PROTEINS: Structure, Function, and Genetics 53:220–228 (2003)

Mechanism for the α-Helix to β-Hairpin Transition

Feng Ding,[1]* Jose M. Borreguero,[1] Sergey V. Buldyrey,[1] H. Eugene Stanley,[1] and Nikolay V. Dokholyan[2]
[1]*Center for Polymer Studies, Department of Physics, Boston University, Boston, Massachusetts*
[2]*Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, School of Medicine, Chapel Hill, North Carolina*

ABSTRACT **The aggregation of α-helix-rich proteins into β-sheet-rich amyloid fibrils is associated with fatal diseases, such as Alzheimer's disease and prion disease. During an aggregation process, protein secondary structure elements—α-helices—undergo conformational changes to β-sheets. The fact that proteins with different sequences and structures undergo a similar transition on aggregation suggests that the sequence nonspecific hydrogen bond interaction among protein backbones is an important factor. We perform molecular dynamics simulations of a polyalanine model, which is an α-helix in its native state and observe a metastable β-hairpin intermediate. Although a β-hairpin has larger potential energy than an α-helix, the entropy of a β-hairpin is larger because of fewer constraints imposed by the hydrogen bonds. In the vicinity of the transition temperature, we observe the interconversion of the α-helix and β-sheet states via a random coil state. We also study the effect of the environment by varying the relative strength of side-chain interactions for a designed peptide—an α-helix in its native state. For a certain range of side-chain interaction strengths, we find that the intermediate β-hairpin state is destabilized and even disappears, suggesting an important role of the environment in the aggregation propensity of a peptide. Proteins 2003;53:220–228.**

Key words: hydrogen bond; amyloid fibril; aggregation; entropy; molecular dynamics

INTRODUCTION

A number of misfolded proteins and peptides aggregate into insoluble fibrils. The aggregation of some of these proteins into amyloid fibrils—amyloidosis—is related to fatal diseases.[1,2] Recently, proteins not implicated in amyloid diseases were found to form fibril structures in vitro under denaturing conditions[3,4] suggesting that the fibril formation is a common feature of destabilized proteins.[5] Regardless of sequences and structures of proteins, the fibrils have similar core structures, mainly composed of β-sheets.[6–8] For example, the native Aβ peptide in Alzheimer's disease[9,10] and prion proteins (PrP^C) in prion diseases[11,12] are α-helix-rich. The aggregation from α-rich proteins or peptides involves a conformational transition from α-helices to β-sheets. A similar transition has also been observed in vitro in some α-helical peptides[13–17] that aggregate into amyloid fibrils by means of changing the environment, such as varying the organic solvent condition,[15] altering the pH,[16] and controlling the redox state.[17] Moreover, similar α-β transitions also occur through the correct folding pathway in proteins with a nonhierarchical folding mechanism.[18] For example, β-lactoglobulin, a predominantly β-sheet protein,[18] is observed to form nonnative α-helical intermediates on folding. Thus, understanding the α-β transition is important for both protein folding and protein aggregation.

The secondary structures of proteins, mainly α-helices and β-sheets, are determined by the amino acid sequence and stabilized by hydrogen bonds. However, under denaturing conditions, proteins with various organization of the secondary structure elements can aggregate into similar β-rich amyloid fibrils. We propose that the α-helix to β-hairpin transition is governed by sequence nonspecific properties of proteins and peptides (i.e., the hydrogen bond network formed between backbones). It has been suggested that sequence nonspecific hydrogen bond interaction among the backbones of proteins is an important factor for aggregation.[19] We hypothesize that the same type of interaction is also the driving force for the α-helix to β-sheet transition.

An overwhelming amount of computational simulations[20–28] and experimental[29–33] and theoretical[34–37] studies have been devoted to α-helix stability and the helix-coil transition. However, the possibility of β-hairpin formation in the folding pathway of peptides/proteins has not been addressed. The self-assembly of β-sheets by polyalanine segments, which usually form α-helices,[20,23] has been observed in silk-like multiblock copolymers.[38] Thus, we aim to identify the presence of metastable β-hairpin intermediate in the folding pathway of a simple polyalanine peptide, an α-helix in its native state.[20,23] Because of the limitations of traditional molecular dynamics simulations, simplified models become crucial in studying protein folding and aggregation.[19–21,39–43] Discrete molecular dynamics,[20,39,40] the combination of simple models and effi-

Grant sponsor: Petroleum Research Fund of the American Chemical Society; Grant number: 37237-AC4; Grant sponsor: National Institutes of Health, National Research Service Award Fellowship; GM20251-01 to N.V.D.

*Correspondence to: Feng Ding, Center for Polymer Studies, Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA 02215. E-Mail: fding@polymer.bu.edu

Received 24 January 2002; Accepted 13 March 2003

cient dynamic simulation algorithms, can access the physical processes in the scale of milliseconds with a single simulation.[44] In contrast, the traditional all-atom molecular dynamics simulations can only resolve the time scale of several nanoseconds in one run or reach several microseconds combining a large number of runs.[22] Therefore, to observe multiple transitions in a single simulation, we use the discrete molecular dynamics algorithm to study a polyalanine peptide.

MATERIALS AND METHODS

Discrete Molecular Dynamics

In general, discrete molecular dynamic (DMD) simulations are based on pairwise spherically symmetrical potentials that are discontinuous functions of an interatomic distance r. Each atom has a specific type—A, B, C, . . . —that determines its interaction with other atoms. Each type is characterized by its mass m. The interaction potential between atoms A and B is a step function of their distance r, characterized by distances $0 < r_{\min}^{AB} < r_1^{AB} \ldots < r_{\max}^{AB}$. If the distance r between two atoms A and B satisfies the inequality $r_i^{AB} < r < r_{i+1}^{AB}$, the pair potential has a value of u_i^{AB}. If $r < r_{\min}^{AB}$, $u^{AB} = \infty$ and $r_{\min}^{AB}$ are the hardcore collision distance. If $r > r_{\max}^{AB}$, $u^{AB} = 0$, and $r_{\max}^{AB}$ are the maximal range of interaction. If atoms A and B are linked by a covalent bond, they interact according to a different potential characterized by values $\tilde{r}_i^{AB}$ and $\tilde{u}_i^{AB}$. In this case, if $r > \tilde{r}_{\max}^{AB}$, $\tilde{u}^{AB} = \infty$, which indicates that the bond is permanent and cannot be broken under any conditions.

In DMD, all atoms move with a constant velocity unless their distance becomes equal to r_i^{AB}. At this moment of time, their velocities change instantaneously. This change satisfies the laws of energy, momentum, and angular momentum conservation. When the kinetic energy of the particles is not sufficient to overcome the potential barrier $\epsilon_i^{AB} = u_{i-1}^{AB} - u_i^{AB}$, the atoms undergo a hardcore reflection with no potential energy change. The main difficulty of this method is the effective sorting and updating of the collision times. However, it is possible to make the speed of the algorithm inversely proportional to $N \ln N$ where N is the total number of atoms.[45] For a sufficiently large number of steps, the method becomes equivalent to a regular MD based on Newtonian dynamics.

The speed of the algorithm also decays linearly with the number of steps in the potential and strongly decays with the density of the system. This method is very effective in simulating proteins (where the density is small and most of the interactions can be modeled by using either a hardcore or a simple square well) and allows us to observe protein folding transitions and aggregations.[19,40–42]

Four-Bead Model

The protein model, using three backbone beads and one side-chain bead to represent each residue[20,43] has been developed to mimic protein backbone structure. Molecular dynamics studies in such a polypeptide system have shown a sharp helix-coil transition,[20] which suggests that it is possible to study the transition from an α-helix to β-sheet in this model system. Thus, we use the four-bead

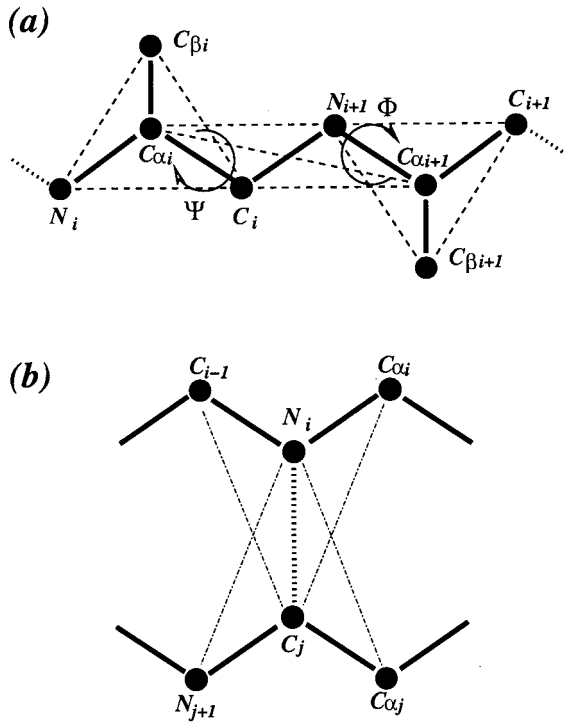

Fig. 1. **a:** Schematic diagram of the four-bead peptide model. The solid thick lines represent the covalent and the peptide bonds. The dashed thin lines denote the effective bonds that are assigned to mimic the tetrahedral constraint of each amino acid and the planar constraint of the peptide bond. **b:** Schematic diagram of the hydrogen bond. The four thin dot-dashed lines connect the auxiliary pairs, and the dashed line represents the hydrogen bond.

TABLE I. Parameters of Bonds and Hardcore Radii Used in Our Simulations

Covalent bond, D^{AB} (Å)		Effective bond, D^{AB} (Å)		Hardcore radius, R (Å)	
N_i, $C_{\alpha i}$	1.455	N_i, $C_{\beta i}$	2.442	C	1.50
$C_{\alpha i}$, $C_{\beta i}$	1.533	N_i, C_i	2.444	N	1.30
$C_{\alpha i}$, C_i	1.510	$C_{\beta i}$, C_i	2.494	C_α	1.85
C_i, N_{i+1}	1.325	$C_{\alpha i}$, N_{i+1}	2.406	C_β	2.20
		$C_{\alpha i}$, $C_{\alpha i+1}$	3.784		
		C_i, $C_{\alpha i+1}$	2.432		

model[20,21,43] to represent amino acids in the peptide. The amino acids are numbered from $k = 1$ (N-terminal) to $k =$ N (C-terminal), where N is the total number of residues. The kth amino acid is composed of nitrogen (N_k), prime carbon (C_k), alpha carbon ($C_{\alpha k}$), and beta carbon ($C_{\beta k}$) atoms [Fig. 1(a)]. In Figure 1(a) the thick lines represent the covalent bonds, and the thin lines denote effective bonds, mimicking the tetrahedral constraint of each amino acid and the planar constraint of the peptide bond. In our simulations, bonds are characterized by $\tilde{r}_{\min}^{AB} = D^{AB}(1 - \sigma)$ and $\tilde{r}_{\max}^{AB} = D^{AB}(1 + \sigma)$, where D^{AB} is the average distance between atoms A and B (listed in Table I) and σ is chosen as 0.02.

First, we study the peptide with only backbone hydrogen bond interaction. The nonbonded atom pairs have either

hardcore collision or hydrogen bond interactions. The hardcore radius R_A of four different types of atoms are listed in Table I ($r_{min}^{AB} = R_A + R_B$). The protein backbone hydrogen bonds are formed between the carbonyl oxygen and amide hydrogen. In the four-bead model, there is no carbonyl oxygen and amide hydrogen, but the position of *O* and *H* can be determined by their neighboring *N*, *G*, and C_α atoms. We model the hydrogen bond formed between the nitrogen N_i of the *i*th amino acid and the prime carbon C_j of the *j*th amino acid. Following the same notation as applied in Res. 20, 21, and 43, the separation along the sequence must satisfy the condition $|i - j| \geq 4$ to form a backbone hydrogen bond between N_i and C_j. It is well known that the hydrogen bond interaction has strong angular dependence (i.e., the hydrogen bonded *CO* and *NH* groups are collinear with each other). The usual pairwise interaction can hardly model this multibody interaction.

Hydrogen Bond Interaction: Reaction Algorithm

We introduce the reaction algorithm to model the hydrogen bond interaction between N_i and C_j. Once N_i and C_j form a hydrogen bond, they change their type into N_i' and C_j', respectively, and cannot form any other hydrogen bonds. Whether the "reaction" $N_i + C_j \rightleftharpoons N_i' + C_j'$ takes place is assessed when the distance between these atoms becomes equal to the hydrogen bond cutoff distance D_{HB} = 4.2 Å. The total potential energy change includes the potential energy gain ϵ_{HB} between N_i' and C_j', and the potential energy changes between the two atoms and their surrounding atoms because of the type changes. Once the kinetic energy is enough to overcome the total potential energy change, the forward reaction happens. Otherwise, the two atoms N_i and C_j do not change their types and undergo original hardcore collision. If the reaction is successful, the atoms change their atom types and interact with other atoms according to the interaction parameters related to their new types.

We implement the angular dependence of hydrogen bonds by assigning an auxiliary interaction between the atom pairs N_i'–$C_{\alpha j}$, N_i'–$N_{j+1}^{[\prime]}$, C_j'–$C_{\alpha i}$, and C_j'–$C_{i-1}^{[\prime]}$ [these four pairs are connected by thin lines in Fig. 1(b); the bracket in the superscript indicates that the atom may or may not have its type changed because of hydrogen bond formation] as

$$V = \begin{cases} \epsilon_{HB}, & d_{min} < d < d_0 \\ \epsilon_{HB}/2, & d_0 < d < d_1 \\ 0, & d_1 < d < d_{max} \\ +\infty, & \text{otherwise} \end{cases} \quad (1)$$

where ϵ_{HB} is the potential energy gain between N_i and C_j, and the parameters d_0, d_1, d_{min}, and d_{max} are chosen to implement the hydrogen bond angular constraints (see Table II). The other interactions involving the N_i and C_j atoms remain unchanged before and after the reaction. The new hardcore collision distance between N_i' and C_j' is assigned at 4.0 Å. Thus, at the lowest energy state of a hydrogen bond, the distance of the four auxiliary pairs is within the distance range of $[d_1, d_{max}]$, and distance of N_i and C_j is within the hydrogen bond range [4.0 Å, 4.2 Å]: the *CO* and *NH* groups are aligned as approximately linear. Parameters d_{min} and d_0 are chosen to allow angular distortion with energy penalizations.

TABLE II. Parameters of the Auxiliary Interactions

Pairs	d_{min}(Å)	d_0(Å)	d_1(Å)	d_{max}(Å)
$N_i', C_{\alpha j}$	4.46	4.66	4.82	5.56
$N_i', N_{j+1}^{[\prime]}$	4.47	4.62	4.78	5.41
$C_j', C_{\alpha i}$	4.40	4.56	4.72	5.39
$C_j', C_{i-1}^{[\prime]}$	4.44	4.62	4.79	5.39

When two atoms N_i and C_j approach each other at the hydrogen bond interaction cutoff distance D_{HB} = 4.2 Å, we evaluate the total potential energy change by checking the four auxiliary interactions. The potential energy change can be $-\epsilon_{HB}$, $-\epsilon_{HB}/2$, 0, $\epsilon_{HB}/2$, . . . $3\epsilon_{HB}$ and ∞, depending on the orientation of the N_i, C_j, and their neighbors. The larger the angular distortion, the higher the potential energy change. Once formed, the four auxiliary pairs will have a high probability of staying in the range of $[d_1, d_{max}]$ with the lowest energy; thus, the orientation of the hydrogen bond is maintained. The thermal fluctuations distort the orientation of the hydrogen bond and large fluctuations may break the hydrogen bond. Once the two atoms N_i' and C_j' come again to the exact distance of D_{HB}, a reverse reaction may happen. We check the potential energy change due to the possible changes of types. The total potential energy change ranges between $-3\epsilon_{HB}$ to ϵ_{HB}, corresponding to different conformations of the hydrogen bond due to thermal fluctuations. Thus, a distorted hydrogen bond will be easier to break. A more realistic modeling of the angular dependence of the hydrogen bond is to increase the number of steps in Eq. 1 to make the interaction potential more continuous. However, the increase of the number of steps decreases the efficiency of the discrete molecular dynamics.

16-mer With Hydrophobic Polar Sequence

We also study the effect of side-chain interactions in a 16-residue model peptide chain, designed to be an α-helix in its native state. The peptide has the following sequence of hydrophobic (H) and polar (P) residues: PPHPPHHP-PHPPHHPP.[21,46] This sequence is derived from a peptide that is designed to be an α-helix in the experiment.[47] In our simulations, the interaction between hydrophobic side-chains (C_β atoms) is modeled as an attractive square well with the cutoff distance D_{HP} = 6.5 Å and the interaction strength ϵ_{HP}; the remaining side-chain interactions are hardcore collisions. The relative strength of the hydrophobic interactions with respect to the hydrogen bond interactions $\rho = \epsilon_{HP}/\epsilon_{HB}$ is a free parameter that can be tuned.

POLYALANINE WITH HYDROGEN BOND INTERACTION ONLY

We study the refolding thermodynamics of the 16-residue polyalanine with only backbone hydrogen bond interactions. We perform discrete molecular dynamics simulations at different temperatures T = 0.09, 0.10, 0.11, 0.12, 0.125, 0.13, 0.14, and 0.15 in units of ϵ_{HB}/k_B (e.g., for

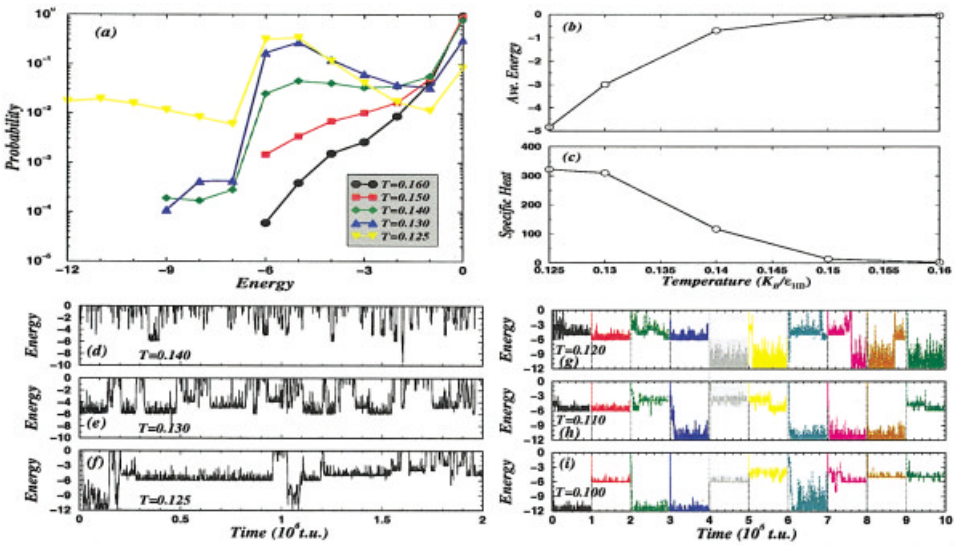

Fig. 2. Thermodynamics of the polyalanine chain with backbone hydrogen bond interaction only. (a) The probability distribution of potential energies, (b) the average energy, and (c) the specific heat at the temperatures where the peptide is not dynamically trapped. The typical energy trajectory at temperatures (d) 0.140, (e) 0.130, and (f) 0.125. At lower temperatures, (g) 0.120, (h) 0.110, and (i) 0.100, the protein is trapped easily and each simulation results in either α-helix or β-hairpin states. For each temperature, the ten different potential energy trajectories of 10^6 time unites (t.u.) are separated by dashed lines and colored differently. At temperature T = 0.120, we transitions are observed from the β-hairpin to the α-helix states (the 8th run) and from the α-helix to the β-hairpin states (the 9th run).

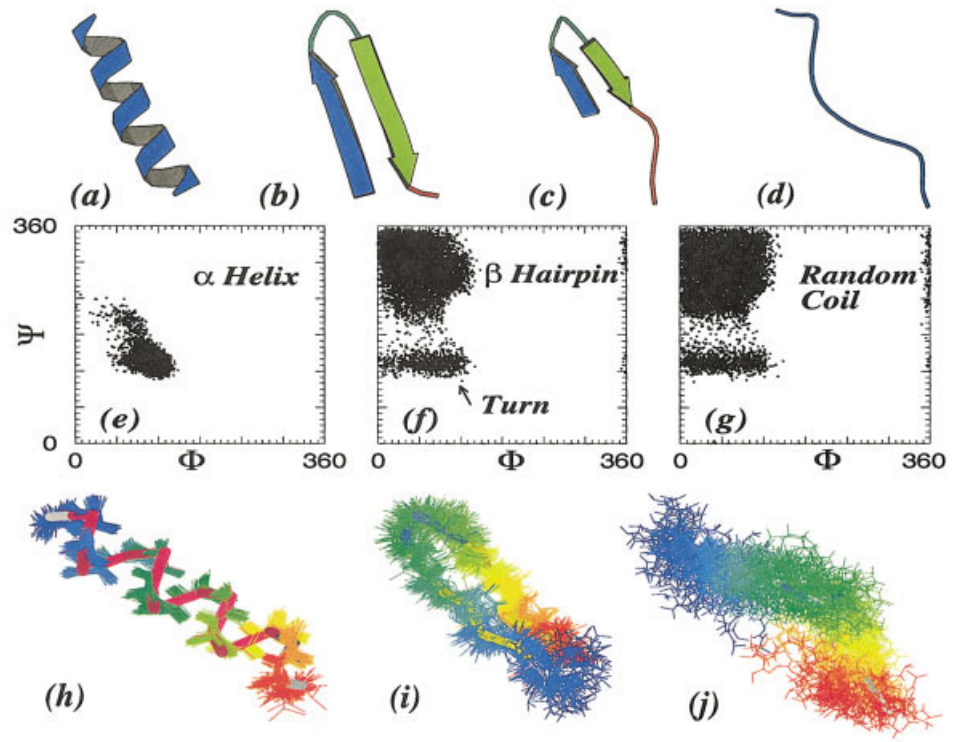

Fig. 3. Typical conformations of (a) an α-helix, (b) the β-hairpin with the β-turn located near the center, (c) an additional β-hairpin conformation with the β-turn positioned differently, and (d) a random coil. The distributions of torsion angles for (e) the α-helix, (f) the β-hairpin, and (g) the random coil states over equilibrated simulations. We align for each of the three states—(h) α-helix, (i) β-hairpin, and (j) random coil—various conformations with respect to a reference conformation using C_α atoms. The reference conformations are shown in backbone representation and the other conformations are displayed as wire frames, with the residues colored in the rainbow order from blue (N-terminal) to red (C-terminal).

ϵ_{HB} = 5 kcal/mol,[48] the temperature T = 0.12 corresponds to 302 K) where k_B is the Boltzmann constant. For each temperature, we perform 10 separate molecular dynamics simulations starting from different random coil conformations (Fig. 2).

At high temperatures, the polyalanine remains at the random coil state, and its average potential energy is close to zero. As we decrease the temperature, the peptide adopts a β-hairpin state [Fig. 2(d)]. For a β-hairpin structure, the lowest potential energy conformations have the β-turn located near the center of the peptide [Fig. 3(b)], and the potential energy is equal to $-6\epsilon_{HB}$. If the turn is positioned differently along the peptide [Fig. 3(c)], the potential energy is higher than $-6\epsilon_{HB}$ because of the smaller number of hydrogen bonds that can be formed. Thus, the occurrence of additional β-hairpin types [Fig. 3(c)] is less probable. At temperature T = 0.13 [(Fig. 2(e)], we observe a reversible random coil to β-hairpin transition, and the probabilities of finding a random coil and a β-hairpin state are approximately equal [Fig. 2(a,e)], so the β-hairpin to random coil transition temperature is $T_{\beta-coil} \approx 0.13$. At temperature T = 0.13, we also detect rare fluctuations with potential energy lower than $-6\epsilon_{HB}$, corresponding to partially formed α-helix states.

As we lower the temperature to T = 0.125, we observe the occurrence of α-helical states [Fig. 2(g) and Fig. 3(a)]. For the 16-residue polyalanine, the complete α-helix has four helix turns and the lowest energy is $-12\epsilon_{HB}$. At T = 0.125, the peptide can either adopt a random coil, an α-helix, or a β-hairpin state. The interconversion between an α-helix and a β-hairpin only takes place if the peptide first unfolds to a random coil state. This is mainly due to the drastic structural difference between these two kinds of conformations and there is no direct pathway between them except via a random coil state. Thus, the α-helix to β-hairpin transition is coupled to the transition between the α-helix to random coil transition. The probability of an α-helix is smaller than that of a β-hairpin at temperature T = 0.125. We expect to observe more α-helix states at lower temperatures. However, at low temperatures, the dynamics of hydrogen bond formation and disruption become slow, and the polyalanine is easily trapped in the local minima of the free energy landscape, dominated by a metastable β-hairpin state. We find that the peptide remains in either an α-helix or a β-hairpin state during the simulation time of 10^6 time units[40] after a quick collapse from the random coil state [Fig. 2(g,h,i)]

We present the distribution of the torsion angles ϕ and ψ for different states in Figure 3(e,f,g). The distributions are in agreement with the Ramachandran plot[49] for secondary structures. We find that the distributions of the β-hairpin and random coil are similar. However, for the random coil state, the torsion angles for each amino acid are fully uncorrelated, whereas for the β-hairpin state, the torsion angles between the hydrogen bonded amino acids are highly correlated. For an α-helix, each residue forms two hydrogen bonds except those near the termini [Fig. 4(a)]; thus, our peptide does not have excessive torsional freedom. On the other hand, for a β-hairpin strand, approximately half of all amino acids do not form any hydrogen bonds [Fig. 4(b)], the peptide chain has a larger value of backbone entropy. Therefore, the β-hairpin has larger hydrogen bond energy and also a larger entropy than the α-helix. To illustrate the backbone flexibility for different states (α-helix, β-hairpin, random coil), we align various conformations with respect to a characteristic structure using C_α atoms for each of these states [Fig. 3(h,i,j)]. We find that the β-hairpin [Fig. 3(i)] is more flexible than the α-helix [Fig. 3(h)]; therefore, the β-hairpin has a higher backbone entropy. (We provide movies for the dynamic motions of an α-helix and a β-hairpin, and one instance of the α-helix to β-hairpin transition: http://www.unc.edu/~dokh/research/AB/home.html). It is of interest that the alignment for the random coil state exhibits a persistent overall topology [Fig. 3(i)], which is possibly due to the excluded volume effect of the residues and is consistent with the finding in Ref. 50. The interplay between energy and entropy allows for the existence of the metastable intermediate state—the β-hairpin.

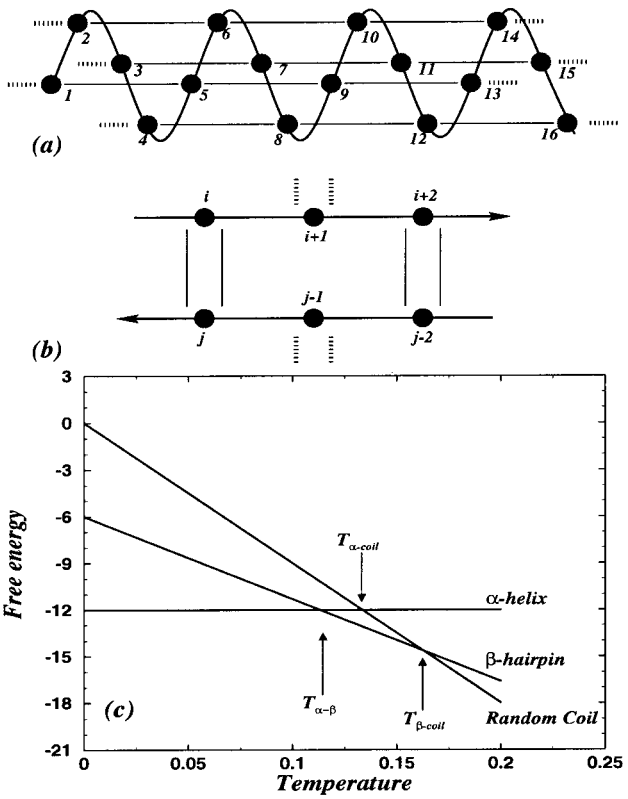

Fig. 4. Schematic diagram of the hydrogen bond pattern for **a:** the α-helix and **b:** the β-hairpin conformations, where the black beads represent each amino acid, the solid lines denote the backbone hydrogen bonds, and the dotted lines denote the free backbone hydrogen bond donor or acceptor. **c:** The free energies for different states versus temperature using the estimated values of the backbone entropies.

Because of the slow dynamics at low temperatures, we cannot accurately identify the α-helix to β-hairpin transition temperature $T_{\alpha-\beta}$. The coexistence of these two states with the random coil state at T = 0.125 suggests that the transition temperature $T_{\alpha-\beta}$ is close to 0.125,

which is in the vicinity of the β-hairpin to coil transition temperature $T_{\beta\text{-coil}} = 0.130$. Each of the four amino acids near the N- or C-termini of an α-helix has one free hydrogen bond donor or acceptor [Fig. 4(a)], similar to the β-hairpin strand where each amino acid on average has one free hydrogen bond donor or acceptor [Fig. 4(b)]. Thus, the potential energy per residue of the α-helix terminal is equal to that of a β-hairpin strand. According to the similar constraints imposed by the hydrogen bonds, we hypothesize that the conformational entropy per residue for the β-hairpin strands and α-helix termini is also similar. Thus, the α-helix termini and β-hairpin strands have similar free energy per residue. Furthermore, the amino acids near the termini have larger free energies than those of the amino acids within the helix, which is consistent with the observation of large fluctuations of the termini even at low temperature. We observe that the polyalanine melting (i.e., the transition into random coil) always starts from the termini. The melting temperature of an α-helix $T_{\alpha\text{-coil}}$ is determined by the free energy per residue between the α-helix termini and the random coil, and the transition temperature from a β-hairpin to random coil $T_{\beta\text{-coil}}$ is determined by the free energy per residue between β-hairpin strands and the random coil. So, $T_{\alpha\text{-coil}}$ is close to $T_{\beta\text{-coil}}$. Because the α-helix to β-hairpin transition is coupled with the α-helix to coil transition (α-helix melting), $T_{\alpha\text{-}\beta} \approx T_{\beta\text{-coil}}$.

To further understand the contribution of backbone entropy to the α-helix to β-hairpin transition, we estimate the backbone entropy for different states. From the alignment of conformations in Figure 3(h,i,j); we find that all conformations fluctuate around the reference structure characteristic to the corresponding state. Assuming that 1) the fluctuations of the structures around the reference structure are Gaussian and 2) the fluctuations of residues are uncorrelated, the conformational entropy can be approximated as $S_x = 3N \ln\langle rmsd_r\rangle_x + S_0$, where $rmsd_r$ is the root-mean-square deviation from the reference structure. The average $\langle\rangle_x$ is taken over conformations out of the corresponding state $\{x\}$: α-helix, β-hairpin, and random coil. S_0 is a constant that can be determined by setting the α-helix as the reference state with $S_\alpha = 0$. Because the four atoms in each residue are constrained to fluctuate as one object, N represents the number of amino acids. To calculate the entropy, we perform equilibrium simulations with hydrogen bonds intact (the α-helix and β-hairpin states) or without forming any hydrogen bonds (the random coil state). We calculate $\langle rmsd_r\rangle_x$ with respect to a selected conformation typical to the corresponding state. The values of estimated entropy for different states are listed in Table III. The transition temperatures can be determined as $T_{\beta\text{-coil}} = 0.162$, $T_{\alpha\text{-coil}} = 0.133$, and $T_{\alpha\text{-}\beta} = 0.115$. These estimated transition temperatures agree with the values determined from simulations. The above assumptions might lead to underestimation of the backbone entropy for the random coil state; therefore, the estimated transition temperatures $T_{\beta\text{-coil}}$ and $T_{\alpha\text{-coil}}$ are higher than the determined values from simulations.

TABLE III. Estimated Values of the Conformational Entropy for Different States[†]

x	$(rmsd_r)_x$ (Å)	Entropy, $S_x(k_B)$	Potential energy, $E(\epsilon_{HB})$
α-Helix (reference state)	0.483	0	−12
β-Hairpin	1.460	53.0	−6
Random coil	3.143	89.9	0

[†]$S_0 = -3N \ln\langle rmsd_r\rangle_\alpha = 34.9$ (N = 16). The transition temperature between two different states can be obtained from the differences of potential energy and entropy, $\Delta E/\Delta S$ [see Fig. 4(c)].

The existence of the metastable β-hairpin state has important implications for aggregation. A β-hairpin conformation that has the exposed hydrogen bond donors or acceptors is capable of further aggregation and can form amyloid fibrils.[51] Most real proteins do not aggregate by folding into the native state without long lifetime intermediates. Proper folding may be enforced by side-chain interactions in the evolutionarily selected sequence. We study a minimal model with hydrophobic side-chain interactions to uncover the propensities of the α-helix to β-hairpin transition for a hydrophobic-polar (HP) sequence, which is designed to be an α-helix in the native state.

MODEL PEPTIDE WITH HYDROPHOBIC-POLAR SEQUENCE

We perform molecular dynamics simulations of the HP peptide with various interaction ratios $\rho = \epsilon_{HP}/\epsilon_{HB}$, from 0.05 to 0.50 with a step of 0.05. For each ratio, we perform simulations at various temperatures. We find that the β-hairpin state becomes less stable as we increase ρ. For small ρ, the thermodynamic property of our HP peptide resembles that of the peptide without specific side-chain interactions. At a characteristic range of ρ values ($0.20 \leq \rho \leq 0.35$), the intermediate β-hairpin state disappears and the peptide folds cooperatively into the native α-helix state (Fig. 5).

For $\rho = 0.25$, the specific heat has a pronounced peak around $T_f = 0.128$ (Fig. 5), indicating a sharp transition specific to a two-state protein. At low temperatures, the peptide adopts the native α-helix structure [Fig. 5(d,e)]. At temperature $T = 0.120$, we also observe some potential energy fluctuations corresponding to the partially unfolded α-helix with the unfolded N- and C-termini. In the vicinity of the transition temperature, the peptide adopts both the α-helix and the unfolded states. The hydrophobic interactions are formed with a higher probability even at high temperatures than the hydrogen bonds because of the larger interaction range ($D_{HP} = 6.5$ Å $> D_{HB} = 4.2$ Å) and the absence of angular dependence of the HP interactions. Snapshots of these unfolded states indicate that the HP peptide adopts "molten globular"[52] conformations [Fig. 5(f)], which have contacts formed between the hydrophobic residues. The average radius of gyration R_g for the molten globular state at temperature $T = 0.135$ is 7.35 Å, compared to the unfolded state $R_g = 10.3$ Å for $\rho = 0$ at the

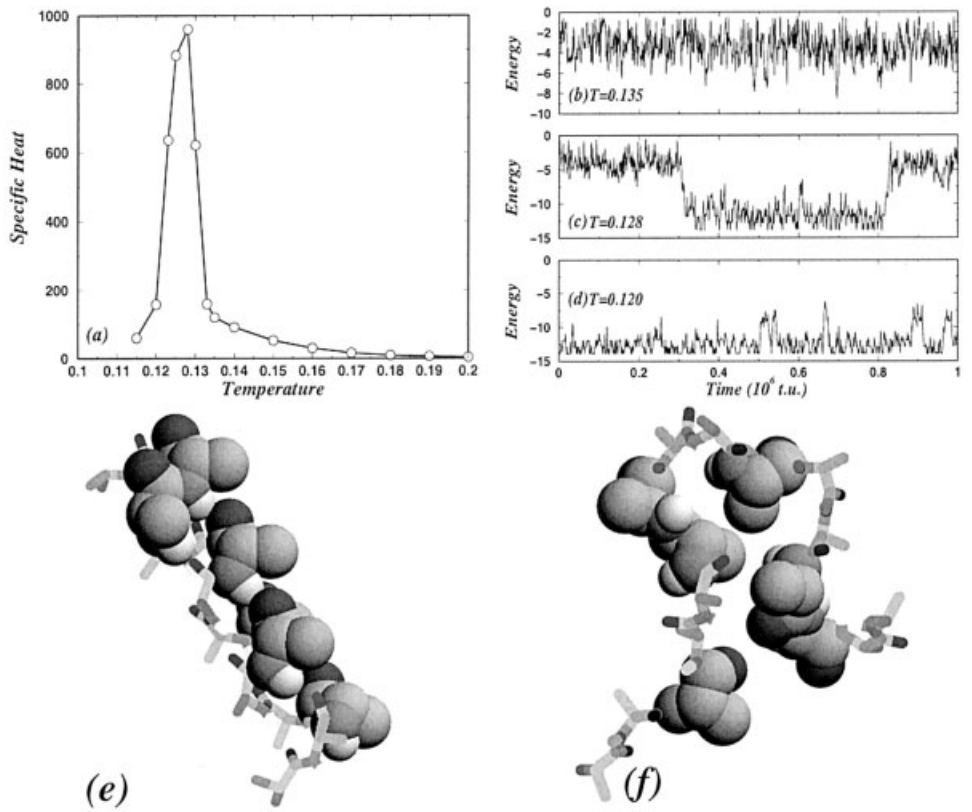

Fig. 5. **a:** Specific heat for the hydrophobic polar peptide with p = 0.25, having a pronounced peak at T_F = 0.128. The typical potential energy trajectories at **b:** T = 0.135, **c:** T = 0.128, and **d:** T = 0.120. The transition is between **e:** the native state and **f:** the molten globular states, where the space-filled amino acids are the hydrophobic atoms.

same temperature. We find that the HP peptide first collapses into the "molten globular" state from the random coil state, which is similar to the coil-globular transition,[52] at a higher temperature than T_F.

Within the characteristic range of ρ values, the difference of the average potential energy between the β-hairpin state and the unfolded state is small, which is not enough to stabilize the β-hairpin state. For example, for ρ = 0.25 (Fig. 5), the average potential energy of the unfolded state is $-4\epsilon_{HB}$ at T_F, and the potential energy of the β-hairpin state $\approx -6\epsilon_{HB}$ is within the range of potential energy fluctuation of the unfolded state [Fig. 5(c)]. However, the potential energy gap between the α-helix state and the unfolded state is $8\epsilon_{HB}$ [(Fig. 5(c)], which is large enough to stabilize the α-helix state.[53]

As we increase ρ, we rarely observe the helix-coil transition (e.g., for ρ = 0.50 the α-helix state is never reached from the unfolded state during the simulation of 10^7 time units). At low temperatures, our HP peptide is frozen in the molten globular states because of the strong hydrophobic interactions.

DISCUSSION

A missing link in understanding the amyloidogenesis of α-helix-rich proteins to β-sheet-rich fibrils is the possible presence of a metastable β-hairpin intermediate state, prone to aggregation.[51] Our results suggest a generic framework that explains why this β-hairpin intermediate is favorable in terms of free energy. Although the potential energy of the β-hairpin state is higher than that of the α-helix state, the entropy of a β-hairpin is significantly larger than that of an α-helix due to fewer constraints imposed by hydrogen bonds. At high temperatures, the free energy of a β-hairpin can be smaller than that of an α-helix. Even though our simulations and discussions are focused on the β-hairpin (an antiparallel two-stranded β-sheet), our analysis is appropriate for both parallel and antiparallel two-stranded β-sheets that are entropically favorable with respect to α-helices.

Our simulations of temperature-driven α-helix to β-hairpin transition are consistent with recent experiments on the solvent-driven conformational transitions.[54,55] By changing the solvents from one type that has a low ability to interact with the backbone peptide groups to another type that has a higher ability,[55] the designed peptides are found to convert from α-helices to β-hairpins. Increasing the ability of the solvent to interact with the backbone, the energy gain to form a backbone hydrogen bond is effectively decreased. Instead of increasing the temperature, the decrease of the hydrogen bond's energy gain drives the conformational transition from α-helix to β-hairpin because of the dominating effect of backbone entropy.

Most proteins in physiological conditions do not aggregate. Proteins with evolutionarily selected sequences avoid aggregation by folding into the native state without metastable intermediate states. In vitro and in vivo experiments show that changes in environmental conditions lead to aggregation.[5,16] The environmental change has a different effect on different types of interactions. Our simulations of HP sequences with various hydrophobic interaction strengths show that if the environmental changes effectively lead to the weakening of the relative side-chain interactions, the peptide or protein may misfold into a metastable β-hairpin intermediate.

Discrete molecular dynamics simulation methodology is a step in simplification of molecular modeling with respect to traditional molecular dynamics simulations. The principal drawback of the discrete molecular dynamics simulations is its difficulty to represent forces. Instead, system's dynamics is realized through ballistic collisions between particles. Interactions between particles are modeled by square-well potentials. Despite its simplicity, discrete molecular dynamics has been proved to be a powerful tool not only to study protein folding thermodynamics[39–42] and kinetics,[41,42,56] but to identify the evasive protein transition state ensembles[41] and to witness aggregation of multiple proteins into amyloid fibrils.[19] The latter two goals have yet to be directly approached with traditional molecular dynamics simulations. In addition, the traditional all-atom molecular dynamics simulations are also a simplification of the quantum mechanics simulations, in which quantum interactions are replaced by approximate Newtonian interactions. The latter, in turn, are approximated by a large number of empirical parameters. The advantage of the discrete molecular dynamics simulations versus traditional molecular dynamics simulations is its ability to resolve larger timescales—10^6 orders of magnitude. The traditional molecular mechanics simulations have similar advantage over quantum mechanics simulations. The traditional molecular dynamics simulations are based on several decades of improving and testing of model force field, whereas applications of discrete molecular dynamics simulations have been limited until recently to colloids and hard spheres. Despite this, we believe that modifying and improving parameters of discrete molecular dynamics simulations for proteins by testing them on simple systems such as the polyalanine chain studied here will eventually lead to models with quantitative predictive power.

ACKNOWLEDGMENT

We thank L. Cruz, B. Kulhman, E.I. Shakhnovich, and B. Urbanc for helpful discussions.

REFERENCES

1. Kelly JW. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. Curr Opin Struct Biol 1998;8:101–106.
2. Tan SY, Pepys MB. Amyloidosis. Histopathology 1994;25:403–414.
3. Guijarro JI, Sunde M, Jones JA, Campbell ID, Dobson CM. Amyloid fibril formation by an SH3 domain. Proc Natl Acad Sci USA 1998;95:4224–4228.
4. Chiti F, Webster P, Taddei N, Clark A, Stefani M, Ramponi G, Dobson CM. Designing conditions for in vitro formation of amyloid protofilaments and fibrils. Proc Natl Acad Sci USA 1999;96:3590–3594.
5. Dobson CM. Protein misfolding, evolution and disease. Trends Biochem Sci 1999;24:329–332.
6. Bonar L, Cohen AS, Skinner M. Characterization of the amyloid fibril as a cross-β protein. Proc Soc Expt Biol Med 1967:131:1373–1375.
7. Eanes ED, Glenner GG. X-ray diffraction studies on amyloid filaments. J Histochem Cytochem 1968;16:673–677.
8. Sunde M, Serpell LC, Bartlam M, Fraser PE, Pepys MB, Blake CCF. The common core structure of amyloid fibrils by synchrotron X-ray diffraction. J Mol Biol 1997;273:729–739.
9. Selkoe DJ. Amyloid beta-protein and the genetics of Alzheimer's disease. J Biol Chem 1996;271:18295–18298.
10. Lansbury PT. A reductionist view of Alzheimer's disease. Acc. Chem. Res. 1996;29:317–321.
11. Prusiner SB. Prion diseases and the BSE crisis. Science 1997;278:245–251.
12. Harrison PM, Bamborough P, Daggett V, Prusiner SB, Cohen FE. The prion folding problem. Curr Opin Struct Biol 1997;7:53–59.
13. Takahashi Y, Ueno A, Mihara H. Design of a peptide undergoing α-β structural transition and amyloid fibrillogenesis by the introduction of a hydrophobic defect. Chem Eur J 1998:4:2475–2484.
14. Fezoui Y, Hartley DM, Walsh DM, Selkoe DJ, Osterhout JJ, Teplow DB. A de novo designed helix-turn-helix peptide forms nontoxic amyloid fibrils. Nat Struct Biol 2000;7:1095–1099.
15. Ono S, Kameda N, Yoshimura T, Shimasaki C, Tsukurimichi E, Mihara H, Nishino N. Supersecondary structure with amphiphilic β-strands probed by phenylalanine. Chem Lett 1995;10:965–966.
16. Cerpa R, Cohen FE, Kuntz ID. Conformational switching in designed peptides: the helix/sheet transition. Fold Design 1996;1:91–101.
17. Schenck HL, Dado GP, Gellman SH. Redox-triggered secondary structure changes in the aggregated states of a designed methionine-rich peptide. J Am Chem Soc 1996;118:12487–12494.
18. Hamada D, Segawa S, Goto Y. Non-native alpha-helical intermediate in the refolding of beta-lactoglobulin, a predominantly beta-sheet protein. Nat Struct Biol 1996;3:868–873.
19. Ding F, Dokholyan NV, Buldyrev SV, Stanely HE, Shakhnovich EI. Molecular dynamics simulation of the SH3 domain aggregation suggests a generic amyloidogenesis mechanism. J Mol Biol 2002;324:851–857.
20. Simth AV, Hall CK. α-Helix formation: discontinuous molecular dynamics on an intermediate-resolution protein model. Proteins 2001;44:344–360.
21. Smith AV, Hall CK. Protein refolding versus aggregation: computer simulations on an intermediate-resolution protein model. J Mol Biol 2001;312:187–202.
22. Ferrara P, Apostolakis J, Caflisch A. Thermodynamics and kinetics of folding of two model peptides investigated by molecular dynamics simulations. J Phys Chem 2000;104:5000–5010.
23. Hansmann UHE, Okamoto Y. Finite-size scaling of helix-coil transitions in poly-alamine studied by multicanonical simulations. J Chem Phys 1999;110:1267–1276.
24. Sung SS, Wu XW. Molecular dynamics simulations of synthetic peptide folding. Proteins 1996;25:202–214.
25. Takano M, Yamato T, Higo J, Suyama A, Nagayama K. Molecular dynamics of a 15-residue poly(L-alamine) in water: helix formation and energetics. J Am Chem Soc 1999;121:605–612.
26. Bertsch RA, Vaidehi N, Chan SI, Goddard WA. Kinetic steps for α-helix formation. Proteins 1998;33:343–357.
27. Hirst JD, Brooks CL. Molecular dynamics simulations of isolated helices of myoglobin. Biochemistry 1995;34:7614–7621.
28. Daggett V, Levitt M. Molecuar dynamics simulations of helix denaturation. J Mol Biol 1992;223:1121–1138.
29. Eaton WA, Munoz V, Thompson PA, Chan CK, Hofrichter J. Submillisecond kinetics of protein folding. Curr Opin Struct Biol 1997;7:10–14.
30. Callender RH, Dyer RB, Gilmanshin R, Woodruff WH. Fast events in protein folding: the time evolution of primary processes. Annu Rev Phys Chem 1998;49:173–202.
31. Ballew RM, Sabelko J, Gruebele M. Direct observation of fast protein filding: the initial collapse of apomuoglobin. Proc Natl Acad Sci USA 1996;93:5759–5764.
32. Shastry MCR, Roder H. Evidence for barrier-limited protein

filding kinetics on the microsecond time scale. Nat Struct Biol 1998;5:385–392.
33. Pascher T, Chesick JP, Winkler JR, Gray HB. Protein folding triggered by electron transfer. Science 1996;271:1558–1560.
34. Schellman JA. The factors affecting the stability of hydrogen-bonded polypeptide structures in solution. J Phys Chem 1958;62: 1485–1494.
35. Zimm BH, Bragg JK. Theory of the phase transition between helix and random coil in polypeptide chains. J Chem Phys 1959;31:526–535.
36. Lifson S, Roig A. On the theory of helix-coil transition in polypeptides. J Chem Phys 1961;34:1963–1974.
37. Munoz V, Serrano L. Elucidating the folding problem of helical peptides using empirical parameters. Nat Struct Biol 1994;1:399–409.
38. Rathore O, Sogah DY. Self-Assembly of β-sheets into nanostructures by poly(alanine) segments incorporated in multiblock copolymers inspired by spider silk. J Am Chem Soc 2001;123:5231–5239.
39. Zhou Y, Karplus M. Folding thermodynamics of a three-helix-bundle protein. Proc Natl Acad Sci USA 1997;94:14429–14432.
40. Dokholyan NV, Buldyrev SV, Stanley HE, Shakhnovich EI. Molecular dynamics studies of folding of a protein-like model. Fold Design 1998;3:577–587.
41. Ding F, Dokholyan NV, Buldyrev SV, Stanley HE, Shakhnovich EI. Direct molecular dynamics observation of protein folding transition state ensemble. Biophys J 2002;83:3525–3532.
42. Borreguero JM, Dokholyan NV, Buldyrev SV, Stanley HE, Shakhnovich EI. Thermodynamic and folding kinetic analysis of the SH3 domain from discrete molecular dynamics. J Mol Biol 2002;318: 863–876.
43. Takada S, Luthey-Schulten Z, Wolynes PG. Folding dynamics with nonadditive forces: a simulation study of a designed helical protein and a random heteropolymer. J Chem Phys 1999;110: 11616–11629.
44. Zhou Y, Karplus M. Folding of a model three-helix bundle protein: a thermodynamic and kinetic analysis. J Mol Biol 1999;293:917–951.
45. Rapaport DC. The art of molecular dynamics simulation. Cambridge University Press: Cambridge, 1997.
46. Guo Z, Thirumalai D. Kinetics and thermodynamics of folding of a de novo designed four-helix bundle protein. J Mol Biol 1996;263: 323–343.
47. Ho SP, DeGrado WF. Design of a 4-helix bundle protein: synthesis of peptides which self-associate into a helical protein. J Am Chem Soc 1987;109:6751–6758.
48. Honig B, Yang AS. Free-energy balance in protein-folding. Adv Protein Chem 1995;46:27–58.
49. Ramakrishnan C, Ramachandran GN. Stereochemistry criteria for polypeptide and protein chain conformations: part II allowed conformations for a pair of peptide units. Biophys J 1965;5:909–933.
50. Pappu RV, Srinivasan R, Rose GD. The flory isolated-pair hypothesis is not valid for polypeptide chains: implications for protein folding. Proc Natl Acad Sci USA 2000;97:12565–12570.
51. Richardson JS, Richardson DC. Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. Proc Natl Acad Sci USA 2002;99:2754–2759.
52. Grosberg AY, Khokhlov AR. Statistical physics of macromolecules. AIP Press: New York, 1994.
53. Shakhnovich EI. Protein design: a perspective from simple tractable models. Fold Design 1998;3:R45–R58.
54. Sha YL, Li YL, Wang Q, Fan KQ, Liu DS, Lai LH, Tang YQ. CD evidence of a peptide ongoing alpha/beta/random transition in different solutions. Protein Pep Lett 1999;6:137–140.
55. Awasthi SK, Shankaramma SC, Paghothama S, Balaram P. Solvent-induced β-hairpin to helix conformational transition in a designed peptide. Biopolymers 2001;58:465–476.
56. Dokholyan NV, Buldyrev SV, Stanley HE, Shakhnovich EI. Identifying the protein folding nucleus using molecular dynamics. J Mol Biol 2000;296:1183–1188.