Self-assembly of polypeptides into left-handedly twisted fibril-like structures

Yan Mu and Yi Qin Gao
Department of Chemistry, Texas A&M University, College Station, Texas 77843, USA
(Received 1 August 2009; published 26 October 2009)

In this paper, we investigated the spontaneous formation of aggregation structures of amyloid-forming peptide (GGVVIA) using a coarse-grained model and Monte Carlo simulations. The effects of concentration and temperature on the formation of different aggregation structures were studied. Three types of aggregation structures, single-layer β sheet, amorphous β-sheet aggregate, and fibril-like structures, were observed in our simulations. The fibril-like structures obtained in simulations have a common cross-β spine structure in which β sheets twist in a left-handed fashion. The averaged twisting angle of the β sheet in the fibril-like structures is $12° \pm 2°$. Moreover, it was found that the peptides in the same β sheets prefer to arrange in a parallel way, which is consistent with the corresponding GGVVIA crystalline structure. On the other hand, it was found that there is a rich family of β-sheet stacking patterns in the fibril-like structures suggesting that the fibril structures are more complex than the corresponding crystalline structure and there exist many local free-energy minima rather than a distinct global minimum.

DOI: 10.1103/PhysRevE.80.041927 PACS number(s): 87.10.Rt

I. INTRODUCTION

The pathological proteins may form elongated unbranched amyloid fibrils, which can cause many fatal diseases including Alzheimer's, Parkinson's, and prion diseases [1–4]. Understanding the formation mechanisms of different protein aggregate structures is critical for developing new drugs and other medical strategies to prevent or slow down the pathological protein aggregation. In recent years, the structures of amyloid fibrils have been studied extensively by many biophysical tools including x-ray diffraction [5–8], NMR [9–14], cryo-electron microscopy [15,16], and atomic force microscopy [17,18]. These experiments revealed that amyloid fibrils formed by different proteins have a common cross-β spine composed of a pair of *twisted* β sheets that are parallel to the axis of fibril, with their strands being perpendicular to this axis [19]. Very recently, Eisenberg and coworkers [20,21] determined atomic-level structures of the cross-β spines formed by ten different short peptides using x-ray diffraction studies. The short fibril-forming peptides derived from their parent proteins can form both fibrils and microcrystals. In their work, it was found that between the two β sheets forming the cross-β spine, there is a completely dry interface where the residue side chains intermesh to form a steric-zipper structure. According to the orientations of peptides in β sheets and the packing styles of the β sheets, the steric-zipper structures can be sorted into eight theoretically possible classes [21].

Computer simulation is a very important supplementary way to study the protein aggregations, which allows us not only to observe the aggregation processes but also to analyze the fibrillar structures formed by different proteins at the atomic level. In recent years, the aggregation processes of various polypeptides including amyloid-β (Aβ), polyalanine, and GNNQQNY peptide have been investigated extensively through computer simulations [22–29]. For example, Farvin *et al.* studied the aggregation mechanism of amyloid $A\beta_{16–22}$ peptides for systems containing up to six $A\beta_{16–22}$ peptides [27]. They found that for the three- and six-chain systems, aggregated structures include high β-sheet content and can have many different configurations. Derreumaux and coworkers [28,29] studied the aggregation of different short peptides KFFE (K: lysine; F: phenylalanine; E: glutamic acid) and NFGAIL (N: asparagine; F: phenylalanine; G: glycine; A: alanine; I: isoleucine; L: leucine) by combining the activation-relaxation technique with an optimized potential for efficient peptide structure prediction. For the KFFE peptide, they also found that four KFFE peptides adopted a variety of oligomeric states (tetramers, trimers, and dimers) with various orientations of the chains in rapid equilibrium [28]. For the aggregation of the NFGAIL peptide, they found that starting from a preformed parallel dimer and ten disordered chains, only amorphous oligomers or more rarely ordered β-sheet structures were formed suggesting that a dimer is not a sufficient seed for avoiding amorphous aggregates and there exists a critical threshold for the seed size of fibrils [29]. Due to the long time scales involved, most computer simulations of aggregations of polypeptides mentioned above have been limited to systems consisting of only a few peptides and it was difficult to obtain stable extensive fibrillar structures and to compare directly with experimental results. To gain information on protein aggregation process for large systems, Nguyen and Hall [30] studied the fibril formations for the systems containing up to 96 coarse-grained polyalanine model peptides. They observed that fibril-like structures were formed at relatively high temperatures and high concentrations and the formation process is nucleation dependent. Moreover, they found that there are two growth mechanisms of aggregation structures: β-sheet elongation and lateral addition. However, the fibrillar structures obtained in their simulations were composed of *flat* β sheets, while experimental results show the β sheets in the fibrils are *twisted* [7,15–18]. This discrepancy could be due to the fact that the simple hydrogen bonding potential used in their simulation does not reflect correctly the hydrogen bond geometry in β sheets. Recently, Bellesia and Shea studied the self-assembly of β-sheet forming model peptides with a more simplified model [31]. They obtained twisted fibrillar structures in simulations with an additional "chirality" de-

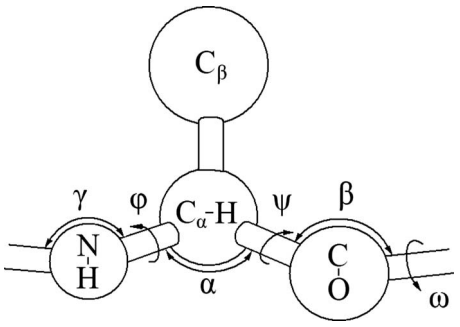

FIG. 1. A Schematic picture of an amino acid residue based on our coarse-grained peptide model. φ and ψ are the two Ramachandran torsional angles. α, β, and γ are the three bond angles.

gree of freedom. In this paper, using an improved coarse-grained model with a more realistic hydrogen bonding interaction potential without any artificial restriction or additional "twisting" degree of freedom, we studied the spontaneous formations of aggregation structures by GGVVIA peptides using the Monte Carlo (MC) simulation technique. Three types of aggregation structures were observed in our simulations. Especially, the fibril-like structures obtained in simulations are composed of left-handedly twisted β sheets that are parallel to the fibril axis with strands being perpendicular to the fibril axis consistent with experimental results. In addition, the effects of concentration and temperature on the formations of aggregation structures are also discussed.

II. MODEL AND METHOD

In the present study, a coarse-grained four-bead model based on earlier work [30,32–34] is used to describe the basic backbone structure of peptide as shown in Fig. 1. Each amino acid residue in the backbone is composed of four united atoms: N, C_α, C, and the side chain C_β. Different from the previous four-bead models [30,32–34], although the hydrogen atom connected to the nitrogen atom and the oxygen atom connected to the carbon atom are embedded in the united atoms N and C, respectively, and implicit for the excluded volume effect, they are explicit in the formation of hydrogen bond in our model [35]. In the current work, to ensure correct hydrogen bond geometry, we used a realistic potential fitted from protein data bank to describe the formation of hydrogen bond [36]. Details of this potential are discussed later. The center-of-mass translation, pivot, and fixed end moves [37] are used to generate the new conformations of the peptide in our MC simulations. Fixed end move is very effective in generating new conformations for individual peptide [37]. It changes the conformation of a polypeptide by rotating a segment between two randomly chosen alpha carbon atoms and keeping all atoms outside this segment fixed. The degrees of freedom of each amino acid are the two Ramachandran torsional angles φ and ψ and the bond angle α. The range of the allowed bond angle α is set to $\{\alpha_0-10°, \alpha_0+10°\}$, where α_0 is the average value of bond angle α and is approximately 111.0°. Furthermore, all bond lengths, bond angles β and γ, and peptide torsional angle ω (180°, corresponding to the *trans* conformation of the amide plane) are fixed [32–34]. All bond lengths and angles used in our simulations are listed in Table I.

TABLE I. Geometry parameters.

Bond lengths	(Å)
N-C_α	1.460
C_α-C	1.510
C-N	1.330
C_α-C_β	1.531
N-C_β	2.44
C-C_β	2.49
Bond angles	**(deg)**
$NC_\alpha C$ (α)	111.0
$C_\alpha CN$ (β)	116.0
CNC_α (γ)	122.0
United-atom diameters	**(Å)**
N	3.300
C_α	3.700
C	4.000
C_β	4.800

Protein folding and aggregation are very complicated processes as results of a balance of many different interactions. Most recent studies suggest that hydrogen bonding and hydrophobic interactions are the most important interactions which are responsible to form and stabilize native conformation of protein and ordered fibril structures [27,30,32–34]. Thus, in order to capture the essential interactions and focusing on the construction of a simple model, in the present study, we only consider the effects of these two most basic interactions on the protein aggregation. The potential function in the current model has the following form:

$$E = E_{ex} + E_{hb} + E_{hp}, \tag{1}$$

where E_{ex}, E_{hb}, and E_{hp} represent excluded volume effects, hydrogen bonding, and hydrophobic interactions, respectively. The excluded volume effect is modeled by a hard-sphere potential and runs over all possible atom pairs except for those consisting of two side chain united atoms. However, for the local hard-sphere interaction between the united atoms separated by three or fewer bonds along the chain, the diameters of these atoms themselves are more appropriate than the effective diameters of united atoms [32]. Thus, we allow the united atoms separated by three or fewer bonds along the chain to overlap by up to 25%. In addition, in order to increase the efficiency of hydrogen bond formation, we allow the overlap between united atoms N and C to up to 15%. The corresponding parameters are also given in Table I.

The hydrogen bonding interaction between peptides plays a crucial role in the formation and stabilization of the β-sheet structures. However, the simple hydrogen bonding potential functions used in the previous coarse-grained models are in-

adequate to accurately reflect the hydrogen bond geometry in β sheets and to capture the left-handedly twisted structural feature of fibril structures due to some artificial requirements and rough definition of the angle-dependent component in the hydrogen bond formation [30,32–34]. Thus, to reflect the right geometry of the fibrillar structure, it is necessary to introduce a more realistic potential function to describe the hydrogen bonding interaction. In the present work, we employed the hydrogen bonding potential developed by Chen *et al.* [36]. This hydrogen bonding potential consists of distance and angle components which can describe the geometry of hydrogen bond accurately. It has the following form:

$$E_{hb} = \varepsilon_{hb} \sum_{ij} u(r_{ij}) v_1(\theta_{1,ij}, \theta_{1,avg}) v_2(\theta_{2,ij}, \theta_{2,avg}) v_3(\theta_{3,ij}, \theta_{3,avg}), \tag{2}$$

where $u(r)$ and $v(\theta)$ are distance and angle components of the potential, respectively:

$$u(r_{ij}) = 5\left(\frac{\sigma_{hb}}{r_{ij}}\right)^{12} - 6\left(\frac{\sigma_{hb}}{r_{ij}}\right)^{10}, \tag{3}$$

$$v_n(\theta_{n,ij}, \theta_{n,avg}) = \begin{cases} \frac{1}{3}[4\cos^2(\theta_{n,ij} - \theta_{n,avg}) - 1], & \theta_{n,ij} < (\theta_{n,avg} + \Delta\theta), n = 1,2,3 \\ 0, & \text{otherwise}, \end{cases} \tag{4}$$

where r_{ij} is the distance between H_i and O_j, and σ_{hb} takes 1.80 Å as the average hydrogen bond length. In order to speed up the simulations, a cutoff radius $r_c = 5.0$ Å is used in the current model. $\theta_{1,ij}$, $\theta_{2,ij}$, and $\theta_{3,ij}$ are defined as ($\pi - \angle$NHO), the angle between $\overrightarrow{NH}$ and $\overrightarrow{OC}$ vectors, and ($\pi - \angle$COH), respectively, and $\theta_{n,avg}$ ($n=1,2,3$) are their average values which are 17.98°, 11.60°, and 26.77°, respectively [38]. Since hydrogen bond interaction is of a very short range, an uncertain deflection angle $\Delta\theta$ is used to increase the efficiency of the hydrogen bond formation. $\Delta\theta$ is taken as 60° in the current model. ε_{hb} is the strength of the hydrogen bonding interaction and set to unity $\varepsilon_{hb}=1$. In the present study, the temperature and all energy parameters are given in units of ε_{hb}. More details of this potential can be found in Ref. [37].

Besides the hydrogen bonding interaction, the hydrophobic interaction between side chains also plays a very important role in the formation and stability of ordered protein structures. In the original four-bead coarse-grained models [30,32–34], side chains are represented by beads and the hydrophobic interaction between them is generally described by the isotropic square-well potentials. This treatment is adequate for amino acids with roughly spherical side chains such as alanine and valine. However, for the peptide sequence GGVVIA considered in the present study, since the amino acid isoleucine has a highly asymmetric side chain, the isotropic potentials are not appropriate to describe the hydrophobic interactions involving the side chains of isoleucine. Therefore, we introduce a simple anisotropic potential to describe the hydrophobic interaction between side chains with different shapes. In the present study, the hydrophobic interaction E_{hp} is given by

$$E_{hp} = \varepsilon_{hp} \sum_{ij} \left[\left(\frac{\sigma_{hpij}^{p(ap)}}{r_{ij}}\right)^{12} - 2\left(\frac{\sigma_{hpij}^{p(ap)}}{r_{ij}}\right)^{6} \right], \tag{5}$$

where r_{ij} is the distance between hydrophobic side chains i and j, and $\sigma_{hpij}^{p(ap)}$ is the distance corresponding to the minimum of the hydrophobic potential. Superscripts p and ap refer to two kinds of relative orientations between a pair of side chains: parallel and antiparallel, which are determined by the orientation vectors of the pair of side chains. In the current model, the orientation vector of the side chain is defined as the vector from the united atom C_α to the side chain C_β. If the angle between the orientation vectors of a pair of side chains is larger than 120°, the relative orientation of this pair of side chains is considered as antiparallel, otherwise, parallel. The values of $\sigma_{hpij}^{p(ap)}$ for different pairs of side chains are listed in Table II. ε_{hp} is the strength of the hydrophobic interaction. In the current model, $\varepsilon_{hp}/\varepsilon_{hb}$ takes a value of 0.20. In addition, all side chains are held in positions relative to the backbone so that all residues are L isomers. In the present simulations, C_β atoms are rigidly attached to the C_α atoms; but considering more realistic

TABLE II. The values of parameter $\sigma_{hpij}^{p(ap)}$ of hydrophobic potential in units of Å.

	V (valine)		I (isoleucine)		A (alanine)	
	σ^{p}	σ^{ap}	σ^{p}	σ^{ap}	σ^{p}	σ^{ap}
V (valine)	5.0	5.0	5.0	5.50	5.0	5.0
I (isoleucine)	5.0	5.50	5.0	6.0	5.0	5.5
A (alanine)	5.0	5.0	5.0	5.50	5.0	5.0

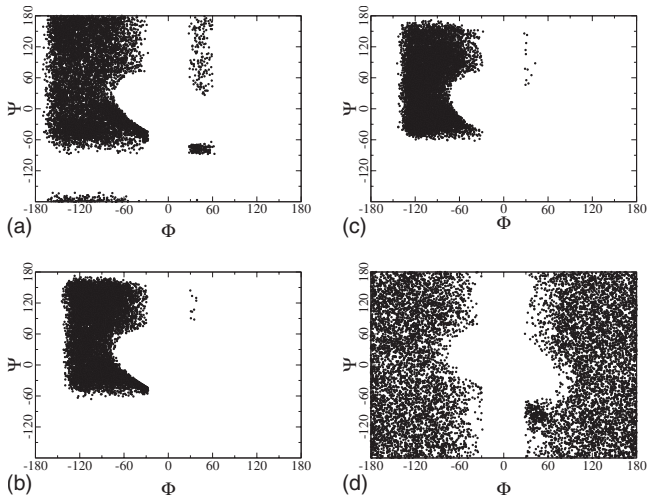

FIG. 2. Ramachandran plots for different amino acids with the current model at temperature $T=0.12$: (a) alanine, (b) valine, (c) isoleucine, and (d) glycine.

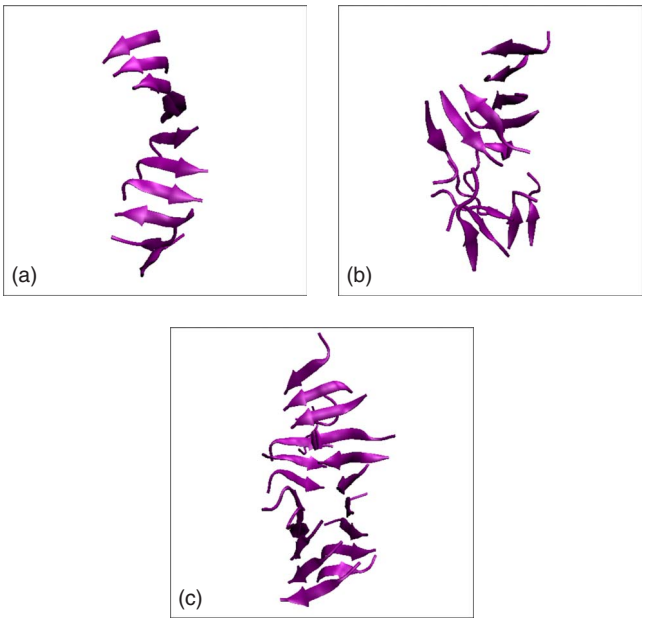

FIG. 3. (Color online) Three typical aggregation structures observed in our simulations: (a) single-layer β sheet, (b) amorphous β-sheet aggregate, and (c) fibril-like structure.

movement of side chains, the distances from C_β to N and C are allowed to fluctuate within 2.5% of their corresponding bond lengths listed in Table I. The Ramachandran plots for different amino acids based on our current model are shown in Fig. 2, which are in good qualitative agreement with the torsional angles φ and ψ distributions of real proteins [38,39].

III. RESULTS AND DISCUSSION

In the present study, all simulations were performed on systems containing 20 or 40 GGVVIA peptides started from random coil conformations. The aggregation structures spontaneously formed by random coil peptides are highly dependent on many environmental conditions including the peptide concentration and temperature. The peptide concentrations considered in the current study are 5.0, 15.0, and 30.0 mM, respectively. Three types of aggregation structures were observed in our simulations: single-layer β sheet, amorphous β-sheet aggregate, and fibril-like structures as shown in Fig. 3. Here, the amorphous β-sheet aggregate is defined as the amorphous structure in which β sheets are in disordered arrangement and the fibril-like structure is defined as the multisheet structure in which β sheets are parallel to the fibril axis and their strands are perpendicular to the fibril axis. To study the effects of temperature on aggregation, we calculated the dependence of the specific heat on temperature as shown in Fig. 4. The results are based on the 20-peptide system and averaged from at least ten independent simulations at each concentration and temperature. In each simulation, 1×10^7 MC steps were run for system equilibration and another 1×10^7 MC steps for data collection. The error bars were derived from the standard deviations. From Fig. 4, it can be seen that the specific heat curve displays a peak which corresponds to the phase transition between aggregated states (including all types of aggregation structures) at low temperatures and random coil state at high temperatures. The phase transition temperature increases as the peptide concentration is increased suggesting that the aggregates are easily formed at relatively high peptide concentrations, which is in accord with Nguyen's simulation results [30]. Furthermore, to study the effects of concentration and temperature on the aggregation structures, we calculated the percentages of peptides forming different aggregation structures, which are shown in Fig. 5. Figure 5(a) shows the percentage of peptides forming single-layered β sheets at different concentrations as a function of temperature. From Fig. 5(a), it can be seen that the percentage of peptides forming single-layered β sheets roughly increases with decreasing temperature suggesting that the single-layered β sheets are more stable at low temperatures. Moreover, the percentage of peptides forming single-layered β sheets decreases as the concentration is increased in the low temperature region indicating that the single-layered β sheets are easily formed and remain at low concentrations. However, due to the strong thermal fluctuation at high temperatures, only small single-layered β sheets (consisting of a few strands) may form and remain at high concentrations resulting in the increase in the percent-

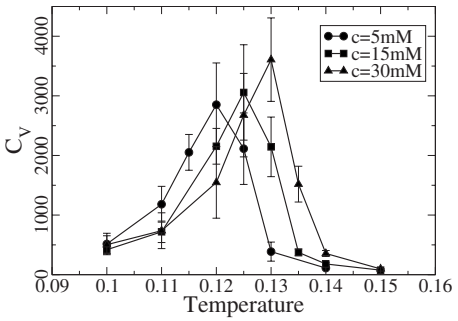

FIG. 4. Specific heat for different concentrations vs reduced temperature.

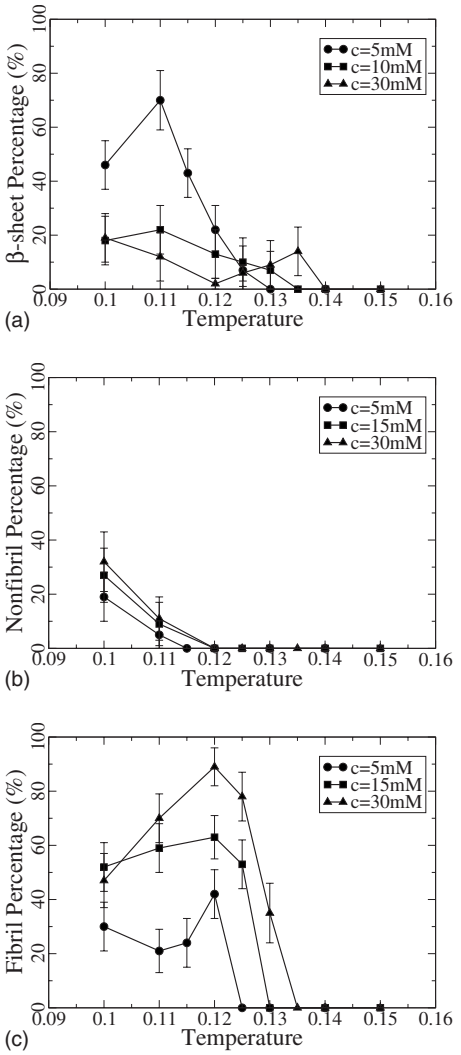

FIG. 5. The percentage of peptides forming different aggregation structures at different concentrations as a function of temperature: (a) the percentage of peptides forming single-layer β sheet, (b) the percentage of peptides forming amorphous β-sheet aggregate (nonfibril structure), and (c) the percentage of peptides forming fibril-like structure.

age of peptides forming single-layered β sheets with increasing the peptide concentration in the high temperature region. On the other hand, the system is easily trapped in amorphous β-sheet aggregate states at low temperatures due to strong hydrophobic interactions among the β sheets especially for the cases of relatively high concentrations. As shown in Fig. 5(b), the percentage of peptides forming amorphous β-sheet aggregates increases with increasing peptide concentration at low temperatures. Figure 5(c) shows the percentage of peptides forming fibril-like structures at different peptide concentrations as a function of temperature. From Fig. 5(c), it can be seen that the percentage of peptides forming fibril-like structures reaches a maximum in the intermediate temperature region and increases as the peptide concentration is increased indicating that the fibril-like structures are more easily formed at intermediate temperatures and high concentrations. In addition, from Fig. 5(c), it can be seen that the critical temperature of forming stable fibril-like structures

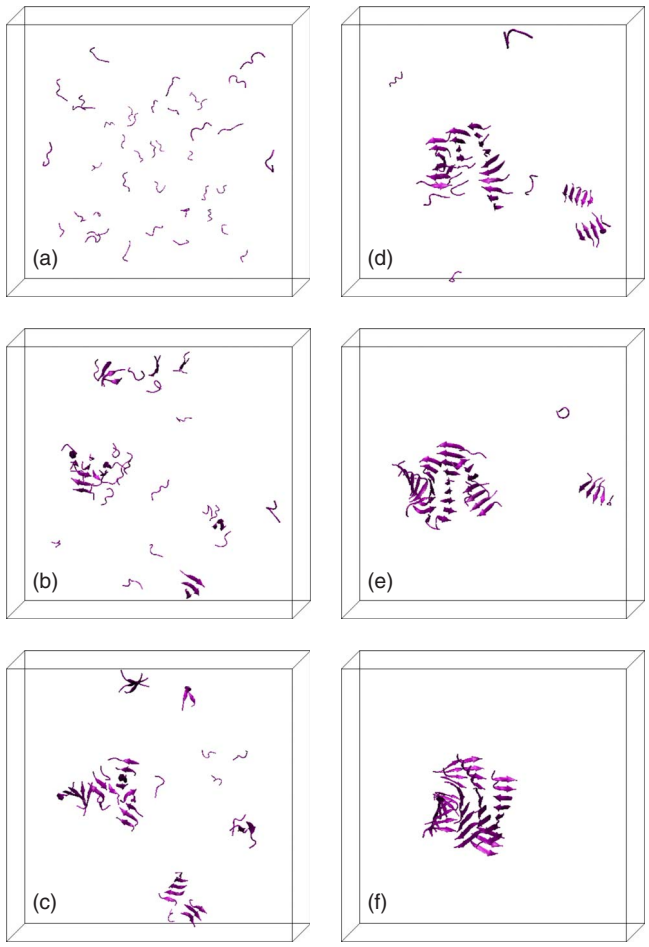

FIG. 6. (Color online) Snapshots of the formation of fibril-like structure in the 40-peptide system at $c=30$ mM and $T=0.12$: (a) MC step=0, (b) MC step=2×10^5, (c) MC step=2×10^6, (d) MC step=1×10^7, (e) MC step=2×10^7, and (f) MC step=3.5×10^7.

increases as the concentration is increased suggesting that the stability of fibril-like structures increases with the increasing peptide concentration. As a result, the systems at relatively high peptide concentrations have a broader temperature region of forming stable fibril-like structures.

To further study the formation and detailed structural features of fibril-like structures, we performed simulations on a 40-peptide system at different concentrations. Figure 6 shows snapshots of a typical formation process of a fibril-like structure in a 40-peptide system at the temperature $T=0.12$ and the peptide concentration $c=30.0$ mM. As shown in Fig. 6, the initial configuration is disordered and all peptides in the system are random coils [Fig. 6(a)]. Staring from the initial state, the peptides aggregate very quickly to form small amorphous aggregates as shown in Fig. 6(b). Subsequently, as shown in Fig. 6(c), the amorphous aggregates in the middle have grown to a big amorphous aggregate mainly consisting of small β sheets through self-organization of the peptides inside and attracting already-formed β sheets to its sides. In the meantime, other amorphous aggregates have collapsed or converted into individual β sheets. Subsequently, the big amorphous aggregate continues to grow by attracting individual peptides to the ends of each β sheet and

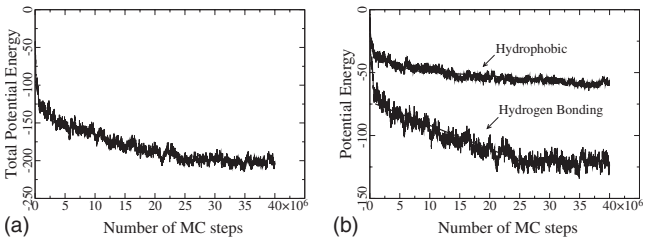

FIG. 7. (a) The total potential energy and (b) hydrogen bonding and hydrophobic potential energies for the 40-peptide system at $c=30$ mM and $T=0.12$ vs Monte Carlo step. All energies are given in units of ε_{hb}. The straight lines are drawn as guide for the eyes.

gradually converts into a left-handedly twisted four-sheet fibrillation nucleus through self-organization of the β sheets inside, as shown in Fig. 6(d). Meanwhile, only two individual β sheets have grown to become longer β sheets by attracting individual peptides to their ends and the others have disappeared. The four-sheet fibrillar nucleus continues to grow by attracting individual peptides to the ends of the β sheets and only one individual β sheet remains as shown in Fig. 6(e). Finally, as shown in Fig. 6(f), a left-handedly twisted four-sheet fibril-like structure is formed. From Fig. 6, it can be seen that there exist two kinds of fibril growth ways: β-sheet elongation (by attracting individual peptides to the ends of each β sheet in the fibril) and β-sheet addition (by attracting already-formed β sheets to the sides of the fibril). These observations are in agreement with Nguyen's simulation results [30], although the current study was performed using MC simulations. Figure 7 shows the total potential energy and hydrogen bonding and hydrophobic potential energies between peptides as a function of Monte Carlo step. From Fig. 7(a), it can be seen that the total potential energy declines very sharply in the very beginning, which subsequently decreases slowly and finally becomes flat, suggesting that the fibrillation process includes three main stages. The beginning short stage ($0\sim1\times10^6$ MC steps) corresponds to the rapid formation of amorphous aggregate by the peptides in random coil configurations. The following long stage ($1\times10^6\sim25\times10^6$ MC steps) corresponds to the fibril formation by self-organizations of peptides and β sheets. In fact, the fibril formation stage consists of two substages: the formation of the fibrillation nucleus ($1\times10^6\sim6\times10^6$ MC steps) and the fibril growth ($6\times10^6\sim25\times10^6$ MC steps). The last stage ($>25\times10^6$ MC steps) corresponds to the stable fibril structure. As shown in Fig. 7(b), the evolution of the hydrophobic potential follows a trend that is similar to the hydrogen bonding potential in the fibrillation nucleus formation stage ($1\times10^6\sim6\times10^6$ MC steps) indicating that both the β-sheet elongation and addition play important roles in the early stage of the fibril formation. However, in the following fibril growth stage ($6\times10^6\sim25\times10^6$ MC steps), the slope of the hydrogen bonding potential curve remains unchanged and the hydrophobic potential becomes largely constant, which suggests that the β-sheet elongation is energetically more favorable than the β-sheet addition and the fibril growth is dominated by the β-sheet elongation in the fibril growth stage. From Fig. 6, it can be seen indeed that once the fibril thickness (the number of β sheets in the fibril) reaches a certain size (four sheets for the case shown in Fig. 6 depending on the concentration in simulations), the continued fibril growth is dominated by the β-sheet elongation along the fibril axis rather than the lateral β-sheet addition. This growth characteristic was found in all simulations.

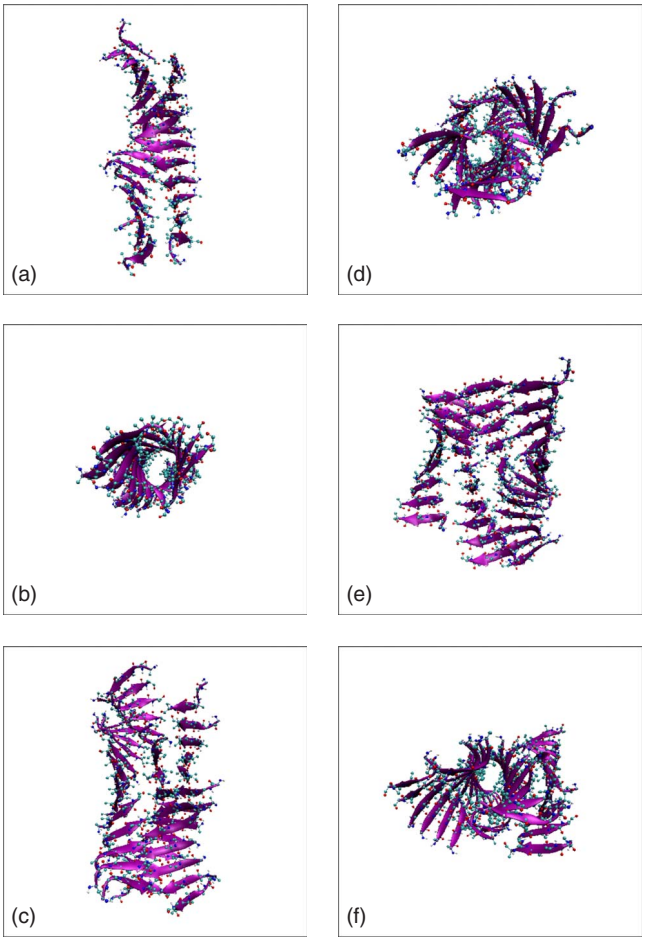

FIG. 8. (Color online) The enlarged figures of the fibril-like structures obtained in simulations at different concentrations viewed from different directions: (a) from the side and (b) along the fibril axis at $c=5$ mM and $T=0.12$; (c) from the side and (d) along the fibril axis at $c=15$ mM and $T=0.12$; (e) from the side and (f) along the fibril axis at $c=30$ mM and $T=0.12$.

Figure 8 shows the enlarged pictures of the fibril-like structures obtained in our simulations from different viewing angles. As shown in Fig. 8, the fibrils are composed of β sheets that are left-handedly twisted around the fibril axis. Here, the orientation vector of each β strand for the sequence GGVVIA is defined as the vector from the third amino acid valine to the last amino acid alanine. The twist angle between two successive β strands is defined as the acute angle between their orientation vectors. The averaged twist angle between successive β strands within each β sheet is $\approx12^\circ\pm2^\circ$ (the β strands at the end of each β sheet were excluded due to the distortion caused by thermal fluctuation), which is in good agreement with the result of Esposito *et al.* [40], who found the average twist angle to be $\approx11^\circ$ by simulating a pair of preformed GNNQQNY β sheets with the

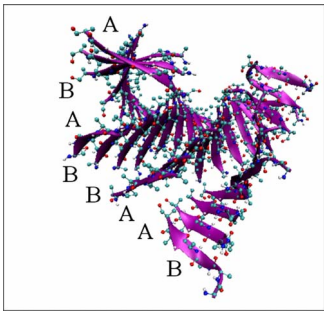

FIG. 9. (Color online) Snapshot showing the stacking patterns of the fibril-like structure formed in the 40-peptide system at $c=30$ mM and $T=0.12$.

GROMACS 3.2 software package. From Fig. 8, it can be seen that most adjacent peptides within each β sheet are in register and arranged in a parallel orientation. The average probability of parallel β-strand formation in the fibril-like structures in our simulations is $\approx 95 \pm 3\%$. This structural characteristic of parallel arrangement of β strands in the fibril-like structures is very similar to that found in the corresponding microcrystalline structure where β strands in each β sheet are exactly arranged in a parallel orientation [21]. In fact, the parallel arrangement of GGVVIA peptides in their aggregate structures is energetically favorable. This is due to the fact that one end of GGVVIA peptide includes the more hydrophobic alanine and isoleucine and the other end is glycine without a hydrophobic side chain; i.e., the hydrophobicity in the sequence GGVVIA is asymmetric. As a result, driven by the preferred hydrophobic interaction between the hydrophobic side chains at one end, the GGVVIA peptides tend to be arranged in parallel.

For the convenience of description, the surface of a β sheet formed by GGVVIA peptides where the side chains of alanines stand out is labeled as an "A" surface and the other is labeled as a "B" surface. In our simulations, three kinds of β-sheet stacking patterns, "A-B," "A-A," and "B-B," were observed in the fibril-like structures as shown in Fig. 9. However, β sheets stack only with A-B surfaces ("face to back") in the GGVVIA microcrystalline structure [21]. In addition, as shown in Fig. 8, the relative orientation of the adjacent β sheets in the fibril-like structures may be parallel or antiparallel, while the neighboring β sheets are arranged only in the antiparallel orientation in the microcrystalline structure [21]. The structural diversity of the fibrils obtained in our simulations may suggest that the fibrils are more complex and have more flexible structures than the corresponding microcrystals; correspondingly, the free-energy landscape of the fibril formation is very rugged and there exist many local free-energy minima rather than a distinct global minimum. Actually, the fibril-like structures in local minima are very stable and can survive many steps in the MC simulations. For example, as shown in Fig. 6, the formed fibril in the simulation is still stable even when the simulation was extended to 6×10^7 MC steps. These conclusions are in accordance with other simulation results in multichain systems [27,28,41] where the free-energy landscape of protein aggregation was found to be very rugged and there exist many local minima.

IV. CONCLUSIONS

In this paper, we studied the spontaneous formation of aggregated structures of GGVVIA peptides using Monte Carlo simulations and a coarse-grained model for amino acids. The hydrogen bonding and hydrophobic interactions play crucial roles in the formation and stabilization of native conformation and aggregated structures of proteins. In the present study, in order to capture correct structural features of fibrils, more realistic hydrogen bonding potential which can produce accurate hydrogen bond geometry in β sheets and hydrophobic potential functions where the side chain shapes of amino acids have been considered are employed. Three types of aggregation structures, single-layer β sheet, amorphous β-sheet aggregate, and fibril-like structures, were observed in our simulations. The growth of fibril in our simulations involves β-sheet elongation and lateral addition. The fibril-like structures have a common cross-β spine structure composed of naturally left-handedly twisted β sheets. The average twisting angle of the β sheets in the fibrils is about $12° \pm 2°$. Most of peptides in the same β sheet are in register and the average probability that neighboring β strands are parallel is $95 \pm 3\%$. The parallel arrangement of the neighboring β strands is in good agreement with the structural feature observed in the crystalline structure [21]. In addition, three possible β-sheet stacking styles, A-B, A-A, and B-B, were observed in the fibril-like structures in our simulations and the relative orientation between neighboring β sheets could be parallel or antiparallel, while the β sheets in the corresponding microcrystalline are stacked only with the A-B fashion with an antiparallel orientation [21]. The existence of different fibril-like aggregation configurations may suggest that the fibrils are more complex than the corresponding crystalline structure, and moreover, the free-energy landscape of fibril formation is very rugged and there exist many local free-energy minima rather than a distinct global minimum. To the best of our knowledge, the arrangement of β sheets in fibril-like structures has not been unambiguously determined and it is of great interest for these simulated structures to be tested by experiments. Although these conclusions were drawn based on our simulations with a coarse-grained model, they provide a vivid and detailed picture for the spontaneous formations of the fibril-like structures by random coil peptides. We have been currently applying the model to study the mechanism of fibril formation for extensive sequences under various conditions to demonstrate the validity of our model for mimicking different peptides. For example, the current model has been used to investigate the solvent effects on the fibril formation for short polypeptide GAAAAG. From an application viewpoint, further simulations are potentially useful in developing new medical strategies to prevent or slow down the formations of amyloid fibrils and constructing new supramolecular self-assembling functional biological and nanomaterials [42].

ACKNOWLEDGMENTS

This work was supported by the Welch Foundation (Grant No. A-1628). Y.Q.G. received partial support from the Searle Foundation.

[1] C. M. Dobson, Trends Biochem. Sci. **24**, 329 (1999).
[2] E. Žerovnik, Eur. J. Biochem. **269**, 3362 (2002).
[3] J. W. Kelly, Nat. Struct. Biol. **9**, 323 (2002).
[4] P. Westermark, FEBS J. **272**, 5942 (2005).
[5] A. J. Geddes, K. D. Parker, and E. Beighton, J. Mol. Biol. **32**, 343 (1968).
[6] E. D. Eanes and G. G. Glenner, J. Histochem. Cytochem. **16**, 673 (1968).
[7] M. Sunde, L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys, and C. Blake, J. Mol. Biol. **273**, 729 (1997).
[8] M. Sunde and C. Blake, Adv. Protein Chem. **50**, 123 (1997).
[9] A. T. Petkova, Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio, and R. Tycko, Proc. Natl. Acad. Sci. U.S.A. **99**, 16742 (2002).
[10] C. P. Jaroniec, C. E. MacPhee, V. S. Bajaj, M. T. McMahon, C. M. Dobson, and R. G. Griffin, Proc. Natl. Acad. Sci. U.S.A. **101**, 711 (2004).
[11] T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert, and R. Riek, Proc. Natl. Acad. Sci. U.S.A. **102**, 17342 (2005).
[12] N. Ferguson, J. Becker, H. Tidow, S. Tremmel, T. D. Sharpe, G. Krause, J. Flinders, M. Petrovich, J. Berriman, H. Oschkinat, and A. R. Fersht, Proc. Natl. Acad. Sci. U.S.A. **103**, 16248 (2006).
[13] K. Iwata, T. Fujiwara, Y. Matsuki, H. Akutsu, S. Takahashi, H. Naiki, and Y. Goto, Proc. Natl. Acad. Sci. U.S.A. **103**, 18119 (2006).
[14] C. Wasmer, A. Lange, H. V. Melckebeke, A. B. Siemer, R. Riek, and B. H. Meier, Science **319**, 1523 (2008).
[15] J. L. Jiménez, E. J. Nettleton, M. Bouchard, C. V. Robinson, C. M. Dobson, and H. R. Saibil, Proc. Natl. Acad. Sci. U.S.A. **99**, 9196 (2002).
[16] N. Rubin, E. Perugia, M. Goldschmidt, M. Fridkin, and L. Addadi, J. Am. Chem. Soc. **130**, 4602 (2008).
[17] A. K. Chamberlain, C. E. MacPhee, J. Zurdo, L. A. Morozova-Roche, H. A. O. Hill, C. M. Dobson, and J. J. Davis, Biophys. J. **79**, 3282 (2000).
[18] K. Lu, J. Jocob, P. Thiyagarajan, V. P. Conticello, and D. Lynn, J. Am. Chem. Soc. **125**, 6391 (2003).
[19] M. Sunde and C. C. F. Blake, Q. Rev. Biophys. **31**, 1 (1998).
[20] R. Nelson, M. R. Sawaya, M. Balbirnie, A. Ø. Madsen, C. Riekel, R. Grothe, and D. Eisenberg, Nature (London) **435**, 773 (2005).
[21] M. R. Sawaya, S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, H. T. McFarlane, A. Ø. Madsen, C. Riekel, and D. Eisenberg, Nature (London) **447**, 453 (2007).
[22] B. Y. Ma and R. Nussinov, Protein Sci. **11**, 2335 (2002).
[23] J. Gsponer, U. Haberthür, and A. Caflisch, Proc. Natl. Acad. Sci. U.S.A. **100**, 5154 (2003).
[24] P. Soto, A. Baumketner, and J. E. Shea, J. Chem. Phys. **124**, 134904 (2006).
[25] Z. Q. Zhang, H. Chen, H. J. Bai, and L. H. Lai, Biophys. J. **93**, 1484 (2007).
[26] B. Strodel, C. S. Whittleston, and D. J. Wales, J. Am. Chem. Soc. **129**, 16005 (2007).
[27] G. Favrin, A. Irback, and S. Mohanty, Biophys. J. **87**, 3657 (2004).
[28] A. Melquiond, G. Boucher, N. Mousseau, and P. Derreumaux, J. Chem. Phys. **122**, 174904 (2005).
[29] A. Melquiond, J. C. Gelly, N. Mousseau, and P. Derreumaux, J. Chem. Phys. **126**, 065101 (2007).
[30] H. D. Nguyen and C. K. Hall, Proc. Natl. Acad. Sci. U.S.A. **101**, 16180 (2004).
[31] G. Bellesia and J. E. Shea, J. Chem. Phys. **126**, 245104 (2007).
[32] S. Takada, Z. Luthey-Schulten, and P. G. Wolynes, J. Chem. Phys. **110**, 11616 (1999).
[33] A. V. Smith and C. K. Hall, Proteins **44**, 344 (2001).
[34] H. D. Nguyen, A. J. Marchut, and C. K. Hall, Protein Sci. **13**, 2909 (2004).
[35] Y. Mu and Y. Q. Gao, J. Chem. Phys. **127**, 105102 (2007).
[36] N. Y. Chen, Z. Y. Su, and C. Y. Mou, Phys. Rev. Lett. **96**, 078103 (2006).
[37] M. R. Betancourt, J. Chem. Phys. **123**, 174905 (2005).
[38] G. N. Ramachandran and V. Sasisekharan, Adv. Protein Chem. **23**, 283 (1968).
[39] S. S. Zimmerman, M. S. Pottle, G. Némethy, and H. A. Scheraga, Macromolecules **10**, 1 (1977).
[40] L. Esposito, C. Pedone, and L. Vitagliano, Proc. Natl. Acad. Sci. U.S.A. **103**, 11533 (2006).
[41] T. Cellmer, D. Bratko, J. M. Prausnitz, and R. Blanch, Proc. Natl. Acad. Sci. U.S.A. **102**, 11692 (2005).
[42] I. Cherny and E. Gazit, Angew. Chem., Int. Ed. **47**, 4062 (2008).