The free energy landscapes of the H1 peptide and temporin L along the first essential eigenvector are reported in Figure 3a and b, respectively.

The two peptides share similar profiles, characterized by downhill, almost barrierless, surfaces with a maximum free energy change of ∼14 and ∼12.5 kJ/mol, respectively. To evaluate the equilibrium distribution of the different secondary structure states along ev 1, we calculated the corresponding equilibrium fraction, $\Omega_{\mathcal{l}}(q)$, where $\mathcal{l}$ indicates the secondary structure state. In Figure 3c and d we report such a property for the β-hairpin and α-helix peptide, respectively, as a function of the position along ev 1. For the H1 peptide the different β-hairpin structures (partially or completely folded) are essentially present in the overall free energy minimum region (around 0.75 nm), whereas the unfolded condition, largely dominant elsewhere, here drops to about 10%. For temporin L, the state characterized by a partially folded condition is dominant within the ev 1 accessible range, except in the left limit (around −1.3) where the probability of the unfolded state is maximum. The completely folded state ($r \geq 2$, $l = 2$) is essentially confined in the overall free energy minimum region (around 0.5 nm).

Finally, to evaluate the convergence of the free energy profile along ev 1, we used the time dependence of the free energy root-mean-square deviation as obtained averaging the free energy deviations over the cells along the first eigenvector and considering as the expectation (reference) profile the one

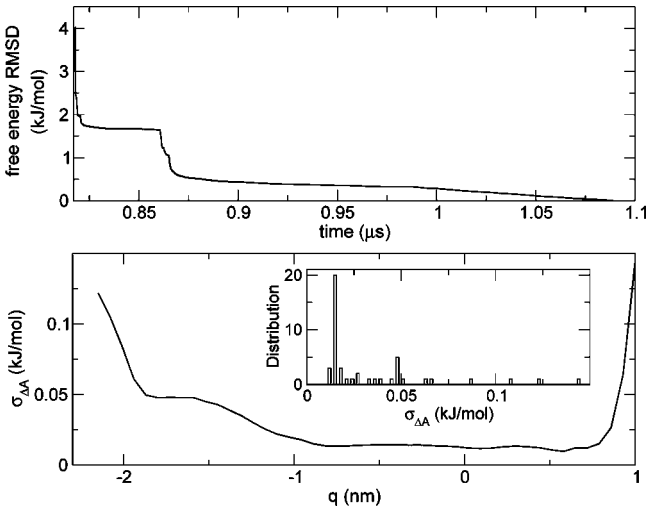


Figure 4. (Panel a) Time dependence of the β-hairpin free energy root-mean-square deviation as obtained averaging the free energy deviations over the cells along the first eigenvector and considering as the expectation (reference) profile the one calculated using the complete trajectory. (Panel b) Free energy standard deviation, $\sigma_{\Delta A}$, of the β-hairpin free energy profile. In the inset the corresponding distribution is reported.

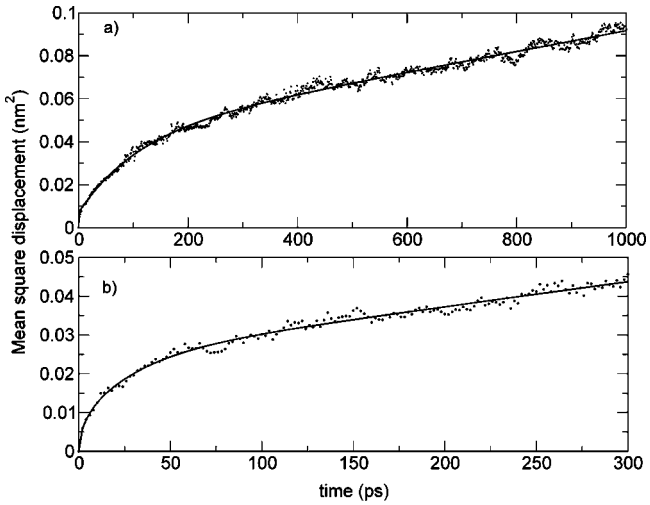


Figure 5. Mean square displacement, as a function of time, along the first essential eigenvector for the H1 peptide (a) and the temporin L peptide (b). The theoretical model (solid line) is parametrized fitting the simulation data (circles).

Table 1. Diffusion Constants and "Relaxation Times" over the First Essential Eigenvector for the Two Peptides[a,b]

	D_0 nm² ps⁻¹	D_∞ nm² ps⁻¹	τ_1 ps	τ_2 ps	τ_3 ps
α-helix	0.026 ± 0.001	$3.2 \times 10^{-5} \pm 0.5 \times 10^{-5}$	<1	4.3 ± 1	30.6 ± 2
β-hairpin	0.026 ± 0.001	$2.4 \times 10^{-5} \pm 0.5 \times 10^{-5}$	<1	7.8 ± 1	113.4 ± 2

[a] D_0 is the short-time diffusion constant; D_∞, the long-time diffusion constant; and τ_1, τ_2, and τ_3, the "relaxation times" of the three switching modes (see Theory section) [b] The noise indicated, as obtained according to the methods section, corresponds to a standard deviation.

calculated using the complete trajectory. In Figure 4a we report this property for the β-hairpin, showing a mean deviation within 1 kJ/mol beyond $\sim$0.86 μs (the α-helical peptide shows a similar trend reaching a mean deviation within 1 kJ/mol beyond $\sim$200 ns). In Figure 4b we also show the statistical noise (free energy standard deviation, $\sigma_{\Delta A}$) of the β-hairpin free energy profile and, in the inset, the corresponding distribution. From the figure it is clear the rather small statistical errors affecting the free energy values, which are larger on the edges of the accessible ev 1 range because of the corresponding lower sampling. Almost identical noise profile and distribution are also observed for the α-helical peptide. Such results indicate that for both peptides the conformational sampling is adequate to obtain rather reliable free energy profiles.

Diffusion Coefficients Evaluation. To evaluate the diffusion coefficients to be used in the DE we used the model described in the theory section based on modeling the velocity auto-correlation function relaxation via a triexponential decay. In Figure 5a and b we show, for the β-hairpin and the α-helix, respectively, the comparison between the mean square displacements obtained by simulations and the theoretical models. Due to the shorter simulation time of the α-helix, the time interval for the simulated mean square displacements is shorter (300 ps instead of 1000 ps). The plots clearly show the high accuracy of the model used in the whole time range. Note that the corresponding χ^2 values are in the range $10^{-5}-10^{-4}$ with correlation coefficients always higher than 0.995 and full fitting convergence is achieved within 2000 steps.

The diffusion coefficients and the "relaxation times" obtained from the model are reported in Table 1. Interestingly, the long time diffusion constants, D_∞, are rather similar for the two peptides, while the relaxation times, providing the time required to switch from the fast to the slow diffusion regime, are significantly different. The slowest relaxation mode, characterized by τ_3, is almost four times faster for the α-helical peptide. Furthermore, it is worth noting that τ_1 and τ_2 of the β-hairpin, as obtained in the present work using a triexponential decay and the longer time interval (1000 ps), almost coincide with the plane averaged values obtained in the previous paper,[16] where a biexponential decay and 100 ps time interval were used ($\sim$1 ps and $\sim$10 ps for τ_1 and τ_2, respectively).

Folding Kinetics. For both peptides we (numerically) solved the corresponding DE (eq 3), starting with the system being completely unfolded, i.e., in a position of ev 1 closed to the extreme left limit of the free energy profile (Figure 3a and b) where the equilibrium probability of the folded state is minimal. In Figure 6 we show the β-hairpin peptide probability density $\rho(q, t)$, as obtained by the DE.

This figure clearly indicates that the β-hairpin barrierless funneled free energy landscape provides a diffusive relaxation occurring within several tens of nanoseconds, involving a metastable kinetic intermediate. Interestingly, such an intermediate is virtually only characterized by the unfolded state (000) and the partially folded secondary structure state (001), defined by the turn and its HB pair. This result is consistent with a zipping model, where the β-hairpin formation initiates from the turn region[27] with, hence, the turn formation the triggering step of β-hairpin folding.

According to eq 1 we may evaluate the unfolded state probability, $P_{000}(t)$, and its complementary probability, $1 - P_{000}(t)$, providing the time dependence of the unfolded global population and of the folded and partially folded one, i.e., $1 - P_{000}(t) = P_{001}(t) + P_{010}(t) + ... + P_{111}(t)$. From Figure 7a, where we report these properties, the results show that the complete folding relaxation occurs within 130−150 ns and, as expected from the almost barrierless free energy profile, the folding kinetics is nonexponential.

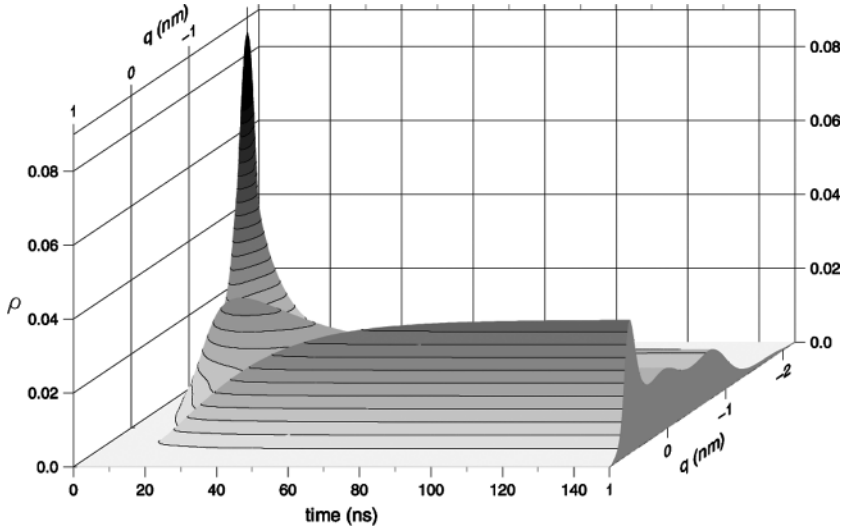


Figure 6. Probability density $\rho(q, t)$, as obtained by the DE 3, as a function of the position q along the first essential eigenvector and as a function of time for the H1 peptide. Note the presence of a metastable kinetic intermediate after about 10 ns.

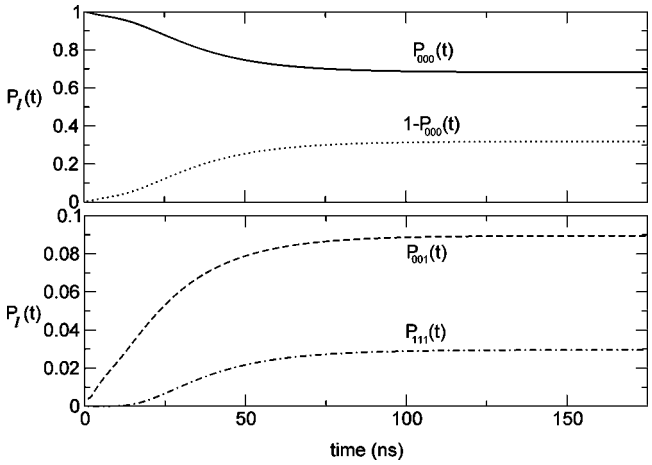


Figure 7. (a) Unfolded state probability $P_{000}(t)$, as obtained by eq 1 and its complementary probability $1 - P_{000}(t)$, providing the time dependence of the unfolded global population and of the folded and partially folded ones for the β-hairpin peptide. (b) Time dependence of the completely and partially folded probabilities, $P_{111}(t)$ and $P_{001}(t)$, respectively, of the β-hairpin peptide.

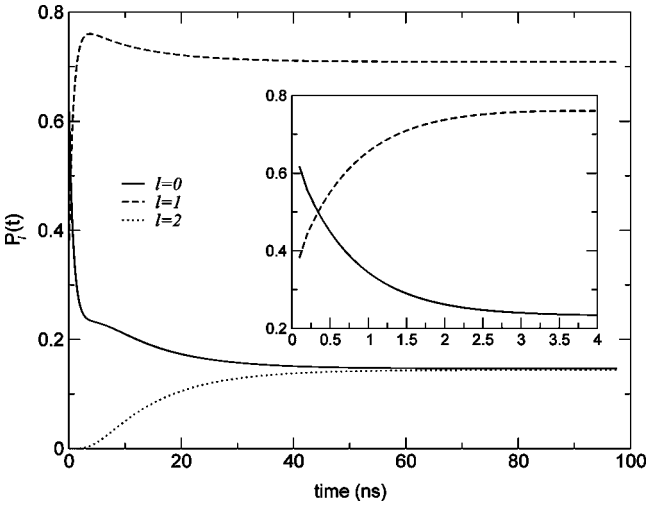


Figure 8. Time dependence of the probabilities, as obtained by eq 1, of the three secondary structure states of the helical temporin L peptide: the unfolded state ($l = 0$), the partially folded state ($l = 1$), and the completely folded state ($l = 2$). In the inset we show the initial kinetic relaxation (within the first 4 ns) of the unfolded and partially folded states.

In Figure 7b we show the time dependence of the complete β-hairpin, $P_{111}(t)$, and of the partially folded state $P_{001}(t)$. It is evident the early buildup of the 001 probability leading to the subsequent growth of the complete folded state (111). The other secondary structure states are also kinetically accessed only after the 001 intermediate has been formed. This result confirms the crucial role played by the turn formation in the folding process which, at least in the present case, is due to the metastable kinetic intermediate observed in the diffusive relaxation along ev 1. The $P_{111}(t)$ curve also provides the complete folding mean time (evaluated by the oblique flexus) which results in ~30 ns.

In Figure 8 we show the equivalent evolution for the temporin L helix peptide, now using the three secondary structure states defined in the methods section, i.e., the unfolded state ($l = 0$), the partially folded state ($l = 1$), and the completely folded state ($l = 2$).

Full folding relaxation, occurring within 70−80 ns, is faster than the β-hairpin one, as expected by the smaller accessible ev 1 range and very similar free energy profile. Also in the folding process of temporin L a kinetic intermediate, characterized by a single helical turn and formed within the first ~4 ns, is present. Such a kinetic intermediate formation, corresponding to the initial rapid unfolded state decay, is hence the essential condition to build up the complete helix. Whether such a kinetic scheme is a specific feature of temporin L or may be generalized to any α-helix peptide cannot be clarified from our data. However, such a fast initial transition (see inset of Figure 8) from the unfolded to the partially folded intermediate state is characterized by an apparently exponential decay, with a time constant of ~860 ps in agreement with previous experimental and computational literature data.[12,13] The folding kinetics for the complete helix formation (see Figure 8) exhibits a non-exponential behavior, in agreement with experimental and computational data,[21,41] with a folding mean time (evaluated by the oblique flexus) of ~8 ns. Such a folding mean time we obtained is in very good agreement with the available experimental data on a fast folding 21-residue α-helical peptide (~16 ns).[42] These results provide strong evidence supporting the

(41) Huang, Z.; Prusiner, S. B.; Cohen, F. E. *Folding Des.* **1996**, *1*, 13−19.

picture that the nucleation process is relatively fast, with a subnanosecond mean time, and the main contribution to the helix formation time is due to the diffusion of the nucleated species to search for the complete helical conformation (in the nanoseconds time scale).[13]

It is worth noting that, differently from the α-helical temporin L, the H1 peptide folding kinetics we obtained is considerably faster than the experimentally evaluated β-hairpin ones, i.e., the 15-residue peptide I (folding mean time 0.76 μs[15]) and the 16 residues GB1 β-hairpin (folding mean time 6 μs[14]). Nevertheless, the fast ∼30 ns folding mean time obtained by our model may be consistent with the extremely fast aggregation kinetics experimentally observed[38] and the very high alanine content reducing the entropic costs for the β-hairpin formation[15] (similarly to alanine-rich helices).

However, to investigate whether our results shed light on a real difference in folding behavior or rather show the limits of the model used, we applied the diffusive model to the C-terminal hairpin of protein GB1. This was accomplished using the free energy barrier (∼18 kJ/mol) provided by Muñoz et al.[14] and assuming the reaction coordinate diffusion coefficient and accessible range identical to the ones of our H1 peptide. Moreover, to simplify the model to be used, we also assumed that the complete unfolded condition (000) is the only state present along the reaction coordinate except at one position corresponding to the folded free energy minimum, where only the complete folded state (111) is present. The corresponding DE solution provides a clear exponential folding kinetics with a transition mean life of about ∼5 μs, in excellent agreement with the experimentally measured value (∼6 μs[14]) as well as with the value (∼5 μs) obtained by a computational procedure based on transition-path sampling,[26] hence confirming the reliability of the theoretical procedure presented.

Finally, the DE solution in the two-dimensional space defined by the first two essential eigenvectors of H1 peptide (see theory and methods sections) provides a folding kinetics with a folding mean life only 22% larger than the value obtained by the one-dimensional solution. This result suggests that the possible kinetic coupling of the essential degrees of freedom, although not negligible, is rather weak and hence the use of the first essential eigenvector as a single folding conformational coordinate is likely to properly describe peptide folding processes. Of course for larger peptides or proteins the choice of a single essential coordinate, as the folding coordinate, might not be always promising, since the first essential eigenvectors may be strongly kinetically coupled.

Conclusions

In this paper we investigated the folding kinetics of β-hairpin and α-helical peptides by modeling such a process as a diffusive relaxation over the conformational free energy surface, as defined by the essential eigenvectors of backbone configurational fluctuations provided by long time scale MD simulations. The results obtained from the two model peptides, the amyloidogenic β-hairpin H1 peptide and the α-helical temporin L, and from the well studied GB1 β-hairpin suggest that this diffusive model may be accurate in determining the essential features of the folding process. However, such a model might not be always applicable, as it is based on a linear regime kinetics for the conformational transitions and a pre-equilibrium of secondary structure interconversions. Deviations from these conditions would clearly imply this model to be inapplicable.

Interestingly, the analysis of the folding kinetics of the two model peptides strongly suggests that, for essentially barrierless conformational free energy landscapes, the folding limiting rate is not the initial secondary structure formation but rather the diffusion in the essential space of partially folded intermediates toward the completely folded conformation. This result is in agreement with recent experimental and computational results.[12,13] Finally, comparison of the folding kinetics of the (β-hairpin) H1 peptide using the first essential eigenvector or the plane defined by the first two essential eigenvectors to define the free energy surface and corresponding diffusion equations clearly shows that only a rather weak kinetic coupling of these conformational coordinates is present.

Acknowledgment. This work was supported by the European Community Training and Mobility Research Network Project "Protein (mis)folding": HPRN-CT-2002-00241. CASPUR (Rome) is acknowledged for the use of its computational facilities.

JA053383F

(42) Williams, S.; Causgrove, T.; Gilmanshin, R.; Fang, K.; Callender, R.; Woodruff, W.; Dyer, R. *Biochemistry* **1996**, *35*, 691−697.
(43) Inouye, H.; Kirschner, D. A. *J. Struct. Biol.* **1998**, *122*, 247−255.
(44) Zhao, H.; Kinnunen, P. K. *J. Biol. Chem.* **2002**, *277*, 25170−25177.