Published in final edited form as:
Proteins. 2011 July ; 79(7): 2132–2145. doi:10.1002/prot.23034.

COMPUTER SIMULATION STUDY OF AMYLOID FIBRIL FORMATION BY PALINDROMIC SEQUENCES IN PRION PEPTIDES

Victoria Wagoner[1], **Mookyung Cheon**[2], **Iksoo Chang**[2], and **Carol Hall**[1]

[1]Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina

[2]Center for Proteome Biophysics, Department of Physics, Pusan National University, Busan, Korea

Abstract

We simulate the aggregation of large systems containing palindromic peptides from the Syrian hamster prion protein SHaPrP 113–120 (AGAAAAGA) and the mouse prion protein MoPrP 111–120 (VAGAAAAGAV) and eight sequence variations: GAAAAAAG, $(AG)_4$, A8, GAAAGAAA, A10, V10, GAVAAAAVAG, and VAVAAAAVAV The first two peptides are thought to act as the Velcro that holds the parent prion proteins together in amyloid structures and can form fibrils themselves. Kinetic events along the fibrillization pathway influence the types of structures that occur and variations in the sequence affect aggregation kinetics and fibrillar structure. Discontinuous molecular dynamics simulations using the PRIME20 force field are performed on systems containing 48 peptides starting from a random coil configuration. Depending on the sequence, fibrillar structures form spontaneously over a range of temperatures, below which amorphous aggregates form and above which no aggregation occurs. AGAAAAGA forms well organized fibrillar structures whereas VAGAAAAGAV forms less well organized structures that are partially fibrillar and partially amorphous. The degree of order in the fibrillar structure stems in part from the types of kinetic events leading up to its formation, with AGAAAAGA forming less amorphous structures early in the simulation than VAGAAAAGAV. The ability to form fibrils increases as the chain length and the length of the stretch of hydrophobic residues increase. However as the hydrophobicity of the sequence increases, the ability to form well-ordered structures decreases. Thus, longer hydrophobic sequences form slightly disordered aggregates that are partially fibrillar and partially amorphous. Subtle changes in sequence result in slightly different fibril structures.

INTRODUCTION

The transmissible spongiform encephalopathies, better known as prion diseases, are fatal neurodegenerative disorders characterized by the misfolding of a normal “prion” protein, PrP^C (cellular protein) into a beta-sheet rich conformation, PrP^{Sc} (scrapie protein) that subsequently aggregates to form amyloid structures. Examples include human Creutzfeldt-Jakob disease, scrapie disease in sheep and goats, kuru which is associated with cannibalism, and Mad Cow disease [1–5]. A common feature of the prion protein sequence across the various species is the presence of regions of short oligopeptide repeat sequences that are sometimes palindromic in nature and commonly between four and ten residues in

Address correspondence to: Carol Hall, Box 7905, COE1, 911 Partners Way, Raleigh, North Carolina 27695-7905. Fax: 919-515-3465; hall@ncsu.edu.

length. These regions have high hydrophobic residue content and are believed to drive aggregation into fibrils [6–14], although the molecular details of this process are not well understood. The short palindromic prion fragments can also form amyloid fibrils in the absence of the flanking residues in the full-length prion protein, suggesting that study of these peptides may help elucidate the mechanism of fibril formation [6–7, 13, 15–16]. In this work, we perform computer simulations of the self-assembly of two simple palindromic aliphatic sequences, one from the Syrian hamster prion protein, residues 113–120 (AGAAAAGA), and one from the mouse prion protein, residues 111–120 (VAGAAAAGAV), and variations of these two sequences. The goal of this work is to provide insight into the molecular motions that underlie the early steps in the assembly of these short palindromic sequences into amyloid fibrils. We also seek to learn how amino acid sequence affects the ability to form ordered structures including fibrils, the types of structures formed along the fibrillization pathway including intermediates like β-sheets and amorphous aggregates, and the kinetics of these processes.

Since it was first postulated in 1991 that the scrapie isoform of the prion protein is a key component in neuronal degeneration [17], interest in how this protein changes conformation and subsequently aggregates has surged. Early experimental studies focused on the polypeptide PrP 27–30, defined to be residues 90–145 of PrP^{Sc}, the protein implicated in prion disease pathogenesis [17–19]. PrP 27–30 can be isolated from scrapie-infected Syrian hamster brains and made synthetically. Conversion of PrP^{C} into PrP^{Sc} occurs only when residues 95–140 are present [20]. A subset of that sequence, PrP106–126 is known to be necessary for β-sheet formation [16, 20] and, since recent hydrogen-deuterium exchange studies on the full PrP fibril structure have confirmed that its hydrophobic core is a β-sheet rich region, that sequence (PrP 106–126) likely plays a role in fibril formation [16, 21–22]. In order to identify which parts of PrP 27–30 are responsible for the α → β transition, and subsequent organization into three-dimensional structures, Nguyen et al. isolated PrP 109–122, PrP 113–120 and PrP 90–145. Electron microscopy and x-ray diffraction measurements revealed that all three sequences formed β-sheets under certain conditions. In particular, the hydrophobic palindromic sequence, PrP 113–120, AGAAAAGA, formed fibrils composed of β-sheets staggered in the H-bonding direction but with a smaller inter-sheet spacing than had previously been reported for other amyloid fibrils. AGAAAAGA, which is common to all mammalian prion peptides (human, sheep, mouse, etc.), has been the subject of numerous investigations focused on understanding what part it plays in PrP fibril formation and toxicity [6, 9, 22–23]. The thought here is that this sequence may be one of the hydrophobic cores that drive PrP aggregation into fibrils [24].

Although the hydrophobic core seems to be necessary for the formation of fibrils in PrP and other disease-related amyloidogenic proteins, it is not clear what role it plays in toxicity. The lack of consensus is illustrated by the following two studies. Jobling et al showed that the presence of MoPrP 106–126 caused a 50% decrease in the number of wild-type neuronal cells, but that substitution of hydrophilic for hydrophobic residues disrupted the peptide's ability to form β-sheets and significantly reduced the loss of neuronal cells [6]. These results suggest that fibril formation is connected with toxicity. Florio et al. found that although PrP 106–126 was completely insoluble in water and had a high Thioflavin T (ThT) binding indicating fibril formation, substitution of alanines for Gly114 and Gly119 increased the PrP 106–126 water solubility, decreased its ThT binding, did not change in its β-sheet content and increased its neurotoxicity. This suggests that the final fibril may not be the neurotoxic agent [22]. Nevertheless it is apparent that some structure along the PrP 106–126 fibril formation pathway is causing cell death and that understanding this pathway could shed light on this perhaps transient structure.

In recent years, molecular-level computer simulation has been used to supplement the information gained from experimental studies. The focus has been on understanding how prion peptides assemble into amyloid fibrils and predicting the fibril's molecular structure. Most of these simulations have been performed on truncated sequences like Syrian hamster prion SHaPrP 113–120[7, 25], yeast prion Sup-35(GNNQQNY) [26–27], and mouse prion MoPrP 106–126[16] because these short peptides are themselves amyloidogenic. The use of atomistic molecular simulations to study the arrangement of β-strands in fibrils was pioneered by Ma and Nussinov. In order to learn more about the nucleation events believed to initiate amyloid formation and growth they performed high temperature (up to 350K) molecular dynamics simulations on two short sequences, A8 (AAAAAAAA) and SHaPrP 113–120 (AGAAAAGA). Examination of various ordered starting configurations containing 3, 4, 6 and 8 peptides showed that as the number of peptides increased, the stability of the conformation increased. They explored how the arrangement of the peptides relative to each other within and between sheets affected conformational stability. Their analysis was based on the assumption that an oligomer that does not disassociate at high temperature must be the nucleus that "seeds" fibril formation. The most stable arrangement of eight peptides for both sequences was an octamer containing two stacked β-sheet layers with antiparallel β-strands within the β-sheet and parallel β-strands between the octamer layers. The A8 octamers were much more stable than the SHaPrP 113–120 octamers.[25] Roder and coworkers combined experiment and simulation to study MoPrP 106–126 which contains the palindromic sequence VAGAAAAGAV[16]. Their H/D exchange experiments on MoPrP 106–126 fibrils showed, as expected, that the hydrophobic core was protected and therefore likely played a role in the structure of the prion fibrils. Room temperature molecular dynamic simulations on six different arrangements of eight MoPrP 106–126 peptides showed that the most stable arrangement was two parallel β-sheet layers composed of parallel β-strands, which is different from the antiparallel β-strand within a β-sheet result of Ma and Nussinov for the hydrophobic sequence SHaPrP 113–120. Many researchers believe that the anti-parallel arrangement of peptides within β-sheets is favored for shorter sequences but the role of sequence in peptide orientation within β-sheets is still unclear [25, 28–29]. In an attempt to elucidate the arrangement of β-strands in PrP amyloid fibrils, Lee and coworkers used experimental data to guide their molecular dynamics simulations of SHaPrP 109–122. Their initial configuration was an octamer composed of two stacked parallel β-sheet layers with each layer composed of four antiparallel β-strands. This configuration remained stable over 2ns[7] and is the same as the stable conformation observed by Ma and Nussinov for SHaPrP 113–120. All of the atomistic simulations indicate that the stable fibrillar structure is composed of two stacked β-sheets.

In this work, we attempt to characterize the prion fibrillization pathway by simulating the spontaneous assembly of prion fibrils formed by the truncated prion sequences SHaPrP 113–120 (AGAAAAGA), MoPrP 111–120 (VAGAAAAGAV) and eight variations of these sequences including: GA_6G (a longer uninterrupted alanine stretch flanked by glycine), $(AG)_4$ (a complete disruption of hydrophobic residues), GAAAGAAA (a mimic of Aβ29-36), A8, VAVAAAAVAV (less flexible than MoPrP 111–120), GAVAAAAVAG (uninterrupted hydrophobic sequence), A10 and V10. Our motivation for studying these variations is to see how the placement of residues of different size and energetics affects the final fibril structure and the kinetics of fibril formation. We combine PRIME20, a new intermediate-resolution model of protein geometry and energetics applicable to all twenty amino acids[30], with discontinuous molecular dynamics (DMD), a fast alternative to traditional molecular dynamics applicable to systems interacting via discontinuous potentials, such as square –well potentials[31–39]. Simulations are performed on 48-peptide homogenous systems containing the eight previously listed peptides starting from random configurations of random coils at high temperatures. The systems are cooled slowly at fixed concentration, c = 5mM, from an initial high temperature to the desired simulation

temperature and then simulated for 50–100 billion events. We explore how changing sequence and temperature affects the pathway of aggregation by monitoring the formation of different structures such as α-helices, β-sheets (dimers, trimers, tetramers, pentamers, hexamers and larger), amorphous aggregates, and fibrils as a function of time.

Highlights of our results are the following. We observe spontaneous fibril formation for most of the sequences studied. In particular AGAAAAGA forms well organized bi-layer fibrillar structures while VAGAAAAGAV forms less well organized structures that are partially fibrillar and partially amorphous. The propensity to form fibrils depends strongly on the temperature. The number of peptides in fibrils is low at low temperatures, increases as temperature increases, reaches a maximum at a "transition" temperature and thereafter abruptly decreases. The peptides exist primarily as amorphous aggregates at the low temperature end of this spectrum and as free monomers above the transition temperature. General trends observed were that as the chain length and the length of the stretch of hydrophobic residues increase, the ability to form fibrils increases. However as the hydrophobicity of the residues in the sequence increases, the ability to form well-ordered structures decreases. Thus, long hydrophobic sequences like VAVAAAAVAV and V10, form slightly disordered aggregates that are partially fibrillar and partially amorphous. Subtle changes in sequence result in different fibril structures. For example, although the prion sequences, AGAAAAGA and VAGAAAAGAV, predominantly form fibrils with two stacked β-sheets, sequences like A8, A10, GA_6G, and GAVAAAAVAG tend to form fibrils with three and even four stacked β-sheets. The presence of more than two valines per sequence led to structures with both fibrillar and amorphous content, at least for the conditions explored in this paper. The strands show no preference for parallel or antiparallel arrangements (either within the sheets or between the sheets) until the chain length exceeds eight residues, at which point a parallel arrangement for these simple aliphatic sequences seems to be preferred. The- intra-strand distance, 4.8Å—5.0Å, is the same for all sequences studied, in qualitative agreement with Nguyen[19] and Shinchuk[11] Fibril growth is also affected by sequence. When valine is present fibrils grow by first forming amorphous aggregates and then β-sheets, but when valine is absent fibrils grow by forming β-sheets that then associate into fibrils. For the sequences containing valine (the largest side chain considered) amorphous aggregates formed early in the simulations are very slow to diminish in size, even after β-sheet formation and fibril association begins. The picture that emerges here is that there is a competition between the formation of amorphous aggregates and β-sheets along the aggregation pathway and that the presence of certain amino acids (valine in this case) diminishes the formation of β-sheets and ultimately the ability to form ordered aggregates.

This paper is organized as follows. In the next section, we describe the peptide model and the simulation method. In the following section, we present the results obtained from simulations of multi-peptide systems at various conditions. The last section is a discussion of our results.

METHODS

Model Peptide and Forces

The truncated prion peptide sequences considered here are: SHaPrP 113–120 (AGAAAAGA) and MoPrP 111–120 (VAGAAAAGAV) and eight variations, GAAAAAAG, AGAGAGAG, GAAAGAAA, A8, VAVAAAAVAV, GAVAAAAVAG, A10 and V10. Our rationale for choosing the eight variations was the following. Interest in short peptides not interrupted by glycine like A8 and A10 stemmed from previous theoretical[15] and experimental studies[11, 40–41] indicating that these sequences easily formed amyloid fibrils. The sequence GAAAGAAA is a mimic of a segment of the Aβ peptide,

Aβ29–36, GAIIGLMV, but is slightly more hydrophobic. Proteins containing poly-valine stretches are known to aggregate and are present in lung surfactant C protein which can, in certain diseases, accumulate in the lung and inhibit normal respiration[42–44]. The remaining four peptides were either inversions of the placement of the glycine residues relative to the parent sequences like GAAAAAAG and GAVAAAAVAG or the addition (AG_4) or subtraction (VAVAAAAVAV) of glycine from the sequence to help ascertain how the flexibility imparted by glycine impacts the aggregation process.

In this work we apply a new implicit-solvent force field PRIME20 to the description of the geometry and energetics of the aliphatic heteropeptide sequences in the prion protein. PRIME20 was recently introduced by Cheon et al.[30] as an extension of PRIME, an implicit solvent intermediate-resolution protein model previously used in simulations of the aggregation of polyalanine and polyglutamine. PRIME was originally developed by Smith and Hall[31–32] and later improved by Nguyen et al.[36]. More recently the PRIME model was extended to the study of polyglutamine peptides[35] illustrating its versatility. In PRIME, the protein backbone is represented by three united atom spheres, one for the amide group (NH), one for the carbonyl group (CO), and one for the alpha-carbon and its hydrogen ($C_\alpha H$). In the original versions of PRIME, each side chain was represented by a single sphere for polyalanine and by a chain of four spheres for polyglutamine. In PRIME20 the twenty possible side chains are modeled as single spheres of unique size, atomic mass and C_α—R bond length. In the truncated prion peptides considered here, the only side chains that appear are alanine and valine; there is no side chain for glycine. All backbone bond lengths and angles are set to realistic values. In order to maintain the *trans*-configuration we fix the consecutive C_α—C_α distance. The side chains are positioned relative to the protein backbone so that all residues are L-isomers. This is critical to producing realistic features for all 20 amino acids and was achieved by carefully adjusting the lengths of the C_α—R bond, and the pseudo-bond. The solvent molecules in our system are modeled implicitly.

All forces between the united atom spheres are modeled with discontinuous potentials, e.g. hard-sphere and square-well interactions. The excluded volume of each of the peptide's four united atom spheres is modeled using a hard sphere interaction. The covalent bond lengths are maintained using a hard sphere interaction that prevents them from moving outside of the range $(1+\delta)l$ to $(1-\delta)l$, where l is the ideal bond length and δ is the tolerance, which is set at 2.375%[45]. Ideal backbone bond angles, C_α—C_α distance, and residue L-isomerization are maintained by imposing a series of pseudobonds whose lengths are also allowed to fluctuate by 2.375%.

In addition to preserving proper protein geometry, we represent one of the most important energetic contributions to protein folding and aggregation− hydrogen bonding. For the simple palindromic aliphatic sequences studied in this paper, hydrogen bonding is represented as a square well attraction of depth ϵ_{HB} and width 4.5Å between the backbone amide and carbonyl groups. Hydrogen bonds are anisotropic in nature so we must constrain their formation to only occur when the NH united atom vector and the CO united atom vector point towards each other and the angle between those vectors is restricted between 120° and 180°[31]. In order to accomplish this we have set the following criteria. A hydrogen bond only forms when: 1) the virtual hydrogen and oxygen atoms (whose location can be calculated at any time) are separated by a distance of 4.5Å, 2) the nitrogen-hydrogen and carbon-oxygen vectors point toward each other within a fairly generous tolerance, 3) neither the NH nor the CO are already involved in a hydrogen bond with a different partner, and 4) the NH and CO are separated by at least three intervening residues along the chain. Further details on the hydrogen bonding model can be found in our earlier work.[34, 36, 38]

The interactions in PRIME20 are all modeled as square well interactions between the spherical units on each amino acid with strength (well depth) and range determined individually for each pair. Since solvent is modeled implicitly these are all effective interactions or potentials of mean force. In PRIME20, the energy parameters that describe the side chain/side chain interactions, the hydrogen bonding interactions between backbone NH and CO and between side chain and side chain are derived in the following way. Briefly, the twenty possible amino acids are classified into 14 groups: [LVI] [F] [Y] [W[[M[[A] [C] [ED] [KR] [P] [ST] [NQ] [H] [G], according to their side chain size, charge, potential for disulfide bond formation, hydrophobicity, and possibility of side chain hydrogen bonding. The matrix of energy parameters between the side chains in the different groups was found by Cheon et al.[30] using a perceptron- learning algorithm with a stochastic component that optimizes the energy gap between 711 known native states from the PDB and decoy structures generated by gapless threading. The number of independent pair-interaction parameters is chosen to be small enough to be physically meaningful yet large enough to give reasonably accurate results in discriminating decoys from native structures. A total of nineteen interaction parameters were used to describe the side chain energetics. In this work we focus on the amino acids alanine, valine and glycine. Alanine and valine are modeled using a three-sphere backbone (for NH, CO and $C_\alpha H$) plus a single sphere side chain. Glycine is modeled solely by a three-sphere backbone; the side chain here is assumed to be embedded in the $C_\alpha H$ united atom sphere. A summary of the side-chain geometric parameters for alanine and valine is provided in Table 1. The energetic parameters for the alanine and valine side chains are $\varepsilon_{AA}/\varepsilon_{HB}=0.086$, $\varepsilon_{VV}/\varepsilon_{HB}=0.198$, $\varepsilon_{AV}/\varepsilon_{HB}=0.141$.

The system temperature is scaled by the hydrogen bonding energy between the backbone NH and CO, ε_{HB}, so that the reduced temperature is $T^* = k_B T/\varepsilon_{HB}$.

Discontinuous Molecular Dynamics

Discontinuous molecular dynamics (DMD) is a variant on standard molecular dynamics that is applicable to systems of molecules interacting via discontinuous potentials (e.g., hard sphere and square-well potentials). Unlike soft potentials such as the Lennard-Jones potential, discontinuous potentials exert forces only when particles collide, enabling the exact (as opposed to numerical) solution of the collision dynamics. This turns the simulation into an efficient event-scheduling algorithm which when applied to an intermediate-resolution model, as opposed to an atomic resolution model, allows sampling of longer time scales and larger systems than traditional atomistic molecular dynamics. The particle trajectories are followed by analytically calculating the time between collisions and then advancing the simulation to the next collision (event) [46–47]. DMD on chain-like molecules is generally implemented using the "bead string" algorithm introduced by Rapaport[48–49] and later modified by Bellemans et al.[45]. Chains of square-well spheres can be accommodated in this algorithm by introducing well-capture, well-bounce, and well-dissociation "collisions" when a sphere enters, attempts to leave, or leaves the square well of another sphere. In this paper, DMD simulations are performed in the canonical ensemble (NVT) using the Andersen thermostat[50] with the initial velocities chosen randomly from a Maxwell-Boltzmann distribution about the desired system temperature. The initial positions of the particles or spheres are chosen randomly while still ensuring that no geometrical constraints are violated.

In DMD simulations of protein aggregation, we determine the number of particles in the system by specifying the concentration which is given by $c = N/L^3$, where N is the number of molecules in the box and L is the simulation box length. Periodic boundary conditions are employed. Since we are simulating large systems at high concentrations in random initial configurations, the box length, L, must be large enough to prevent the macromolecules from interacting with themselves but still allow them to interact with their periodic image. We set

L = 252Å in this study. The simulation proceeds according to the following schedule: identify the first event (e.g., a collision), move forward in time until that event occurs, calculate new velocities for the pair of spheres involved in the event and calculate any changes in system energy resulting from hydrogen bond events or hydrophobic interactions, find the second event, and so on. Types of events include excluded volume events, bond events, and square-well hydrogen bond and hydrophobic interaction events. An excluded volume event occurs when the surfaces of two hard spheres collide and repel each other. A bond (or pseudobond) event occurs when two adjacent spheres attempt to move outside of their assigned bond length and the two particles feel an infinite repulsion that forces them back into their assigned bond length. Square-well events include well-capture, well-bounce, and well-dissociation "collisions" when a sphere enters, attempts to leave, or leaves the square well of another sphere. For more details on DMD simulations with square-well potentials, see articles by Alder and Wainwright[46] and Smith et al.[47].

A total of ten sequences are studied in this work; each system contains 48 peptides at concentration c = 5mM. The simulations are started at high temperature to ensure a random initial configuration and then slow-cooled to the temperature of interest to minimize kinetic trapping. Slow-cooling is achieved by decreasing the temperature in discrete steps starting from a high temperature until the desired simulation temperature is reached. Once the temperature is interest is reached, it is maintained using the Andersen thermostat[50]; in this method all the particles undergo random infrequent "events" or "collisions" with a ghost particle that reassigns the particle's velocity randomly from a Maxwell-Boltzmann distribution centered at the simulation temperature. Five simulations starting from different configurations are run for each sequence at the given temperature and concentration (state). Error bars are taken to be the standard deviation at each state. All simulations are run for an average of 50–100 billion collisions, depending on simulation conditions, sequence, temperature and concentration. We can perform 500 million collisions per hour on a 2.4 GHz AMD Opteron™ workstation, which means that our simulations require 100 to 200 hours.

In this study, we monitor and analyze the formation of β-strands, β-sheets, amorphous aggregates and fibrils. We also check to see if the β-strands in a β-sheet are arranged in a parallel or anti-parallel configuration. The criteria for assigning the types of structures formed are the following. If each peptide in a group of peptides has at least two inter-peptide hydrogen bonds or hydrophobic interactions with a neighboring peptide in the same group, then that group is classified as an aggregate. Aggregates can be either ordered or amorphous. If an aggregate contains β-sheets or fibrils, we classify it as an ordered aggregate. If each peptide in a group of peptides has at least n/2, where n = chain length, inter-peptide β-hydrogen bonds to a particular neighboring peptide in the group, we classify this group as a β-sheet. (A β-hydrogen bond is a hydrogen bond between two residues whose backbone angles are in the β-region of the Ramachandran *plot*.). The n/2 definition emerges from the fact that if two strands within a β-sheet are perfectly aligned then the number of hydrogen bonds between them equals the chain length of the peptide. If at least two β-sheet structures form inter-sheet hydrophobic interactions (at least four hydrophobic interactions per peptide per β-sheet), we classify this as a fibril; otherwise, we classify this and isolated β-sheets as non-fibrillar β-sheet structures. If an aggregate does not contain β-sheets but the peptides in the aggregate have any hydrophobic contacts, then the aggregate is considered amorphous. If an aggregate contains peptides with less than n/2 inter-peptide β-hydrogen bonds between neighboring chains, then this is also considered to be an amorphous aggregate.

RESULTS

Here we describe the results from DMD simulations on 48-peptide systems containing SHaPrP 113–120 (AGAAAAGA), MoPrP 111–120 (VAGAAAAGAV) and eight variations on these sequences: $(AG)_4$, GA_6G, GAAAGAAA, A8, GAVAAAAVAG, VAVAAAAVAV, A10, V10.

Ordered vs. Disordered Structures

Table 2 is a summary of the average types of structures formed at the end of the simulation runs for each sequence over the range of temperatures where fibril formation is most likely to occur. Both SHaPrP 113–120 (AGAAAAGA) and MoPrP 111–120 (VAGAAAAGAV) form fibrillar structures; this is the first time that the spontaneous formation of fibrils has been observed in simulations of large systems containing these sequences. Averages at each temperature are taken over five independent runs started from different initial configurations. The table is organized so that the Syrian hamster prion AGAAAAGA and its variations are listed on the top half and the mouse prion VAGAAAAGAV and its variations are listed in the bottom half. Within each half, the sequences are listed from least likely to most likely to form ordered structures, with the parent sequence listed at the top. The optimum temperature range for fibril formation for the AGAAAAGA and variations is $0.14 \leq T^* \leq 0.16$ and for VAGAAAAGAV is $0.16 \leq T^* \leq 0.18$.

The percentage of AGAAAAGA peptides in fibrils and in amorphous aggregates is plotted versus temperature in Figure 1. The number of peptides in fibrils is low at low temperatures, increases as temperature increases, reaches a maximum at a temperature above which fibrils no longer form and thereafter abruptly decreases. The peptides exist primarily as amorphous aggregates or β-sheets at the low temperature end of this spectrum and as free monomers above the maximum fibril-forming temperature. The existence of amorphous aggregates at low temperatures is a kinetic effect resulting from competing tendencies to form folded α-helix and extended β-strand states. The dependence of the types of structures observed (amorphous aggregates, β-sheets, fibrils and monomers is consistent with the recent phase diagram of Auer and Kashchiev based on simulations of a coarse-grained model[51]. The maximum fibril-forming temperature for AGAAAAGA is $T^*=0.155$; at $T^* = 0.17$, the peptides are primarily random coils. This trend with temperature is characteristic of the other sequences related to AGAAAAGA that form fibrils. The percentage of VAGAAAAGAV peptides in fibrils and in amorphous aggregates is plotted versus temperatures in Figure 2. As we saw with AGAAAAGA in Figure 1, the number of VAGAAAAGAV peptides in fibrils is low at low temperatures, increases as temperature increases, reaches a maximum at a certain temperature above which fibrils no longer form and thereafter abruptly decreases. The VAGAAAAGAV peptides also prefer to be in amorphous aggregates or β-sheets at low temperatures and in free monomers at high temperatures. The maximum fibril-forming temperature for VAGAAAAGAV is $T^*=0.18$; at $T^*=0.19$ the peptides are primarily random coils although some amorphous aggregates remain. The main difference between the VAGAAAAGAV and AGAAAAGA is that VAGAAAAGAV is more prone to form amorphous aggregates and forms substantial numbers of them at temperatures above which fibrils are no longer formed. The ability to form amorphous aggregates is characteristic of all the peptides studied containing valine residues.

The percentage of peptides in fibrils shown in Table 2 for each sequence correlates with both the length of the hydrophobic stretch and the overall chain length of the peptide. We begin our discussion of this by first focusing on the sequences in the top half of the table. At least half of the peptides in the parent sequence AGAAAAGA were in fibrils at the end of the runs with the remaining peptides in β-sheets, indicating ordered structures. As the

hydrophobic stretch increases from four (AGAAAAGA) to six (GAAAAAAG) residues, the percentage of peptides in fibrils increases slightly. As the number of hydrophobic residues in a row increases from six to eight (A8) the sequence continues to form fibrils but there were still many free monomer chains available, indicating that the system may not have yet reached a stable equilibrium. Sequences with a run of less than four hydrophobic residues either formed a small amount of fibrils (GAAAGAAA) or were unable to form fibrils $(AG)_4$. We are not certain why the sequence GAAAGAAA is a weak fibril former (although it readily forms β-sheets) because it is a mimic of Aβ 29–36 (GAIIGLMV) which is believed to aggregate into amyloid fibrils. Two possible explanations are: (1) the simulation has not reached a stable equilibrium and maybe the β-sheets would aggregate into amyloid fibrils if given enough time, and (2) the reduced hydrophobicity of GAAAGAAA compared to GAIIGLMV could delay fibrillization or the overall ability to form fibrils. Two possible reasons that $(AG)_4$ is not a strong fibril former are: (1) there are not enough side chains available to energetically drive the stacking of the β-sheets (fibrillization), and (2) the increased flexibility provided by the glycine residues makes it less likely that the peptides will zip up into multi-mer structures. Finally, it is interesting to note that just the addition of a single alanine residue to the center of the sequence (compare AGAAAAGA to GAAAGAAA) leads to a higher percentage of peptides in fibrils and a faster fibril formation time

We next consider the trends displayed in the second half of Table 2 for the mouse prion peptide and its variations. The mouse prion peptide VAGAAAAGAV is very similar to the Syrian hamster prion peptide AGAAAAGA but has two additional flanking valine residues. Although VAGAAAAGAV formed fibrils in our simulations, it also formed amorphous aggregates, suggesting that the flanking valine residues destabilize the highly ordered hydrophobic core but are strong enough to still cause peptide association. Moving down the table to GAVAAAAVAG, (where the glycine and valine residues in VAGAAAAGAV have been exchanged) we observed an increase in the percentage of peptides in fibrils and a decrease in the percentage of peptides in amorphous aggregates. In fact, GAVAAAAVAG formed the most-well-ordered structures of all the peptides simulated. As we continue down the table we come to A10; this sequence readily formed fibrils but there were still many free monomer chains available, indicating that, just as we observed for A8, A10 may not have yet reached a stable equilibrium. Moving down the table further to VAVAAAAVAV (where the glycine residues in GAVAAAAVAG have been replaced by valines), we see an overall decrease in order; in fact, the peptides were as likely to be in amorphous aggregates as in β-sheets or fibrils. In this case the freedom of the peptide backbone to adopt certain types of conformations has been reduced due to the increased bulkiness and hydrophobicity of the side chains. The replacement of all alanines in A10 to valine, VVVVVVVVVV, also produced a mixture of amorphous aggregates and fibrils as we observed for VAVAAAAVAV.

Subtle changes in sequence can result in significantly different final structures. This is illustrated in Figure 3 which shows snapshots of the final structures at the end of simulation runs for two of the variations on MoPrP 111–120 (a) GAVAAAAVAG at $T^*=0.17$ and (b) VAVAAAAVAV at $T^*=0.17$. For ease of viewing we have portrayed the backbone as ribbons and the side chains in ball and stick format. The fibril for GAVAAAAVAG is a stack of two beta sheets. It is interesting to see that the simple substitution of valine for glycine on the end caps of the sequence to form VAVAAAAVAV causes increased formation of amorphous aggregates with some β-sheet content. It is possible that the structure formed by VAVAAAAVAV is only transient on the pathway to fibril formation or perhaps we are witnessing the small spherical oligomers believed to be off pathway of the fibrillization process. What is clear here is that simple substitutions in primary sequence can result in changes in structure not readily observable in experiment.

Fibril Structure

Table 3 summarizes the characteristics of the fibrils formed by those sequences that form well-ordered structures (fibrils) in our simulations. The table lists the number of fibrils produced during the simulation, the number of β-sheets per fibril, the number of β-strands per β-sheet, the ratio of parallel to anti-parallel strands within each β-sheet, and the average intra-strand distance between peptides within a β-sheet. This type of information can be difficult to ascertain from experiment alone. Given the small system size considered here, 48 peptides, it is not surprising that only one fibril is formed. Recent simulations in our group have indicated that tripling the system size to 192 chains allows several fibrils to form. For the purposes of the study, we are interested in the arrangement of peptides within a single fibril. As we increase the length of the hydrophobic stretch of residues we see the number of beta sheets in a fibril increase from two to three. This can be explained by appealing to the Gibbs-Curie-Wulff theorem: the ratio of the number of β-sheets to the number of β-strands within a sheet is proportional to the ratio of the inter-sheet energy to the intra-strand energy[52,53]. In fact only two sequences prefer stacks of two β-sheets, the AGAAAAGA (SHaPrP 113–120)[6, 15, 19] and VAGAAAAGAV (MoPrP 111–120)[16] which is consistent with recent experimental and simulation studies of these two peptides. We also calculated the average number of strands per sheet for each sequence. In general, if fewer sheets/fibril are present, then the higher the number of strands per sheet.

There has been considerable discussion about how the peptide strands arrange within a β-sheet. Do short strands show a preference for anti-parallel versus parallel? When peptides form parallel or anti-parallel alignment at early stage, it is kinetically driven; rearrangement might occur due to thermodynamic energy difference. We measured the number of strands arranged in parallel or anti-parallel and then calculated the (parallel: anti-parallel) ratio. For the shorter sequences related to SHaPrP 113–120, (AGAAAAGA) we see an equal preference for parallel and anti-parallel arrangements; this is likely due to the palindromic nature of the sequences which means parallel and antiparallel arrangements have the same hydrogen-bonding and hydrophobic interaction energies, i.e. there is no physical reason why these short, un-charged peptides in a “neutral” environment should prefer either orientation. For the longer sequences we see a slight shift towards preference for parallel arrangement of strands within the sheet, particularly for sequences with longer stretches of uninterrupted hydrophobic residues. We do not know why this is so. The only feature of our peptide model that imparts directionality is backbone geometry in PRIME20 and shape-complementary of side-chains, which might need further discussion in the future.

The distance between the β-strands within each sheet, the intra-strand distance, is of interest because it can be measured using X-ray diffraction. For the short peptides containing only alanine and glycine we observe distances of approximately 4.8Å which is slightly higher than the expected 4.7 Å observed for the cross-β diffraction pattern of the model cross-β x-ray amyloid structure[54–56]. The intra-strand distance is taken to be the distance between corresponding residues on neighboring strands within each β-sheet. This assumes that the peptides are in register which is, in fact what we observe. For sequences containing valine, the intra-strand distance increases slightly from 4.8Å to 5Å, again suggesting that at least for the MoPrP 111–120 sequence there is still a lot of fluctuation in the backbone of the peptide. This may be due to the disturbance posed by the flanking valine residues which disrupts the tight close-packed β-sheets observed experimentally.

Simulations provide us with a molecular-level snapshot of how the peptides arrange themselves into amyloid fibrils. For the sequences that form ordered structures, such as AGAAAAGA, the positions of the side chains within the interface between the β-sheets is of particular interest. In Figure 4(a) we show a simulation snapshot of the fibril formed by AGAAAAGA for a case in which the fibril contains three β-sheets and a close-up that

shows how the strands of one sheet alternate or stagger relative to the strands in the neighboring sheet. This staggered conformation is consistent with x-ray diffraction[19]. Figure 4(b) shows a rotated view of the snapshot of Figure 4(a) looking down the fibril axis; the side chains on the peptide form a steric zipper between the β-sheets, holding the sheets together and forming the cross-β structure. The four center alanine side chains lie above and below the plane of the beta sheet and interact with the alanine side chains on the neighboring sheets. The phi-psi angles for this structure were found to be primarily in the β-region of the Ramachandran plot, with the degree of flexibility expected for the presence of glycine (data not shown). The intra-strand distance, conformational stagger, and side-chain inter-digitation in the fibrillar structures formed in our simulations of the AGAAAAGA peptide are generally in agreement with those found experimentally[19] and predicted using molecular dynamics[15].

Kinetics along the Fibril Formation Pathway

One of the benefits of DMD-PRIME20 computer simulation is that we are able to monitor all of the steps, including intermediate creation and dissolution, along the fibril formation pathway. This is useful because it is difficult to isolate experimentally the soluble disordered intermediate species that populate the aggregation pathway experimentally. Figure 5 shows snapshots from a simulation of 48 chains of Syrian Hamster prion peptide AGAAAAGA at T*=0.15 from free monomers, to β-sheet formation and then β-sheet association into a fibril. Figure 6(a) shows population data corresponding to the simulation in Figure 4(a) on the number of AGAAAAGA monomers, β-sheets, amorphous aggregates and fibrils as a function of time. At t=10, all of the peptides are free monomers (see Figure 5 (top left) and blue curve in Fig 6(a)). At t = 400, small β-sheets are starting to form; very few if any amorphous aggregates form before β-sheet formation occurs. (See Figure 5(top right) and red curve in Figure 6(a)). At t =1000, fibrils begin to appear and the fraction of β-sheets starts to diminish as the peptides associate to form fibrils. (See Figure 5 (bottom left) and purple curve in Figure 6(a).) At the end of the simulation, t=1700, all of the free β-sheets are now in a fibril structure (See Figures 5 (bottom right) and 6(a)).

The assembly mechanism for the VAGAAAAGAV peptides is quite different. Figure 7 shows snapshots over the course of a simulation of 48 VAGAAAAGAV peptides at T*=0.18 starting from an initial configuration of random coils. As described in detail below, the fibril assembly of VAGAAAAGAV proceeds through the formation of amorphous aggregates in contrast to the fibril assembly of AGAAAAGA which proceeds primarily through the formation of β-sheets. Figure 6(b) shows population data corresponding to the simulation in Figure 7 on the number of VAGAAAAGAV monomers, β-sheets, amorphous aggregates and fibrils as a function of time. At t=30, all of the peptides are free monomers (see Figure 7 (top left) and blue curve in Fig 6(b)). Figure 6(b) shows that by t=250, the number of monomers (blue curve) has decreased in favor of association into amorphous aggregates (green curve). This could be because the valine end groups are pulling the peptides together but the energy is not strong enough to overcome the entropic loss associated with the formation of β-sheets. Between t=250 and t=500, the peptides tend to form amorphous aggregates (green curve) rather than the more-ordered β-sheets (red curve). This is apparent in Figure 7 at time t=350 which shows three relatively indistinct globules of peptides. As the number of amorphous aggregates grows, the energy of the system eventually becomes large enough to overcome the meta-stable kinetic trap to form the more-ordered β-sheet structures and so β-sheets start to form at t=500 (red line). The competition between β-sheets and amorphous aggregates continues throughout the simulation, even after fibrils start to form (purple line) at t=750. This is illustrated in Figure 7 at time t=1500, which shows that the pink and blue β-sheets have associated to form a fibril but that a small cluster of amorphous aggregates still remains (shown as green peptides). This persistence of

the amorphous state over time explains why VAGAAAAGAV makes fibrils to a lesser extent than the other sequences simulated in this paper. Finally at time, t=1800, most of the peptides are in a fibril composed of two β-sheets (blue and pink in Figure 7) but about twenty percent of the peptides are either in amorphous aggregates or exist as free monomers (see Figure 6(b)).

The connection between aggregation kinetics and the structure of the fibril that is ultimately formed can be understood by considering the formation of fibrils by one of the variations on the MoPrP, GAVAAAAVAG. As shown in Table 2, this sequence formed the highest percentage of fibrils of all of the sequences considered. In the simulations for this sequence, the first non-monomer species to appear are the amorphous aggregates and then, after a long time, the β-sheets. The β-sheets very quickly associate into fibrils and the amorphous aggregates begin to disappear, suggesting that they are becoming part of the ordered fibril structure. In contrast, VAGAAAAGAV peptides, as we saw in Figure 7, seem to have an equal preference for β-sheets and amorphous aggregates and do not exhibit a significant drop-off in amorphous aggregate population even after fibril formation begins. Thus the key difference between the kinetics of VAGAAAAGAV and GAVAAAAVAG is that as soon as GAVAAAAVAG begins to form β-sheets, the amorphous aggregates decrease and well ordered fibrils arise, whereas for VAGAAAAGAV the amorphous aggregates never completely disappear causing a relative dearth of nicely structured material (β-sheets) around which to build nicely ordered fibrils.

DISCUSSION

Spontaneous fibril formation was observed in discontinuous molecular dynamics simulations of 48 peptide systems containing various sequences related to the short palindromic prion peptides, SHaPrP 113–120 and MoPrP 111–120 starting from a configuration of random coils. This was accomplished using our new implicit-solvent force field, PRIME20. The ten different sequences (AGAAAAGA, GAAAGAAA, $(AG)_4$, GAAAGAAA, A8, A10, V10, VAGAAAAGAV, GAVAAAAVAG and VAVAAAAVAV) formed structures, including fibrils, with varying degrees of order, depending on the chain length and hydrophobicity over a range of temperatures. At higher temperatures, the peptides remained as random coils and at lower temperatures the peptides “clumped” together into amorphous aggregates. To our knowledge this is the first time that the spontaneous formation of AGAAAAGA and VAGAAAAGAV fibrils has been observed in a simulation

The structure of the fibrils formed and the kinetics along the aggregation pathway for the two prion sequences SHaPrP 113–120 (AGAAAAGA) and MoPrP 111–120 (VAGAAAAGAV) were significantly different. Overall the fibrils formed by AGAAAAGA were more ordered than those formed by VAGAAAAGAV. We believe this is a consequence of the different aggregation pathways for the two sequences. AGAAAAGA initially forms small β-sheets; these elongate by monomer addition and then associate into fibrils. VAGAAAAGAV initially forms amorphous aggregate and forms β-sheets only after a long time; the number of amorphous aggregates and beta sheets remain on par throughout the simulation. This, leads to a relative dearth of β-sheets that can act as seeds for fibril formation, essentially preventing the mouse prion peptides from forming fibrils as nicely structured as the Syrian hamster prion peptide. The relationship between the type of pathway, final structure and peptide sequence can be better understood by reference to the concept of the condensation-ordering transition [57–59], in which disordered oligomers formed by condensation at an early stage are transformed to ordered structures at a later stage. For AGAAAAGA(SHaPrP113–120), the monomers form oligomers which rapidly form single β-sheets. This condensation step is driven primarily by the strong hydrogen bonding energy.

Thereafter the β-sheets merge and structurally reorganize to form ordered proto-filaments (the ordering step). However for VAGAAAAGAV(MoPrP111–120), the monomers rapidly form disordered oligomers. This condensation step is driven by both hydrogen bonding and by hydrophobic interactions which are stronger than in the AGAAAAGA case. Thereafter the oligomers merge and structurally reorganize to from ordered proto-filaments (the ordering step).

The structures of the fibrils formed by each sequence were characterized. Most of the sequences that were able to form fibrils, ended up with 2–3 sheets. The shorter sequences showed no preference for anti-parallel or parallel arrangement. We observed an intra-strand distance between 4.7—5.0 Å which is slightly higher than the 4.7Å predicted for the cross-β x-ray model [54,56,60]. We do not, as yet, have a systematic manner to calculate the inter-sheet distance but a rough estimate from the molecular structure file suggests an inter-sheet distance of ~9Å which is consistent with earlier estimates[19]. We were also able to observe the steric zipper pattern of the side chains between the β-sheets in the fibril. The differences observed in the fibrillization pathway and fibril structure between the very similar Syrian hamster and mouse prion peptides (and their variants) lends further support to the importance of peptide sequence in the fibrillization process.

The “take home” message for this paper is that it is not just thermodynamics that determines whether or not a particular sequence forms a fibrillar structure. Kinetics plays an equally important role in that the sequence dictates the types of intermediates that appear along the aggregation pathway which in turn determine the type of structure that appears in the final stages of the aggregation process.

It is important to point out that although PRIME20 is a relatively comprehensive intermediate-resolution model, some refinements may need to be considered before we can make definitive conclusions regarding its ability to capture the role of valine in protein aggregation. In fact, additional simulations are underway to determine how well PRIME20 accounts for the subtle differences between the twenty amino acid residues especially in regards to the essential physics of fibril formation.

Acknowledgments

This work was supported by the National Institutes of Health, USA under grant GM56766 and EB006006 (MC,CKH) and National Creative Research Initiatives (Center for Proteome Biophysics) of National Research Foundation/Ministry of Education, Science and Technology, Korea, Grant No. 2008-0061984 (MC,IC). The authors are grateful to E. Phelps whose kinetic model of aggregation helped to clarify the fibrillization processes of the sequences studied in this paper and J. Maury-Evertsz for help with visualization software.

REFERENCES

1. Kelly JW. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. Curr. Opin. Struct. Biol. 1998; 8:101–106. [PubMed: 9519302]
2. Prusiner SB. Novel proteinaceous infectious particles cause scrapie. Science. 1982; 216:136–144. [PubMed: 6801762]
3. Prusiner SB. Prion Diseases and the BSE Crisis. Science. 1997; 278:245–251. [PubMed: 9323196]
4. Prusiner SB. Prions. Proc Natl Acad Sci USA. 1998; 95:13363–13383. [PubMed: 9811807]
5. Clarke AR, Jackson GS, Collinge J. The molecular biology of prion propagation. Phil. Trans. R. Soc. Lond. B. 2001; 356:185–195. [PubMed: 11260799]
6. Jobling MF, Stewart LR, White AR, McLean C, Friedhuber A, Maher F, Beyreuther K, Masters CL, Barrow CJ, Collins SJ, et al. The Hydrophobic Core Sequence Modulates the Neurotoxic and Secondary Structure Properties of the Prion Peptide 106–126. J Neurochemistry. 1999; 73:1557–1565.

7. Lee S-W, Mou Y, Lin S-Y, Chou F-C, Tseng W-H, Chen C-h, Lu C-YD, Yu SS-F, Chan JCC. Steric Zipper of the Amyloid Fibrils formed by residues 109–122 of the Syrian Hamster Prion Protein. Journal of Molecular Biology. 2008; 378:1142–1154. [PubMed: 18423487]
8. Grosset A, Moskowitz K, Nelsen C, Pan T, Davidson E, Orser CS. Rapid presymptomatic detection of PrPSc via conformationally responsive palindromic PrP peptides. Peptides (New York, NY, United States). 2005; 26:2193–2200.
9. Gasset M, Baldwin MA, Lloyd DH, Gabriel JM, Holtzman DM, Cohen F, Fletterick R, Prusiner SB. Predicted alpha-helical regions of the prion protein when synthesized as peptides form amyloid. Proc. Natl. Acad. Sci. USA. 1992; 89:10940–10944. [PubMed: 1438300]
10. Ziegler J, Viehrig C, Geimer S, Rosch P, Schwarzinger S. Putative aggregation initiation sites in prion protein. FEBS Lett. 2006; 580:2033–2040. [PubMed: 16545382]
11. Shinchuk LM, Sharma D, Blondelle SE, Reixach N, Inouye H, Kirschner DA. Poly-(L-alanine) expansions form core beta-sheets that nucleate amyloid assembly. Proteins: Structure, Function and Bioinformatics. 2005; 61:579–589.
12. Tessier PM, Lindquist SL. Unraveling infectious structures, strain variants and species barriers for the yeast prion [PSI+]. Nature Structural & Molecular Biology. 2009; 16:598–605.
13. Walsh P, Simonetti K, Sharpe S. Core structure of amyloid firbrils formed by residues 106–126 of the Human Prion Protein. Structure. 2009; 17:417–426. [PubMed: 19278656]
14. Walsh P, Yau J, Simonetti K, Sharpe S. Morphology and Secondary Structure of Stable beta-oligomers fromed by amyloid peptide PrP(106–126). Biochem. 2009; 48:5779–5781. [PubMed: 19476383]
15. Ma B, Nussinov R. Molecular dynamics simulations of alanine rich b-sheet oligomers: Insight into amyloid formation. Prot. Sci. 2002; 11:2335–2350.
16. Kuwata K, Matumoto T, Cheng H, Nagayama K, James TL, Roder H. NMR-detected hydrogen exchange and molecular dynamics simulations provide structural insight into fibril formation of prion protein fragment 106–126. Proc Natl Acad Sci USA. 2003; 100:14790–14795. [PubMed: 14657385]
17. Prusiner SB. Molecular biology of prion diseases. Science. 1991; 252:1515–1522. [PubMed: 1675487]
18. Inouye H, Kirschner DA. Polypeptide chain folding in the hydrophobic core of hamster scrapie prion: analysis by X-ray diffraction. J Struct Biol. 1998; 122:247–255. [PubMed: 9724626]
19. Nguyen JT, Inouye H, Baldwin MA, Fletterick RJ, Cohen FE, Prusiner SB, Kirschner DA. X-ray diffraction of scrapie prion rods and PrP peptides. J Mol Biol. 1995; 252:412–422. [PubMed: 7563061]
20. Muramoto T, Scott M, Cohen F, Prusiner S. Recombinant scrapie-like prion protein of 106 amino acids is soluble. Proceedings of the National Academy of Sciences of the United States of America. 1996; 93:15457–15462. [PubMed: 8986833]
21. Lu X, Wintrode PL, Surewicz WK. Beta-sheet core of human prion protein amyloid fibrils as determined by hydrogen/deuterium exchange. Proc Natl Acad Sci U S A. 2007; 104:1510–1515. [PubMed: 17242357]
22. Florio T, Paludi D, Villa V, Principe DR, Corsaro A, Millo E, Damonte G, D'Arrigo C, Russo C, Schettini G, et al. Contribution of two conserved glycine residues to fibrillogenesis of the 106–126 prion protein fragment. Evidence that the soluble variant of the 106–126 peptide is neurotoxic. Journal of Neurochemistry. 2003; 85:62–72. [PubMed: 12641727]
23. Kupfer L, Hinrichs W, Groschup MH. Prion Protein Misfolding. Current Molecular Medicine. 2009; 9:826–835. [PubMed: 19860662]
24. Teplow DB. Structural and kinetic features of amyloid beta-protein fibrillogenesis. Amyloid. 1998; 5:121–142. [PubMed: 9686307]
25. Ma B, Nussinov R. Stabilities and conformations of Alzhemier's b-amyloid peptide oligomers (Ab16-22, Ab16-35, and Ab10-35):Sequence Effects. Proc Natl Acad Sci USA. 2002; 99:14126–14131. [PubMed: 12391326]
26. Zheng J, Ma B, Tsai C-J, Nussinov R. Structural stability and dynamics of an amyloid-forming peptide GNNQQNY from the yeast prion Sup-35. Biophys J. 2006; 91:824–833. [PubMed: 16679374]

27. Balbirnie M, Grothe R, Eisenberg DS. An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated beta-sheet structure for amyloid. Proc Natl Acad Sci U S A. 2001; 98:2375–2380. [PubMed: 11226247]
28. Antzutkin ON, Balbach JJ, Leapman RD, Rizzo NW, Reed J, Tycko R. Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of beta-sheets in Alzheimer's beta-amyloid fibrils. Proc. Natl. Acad. Sci. USA. 2000; 97:13045–13050. [PubMed: 11069287]
29. Benzinger TL, Gregory DM, Burkoth TS, Miller-Auer H, Lynn DG, Botto RE, Meredith SC. Two-dimensional structure of beta-amyloid(10–35) fibrils. Biochem. 2000; 39:3491–3499. [PubMed: 10727245]
30. Cheon M, Chang I, Hall CK. Extending the Prime Model for Protein Aggregation of All Twenty Amino Acids. PRoteins: Structure, Function, and Bioinformatics. 2010; 78:2950–2960.
31. Smith AV, Hall CK. a-Helix Formation: Discontinuous Molecular Dynamics on an Intermediate Resolution Model. Protein: Structure, Function and Genetic. 2001; 44:344–360.
32. Smith AV, Hall CK. Protein Refolding Versus Aggregation: Computer Simulations on an Intermediate Resolution Model. J. Mol. Biol. 2001; 312:187–202. [PubMed: 11545596]
33. Marchut AJ, Hall CK. Effects of chain length on the aggregation of model polyglutamine peptides: molecular dynamics simulations. Proteins. 2007; 66:96–109. [PubMed: 17068817]
34. Marchut AJ, Hall CK. Spontaneous formation of annular structures observed in molecular dynamics simulations of polyglutamine peptides. Comput Biol Chem. 2006; 30:215–218. [PubMed: 16678490]
35. Marchut AJ, Hall CK. Side-chain interactions determine amyloid formation by model polyglutamine peptides in molecular dynamics simulations. Biophys J. 2006; 90:4574–4584. [PubMed: 16565057]
36. Nguyen HD, Marchut AJ, Hall CK. Solvent effects on the conformational transition of a model polyalanine peptide. Prot. Sci. 2004; 13:2909–2924.
37. Nguyen HD, Hall CK. Phase diagrams describing fibrillization by polyalanine peptides. Biophys J. 2004; 87:4122–4134. [PubMed: 15465859]
38. Nguyen HD, Hall CK. Molecular dynamics simulations of spontaneous fibril formation by rando-soil peptides. Proc Natl Acad Sci USA. 2004; 101:16180–16185. [PubMed: 15534217]
39. Nguyen HD, Hall CK. Kinetics of fibril formation by polyalanine peptides. J Biol Chem. 2004; 280:9074–9082. [PubMed: 15591317]
40. Blondelle SE, Forood B, Houghten RA, Perez-Paya E. Polyalanine-based peptides as models for self-associated b-pleated-sheet complexes. Biochemistry. 1997; 36:8393–8400. [PubMed: 9204887]
41. Perez-Paya E, Forood B, Houghten RA, Blondelle SE. Structural characterization and 5'-mononucleotide binding of polyalanine b-sheet complexes. J. Mol. Recognit. 1996; 9:488–493. [PubMed: 9174929]
42. Gustafsson M, Griffiths WJ, Furusjo E, Johansson J. The palmitoyl groups of lung surfactant protein C reduce unfolding into a fibrillogenic intermediate. Journal of Molecular Biology. 2001; 310:937–950. [PubMed: 11453699]
43. Hosia W, Johansson J, Griffiths WJ. Hydrogen/deuterium exchange and aggregation of a polyvaline and an polyleucine alpha-helix investigated by matrix-assisted laser desorption ionization mass spectrometry. Molecular and Cellular Proteomics. 2002; 1.8:592–597. [PubMed: 12376574]
44. Szyperski T, Vandenbussche G, Curstedt T, Ruysschaert J-M, Wuthrich K, Johansson J. Pumonary surfactant-associated polypeptide C in a mixed organic solvent transforms from a monomeric alpha-helical state into insoluble beta-sheet aggregates. Protein Science. 1998; 7:2533–2540. [PubMed: 9865947]
45. Bellemans A, Orbans J, Belle DV. Molecular dynamics of rigid and non-rigid necklaces of hard disks. Mol. Phys. 1980; 39:781–782.
46. Alder BJ, Wainwright TE. Studies in molecular dynamics, I: General method. J Chem Phys. 1959; 31:459–466.
47. Smith SW, Hall CK, Freeman BD. Molecular Dynamics for Polymeric Fluids Using Discontinuous Potentials. Journal of Computational Physics. 1997; 134:16–30.

48. Rapaport DC. Molecular dynamics study of polymer chains. J. Chem. Phys. 1979; 71:3299–3303.
49. Rapaport DC. Molecular dynamics simulation of polymer chains with excluded volume. J. Phys. A. 1978; 11:L213.
50. Andersen HC. Molecular dynamics simulation at constant temperature and / or pressure. J. Chem. Phys. 1980; 72:2384–2393.
51. Auer S, Kashchiev D. Phase Diagram of α-Helical and β-Sheet Forming Peptides. Phys Rev Lett. 2010; 104:168105. [PubMed: 20482086]
52. Kashchiev D, Auer S. Nucleation of amyloid fibrils. J. Chem. Phys. 2010; 132:215101. [PubMed: 20528047]
53. Zhang J, Muthukumar M. Simulation of nucleation and elongation of amyloid fibrils. J Chem Phys. 2009; 130:035102. [PubMed: 19173542]
54. Serpell LC. Alzheimers amyloid fibrils: Structure and assembly. Biochimica et Biophysica Acta. 2000; 1502:16–30. [PubMed: 10899428]
55. Serio TR, Cashikar AG, Kowal AS, Sawicki GJ, Moslehi JJ, Serpell L, Arnsdorf MF, Lindquist SL. Nucleated conformational conversion and the replication of conformational information by a prion determinant. Science. 2000; 289:1317–1321. [PubMed: 10958771]
56. Sunde M, Serpel LC, Bartlam M, Fraser PE, Pepys MB, Blake CCF. Common core structure of amyloid fibrils by synchotron X-ray diffraction. J. Mol. Biol. 1997; 273:729–739. [PubMed: 9356260]
57. Auer S, Meersman F, Dobson CM, Vendruscolo M. A Generic Mechanism of Emergence of Amyloid Protofilaments from Disordered Oligomeric Aggregates. PLoS ComP Biol. 2008; 4:e1000222.
58. Pellarin R, Caflisch A. Interpreting the Aggregation Kinetics of Amyloid Peptides. J Mol Bol. 2006; 360:882–892.
59. Bellesia G, Shea J-E. Diversity of kinetic pathways in amyloid fibril formation. J Chem Phys. 2009; 131:111102. [PubMed: 19778093]
60. Blake C, Serpell LC. Synchotron X-ray studies suggest that the core of transthyretin amyloid fibril is a continuous b-sheet helix. Structure. 1996; 4:989–998. [PubMed: 8805583]

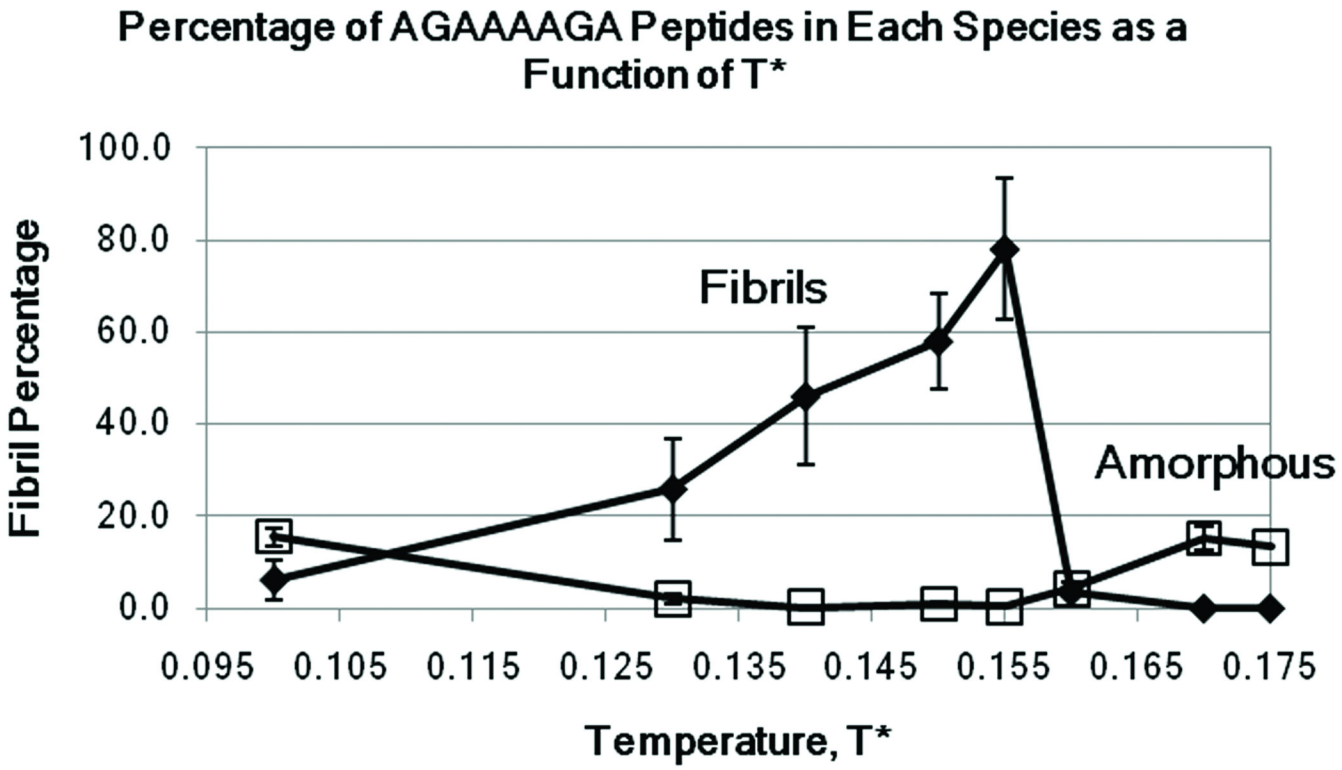

Figure 1.
Percentage of AGAAAAGA peptides in a fibril (filled diamond) and in amorphous aggregates (empty square) as a function of temperature, T*.

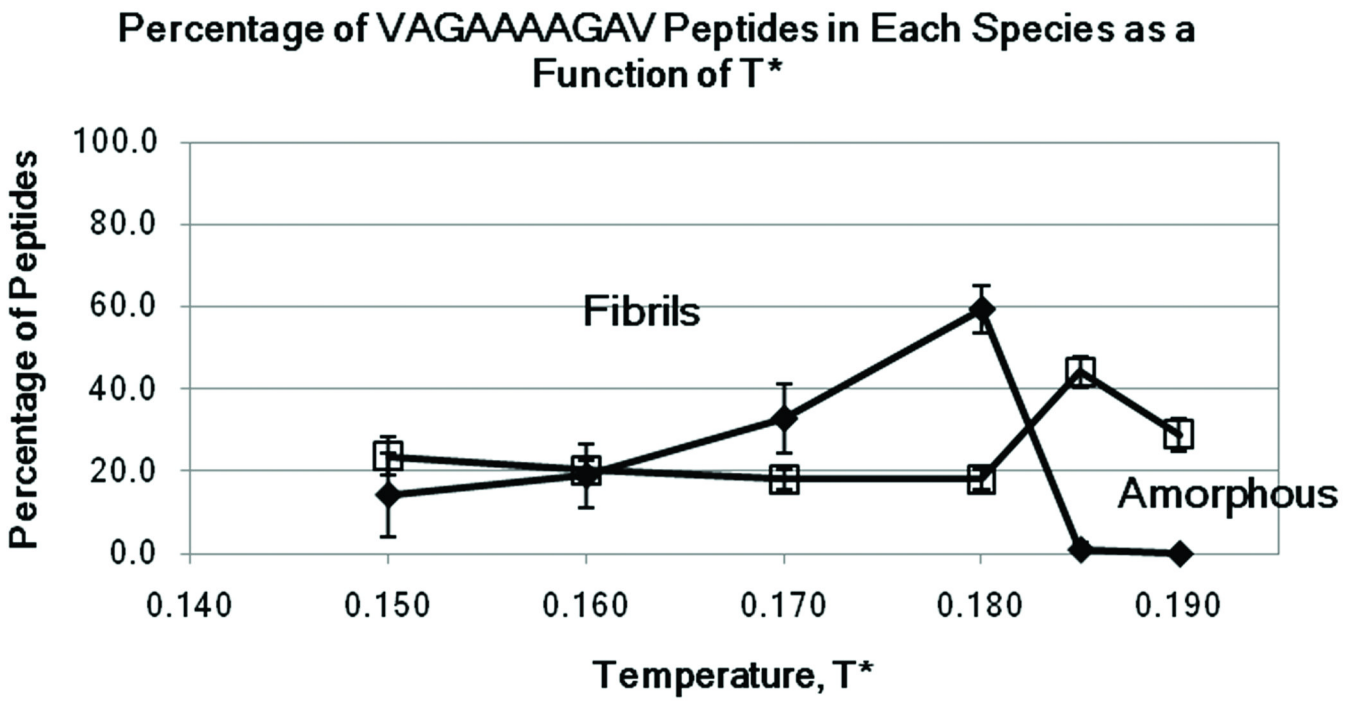

Figure 2.
Percentage of VAGAAAAGAV peptides in a fibril (filled diamond) and in amorphous aggregates (empty square) as a function of temperature, T*.

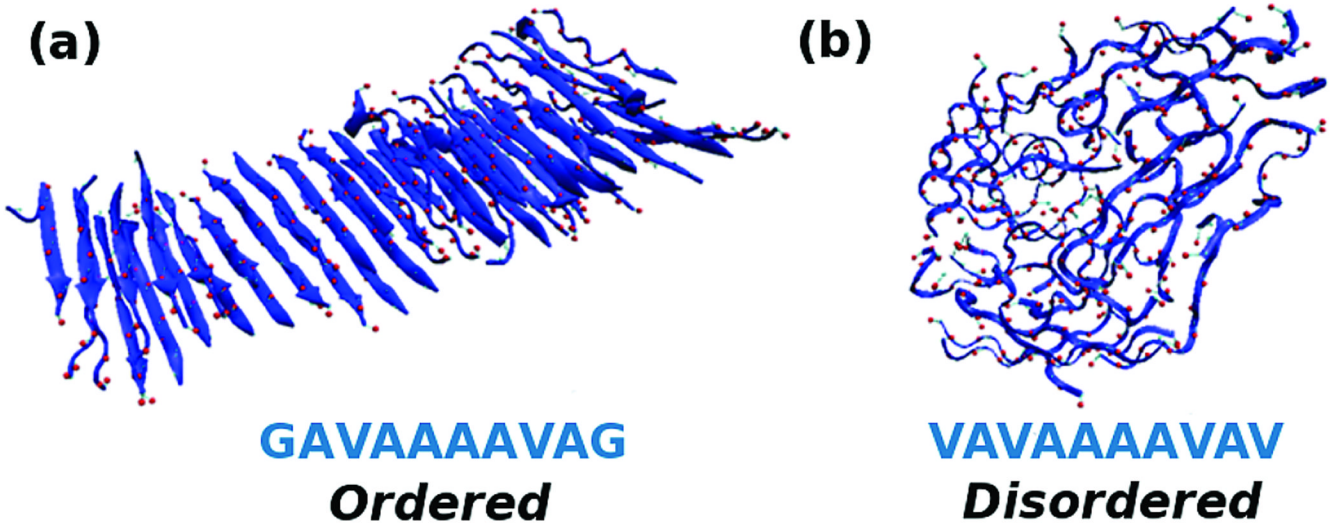

Figure 3.
Simulation snapshots of GAVAAAAVAG (left) and VAVAAAAVAV (right) at T*=0.17. Snapshots rendered using VMD.

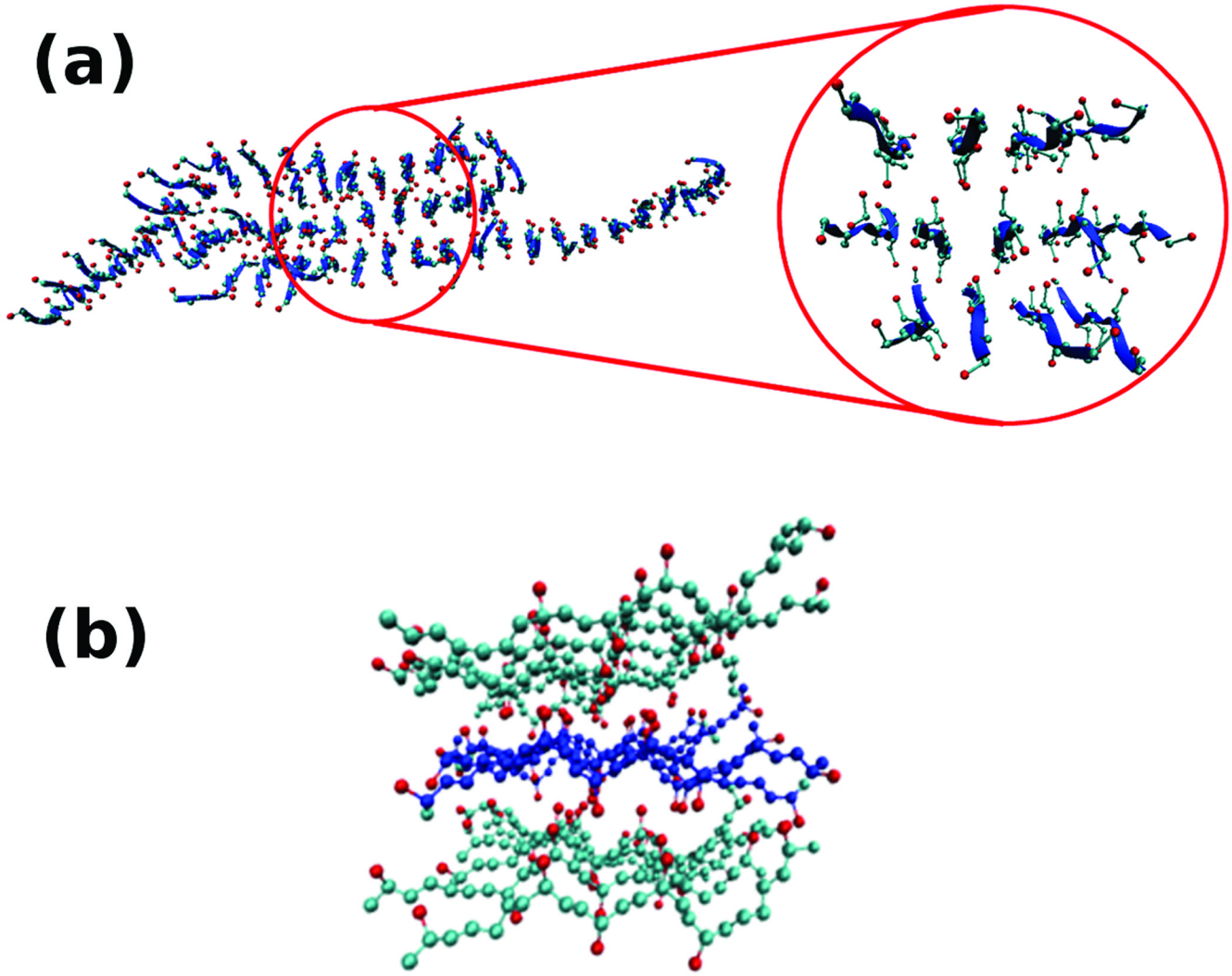

Figure 4.
(a) Simulation snapshot of the SHaPrP 113–120 AGAAAAGA fibril and close up of the arrangement of side chains at the interface between the sheets. The alanine side chains are shown in red. (b) Snapshot of the AGAAAAGA, rotated image of Figure 4(a).

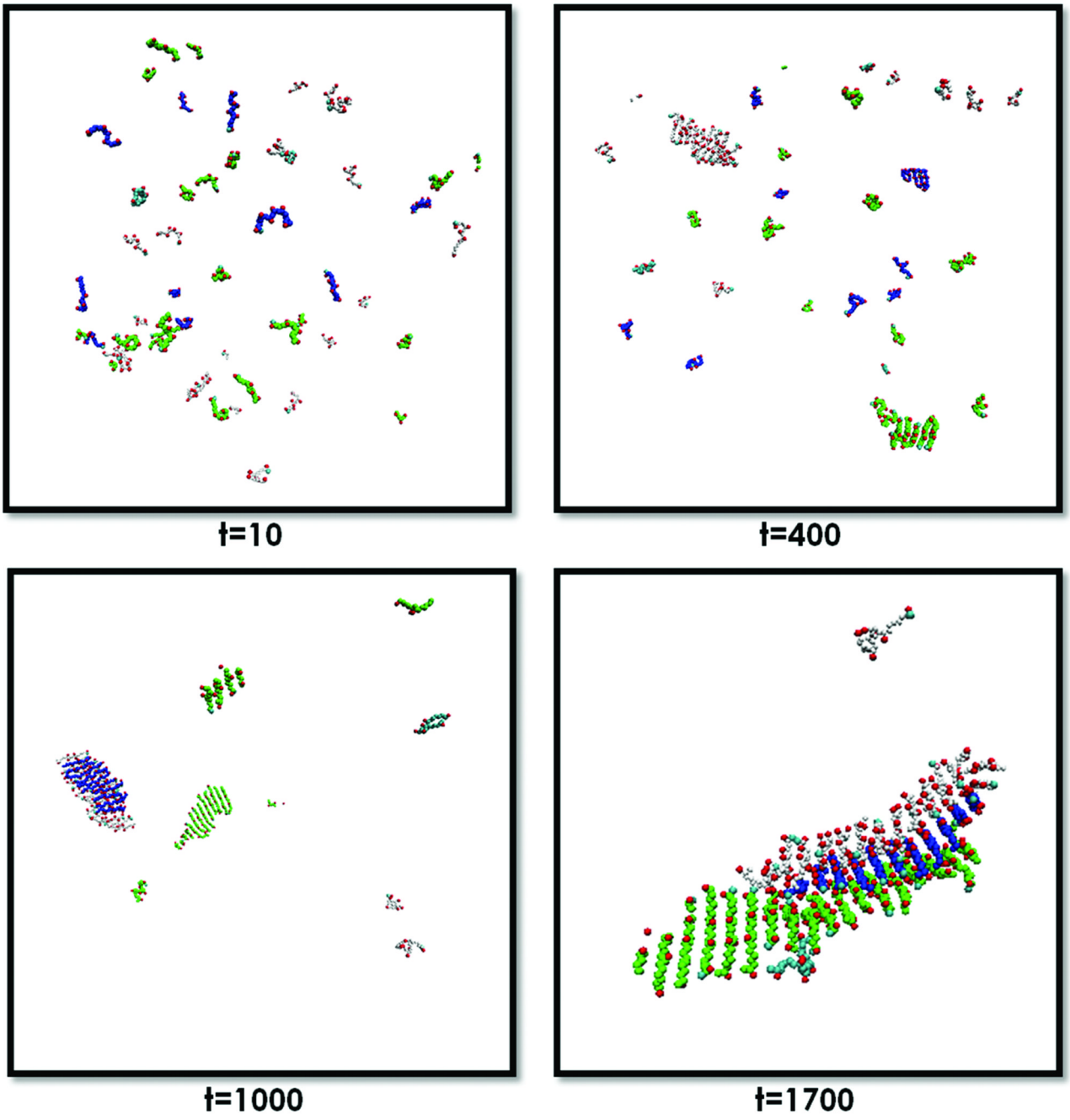

Figure 5.
Simulation snapshots for SHaPrP 113–120 at T*=0.15, shown at time, t=10 (very early), t=400, t=1000 and at the end of the simulation, t=1700.

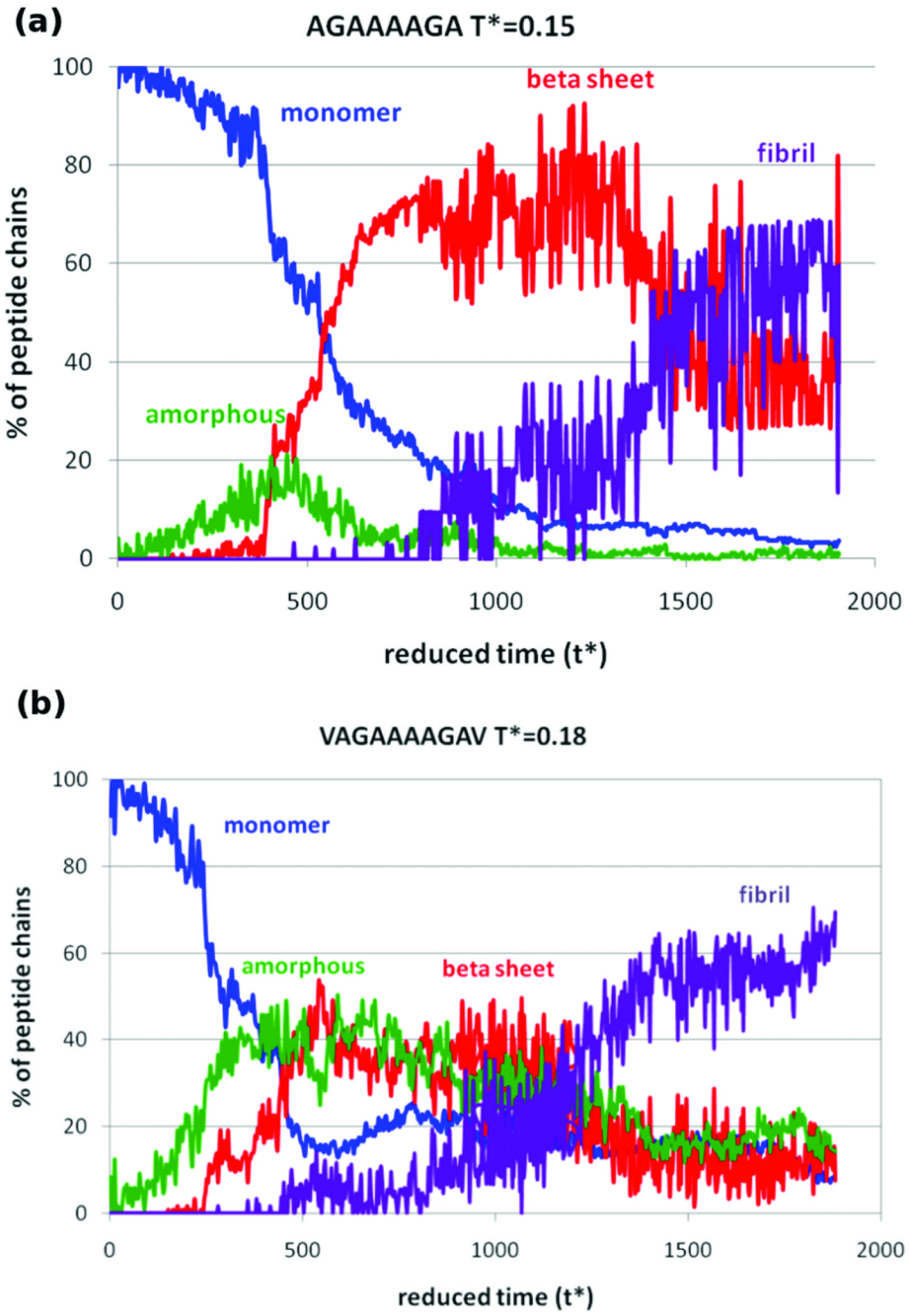

Figure 6.
Percentage of peptides in various conformations (monomers, β-sheets, amorphous and fibrils for (a) AGAAAAGA at T*=0.15 and (b) VAGAAAAGAV at T*=0.18

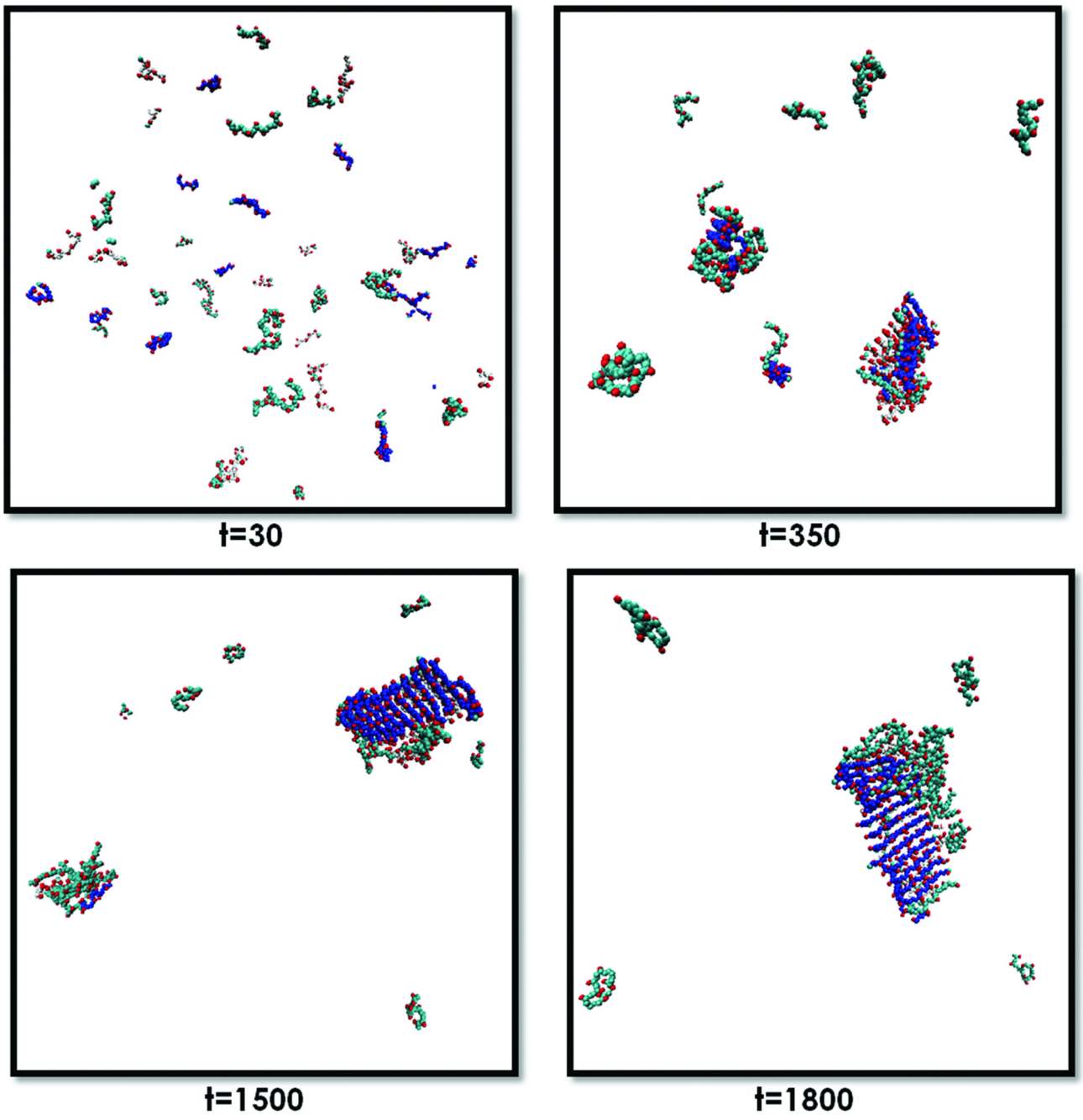

Figure 7.
Simulation snapshots for MoPrP 111–120 at T*=0.18, shown at time, t=30 (very early), t=350, t=1500, and at the end of the simulation, t=1800.

Table 1

PRIME20 Geometric and Energetic Parameters for alanine, and valine. Distances are in Å unit and masses are in unit mass of CH_3 (15amu=1.0)

Identity	Abbr.	R-Cα bond	σ	λσ	m
Alanine	A	1.600	2.700	5.400	1.00
Valine	V	2.002	3.300	6.306	2.87

Table 2

Types of structures formed by each sequence at inter-mediate temperatures. The percentages of peptides that are monomers, beta-sheets, amorphous aggregates and fibrils are indicated.

	T*	%Mono	%Beta	%Amor	%Fibril	Structure
AGAAAAGA	0.14–0.155	1.4	39.2	0.6	58.8	Ordered
GAAAAAAG	0.15–0.16	3.8	35.7	0.5	59.9	Ordered
AAAAAAAA	0.15–0.16	18.3	29.1	3.4	49.2	Ordered
GAAAGAAA	0.14–0.15	2.4	64.5	0.7	32.4	Slightly Ordered
AGAGAGAG	0.14–0.15	22.6	74.6	2.8	0.0	Disordered
VAGAAAAGAV	0.17–0.18	6.9	29.0	18.0	46.0	Slightly Ordered
GAVAAAAVAG	0.16–0.17	6.3	15.1	4.1	74.5	Ordered
AAAAAAAAAA	0.16–0.17	22.1	14.1	3.5	60.4	Ordered
VAVAAAAVAV	0.17–0.18	4.4	31.3	31.0	33.3	Slightly Ordered
VVVVVVVVVV	0.17–0.18	0.9	20.5	44.5	34.1	Slightly Ordered

Table 3

Structural characteristics of fibrils formed by AGAAAAGA, GAAAAAAG, A8, VAGAAAAAGAV, GAVAAAAGAV, and A10.

	# Fibrils	# Sheets /Fibril	# Strands /Sheet	Parallel/ Antiparallel	Intrastrand Distance
AGAAAAGA	1	2	16±7	0.90±0.2	4.80±0.4
GAAAAAAG	1	2–3	14±4	1.20±0.3	4.76±0.6
AAAAAAAA	1	2–3	15±3	0.99±0.2	4.83±0.4
VAGAAAAGAV	1	2	12±1	1.30±0.8	5.03±0.6
GAVAAAAVAG	1	2–3	13±2	1.90±0.2	4.83±0.5
AAAAAAAAAA	1	2–4	9±1	1.80±0.5	4.95±0.4