Free Energy Landscapes for Amyloidogenic Tetrapeptides Dimerization

A. Baumketner and J.-E. Shea
Department of Chemistry and Biochemistry, University of California, Santa Barbara, California

ABSTRACT The oligomerization of four peptide sequences, *KFFE*, *KVVE*, *KLLE*, and *KAAE* is studied using replica-exchange molecular dynamics simulations with an atomically detailed peptide model. Previous experimental studies reported that of these four peptides, only those containing phenylalanine and valine residues form fibrils. We show that the fibrillogenic propensities of these peptides can be rationalized in terms of the equilibrium thermodynamics of their early oligomers. Thermodynamic stability of dimers, as measured by the temperature of monomer association, is seen to be higher for those peptides that are able to form fibrils. Although the relative high and low stabilities of the *KFFE* and *KAAE* dimers arise from their respective high and low interpeptide interaction energies, the higher stability of the *KVVE* dimer over the *KLLE* system results from the smaller loss of configurational entropy accompanying the dimerization of *KVVE*. Free energy landscapes for dimerization are found to be strongly sequence-dependent, with a high free energy barrier separating the monomeric and dimeric states for *KVVE*, *KLLE*, and *KAAE* sequences. In contrast, the most fibrillogenic peptide, *KFFE*, displayed downhill assembly, indicating enhanced kinetic accessibility of its dimeric states. The dimeric phase for all peptide sequences is found to be heterogeneous, containing both antiparallel β-sheet structures that can grow into full fibrils as well as disordered dimers acting as on- or off-pathway intermediates for fibrillation.

INTRODUCTION

A number of debilitating diseases of the central nervous system, such as Parkinson's, Jacob-Creutzfeld, and Alzheimer's disease, are linked to the misfolding and subsequent aggregation of proteins into insoluble fibrils (1–3). The precise role of these fibrils in the etiology of the disease is unclear, with increasing evidence pointing to strong cytotoxicity of smaller prefibrillar oligomers. Much of the research on aggregation over the last decade has focused on a structural characterization of fibrils with particular emphasis on the factors that contribute to the overall stability of these aggregation end-products. A number of attempts have been made to rationalize fibrillation of proteins and peptides in terms of their amino acid sequences. Hydrophobic tail-to-tail contacts, for instance, have been shown to stabilize octanoyl-mutated Aβ16-22 peptides (4), locking them into the parallel β-sheet orientation. The possible role of electrostatic interactions in enhancing the stability of amyloid fibrils through salt bridges has been investigated for a number of amyloidogenic peptides (5–7). Recent computer simulations have also sought to characterize the stability (8–11) of fiberlike, β-sheet structures. Although providing invaluable information about the nature of the resulting fibrils, the above structural studies offer limited insight into how and why aggregation proceeds. A detailed molecular mechanism of fibril formation, which takes into account both the disaggregated as well as aggregated states, has yet to emerge from such studies.

The aim of this article is to use fully atomic molecular dynamics simulations to investigate the process of aggregation in its entirety, starting from the monomeric state all the way through to the formation of oligomeric species. We focus on the simplest aggregation process, that of dimerization, and consider four tetrapeptide sequences experimentally shown to form fibrils (12). Our analysis is centered on the study of the underlying energy landscape, which offers a comprehensive picture of both the thermodynamics and kinetics of aggregation. Aggregation can be considered a nucleation-growth process, in which the formation of a critical nucleus allows fibrillation to proceed (13). Under conditions of perfect thermodynamic equilibrium, this critical nucleus corresponds to an aggregate of a size such that the enthalpic gain of forming the assembly balances the associated loss of entropy. Aggregation can be viewed as mostly governed by equilibrium thermodynamics (14), so long as the monomers and various oligomers and fibrils correspond to distinct physical phases and not to kinetically trapped metastable phases (15). Peptide aggregation proceeds directly from a natively unfolded state, and the free energy barrier between this unstructured native state and the nucleus governs the ease with which aggregation can occur. Aggregation appears to be an intrinsic property of polypeptides chains and given appropriate external conditions (temperature, concentrations, etc.), any peptide can form aggregates (16). Whether or not the aggregate forms depends on the nature of the free energy surface under the given set of conditions considered.

In this work, we find that the fibrillation propensities of model aggregating peptides can be predicted from the thermodynamic stability and kinetic accessibility of early oligomers. We consider dimers of the four tetrapeptide

Submitted January 7, 2005, and accepted for publication May 31, 2005.

Address reprint requests to J. E. Shea, E-mail: shea@chem.ucsb.edu.

A. Baumketner's permanent address is the Institute for Condensed Matter Physics, 1 Svientsitsky Str., Lviv 79011, Ukraine. E-mail: andrij@icmp.lviv.ua.

0006-3495/05/09/1493/11 $2.00

doi: 10.1529/biophysj.105.059196

sequences, *KFFE*, *KVVE*, *KLLE*, and *KAAE*, studied by Tjernberg et al. (12). The peptides are modeled using atomically detailed representation (17) and an accurate implicit solvation method (18). Spherical cages with repulsive walls are introduced to encapsulate the peptides to enforce constant concentration conditions. Molecular dynamics simulations are carried out at a range of temperatures according to the replica-exchange algorithm (19). At a characteristic temperature T_a, specific to each system, the peptides undergo an association transition from monomeric states at $T > T_a$ into a dimeric state at $T < T_a$. This transition temperature T_a is used as a gauge of the thermodynamic stability of the dimers relative to the monomers at a given temperature T. The higher the value of T_a, the greater the stability of the dimer. Our simulations predict that the sequences can be ranked in terms of stability as follows: *KFFE* > *KVVE* > *KLLE* > *KAAE*. This ranking correlates well with the experimental findings of Tjernberg et al. (12), which indicate that only the *KFFE* and *KVVE* peptides are able to form fibrils, whereas *KLLE* and *KAAE* peptides do not fibrillate under the same set of conditions.

To gain further insight into the mechanism of peptide assembly, we construct free energy surfaces for dimerization. We find that the three least fibrillogenic sequences, *KVVE*, *KLLE*, and *KAAE*, must undergo a cooperative two-state transition to dimerize, a process that requires overcoming a significant free energy barrier separating monomeric and dimeric states. In contrast, the *KFFE* free energy surface is downhill in nature, with no significant barrier present. As barrierless transitions imply faster rates, the *KFFE* sequence emerges as having not only the most thermodynamically stable dimer, but also the most kinetically accessible one. We propose new experiments to probe the peptide rankings, the reasons for which fibrillation of other tetrapeptide sequences is inhibited, and to identify experimental conditions under which each peptide sequence can fibrillate.

METHODS AND MODELS

Protein model

The *KFFE*, *KVVE*, *KLLE*, and *KAAE* peptides studied in this work were modeled using the CHARMM19 united-atom force field (17). Covalent bonds involving hydrogen atoms were kept constant using the constrained dynamics SHAKE algorithm (20). The dynamics were propagated using Langevin method with the friction constant $\eta = 20\ \mathrm{ps}^{-1}$ and a time step δt of 2 fs. No cutoffs for the long-range electrostatic and van der Waals interactions were used. Solvent effects were taken into account implicitly, through the Generalized Born solvation method (18) parameterized for peptides. The force-field/solvation model scheme employed in this article has been successfully used by Jang et al. (21) to fold peptides of various native-state geometries. The Generalized Born method accounts only for the electrostatic component of the solvation free energy, but not the nonpolar (hydrophobic) (22) component. This approximation is justified for small peptides such as the ones considered in this work. Indeed, the solvent-accessible surface area, commonly used to model hydrophobic effects, is not suitable at small length scales; the strength of the hydrophobic effect scales with the volume and not the surface area of the solute (23,24) for interacting objects of a diameter 10 Å or less. Only for objects larger than 10 Å does the hydrophobic effect become increasingly stronger and dependent on the surface area terms (23). In addition, the hydrophobic effect at small length scales is mostly entropic in nature and, as a result, weaker than the other types of interactions (23). Since the peptides investigated here are smaller than 10 Å in size in the dimeric state, the hydrophobic solvation free energy term can be neglected. We note that the hydrophobic solvation term is very important at larger length scales and will play a role in bringing pre-assembled β-sheets together to form fibrils.

Confinement potential

The peptides were confined to model constant concentration conditions and prevent the peptides from drifting away. We used spherical cages with repulsive walls developed by Klimov et al. (25) in the context of chaperone-assisted protein-folding. Inside a cage of radius R, an atom at position $\vec{r}$ from the origin of the reference frame is subject to a confining interaction,

$$U_c(\vec{r}) = 4e\frac{\pi R}{5r}\left[\left(\frac{\sigma}{r-R}\right)^{10} - \left(\frac{\sigma}{r+R}\right)^{10}\right], \qquad (1)$$

where parameters σ and e were assigned values 1 Å and 1 kCal/mol, respectively, for all atoms of the peptides, including hydrogens. The above potential energy is obtained by integrating the standard $1/r^{12}$ Lennard-Jones repulsion term over the surface of the sphere. We do not expect the details of the confining potential used here to qualitatively affect our results and conclusions.

The radius of the cage was chosen to be $R = 17$ Å for all peptide systems studied. This corresponds to a molecular concentration of 160 mM which is approximately three orders-of-magnitude higher than the concentration of 200–300 μM at which the experiment of Tjernberg et al. (12) was reported. Simulations performed at larger R could produce a better agreement with the experiment, but would be prohibitively expensive computationally, due to the need to sample the increased volume of the configuration space that corresponds to monomeric states. Such simulations would not change our main conclusions regarding the dimeric phase, since we explicitly verified that the dimers of all four peptides studied here are not affected by the presence of the spherical constraints. Using the same sphere radius R for all four peptide sequences allows us to test their aggregation propensities relative to one another and predict which peptides will form stable dimers under a given set of conditions. We also explicitly verified that the conformational statistics of the monomers of all four peptides studied are not affected either by the repulsive potential of the encapsulating sphere or by the presence of the second peptide.

Replica-exchange molecular dynamics

Replica-exchange (REX) molecular dynamics simulations (19,26) were used to ensure an efficient exploration of conformational space. The REX protocol is specifically designed to overcome trapping in potential energy minima, a problem that typically hampers conventional constant temperature simulations. Within the REX protocol, simulations of a number of identical copies (replicas) of the original system are run in parallel. Temperature is treated as a dynamical variable and assigned on a per-replica basis. Periodically, every τ time steps, attempts to swap replicas n and m, simulated at inverse temperatures β_i and β_j, respectively ($j = i \pm 1$), are made. The swaps are accepted with the probability $min\ \{1, e^{\Delta\beta\Delta E}\}$, where $\Delta E = E_n - E_m$, $\Delta\beta = \beta_j - \beta_i$ and E_k denotes the potential energy of replica k. The procedure allows for unrestrained walks for each replica up and down in temperature, from a minimum value T_{min} to maximum value T_{max}. This results in overall accelerated dynamics associated with facilitated crossing of potential energy barriers at elevated temperatures.

For the four peptide sequences studied in this work, the parameter T_{max} was adjusted so that each replica underwent multiple associations/dissociations over the course of the simulation. The minimum temperature

T_{min} was chosen in short trial runs such that it lies in the vicinity of the estimated transition temperature for assembly T_a. Specifically, T_{min} and T_{max} were chosen to be 325 and 500 K for the *KFFE* peptide, 270 and 800 K for the *KVVE* and *KLLE* peptides, and 260 and 724 K for the *KAAE* peptide. Six independent replicas, at temperatures exponentially spaced between T_{min} and T_{max}, were considered for the phenylalanine mutant, and eight replicas for the other three peptides. Acceptance ratio for the swaps among replicas was seen to be uniform in the considered temperature range and never <30%. Replica swaps were attempted every 2000 time steps for the leucine mutant and 1000 time steps for the other peptides. A total of 200,000 swap attempts were made for the *KFFE* and *KAAE* peptides, 250,000 for the *KVVE* peptide, and 100,000 for the *KLLE* peptide. This resulted in the total simulation time of 400 ns for the phenylalanine, leucine, and alanine mutants and 500 ns for the valine mutant. We note that the relaxation time observed in our simulations (as judged from the fluctuations of Rg) would be 50–100 times longer, had we used the conventional constant temperature algorithm. Hence the use of the REX protocol was essential for these simulations, which took on the order of one week per system to complete.

The energy distributions and joint energy-Rg distributions collected during the simulations at varying temperature were used in the multiple histogram reweighting technique (27) to extract the canonical-ensemble averages from the REX trajectories. Mean potential energy, Rg, specific heat, and fluctuations of Rg were computed in this way as a function of temperature.

RESULTS

Thermodynamics of dimerization

All four peptide sequences can undergo a structural transition from monomeric phase into the dimeric phase. Dimerization is associated with a decrease in the overall size of the system (as measured by the radius of gyration Rg of the two peptides). The Rg and mean-squared deviation of Rg are plotted as a function of temperature in Fig. 1 *a* for all four peptide sequences. At high temperatures (700 K), all four peptides occupy conformations with a similar large Rg = 8–9 Å. The size of the system is controlled at these temperatures exclusively by the confining sphere. As the temperature decreases, the distance between the peptides starts to diminish, driven by favorable intermolecular interactions, reaching Rg ~4–5 Å at T ~300 K. The manner in which the reduction of the molecular size occurs is specific to each peptide sequence. At a transition midpoint temperature T_a the peptide can be found with equal probability either in monomeric or dimeric conformations and, as a result, the deviation of the molecular size ΔRg is maximal. Below this temperature T_a, which we will refer to as the association temperature, the dimeric conformations with low Rg become more populated than the monomeric states. Conversely, for $T > T_a$ the monomeric states with large Rg dominate the conformational statistics. The difference ΔT between a given temperature T and the temperature at which the association occurs is a measure of the thermodynamic stability of the dimeric phase. Large values of ΔT correspond to systems that form stable dimers at this temperature T. Small (approaching zero) values of ΔT indicate that the dimers are not very stable at the temperature T and can easily break apart. Under such conditions the system can exist as a mixture of dimers and monomers. The mean-square deviation of Rg as a function of temperature is shown in Fig. 1 *b* for all four peptide sequences studied in this work. The association temperature T_a is identified from this figure as the maximum in $\langle \Delta Rg^2 \rangle (T)$. The association temperature $T_a = 325$ K of the phenylalanine mutant is higher than the T_a of all other peptides. The transition temperatures of the others are, in decreasing order: T_a = 285 K for the valine mutant, 275 K for the leucine mutant, and 225 K for the alanine mutant. In terms of thermodynamic stabilities of the dimers at a given temperature and concentration, the order of the transition temperatures implies that the sequences can be ranked in order of increasing dimer stability as follows: *KAAE* < *KLLE* < *KVVE* < *KFFE*.

We now turn to an in-depth analysis of the equilibrium thermodynamics of association to rationalize, at a microscopic level, the ordering of dimer stabilities. At the temperature of association $T = T_a$, both the monomeric and dimeric phases have equal free energy $F_d = F_m$. Hence,

$$T_a = \Delta U / \Delta S. \quad (2)$$

ΔU in this expression denotes the gain in potential energy due to dimerization and ΔS is the loss of entropy associated with it. Monomers are favored by entropy and are thus predominantly populated at high temperatures ($T > T_a$), whereas the dimeric phase dominates at low temperatures ($T < T_a$) due to favorable potential energy contributions.

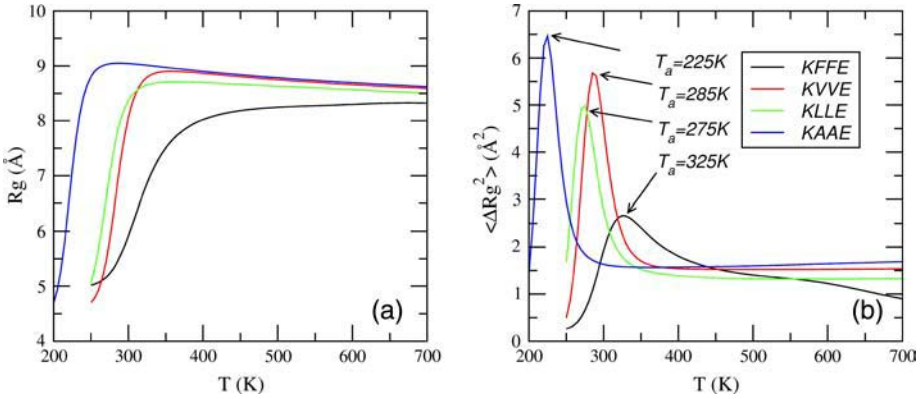

FIGURE 1 Temperature-dependence of the radius of gyration Rg (*a*) and the mean-square deviation $\langle \Delta Rg^2 \rangle$ (*b*) computed in replica-exchange simulations for dimers of the four amyloidogenic peptides *KFFE*, *KVVE*, *KLLE*, and *KAAE*. Rg was evaluated over positions of α-carbons. The temperature of association is identified as the maximum in $\langle \Delta Rg^2 \rangle (T)$.

The potential energy term can be decomposed into a contribution from the monomers self-interaction ΔU_{mon}, and from the interaction energy between monomers ΔU_{int}:

$$\Delta U = \Delta U_{\text{mon}} + \Delta U_{\text{int}}. \quad (3)$$

The self-interaction or intrachain term ΔU_{mon} accounts for the interaction energy among the atoms of only one of the two peptides considered in the simulations, irrespective of whether this peptide is part of a dimer or remains in a monomeric state. An analysis of ΔU_{mon} as a function of the radius of gyration (Rg) reveals that the intrachain energy is the same at both large and small Rg for all four peptides, implying that the transition into dimeric states is not accompanied by a substantial change in potential energy of monomers. The total dimerization potential energy ΔU is hence mostly due to the intermonomer interaction energy ΔU_{int}.

Free energy surfaces as a function of ΔU_{int} and the radius of gyration are plotted in Fig. 2 for each peptide near its transition temperatures ($T = 285$ K for *KVVE*, $T = 270$ K for *KLLE*, $T = 240$ K for *KAAE*, and $T = 325$ K for *KFFE*). The surfaces reveal a clear energetic benefit for monomers to populate dimeric conformations. Comparing the monomer interaction energy at $Rg \sim 9$ Å and 5 Å we find that the dimers of *KVVE* and *KLLE* peptides are stabilized by ~15 kCal/mol, those of *KFFE* peptides by ~16 kCal/mol, and those of *KAAE* by 10 kCal/mol. These data suggest that the peptides can be ranked in order of increasing interaction energy ΔU as *KAAE* < *KVVE* ~ *KLLE* < *KFFE*. Consequently, since the peptides rich in phenylalanine and alanine residues display the highest and lowest interaction energy, respectively, we conclude that the observed highest/lowest relative stabilities of their dimers are due to the potential energy contributions to the free energy. Additional detailed analysis of the different components of the interaction energy between *KFFE* monomers reveals that the terms pertaining to the *PHE* atoms contribute to the greatest extent to the favorable total energy.

This observed ordering of the intermonomer potential energy does not explain, however, why the association temperature is higher for the valine-based peptides than for the leucine-based system. Since both the *KVVE* and *KLLE* systems have comparable ΔU values, the different stabilities of their dimers must be a result of the relative entropic contributions to T_a. The total loss of entropy due to dimerization is composed of translational ΔS_{tr} and conformational ΔS_c contributions,

$$\Delta S = \Delta S_c + \Delta S_{\text{tr}}. \quad (4)$$

The loss of translation entropy occurs because the cavity volume V^2 available to two monomers moving independently is reduced to V once they form a dimer. Assuming that the molecular volume V_p is approximately equal for all four peptides, then $\Delta S_{\text{tr}} \sim -log(V - V_p)$ becomes independent of peptide type, as long as the cavity volume V is kept constant. The conformational entropy ΔS_c is then responsible for the variations in T_a observed for the valine- and leucine-based systems. Since ΔS_c, which measures the number of molecular conformations available to the system under a given set of conditions, is not directly available from simulations, we must turn to a qualitative analysis of ΔS_c. We assess the magnitude of ΔS_c from projections of the conformational space available to the molecules onto a subspace of selected configurational variables. We use ϕ/ψ dihedral angles as the main descriptors of conformational states. Free energy maps

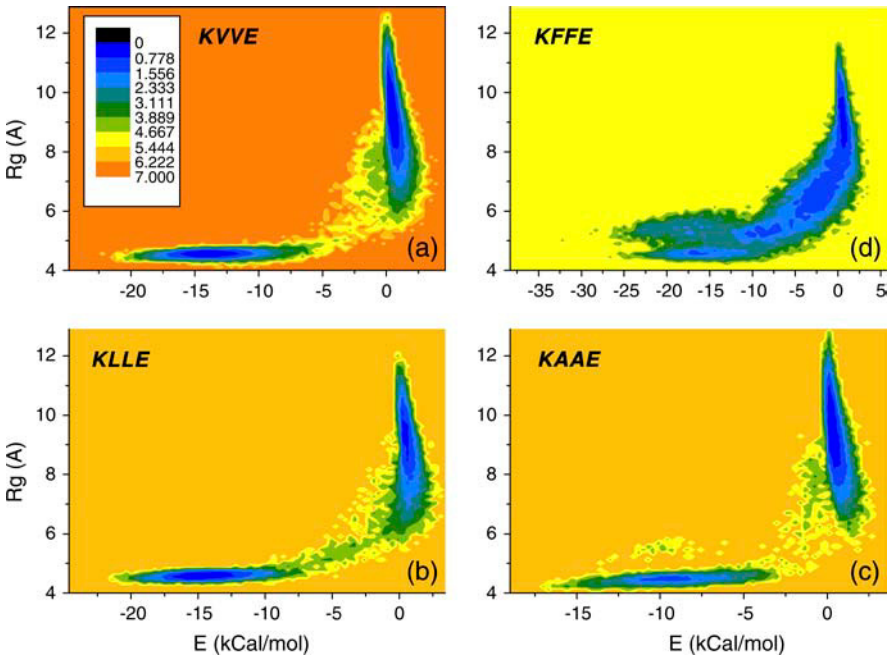

FIGURE 2 Free energy surface for the *KVVE*, *KFFE*, *KLLE*, and *KAAE* dimers, plotted as a function of the interaction energy between the two monomers and radius of gyration of the system, at temperatures near the respective transition temperatures of each peptide (*a*) $T = 270$ K, (*b*) $T = 285$ K, (*c*) $T = 240$ K, and (*d*) $T = 325$ K. The free energy is shown in units of k_BT.

as a function of ϕ/ψ angles for the two Leu residues of the *KLLE* sequence are shown in Fig. 3. Considering that dimers are formed when $Rg < 5$ Å, we computed the maps for monomeric and dimeric states separately. The peptide conformational entropy is assessed from these maps from the surface area of the occupied regions. Fig. 3 reveals that there are three main, highly populated regions in the ϕ/ψ maps of Leu-1 and Leu-2 residues for the peptide in monomeric form. Region I is centered at coordinates (−100,125) and corresponds to extended, β-strandlike conformations. Region II is found at the minimum corresponding to α-helical conformations (−100,−50). Region III, located at (50,−100), corresponds to conformations without well-defined secondary structure. Upon transition to the dimeric states, the following two features emerge from the free energy maps: Firstly, the surface area under region I (its volume) becomes smaller when the transition is made. Secondly, the random-coil minimum (III) and α-helical minimum (II) completely disappear for Leu-1 whereas the volume of region II is significantly reduced for Leu-2. We interpret these changes in the ϕ/ψ maps as indicative of conformational restrictions imposed on the monomeric *KLLE* peptides due to their integration into dimers. Based on the data presented in Fig. 3, we conclude that dimerization of *KLLE* peptides is accompanied by a reduction in conformational entropy, $\Delta S_c^L < 0$. A similar map to Fig. 3 was obtained for the *KVVE* sequence.

The precise extent of entropy loss during dimerization depends on the relative population of minima I, II, and III in the free energy map. The more populated the minimum I for monomeric states, the lesser the resulting entropy loss. We find that for the leucine-rich peptide, the β-strand minimum I is populated up to 90% of the time. For the valine-based peptide this number is slightly higher, 93%, indicating a lower loss of entropy associated with the dimer transition, $|\Delta S_c^L| > |\Delta S_c^V|$. Returning to Eq. 2, we now see the implications of this relationship for the ordering of the dimerization temperature of the peptides. Recall that the interaction energy between monomers of *KVVE* and *KLLE* peptides is almost equal, $\Delta U^L \sim \Delta U^V$. Therefore, the above ordering of the entropic contributions implies that $T_a^V > T_a^L$, in agreement with our simulations results presented in Fig. 1.

Since the difference between β-structure contents of monomeric *KVVE* and *KLLE* is small (on the order of 3%), we undertook further tests to verify that such a small difference can indeed bring about a sizable change in dimerization temperature. In particular, we sought to determine the value of the transition temperature of the *KLLE* sequence if the monomers were forced to adopt β-strandlike conformations. This was done by performing simulations with parameters identical to those described in Methods and Models for the original *KLLE* system but using additional biasing potentials. To limit conformational sampling to the region of the β-strand minimum I shown in Fig. 3, we employed quadratic energy terms centered around (−90,120) for ϕ/ψ angles of Leu[1] and around (−100,130) for Leu[2]. As a result, the non-β-strand minima II and III in the monomeric ϕ/ψ map for this new system were completely eliminated, with the free energy map of dimers undergoing changes mostly related to the shape of the populated regions. The biasing potential did not affect the intermonomer interaction energy, which remained near 15 kCal/mol. Based on the newly generated ϕ, ψ map we concluded that the loss of conformational entropy due to dimerization was significantly

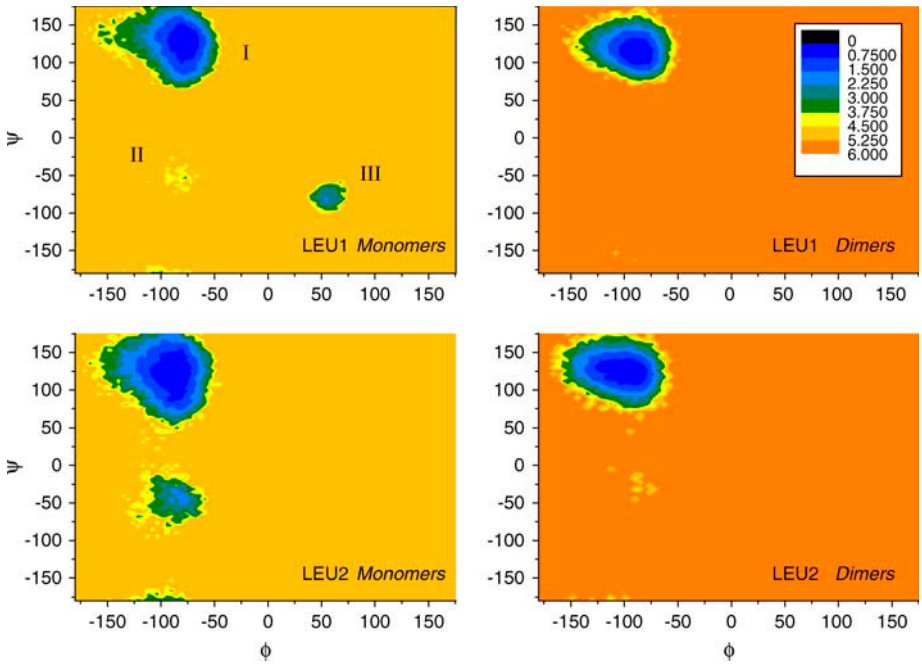

FIGURE 3 Free energy map in ϕ/ψ dihedral angles for the Leu residues of *KLLE* di-peptide system, plotted at $T = 270$ K. The data are shown for both monomeric and dimeric states of the peptide, assuming that dimers are characterized by gyration radius <5 Å. The loss of conformational freedom due to dimerization is readily apparent.

reduced, if not completely eliminated for the modified system. Interestingly, statistical analysis of the radius of gyration at varying temperature revealed that the association transition temperature had risen ~10 K as a result of the biasing, from 275 K for the original system to 285 K for the peptide with biased sampling. This is the same difference in T_a as the one observed for the *KVVE* and *KLLE* systems. This finding further confirms that the conformational entropy contribution to the free energy is responsible for the relative stability of *KLLE* and *KVVE* dimers.

Fibril-productive β-sheet conformations compete at low temperature with amorphous aggregates

The nature of the dimeric phase for each peptide sequence is analyzed to gain insight into possible mechanisms for fibrillogenesis. Different dimeric conformations can give rise to different scenarios for self-assembly of larger oligomers and a structural characterization of these conformations can provide important clues about their ability to generate amyloid structures. Certain conformations may serve as seeds for the subsequent growth of larger oligomers whereas others may be dead-end products, remaining in a dimeric form or possibly dissociating, allowing for their later assembly with preexisting fibrils.

The aggregated phase for all four tetrapeptide sequences is found to be inhomogeneous, presenting both β-sheet conformations that can lead to fibrillation as well as dimers that do not possess any well-defined secondary structure. These amorphous dimers, which constitute the majority of associated conformations, differ in both shape and potential energy. Fig. 4 shows a plot of potential energy as a function of the RMS deviation from the ideal β-sheet conformation for the case of the *KFFE* peptide sequence. (Similar plots are obtained for the other sequences.) All conformations are grouped into four distinct clusters, based on the hierarchical clustering method implemented in the MMTSB tools set (28). The first cluster, cluster C1, contains ideal antiparallel β-sheet conformations (see Fig. 5) as well as conformations in which one of the strands is flipped over, as shown in Fig. 6.

The ideal β-sheet dimers have a 5–10 kCal/mol higher potential energy than their misshaped variants. It is clear from Figs. 5 and 6 that only the dimers in the regular β-sheet conformation are able to immediately bind a third chain in a β-sheet hydrogen-bonding conformation. The distorted conformations (Fig. 6) lack the necessary pattern of hydrogen bond donors and acceptors on one side of the complex to extend the sheet. Both the ideal and distorted β-sheets have higher potential energy than the conformations belonging to the three remaining clusters (C2, C3, and C4). Parallel orientations of the β-sheets are not observed in any cluster. The conformations from clusters C2, C3, and C4 are very dissimilar to each other as well as from a β-sheet conformation. The average C_α RMS deviation among the centroids of all the four clusters is ~4 Å. A feature shared by non-β conformations is that they all possess bends in the individual monomeric strands that allow for a larger number of hydrogen bonds to be formed than in the β-sheet case. The centroid of the cluster with the lowest potential energy (cluster C4) is displayed in Fig. 7. Compared to the ideal β-sheet configuration, which forms four hydrogen bonds, the C4 centroid has six intact bonds and thus possesses a lower potential energy. The structure is stabilized by interactions between the positively charged side chain of the GLU residue of one of the monomers with the negatively charged LYS side chains of either the same monomer or the other monomer. As is the case for the distorted β-sheet structures,

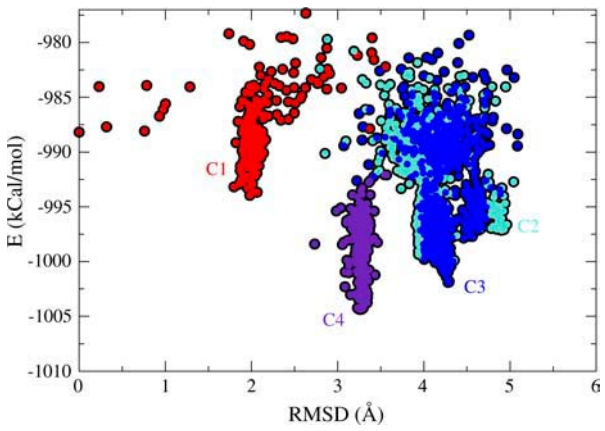

FIGURE 4 Clustering of the energy-minimized dimeric conformations obtained in the replica-exchange molecular dynamics simulations of the *KFFE* peptide. The data are shown as a function of potential energy and C_α RMS deviation from the β-sheet conformation (represented in Fig. 5). The clustering was carried out using the hierarchical clustering method as implemented in the MMTSB tools set (28).

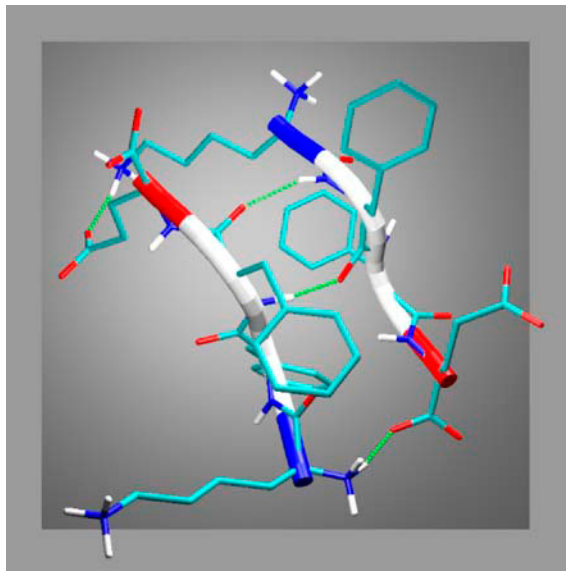

FIGURE 5 Antiparallel β-sheet conformation observed in replica-exchange simulations of the phenylalanine tetrapeptide *KFFE*.

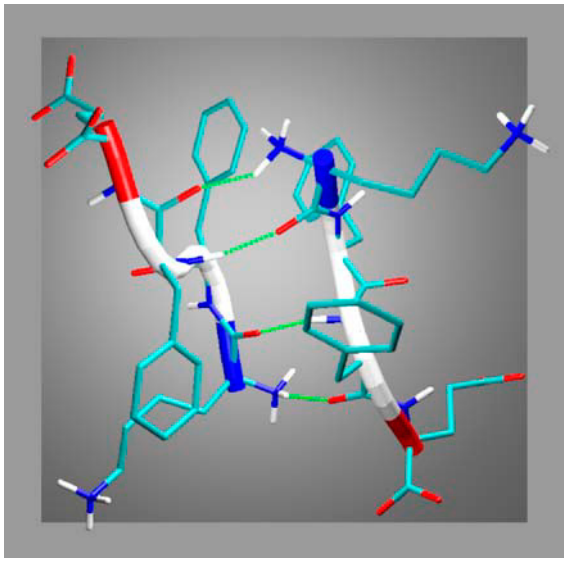

FIGURE 6 A β-sheet conformation in which two strands are aligned parallel but one of them is flipped over (cluster C1, *KFFE* tetrapeptide). Compare to the regular β-sheet conformation in Fig. 5.

the amorphous conformations are kinetic intermediates that need to rearrange before adopting a β-sheet structure consistent with the cross-β structural model of amyloid fibrils.

Downhill free energy for dimerization of the *KFFE* peptide sequence

Free energy surfaces projected onto appropriate reaction coordinates can provide important information about both the thermodynamics and kinetics of peptide association.

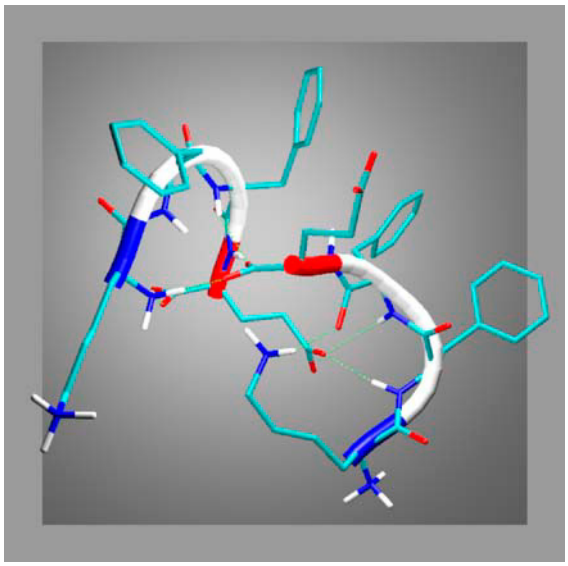

FIGURE 7 Centroid of the dimeric cluster C4 with the lowest potential energy observed in this work for the *KFFE* peptide using replica-exchange simulations.

Thermodynamic preferences for the dimeric phases can be quantified in terms of free energy differences between the monomeric and dimeric states, whereas the rate of dimerization is governed by the height of the barrier separating the assembled and disassembled states (29).

The free energy surface for the *KFFE* tetrapeptide dimerization is plotted at the transition temperature $T_a = 325$ K as a function of the potential energy and the radius of gyration in Fig. 8 *a*. At this temperature, the monomeric conformations ($Rg > 6$ Å) coexist with the dimeric phase ($Rg < 6$ Å). The structural diversity of the dimeric phase, discussed in the previous section (see Fig. 4), is also visible, with dimeric conformations with $Rg < 5$ Å separated by a very small barrier (less than a k_BT) from the other dimers. Oligomerization of this peptide is accompanied by a gradual decrease of the gyration radius as the temperature is lowered. This can be seen in Fig. 8 *b*, which depicts the free energy map at a high temperature ($T = 500$ K) at which the monomers constitute the dominant phase. At this temperature, dimers with $Rg < 5$ Å are unstable by $\sim 7\ k_BT$.

The free energy maps of the phenylalanine peptide are in stark contrast with those of the other peptides studied in this work. As an illustration, Fig. 9 shows the free energy surface for the *KLLE* sequence at $T = 298$ K (slightly above its T_a of 275 K). The free energy map contains two minima corresponding to monomeric and dimeric phases, respectively, separated by a barrier of 4 k_BT. The leucine tetrapeptide, unlike the phenylalanine variant, displays the two-state behavior characteristic of cooperative transitions. The presence of a significant barrier implies that the leucine dimers will form more slowly than their phenylalanine counterparts.

The mechanisms leading to the formation of dimers of *KFFE* and of *KLLE* were probed by considering free energy maps plotted as a function of the C_α radius of gyration and the radius of gyration computed over the phenylalanine (Fig. 10 *a*) and leucine (Fig. 10 *b*) atoms at their respective transition temperatures. The *Rg–Rg* map for the *KFFE* sequence reveals that the dimerization of the phenylalanine-containing peptides proceeds in two phases. In the first step, the mutual separation among the phenylalanine atoms is dramatically reduced from 12 to ~5 Å, whereas the overall size of the system remains large (up to 8 Å). In the second step, the backbone C_α atoms come into contact, cementing the dimeric conformations and reducing the total size of the system from a radius of gyration of ~5 Å. This sequence of events during oligomerization indicates that the dimerization of *KFFE* monomers is initiated by *PHE–PHE* side chain coming into contact, followed by interactions of the peptide backbones (hence the off-diagonal appearance of Fig. 10 *a*. The early formation of the *PHE–PHE* contacts, stabilized by van der Waals interactions, effectively removes the free energy barrier to dimerization, a barrier which would otherwise hinder monomer assembly. In contrast to *KFFE*, the dimerization of *KLLE* displays a *Rg–Rg* free energy map (Fig. 10 *b*) with a diagonal shape. In this instance, the Leu

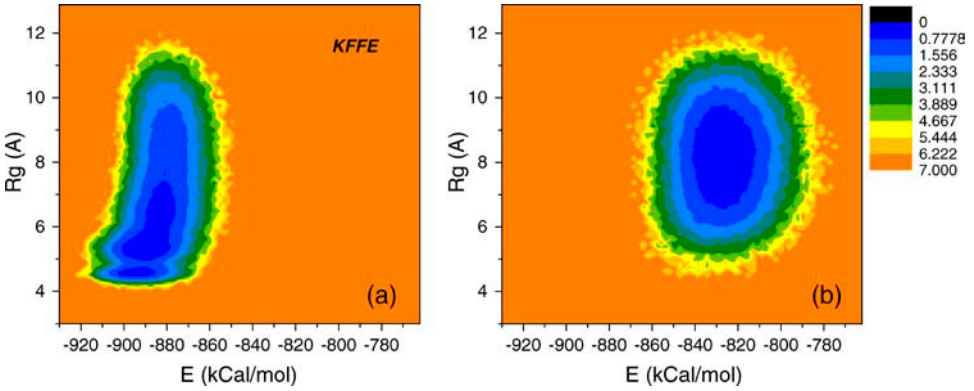

FIGURE 8 Free energy surface for the *KFFE* dimer plotted relative to its lowest value in units of k_BT (*a*) near T_a and (*b*) above T_a.

side-chain atoms and the monomer backbones undergo a simultaneous transition into collapsed dimeric states. The monomeric and dimeric phases are separated by a free energy barrier at the transition state region (Rg ~7–8 Å), as the Leu residues do not interact in a stabilizing manner to form pre-dimer intermediate conformations (as was the case for the *KFFE* sequence). The *KLLE* and *KFFE* peptides clearly follow different oligomerization pathways, with the *PHE–PHE* interactions emerging as the main driving force for *KFFE* assembly.

DISCUSSION AND CONCLUSIONS

With the goal of gaining further insight into the general principles underlying amyloid formation, we have investigated the equilibrium thermodynamics of dimerization of four tetrapeptide sequences, *KFFE*, *KVVE*, *KLLE*, and *KAAE*, two of which, *KFFE* and *KVVE*, were shown to fibrillate, and the other two, *KLLE* and *KAAE*, to remain soluble under the same experimental conditions (12). In our analysis of the peptide dimers, we observe some clear thermodynamic trends which appear to be in good agreement with experiment. The applicability of our results to predictions regarding the full fibrils is, however, contingent on the experimental critical nucleus being small. As a general rule, simulations addressing questions of peptide fibril formation or growth need to consider a number of peptides in the simulation cell that is at least as large as the size of the critical nucleus. Although the size of the critical nucleus for aggregation of the peptides studied in this work is not known, it is conceivable that it can be made quite small, by carefully tuning the experimental conditions—in particular, temperature and peptide concentration. In fact, many aggregating peptides are believed to possess critical nuclei as small as dimers, which implies a completely downhill aggregation scenario. In the initial fibrillization stages of Aβ14-23 (5) and β-2 microglobulin (30), for instance, stable dimeric intermediates are populated. Isolated Aβ40/42 dimers have, however, been shown to generate amorphous aggregates rather than fibrils (31). In assessing the predictive power of our results, one can argue that, qualitatively, an amyloidogenic peptide that can exist as a stable dimeric β-sheet should also form stable larger β-sheets, since the formation of both dimer and fibril is driven by essentially the same energy-entropy terms. In the event that the amorphous dimers are on-, rather than off-pathway intermediates, the observed agreement between simulation and experiment becomes not merely coincidental, and stability trends for the KXXE dimers can be directly extrapolated to rationalize the stability of fibrils. The full extent to which the thermodynamics of small oligomers can be used to quantitatively predict aggregation propensities of fibrils is the subject of our future investigations involving larger-size oligomers.

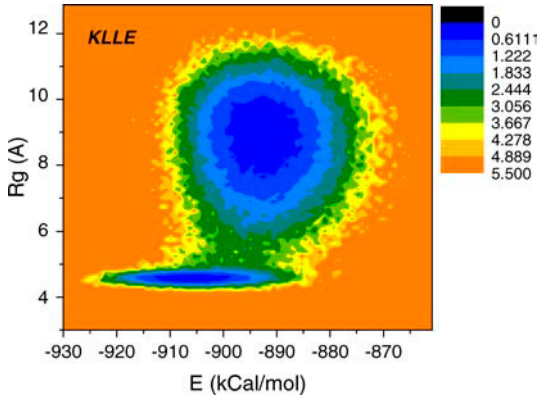

FIGURE 9 Free energy surface for the *KLLE* dimer plotted relative to its lowest value in units of k_BT near its T_a.

The dimerization trends found in our simulations are in good agreement with experimental fibrillation propensities. The *KFFE* and *KVVE* sequences have higher association temperatures T_a than the other two peptides, implying that at a temperature T below T_a of all mutants, the *KFFE* and *KVVE* dimers are more thermodynamically stable than the *KLLE* and *KAAE* dimers. This is consistent with the results of Tjernberg et al. (12), which show that only peptides containing phenylalanine and valine are capable of forming fibrils. We note that the fact that the *KLLE* and *KAAE* peptides were not seen to aggregate under the same experimental conditions as the *KFFE* and *KVVE* peptides does not

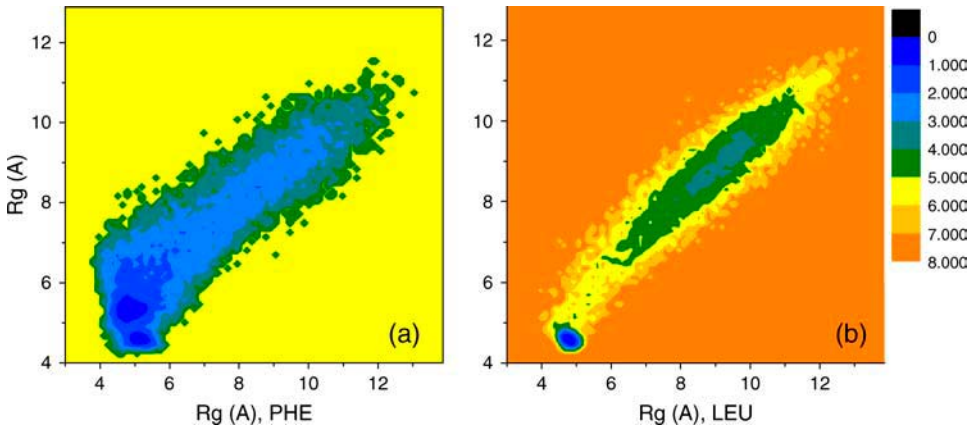

FIGURE 10 Free energy surfaces explaining the mechanism of dimerization for (*a*) *KFFE* and (*b*) *KLLE* peptides. The surfaces are shown as a function of the C_α radius of gyration (*Y* axis) and the radius of gyration calculated over atoms involved in two central residues (phenylalanines for *KFFE* and leucines for *KLLE*).

mean that these peptides are incapable of aggregation. It simply indicates that experiments were performed in a regime where the fibrils are neither thermodynamically stable nor kinetically accessible. Based on the association temperature trend for dimerization seen in our simulations, we predict that performing experiments at lower temperatures (higher concentrations) could lead to fibrillation of these peptides.

Our simulations reveal that the free energy landscape of aggregation is strongly dependent on the sequence of the involved peptide. Although the peptide mutants containing leucine, alanine, or valine need to overcome a significant barrier on the way to an aggregate state, the phenylalanine peptide undergoes association in an noncooperative manner, with a free energy surface devoid of barriers. Its radius of gyration decreases gradually as the temperature is lowered, a signature of a continuous, second-order-like, rather than a discontinuous, first-order-like transition. The implication of this type of the free energy landscape is that dimer conformations are more easily accessible for the *KFFE* peptide than for the other peptide. The distinctive character of the free energy map observed in our simulations of the *KFFE* sequence is consistent with the recent reports (32) ascertaining that phenylalanine residues play an exceptionally important role at inducing fibril formation in peptides. A more detailed analysis on the factors that contribute to this fibrillogenic property of phenylalanines will be presented elsewhere.

Our results indicate that by both thermodynamic and kinetic criteria, the phenylalanine mutant exhibits features favoring aggregation that set it apart from the experimentally nonaggregating sequences. The barrier observed for the *KVVE*, *KLLE*, and *KAAE* peptides may play a dual role in the fibrillation reaction: 1), it may slow down formation of dimers from monomers at the initial stages of the process; and 2), it may hinder late-stage re-equilibration of the system by preventing rapid disintegration or rearrangement of the dimers on their way to form larger aggregates. In light of the strong sequence-dependence of the free energy surfaces for aggregation that emerges from our simulations, we suggest that future research in peptide aggregation should be directed at understanding how the sequence characteristics of aggregating peptides affect the free energy barrier for aggregation. Fibrillation propensities are indeed determined not only by the free energy differences between reactants (monomers) and product (aggregates) but perhaps more importantly by the kinetic accessibility of the product. How quickly (and in fact whether) aggregation can proceed depends primarily on the free energy barrier separating the reactants from the products. Ultimately, it is those conformations belonging to this transition state ensemble that will determine whether fibril formation will occur.

Our simulations, in conjunction with the work of Tjernberg et al. (12) on the same peptides, shed interesting new light into the role of hydrophobic interactions in aggregation propensities. Hydrophobic forces are often considered to play a dominant role in polypeptide aggregation propensities (33). Mutations of hydrophobic residues (34) into polar ones, as well as reducing the temperature such that the hydrophobic forces are significantly weakened (35), have both been shown to strongly inhibit aggregation. More recent articles indicate that hydrophobicity may actively participate in the fine-tuning of the fibrillar structure (36), quantitatively defining amyloidogenic propensities (15). According to hydrophobicity scales (37), the side chain of leucine residues is more hydrophobic than that of valine residues and one would hence expect a stronger interaction energy in oligomers of *KLLE*. Yet, the valine-containing tetrapeptide (*KVVE*) has been shown experimentally to form fibrils whereas the *KLLE* remains soluble (12). We find in our simulations that entropic rather than energetic factors govern the relative stabilities of the *KVVE* and *KLLE* dimers. The valine-rich peptide is seen to populate more β-strand conformations than its leucine-containing counterpart. As a result, oligomerization of the *KVVE* system is accompanied by a smaller loss of conformational entropy, which, in turn, promotes higher thermodynamic stability for the oligomers.

The dimeric phase for all four sequences studied was found to be heterogeneous, with β-sheets conformations competing at low temperatures with structurally disordered conformations. A similar diversity of structures has been reported in simulations of hairpins (38) and other β-sheet-forming

peptides (39). Only ordered β-sheets have the correct hydrogen-bonding properties consistent with the cross-β structure found in amyloid fibrils. Although not in the native β-sheet secondary structure arrangement, stable amorphous dimers can potentially act as on-pathway kinetic intermediates for fibrillization. Although the formation of an ideal β-sheet structure may turn out to be the main nucleation event, it is indeed also conceivable that the energetically favored on-pathways amorphous dimers may form initially, followed by a phase transition into more stable fibrillar structures. A number of recent experiments nonetheless indicate that soluble amorphous oligomers also act as off-pathway species for the formation of fibrils. A mutant of the Aβ14-23 peptide was seen to bring fibril formation to a complete halt through over-stabilization of dimeric conformations (5). A dimeric intermediate of Aβ1-42 obtained from the brain of a patient with Alzheimer's disease, as well as in an in vitro study, generated only amorphous aggregates and not fibrils (31). Similarly, the fiber assembly of a synthetic peptide lacking sufficient number of charged residues was shown to be strongly inhibited by an overwhelming competition from the amorphous aggregates (40). In addition, nitration has been shown to inhibit the fibrillation of α-synuclein by forming stable soluble oligomers (41).

We speculate that the reason for nonfibrillation of another set of tetrapeptides studied by Tjernberg et al. (12), those with neutralized termini, is due to a kinetical partitioning into non-β-sheet structures rather than to insufficient stability of the fibers. This hypothesis can be easily tested experimentally through temperature-dependent studies. High temperatures should destabilize any amorphous dimers and provide additional monomers for fibril growth. We would therefore expect fibrils to become more abundant at higher temperatures if kinetic traps are responsible for fibril inhibition. Fibril formation should decrease at higher temperatures, on the other hand, if lack of stabilizing interactions is the cause for the absence of fibrils under physiological conditions. Elucidating the precise reasons for fibril inhibition will further our understanding of aggregation and may contribute to the development of therapeutic strategies for diseases in which amyloid fibrils are implicated.

We thank W. Guo, M. Friedel, and P. Soto for helpful discussions.

We gratefully acknowledge the support of the National Science Foundation Career Award No. 0133504, the A. P. Sloan Foundation, the David and Lucile Packard Foundation, and the Institute for Collaborative Biotechnologies through grant No. DAAD19-03-D-0004 from the U.S. Army Research Office. Part of the simulations were performed using the National Science Foundation TERAGRID facilities, through allocation grant No. MCB040061N.

REFERENCES

1. Selkoe, D. J. 2003. Folding proteins in fatal ways. *Nature*. 426:900–904.
2. Temussi, P. A., L. Masino, and A. Pastore. 2003. From Alzheimer to Huntington: why is a structural understanding so difficult? *EMBO J.* 22:355–361.
3. Thirumalai, D., D. K. Klimov, and R. I. Dima. 2003. Emerging ideas in the molecular basis of protein and peptide aggregation. *Curr. Opin. Struct. Biol.* 13:1–14.
4. Balbach, D. J., J. J. Gordon, R. Tycko, and S. C. Meredith. 2004. Increasing the amphiphilicity of an amyloidogenic peptide changes the β-sheet structure in the fibrils from antiparallel to parallel. *Biophys. J.* 86:428–434.
5. Tjernberg, L. O., D. J. E. Callaway, A. Tjernberg, S. Hahne, C. Lilliehöök, L. Terenius, J. Thyberg, and C. Nordstedt. 1999. A molecular model of Alzheimer amyloid β-peptide fibril formation. *J. Biol. Chem.* 274:12619–12625.
6. Balbach, J. J., Y. Ishii, O. N. Antzutkin, R. D. Leapman, N. W. Rizzo, F. Dyda, J. Reed, and R. Tycko. 2000. Amyloid fibril formation by Aβ_{16-22}, a seven-residue fragment of the Alzheimer's β-amyloid peptide, and structural characterization by solid state NMR. *Biochemistry.* 39:13748–13759.
7. Fändrich, M., and C. M. Dobson. 2002. The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.* 21:5682–5690.
8. Klimov, D. K., and D. Thirumalai. 2003. Dissecting the assembly of Aβ_{16-22} amyloid peptides into antiparallel β sheets. *Structure.* 11:295–307.
9. Ma, B., and R. Nussinov. 2002. Stabilities and conformations of Alzheimer's β-amyloid peptide oligomers (Aβ_{16-22},Aβ_{10-35}): sequence effects. *Proc. Natl. Acad. Sci. USA.* 99:14126–14131.
10. Massi, F., D. Klimov, D. Thirumalai, and J. E. Straub. 2002. Charge states rather than propensity for β-structure determine enhanced fibrillogenesis in wild-type Alzheimer's β-amyloid peptide compared to e22q Dutch mutant. *Protein Sci.* 11:1639–1647.
11. Gsponer, J., U. Haberthür, and A. Caflicsh. 2003. The role of side-chain interactions in the early steps of aggregation: molecular dynamics simulations of an amyloid-forming peptide from the yeast prion sup35. *Proc. Natl. Acad. Sci. USA.* 100:5154–5159.
12. Tjernberg, L. O., W. Hosia, N. Bark, J. Thyberg, and J. Johansson. 2002. Charge attraction and β propensity are necessary for amyloid fibril formation from tetrapeptides. *J. Biol. Chem.* 277:43243–43246.
13. Rochet, J.-C., and P. T. Lansbury, Jr. 2000. Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol.* 10:60–68.
14. Uversky, V. N., and A. L. Fink. 2004. Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochim. Biophys. Acta.* 1698:131–153.
15. Chiti, F., N. Taddei, F. Baroni, C. Capanni, M. Stefani, G. Ramponi, and C. M. Dobson. 2002. Kinetic partitioning of protein folding and aggregation. *Nat. Struct. Biol.* 9:137–143.
16. Chiti, F., P. Webster, N. Taddei, A. Clark, M. Stefani, G. Ramponi, and C. M. Dobson. 1999. Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci. USA.* 96:3590–3594.
17. Neria, E., S. Fischer, and M. Karplus. 1996. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* 105:1902–1921.
18. Dominy, B. N., and C. L. Brooks III. 1999. Development of a generalized Born model parameterization for proteins and nucleic acids. *J. Phys. Chem. B.* 103:3765–3773.
19. Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314: 141–151.
20. Allen, M. P., and D. J. Tildesley. 1986. Computer Simulation of Liquids. Clarendon Press, Oxford.
21. Jang, S., S. Shin, and Y. Pak. 2002. Molecular dynamics study of peptides in implicit water: ab initio folding of β-hairpin, β-sheet and $\beta\beta\alpha$-motif. *J. Am. Chem. Soc.* 124:4976–4977.
22. Roux, B. 2001. Implicit solvent models. *In* Computational Biochemistry and Biophysics. Marcel Dekker, New York.
23. Lum, K., D. Chandler, and J. D. Weeks. 1999. Hydrophobicity at small and large length scales. *J. Phys. Chem. B.* 103:4570–4577.

24. Rein ten Wolde, P., S. X. Sun, and D. Chandler. 2001. Model of a fluid at small and large lengthscales and the hydrophobic effect. *Phys. Rev. E.* 65:011201-1–011201-9.

25. Klimov, D. K., D. Newfield, and D. Thirumalai. 2002. Simulations of β-hairpin folding confined to spherical pores using distributed computing. *Proc. Natl. Acad. Sci. USA.* 99:8019–8024.

26. Pitera, J. W., and W. Swope. 2003. Understanding folding and design: replica-exchange simulations of ''Trp-cage'' miniproteins. *Proc. Natl. Acad. Sci. USA.* 100:7587–7592.

27. Ferrenberg, A. M., and R. H. Swendsen. 1989. Optimized Monte Carlo data analysis. *Phys. Rev. Lett.* 63:1195–1198.

28. Feig, M., J. Karanicolas, and C. L. Brooks III. 2001. MMTSB Tool Set. MMTSB NIH Research Resource, The Scripps Research Institute, La Jolla, CA.

29. Hänggi, P., P. Talkner, and M. Borkovec. 1990. Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* 62:251–341.

30. Eakin, C. M., F. J. Attenello, C. J. Morgan, and A. D. Miranker. 2004. Oligomeric assembly of native-like precursors precedes amyloid formation by β-2 microglobulin. *Biochemistry.* 43:7808–7818.

31. Roher, A. E., M. O. Chaney, Y.-M. Kuo, S. D. Webster, W. B. Stine, L. J. Haverkamp, A. S. Woods, R. J. Cotter, J. M. Tuohy, G. A. Krafft, B. S. Bonnell, and M. R. Emmerling. 1996. Morphology and toxicity of Aβ-(1–42) dimer derived from neuritic and vascular amyloid deposits of Alzheimer's disease. *J. Biol. Chem.* 271:20631–20635.

32. Azriel, R., and E. Gazit. 2001. Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide. *J. Biol. Chem.* 276:34156–34161.

33. Horwich, A. 2002. Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions. *J. Clin. Invest.* 110:1221–1232.

34. Hilbich, C., B. Kisters-Woike, J. Reed, C. L. Masters, and K. Beyreuther. 1991. Aggregation and secondary structure of synthetic amyloid βa4 peptides of Alzheimer's disease. *J. Mol. Biol.* 218:149–163.

35. Mitraki, A., J.-M. Betton, M. Desmadril, and J. M. Yon. 1987. Quasi-irreversibility in the unfolding-refolding transition of phosphoglycerate kinase induced by guanidine hydrochloride. *Eur. J. Biochem.* 163:29–34.

36. Tycko, R. 2004. Progress towards a molecular-level structural understanding of amyloid fibrils. *Curr. Opin. Struct. Biol.* 14:96–103.

37. Creighton, T. E. 1993. Proteins: Structures and Molecular Properties. W.H. Freeman and Co., New York.

38. Higo, J., O. V. Galzitskaya, S. Ono, and H. Nakamura. 2001. Energy landscape of a β-hairpin peptide in explicit water studied by multicanonical molecular dynamics. *Chem. Phys. Lett.* 337:169–175.

39. Hwang, W., S. Zhang, R. D. Kamm, M. Karplus. 2004. Kinetic control of dimer structure formation in amyloid fibrillogenesis. *Proc. Natl. Acad. Sci. USA.* 101:12916–12921.

40. López de la Paz, M., K. Goldie, J. Zurdo, E. Lacroix, C. M. Dobson, A. Hoenger, and L. Serrano. 2002. De novo designed peptide-based amyloid fibrils. *Proc. Natl. Acad. Sci. USA.* 99:16052–16057.

41. Yamin, G., V.-N. Uversky, and A. L. Fink. 2003. Nitration inhibits fibrillation of human alpha-synuclein in vitro by formation of soluble oligomers. *FEBS Lett.* 542:147–152.