Assembly and Kinetic Folding Pathways of a Tetrameric β-Sheet Complex: Molecular Dynamics Simulations on Simplified Off-Lattice Protein Models

Hyunbum Jang,* Carol K. Hall,* and Yaoqi Zhou†
*Department of Chemical Engineering, North Carolina State University, Raleigh, North Carolina 27695-7905; and
†Department of Physiology and Biophysics, State University of New York at Buffalo, Buffalo, New York 14214

ABSTRACT We have performed discontinuous molecular dynamic simulations of the assembly and folding kinetics of a tetrameric β-sheet complex that contains four identical four-stranded antiparallel β-sheet peptides. The potential used in the simulation is a hybrid Go-type potential characterized by the bias gap parameter g, an artificial measure of a model protein's preference for its native state, and the intermolecular contact parameter η, which measures the ratio of intermolecular to intramolecular native attractions. The formation of the β-sheet complex and its equilibrium properties strongly depend on the size of the intermolecular contact parameter η. The ordered β-sheet complex in the folded state and nonaligned β-sheets or tangled chains in the misfolded state are distinguished by measuring the squared radius of gyration R_g^2 and the fraction of native contacts Q. The folding yield for the folded state is high at intermediate values of η, but is low at both small and large values of η. The folded state at small η is liquid-like, but is solid-like at both intermediate and large η. The misfolded state at small η contains nonaligned β-sheets and tangled chains with poor secondary structure at large η. Various folding pathways via dimeric and trimeric intermediates are observed, depending on η. Comparison with experimental results on protein aggregation indicates that intermediate η values are most appropriate for modeling fibril formation and small η values are most appropriate for modeling the formation of amorphous aggregates.

INTRODUCTION

We have been engaged in a computational study aimed at understanding the basic physical principles underlying protein aggregation, a cause or associated symptom of a number of fatal protein deposition diseases (Clark and Steele, 1992; Eaton and Hofrichter, 1990; Gallo et al., 1996; Massry and Glasscock, 1983; Moore and Melton, 1997). An example is Alzheimer's disease (Selkoe, 1991; Simmons et al., 1994), which is thought to be caused by the abnormal deposition of insoluble fibril (amyloid) plaques in the extracellular space of brain tissue during aging. The major component of these plaques is a 39- to 43-residue β-peptide, called β-amyloid peptide (Aβ), whose predominant secondary structure in the fibril is a β-sheet. Although extensive experimental studies (Benzinger et al., 1998, 2000; Burkoth et al., 1998; Esler et al., 1996, 2000; Lazo and Cowning, 1998; Lynn and Meredith, 2000; Sunde et al., 1997; Zhang et al., 2000) have been conducted on the fibrillization of β-amyloid peptide (Aβ) and other amyloid forming peptides, the causes of amyloid fibril formation are largely unknown and the fundamental mechanisms underlying this aggregation are still under investigation. Recent experimental observations that many different proteins are capable of forming fibrils under slightly denatured concentrated conditions have led leaders in the field to suggest that fibril formation is an intrinsic property of polypeptides (Booth et al., 1997; Chiti et al., 1999). The idea here is that fibril formation is not strongly dependent upon sequence but is instead a consequence of the ability of peptides to experience hydrogen bonding and hydrophobic interactions. This suggests that computer simulations of multipeptide systems that take these types of interactions into account could offer insights into the basic principles behind fibril formation in peptides.

The long-term goal of our work is to reveal the molecular-level mechanisms underlying protein aggregation and fibril formation with particular focus on the role played by β-strands. Because protein aggregation often involves β-strand association, or the conversion of α-helices to β-strands (Benzinger et al., 1998, 2000; Burkoth et al., 1998; Esler et al., 1996, 2000; Lazo and Cowning, 1998; Lynn and Meredith, 2000; Sunde et al., 1997; Zhang et al., 2000), we have focused part of our efforts on the development of a model fibril made up of β-sheets. To study protein aggregation by computer simulation, one must devise protein models that capture not only the basic physical features of real proteins, but also allow the simulation of many proteins with current computer capability. Our approach is to use low-resolution or simplified protein models; such models are best suited for simulating multiprotein systems because they allow us to study folding behavior over relatively long timescales. Investigations based on these models have already provided numerous insights into the thermodynamic and kinetic properties of protein folding for isolated chains (Chan and Dill, 1994; Dokholyan et al., 1998, 2000; Guo and Brooks, 1997; Guo and Thirumalai, 1995, 1996; Gupta and Hall, 1997; Gupta et al., 1999; Jang et al., 2002a,b; Kolinski et al., 1995, 1999; Lau and Dill, 1989; Miller et al., 1992; Nymeyer et al., 1998; Pande and Rokhsar, 1998; Shea et al., 2000; Skolnick and

Submitted March 17, 2003, and accepted for publication September 8, 2003.
Address reprint requests to Carol K. Hall, E-mail: hall@turbo.che.ncsu.edu.

0006-3495/04/01/31/19 $2.00

Kolinski, 1991; Zhou and Karplus, 1997a,b; 1999) and of protein aggregation for multichain systems (Bratko and Blanch, 2001; Dill and Stigter, 1995; Dima and Thirumalai, 2002; Gupta and Hall, 1998; Harrison et al., 1999, 2001).

We recently introduced three minimalist models of four-strand antiparallel β-strand peptides: the β-sheet, the β-clip, and the β-twist (Jang et al., 2002a,b). In our study of the folding thermodynamics of these three models (Jang et al., 2002a), discontinuous molecular dynamics (DMD) simulations (Alder and Wainwright, 1959; Rapaport, 1978; Smith et al., 1996) were performed to determine how the thermodynamic properties of an isolated peptide vary with temperature. Despite these models' simplicity, they undergo a complex set of protein transitions similar to those observed in experimental studies on real proteins (Ptitsyn, 1995). Starting from high temperature, these transitions include a collapse transition, a disordered-to-ordered globule transition, a folding transition, and a liquid-to-solid transition. In our study of the folding kinetics for the same models (Jang et al., 2002b), the β-sheet exhibits a fast-track folding pathway without becoming trapped in any intermediate. In contrast, the β-clip and β-twist exhibit multiple folding pathways that include trapping in intermediates and direct folding to the native state. The folding speed of the model proteins with different native state topologies strongly depends on the contact order in the native state (Plaxco et al., 1998). The thermodynamics and kinetics results presented in our previous papers set the stage for this work in which we examine the folding and assembly of several of the model β-sheet proteins into fibrils.

In this article, we investigate the equilibrium properties and the folding kinetics of a tetrameric β-sheet complex. The β-sheet complex consists of four identical four-stranded antiparallel β-sheet peptides, with each peptide containing 39 connected residues (beads). Nonbonded beads interact both intramolecularly and intermolecularly through a hybrid Go-type potential (Go and Taketomi, 1978, 1979; Taketomi et al., 1975; Ueda et al., 1978) modeled as a square-well or square-shoulder potential, depending on the values of the bias gap parameter g and the intermolecular contact parameter η. The bias gap g measures the difference in interaction strength between the intramolecular native and nonnative contacts; the larger it is, the more the native state is favored over the nonnative state. Intermediate values of the bias gap are thought to be the most representative of real proteins in their equilibrium and dynamic behavior. The intermolecular contact parameter η measures the ratio of intermolecular to intramolecular native attractions.

Our model of the β-sheet complex was designed in part to mimic the small Aβ oligomers that are observed (albeit indirectly) in the early stages of fibril formation (Harper et al., 1997). These oligomers are now widely believed to serve as the nuclei that seed the growth of the fibrils that characterize Alzheimers's and other amyloid diseases (Pallitto and Murphy, 2001). In fact recent evidence suggests that it is these oligomers, rather than the fully formed fibrils, that are the toxic species in Alzheimers's disease (Kirkitadze et al., 2002). Thus it is of interest to understand the mechanisms associated with oligomer complex nucleation. Our model was designed in part to mimic real Aβ oligomer complex formation in several respects. First, our peptides are 39 residues long and Aβ is between 39 and 42 residues long. Second, Aβ has several extended stretches of hydrophobic residues (Aβ(17-21), Aβ(32-42)); our peptide has essentially four stretches of residues that act with an external hydrophobic interaction. Third, the Aβ peptides in Aβ fibrils (and indeed the peptides in all fibrils) experience intramolecular hydrogen bonding within the β-sheets and intermolecular hydrophobic interactions between the β-sheets. The model peptides in our β-sheet complex experience intramolecular interactions within the sheets that are reminiscent of hydrogen bonding; these interactions are characterized by the Go-model bias gap parameter, g. They also experience intermolecular interactions between the sheets that are reminiscent of hydrophobic interactions; these interactions are characterized by the intermolecular contact parameter, η. The Go-model bias gap parameter is set to an intermediate value ($g = 0.9$) because Go models are considered to be the most realistic at these parameter values. The intermolecular contact parameter, η, which essentially measures the ratio of the intermolecular (hydrophobic) interaction and the intramolecular (hydrogen bonding) interaction, is selected in the range between $0.2 \leq \eta \leq 1$. The reasons for considering a range of intermolecular contact parameter values are: 1), the best values for the relative strengths of the hydrophobic and hydrogen bonding interactions in simplified models are a matter of debate (Pace et al., 1998), and 2), intermolecular hydrophobic interaction strengths depend upon the intrinsic conditions of the protein (sequence and number of hydrophobic residues) and on external conditions (pH and temperature) (Fraser et al., 1991a,b; Kowalewski and Holtzman, 1999; Snyder et al., 1994). Furthermore, by exploring how the aggregation mechanism varies with the intermolecular contact parameter we provide insights to other research workers interested in modeling aggregation phenomena.

Discontinuous molecular dynamics simulations (Alder and Wainwright, 1959; Rapaport, 1978; Smith et al., 1996) were performed on the model systems at an intermediate value of the bias gap ($g = 0.9$) for different intermolecular contact parameters ($0.2 \leq \eta \leq 1$). At each value of η, at least 50 independent trajectories starting from different initial configurations were monitored and used in the averaging. Four separated random coils generated at high temperature were used as the initial configuration in the simulations. The equilibrium properties of folded and misfolded β-sheet complexes were investigated. Thermodynamic averages were taken by accumulating the last quarter of the simulation data because the β-sheets assembled into the complex structure early in the simulation and were stable once they

were assembled. The kinetics of fibril formation was investigated by monitoring the progress parameters measuring the complex formation, including the squared radius of gyration of the entire system, R_g^2, the squared radius of gyration for the nth chain, R_{gn}^2, the fraction of total native contacts, Q_{total}, the fraction of intramolecular native contacts, Q_{intra}, the fraction of intermolecular native contacts, Q_{inter}, and the fraction of intermolecular native contacts for the nth chain, Q_n^{inter}.

Highlights of our results are the following. We find that the equilibrium properties and the kinetic formation of the β-sheet complex strongly depend on the size of the intermolecular contact parameter η. For small η, the folded state is an ordered β-sheet complex in a liquid-like phase. Most folding trajectories follow the path: four monomers $\rightarrow$ dimer and two monomers $\rightarrow$ trimer and monomer or two dimers $\rightarrow$ tetramer. The folding yield is low. Secondary structure begins to form before the β-sheet complex assembles. Many trajectories fold into a misfolded state that contains two well-aligned eight-stranded β-sheets, albeit a highly ordered structure in terms of fibril growth. For intermediate η, the folded state is a highly ordered β-sheet complex whose interior β-sheets are solid-like, and whose outer β-sheets have both solid-like and liquid-like characteristics, with the entire structure being essentially solid-like. Most folding trajectories follow the path: four monomers $\rightarrow$ dimer and two monomers $\rightarrow$ trimer and monomer $\rightarrow$ tetramer. The folding yield is very high. For large η, the folded state is a highly ordered β-sheet complex in a solid-like phase that is physically almost the same as the global energy minimum structure. Most folding trajectories follow the path: four monomers $\rightarrow$ dimer and two monomers $\rightarrow$ two dimers $\rightarrow$ tetramer. The folding yield is very low. The misfolded state contains tangled chains with poor secondary structure and is observed in a trajectory with early trimer formation. Comparison with experimental results on protein aggregation indicates that intermediate η values (where the majority of folding trajectories lead to the highly ordered β-sheet complex) are most appropriate for modeling fibril formation, whereas small η values (where the majority of folding trajectories lead to the misfolded state) are most appropriate for modeling the formation of amorphous aggregates.

In the following section, a full description of the model and the simulation method are presented. Next is the Results section, which is divided into seven subsections; the first presents the results for the equilibrium properties of the β-sheet complex, the second presents the results for the folding kinetics, the third through fifth present the results for the kinetics of the formation of the β-sheet complex at $\eta = 0.2$, 0.5, and 0.8, respectively, the sixth describes the misfolded states, and the seventh discusses the folding pathways for $\eta = 0.2$ and 0.5. The article concludes with a discussion of the key findings.

MODELS AND SIMULATION METHOD

We consider an off-lattice protein model of a β-sheet complex that consists of four identical four-stranded antiparallel β-sheet peptides. The global energy minimum structure for the model β-sheet complex is shown in Fig. 1. Each color represents a different β-sheet chain: chain 1 (*blue*), chain 2 (*red*), chain 3 (*green*), and chain 4 (*yellow*). A topology diagram representing a top view of the complex is shown at the bottom of the figure. In the topology diagrams, thick lines indicate chain connectivity at the top of the strands, and thin lines indicate chain connectivity at the bottom of the strands. It can be seen from the topology diagram that the β-sheet is not perfectly planar because it has an inflection in the middle. The β-sheet complex consisting of the kinked β-sheets has more intermolecular native contacts than that consisting of planar β-sheets. In the global energy minimum state, the total number of intramolecular native contacts, N_{intra}^{gem}, the total number of intermolecular native contacts, N_{inter}^{gem}, and the reduced squared radius of gyration for the entire system, R_g^2/σ^2 for the β-sheet complex are summarized in Table 1. For a monomer β-sheet in the global energy minimum state, the table also shows the number of intramolecular native contacts for the nth chain, $N_{n\,intra}^{gem}$, where n is the chain number ($n = 1$-4), the number of intermolecular native contacts for the nth chain, $N_{n\,inter}^{gem}$, and the reduced squared radius of gyration for the nth chain, R_{gn}^2/σ^2. Note that $N_{intra}^{gem} = \sum_{n=1}^{4} N_{n\,intra}^{gem}$, but $N_{inter}^{gem} \neq \sum_{n=1}^{4} N_{n\,inter}^{gem}$, because the intermolecular native contacts for some of the chains are shared, i.e., the interior β-sheets (*red* and *green*) have twice the number of intermolecular native contacts as the outer β-sheets (*blue* and *yellow*) in the global energy minimum state.

The model β-sheet complex contains a total of 156 (M) beads, each representing an amino acid residue that can be regarded as being localized at the C_α atom. Each β-sheet monomer contains 39 (M_n) connected beads, so that $M = 4 \times M_n$. Nonbonded beads can interact with each other through a square-well or square-shoulder potential (Zhou et al., 1996, 1997),

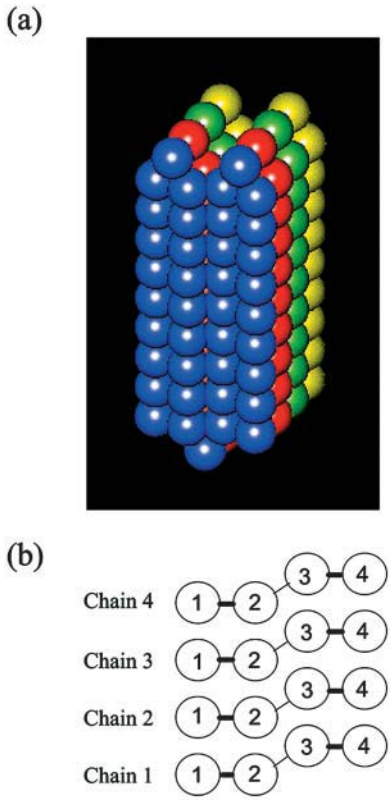

FIGURE 1 Bead (*a*) and topology (*b*) diagrams of the global energy minimum structures for the model β-sheet complex.

TABLE 1 Total number of intramolecular, N_{intra}^{gem}, and intermolecular, N_{inter}^{gem}, native contacts and the reduced squared radius of gyration, R_g^2/σ^2, for a β-sheet complex in the global energy minimum state

β-sheet complex		Each β-sheet (n = 1–4)	
N_{intra}^{gem}	300	$N_{n\,intra}^{gem}$	75
N_{inter}^{gem}	522	$N_{n\,inter}^{gem}$	174 (n = 1,4), 348 (n = 2,3)
R_g^2/σ^2	10.3495	$R_{g\,n}^2/\sigma^2$	9.0995

Also shown are the number of intramolecular native contacts for the nth chain, $N_{n\,intra}^{gem}$, the number of intermolecular native contacts for the nth chain, $N_{n\,inter}^{gem}$, and the reduced squared radius of gyration for the nth chain, $R_{g\,n}^2/\sigma^2$, for an individual β-sheet in the global energy minimum state.

$$u_{ij}(r) \equiv \begin{cases} \infty, & r \le \sigma \\ B_{ij}\varepsilon, & \sigma < r \le \lambda\sigma \\ 0, & r > \lambda\sigma, \end{cases} \tag{1}$$

where σ is the bead diameter, ε is an energy parameter with $\varepsilon < 0$, $\lambda\sigma$ is the square-well diameter with $\lambda = 1.5$ throughout, and r is the distance between residues i and j. The quantity $B_{ij}\varepsilon$, the square-well depth or square-shoulder height, represents the interaction strength between nonbonded residue pair i and j and is defined as

$$B_{ij}\varepsilon = \begin{cases} B_N^{intra}\varepsilon, & \text{Intramolecular native contacts} \\ B_O\varepsilon, & \text{Nonnative contacts} \\ B_N^{inter}\varepsilon, & \text{Intermolecular native contacts,} \end{cases} \tag{2}$$

where B_N^{intra}, B_O, and B_N^{inter} are measures of the relative strengths of the energies associated with the intramolecular native, nonnative, and intermolecular native pair interactions in this Go-type potential (Go and Taketomi, 1978, 1979; Taketomi et al., 1975; Ueda et al., 1978). Within a chain, nonbonded pairs of beads that are in contact in the global energy minimum structure experience an attractive interaction, i.e., $B_N^{intra} < 0$, when their square-wells overlap. On the other hand, within a chain and in different chains, nonbonded pairs of beads that are not in contact in the global energy minimum structure experience either an attractive interaction ($B_O < 0$) or a repulsive interaction ($B_O > 0$). The parameter B_O can be either positive or negative depending on the size of the bias gap parameter g,

$$B_O = (1-g)B_N^{intra}, \quad (B_N^{intra} < 0,\ g > 0), \tag{3}$$

where g is the bias gap (Jang et al., 2002a,b; Zhou and Karplus, 1997a,b; 1999). The bias gap measures the ratio of the interaction strength between the intramolecular native contacts and nonnative contacts. Note that for $g > 1$, $B_O > 0$; in this case nonnative contacts are repulsive so that the native state structure is strongly favored over any nonnative state structure. For $0 < g < 1$, $B_O < 0$; all nonbonded contact pairs are attractive but the intramolecular native contacts are always more favorable than the nonnative contacts. For $g = 0$, the intramolecular native and nonnative contacts are equally favorable, $B_O = B_N^{intra}$; the model reduces to a homopolymer. The bias gap is an artificial measure of a model protein's preference for its native state; in a real protein this preference for the native state might be measured for example by the energy difference between the native and nonnative state. In different chains, bead pairs that are in contact in the global energy minimum structure experience an attractive interaction, i.e., $B_N^{inter} < 0$, when their square-wells overlap. The parameter B_N^{inter} measures the interaction strength between beads in different chains and depends on the size of the intermolecular contact parameter η,

$$B_N^{inter} = \eta B_N^{intra}, \quad (0 < \eta \le 1), \tag{4}$$

where η is the intermolecular contact parameter. The quantity η measures the ratio of the interaction strength between the intermolecular and intramolecular native contacts. For $\eta = 1$, the intramolecular and intermolecular native contacts are equally favorable, i.e., $B_N^{inter} = B_N^{intra}$. For $\eta < 1$, the intramolecular native contacts are more favorable than the intermolecular native contacts, i.e., $B_N^{inter} > B_N^{intra}$. However, nonnative contacts are always less favorable than the intramolecular and intermolecular native contacts, so that $B_N^{intra} \le B_N^{inter} < B_O$ for $g > 0$. In this model, the intramolecular native interaction, intermolecular native interaction, and nonnative interaction might be regarded as mimicking hydrogen bonding, hydrophobic interactions, and van der Waals interactions, respectively, because in real fibrils, the molecular interactions along the sheets (generally hydrogen bonds) are different from the molecular interactions between the sheets (generally hydrophobic interactions).

The interaction between two bonded beads, i and $i + 1$, is given by an infinitely deep square-well potential,

$$u_{i,i+1}^{bond}(r) \equiv \begin{cases} \infty, & r < (1-\delta)\sigma \\ 0, & (1-\delta)\sigma \le r \le (1+\delta)\sigma, \\ \infty, & r > (1+\delta)\sigma, \end{cases} \tag{5}$$

where δ is a flexibility parameter that controls the bond length. The bond length between neighboring beads can be varied over a small distance between the infinitely high potential barriers at $(1-\delta)\sigma$ and $(1+\delta)\sigma$. This method for creating a flexible bond length was introduced by Bellemans et al. (1980). The flexible bond-length parameter is set to $\delta = 0.1$ through the paper.

The assembly and folding kinetics of the β-sheet complex are simulated using the discontinuous molecular dynamics (DMD) algorithm (Alder and Wainwright, 1959; Rapaport, 1978; Smith et al., 1996). This algorithm is very efficient for calculating the properties of systems containing hard or square-well spheres or chains (Smith et al., 1996; Zhou et al., 1997) and is significantly faster than standard molecular dynamics on Lennard-Jones chains. The algorithm proceeds by searching for the next collision time and collision pair, advancing all beads in time to the next collision event, and then calculating the velocity change of the colliding pair. During any collision process, the total energy of the system is conserved, so that the change in kinetic energy is equivalent to the negative of the change of potential energy. The DMD simulations are conducted in the canonical ensemble with constant number of particles, volume, and temperature; periodic boundary conditions are applied in all directions. The length of the cubic box is twice the length of a fully extended chain, preventing that chain from interacting with more than one image of another chain at one time. To maintain constant temperature, the Andersen stochastic collision method (Anderson, 1980) is used. In this method the system's particles collide with imaginary or ghost particles that serve as an effective heat bath. In the collision with the imaginary particles, the velocity of each bead is reassigned from a Maxwell-Boltzmann distribution at the simulation temperature. A reduced ghost particle density of $n_g^* \equiv n_g\sigma^3 = 0.1$, where n_g is the number density of ghost particles, was used. As a result, 2% of the collisions were ghost particle collisions. Because the value of the ghost particle density can alter the kinetic results, we use the same value of the ghost particle density for all kinetic simulations.

The system is quenched to the temperature of interest from the denatured or random coil state equilibrated at a high temperature. The initial configurations for the kinetic simulations were generated from equilibrium simulations of the random coil state at $T^* = 2.0$. The temperature used in this paper is a dimensionless reduced temperature, $T^* \equiv k_BT/\varepsilon$, where k_B is Bolzmann's constant, T is the absolute temperature, and ε is the energy parameter. The coordinates and velocities of the beads in the random coil state were obtained every 10^5 collisions during the equilibrium simulation to create independent initial conformations for the quenching studies. This large collision interval ensures that the sampled configurations are independent. After quenching, the four random coils undergo a conformational change toward an energy minimum, i.e., the highly ordered β-sheet complex (native state) or a local energy minimum, i.e., a misfolded state. All chains are identical, so that any chain can fold into any location within the β-sheet complex. We investigated the time series of folding behavior from the random coil state that was generated at $T^* = 2.0$ to the native state,

keeping the temperature constant at $T^* = 0.2$. Our previous thermodynamic study on the isolated β-sheet models (Jang et al., 2002a) showed that at $T^* = 2.0$ and $T^* = 0.2$, the β-sheet peptide is in a random coil state and the native state, respectively.

Simulations were performed for intermolecular contact parameters in the range of $\eta = 0.2$ to 1.0 at a bias gap $g = 0.9$. The intermediate value of the bias gap $g = 0.9$ is thought to be the most representative of a real protein in its equilibrium and dynamic behavior (Jang et al., 2002a,b). At this bias gap, all nonbonded pairs of beads are attracted with square-well depths $B_N^{intra}\varepsilon = -1$, $B_O\varepsilon = -0.1$, and $-0.2 < B_N^{inter}\varepsilon < -1$ for the intramolecular native contact, nonnative contact, and intermolecular native contact pairs, respectively. For each value of η, a number of quenching simulations were performed over at least 50 independent trajectories that start from different initial configurations. The total time elapsed during a simulation of a given number of collisions strongly depended on the size of the intermolecular contact parameter η. The time used in the simulations is expressed in terms of reduced time (Zhou and Karplus, 1999), $t^* = t\sqrt{\varepsilon/m\sigma^2}$, where t is the real time, ε is the energy parameter, m is the mass of the bead, and σ is the bead diameter. For all values of η, the simulations were performed for 4×10^8 collisions, which is approximately equivalent to $t^* = 3.4 \times 10^5$, 1.9×10^5, and 1.6×10^5 for $\eta = 0.2$, 0.5, and 0.8, respectively. This indicates that the simulation speed for large η is faster than for small η.

The equilibrium properties of the β-sheet complex were also investigated once oligomerization was achieved. Oligomers, whether folded or misfolded, were typically seen after 1×10^8 collisions for all values of η, indicating that the β-sheets assembled into the oligomer early in the simulation and were stable once they were assembled. Equilibrium averages were taken by accumulating the simulation data over the last quarter of the collisions (i.e., after 3×10^8 collisions) to ensure system equilibration. Sampling was taken over the folding or misfolding trajectories separately. For the folding trajectories, the sampled equilibrium ensembles are well within a Gaussian distribution with a narrow width, because the folded oligomers are structurally similar to each other. For the misfolding trajectories, the sampled equilibrium ensembles are also well within a Gaussian distribution with a wide width, because the misfolded oligomers are structurally slightly different from each other but the majority of the misfolding trajectories end up in some common misfolded state.

The progression of a conformation toward the native state was monitored by introducing the time-dependent fraction of total native contacts formed (Lazaridis and Karplus, 1997; Šali et al., 1994), $Q_{total}(t)$, defined by

$$Q_{total}(t) = \frac{N_{total}(t)}{N_{total}^{gem}}, \tag{6}$$

where $N_{total}(t)$ is the total number of native contacts at a given time, with $N_{total}(t) \equiv N_{intra}(t) + N_{inter}(t)$, and N_{total}^{gem} is the total number of native contacts in the global energy minimum state, with $N_{total}^{gem} \equiv N_{intra}^{gem} + N_{inter}^{gem}$. The fraction of total native contacts is a weighted sum of the fraction of intramolecular native contacts, $Q_{intra}(t)$, and the fraction of intermolecular native contacts, $Q_{inter}(t)$, by

$$Q_{total}(t) = (1 - \kappa)Q_{intra}(t) + \kappa Q_{inter}(t), \tag{7}$$

where $Q_{intra}(t) \equiv N_{intra}(t)/N_{intra}^{gem}$, $Q_{inter}(t) \equiv N_{inter}(t)/N_{inter}^{gem}$, and κ is a constant, $\kappa = N_{inter}^{gem}/N_{total}^{gem} = 0.6350$. The assembly of the complex was monitored by introducing the time-dependent fraction of intermolecular native contacts for the nth chain, $Q_n^{inter}(t)$, defined by

$$Q_n^{inter}(t) = \frac{N_{n\,inter}(t)}{N_{n\,inter}^{gem}}, \tag{8}$$

where $N_{n\,inter}(t)$ is the number of intermolecular native contacts for the nth chain at a given time. The value used for the denominator in Eq. 8 is that for the interior β-sheets at $n = 2$ or 3, because the β-sheet has a maximum number of intermolecular native contacts when it is located in the interior of the β-sheet complex (see Table 1).

To determine the separation between the chains before they assemble, the size of the entire system after oligomerization, and the compactness of the individual chains during the simulations, the time-dependent squared radius of gyration for the entire system, $R_g^2(t)$, and for the nth chain, $R_{g\,n}^2(t)$, were monitored where

$$R_g^2(t) \equiv \frac{1}{M}\sum_{i=1}^{M} \mathbf{r}_i^2(t), \tag{9}$$

and

$$R_{g\,n}^2(t) \equiv \frac{1}{M_n}\sum_{k=1}^{M_n} \mathbf{r}_{nk}^2(t), \tag{10}$$

with M_n equal to the number of beads for each β-sheet monomer. The time-dependent bead positions, $\mathbf{r}_i(t)$ ($i = 1$ to M), were obtained by shifting the center of mass coordinates of the whole system to the origin. Likewise, the time-dependent bead positions for the nth chain, $\mathbf{r}_{nk}(t)$ ($k = 1$ to M_n), were obtained by shifting the center of mass coordinates of each chain to the origin. To determine the degree of secondary structure formation in the β-sheet complex, the time-dependent squared radius of gyration for the nth chain, $R_{g\,n}^2(t)$, is averaged over the N chains, where N is the total number of the chains, i.e., $N = 4$, to give the mean squared radius of gyration for the individual chains, $\langle R_{g\,n}^2(t)\rangle_n$, defined by

$$\langle R_{g\,n}^2(t)\rangle_n \equiv \frac{1}{N}\sum_{n=1}^{N} R_{g\,n}^2(t) = \frac{1}{M}\sum_{n=1}^{N}\sum_{k=1}^{M_n} \mathbf{r}_{nk}^2(t). \tag{11}$$

We have studied the equilibrium properties of the β-sheet complex after oligomerization. The equilibrium averages were determined by accumulating the last one-quarter of the quenching simulation data, because a β-sheet complex occurs early in the simulation and is stable once it is assembled. The reduced internal energy can be calculated from

$$E^* \equiv \frac{\langle E\rangle_{cf}}{\varepsilon}, \tag{12}$$

where $\langle\ \rangle_{cf}$ denotes the configurational average over the last one-quarter of the total collisions. The degree of bead mobility can be characterized by the root-mean-squared (rms) pair separation fluctuation, Δ_B, defined by

$$\Delta_B \equiv \frac{2}{M(M-1)}\sum_{i<j}\left(\frac{\langle r_{ij}^2\rangle_{cf}}{\langle r_{ij}\rangle_{cf}^2} - 1\right)^{1/2}, \tag{13}$$

where r_{ij} denotes the separation between beads i and j in the β-sheet complex. To characterize the phase of the β-sheet complex, the Lindemann disorder parameter for bead i, Δ_{Li}, is defined to be the root-mean-squared fluctuation for bead i,

$$\Delta_{Li} \equiv \left(\frac{\langle(|\mathbf{r}_i| - \langle|\mathbf{r}_i|\rangle_{cf})^2\rangle_{cf}}{\sigma^2}\right)^{1/2}, \tag{14}$$

and the Lindemann disorder parameter, Δ_L, is defined to be

$$\Delta_L = \left(\sum_i \frac{\Delta_{Li}^2}{M}\right)^{1/2}, \tag{15}$$

where $\mathbf{r}_i$ is the position of bead i ($i = 1$ to M) and $(|\mathbf{r}_i| - \langle|\mathbf{r}_i|\rangle_{cf})^2$ is the mean-squared fluctuation. Because the entire structure of the assembled β-sheet complex continuously translates and rotates over space even in equilibrium, this causes a large value of the root-mean-squared deviation. In the calculation of the mean-squared fluctuation, the translations are removed by shifting the system's center of mass coordinates to the origin. The rotations are removed by rotating the system so as to minimize the root-mean-squared deviation of the new coordinates from the coordinates set at

the beginning of the calculation (Kabsch, 1976). The Lindemann disorder parameter is often used to analyze the liquid-to-solid transition. A form of the Lindemann criterion (Lindemann, 1910) for melting is applied to the systems here to determine if they are in the solid phase or the liquid phase. A system is regarded as a solid if its Lindemann disorder parameter, Δ_L, is in the range of 0.1–0.15, whereas a system with $\Delta_L > 0.15$ is considered to be liquid (Bilgram, 1987; Löwen, 1994; Stillinger, 1995; Zhou et al., 1999).

RESULTS

We performed DMD simulations to investigate the equilibrium properties and the folding kinetics of the tetrameric β-sheet complex. Four separate random coils were quenched from a temperature $T^* = 2.0$ to $T^* = 0.2$, the temperature at which the peptides are in their β-sheet native state. After quenching, the four monomer chains undergo a conformational change toward the highly ordered β-sheet complex (folded state) or to a partially folded or disordered (misfolded) state. The equilibrium properties of the assembled systems were investigated as a function of the intermolecular contact parameter η, in the range $0.2 \leq \eta \leq 10$. At each value of η, at least 50 independent trajectories starting from different initial configurations were monitored and used in the averaging. All simulations were performed at a bias gap $g = 0.9$.

The dependence of the folding yield on the intermolecular contact parameter, η, is shown in Fig. 2. The folding yield is defined to be the ratio of the number of trajectories leading to the folded state and the total number of trajectories at each value of η. The folded and misfolded states were determined by visualizing the assembled tetramer structures with graphic tools and by measuring the Q_n^{inter} values in the tetramer state. In particular, the values of Q_n^{inter} for the two inner chains in the folded state should be twice the size of Q_n^{inter} for the two outer chains as was seen in the global energy minimum state. It can be clearly seen that the folding yield is high at intermediate values of η ($\eta = 0.4$, 0.5, 0.6, and 0.7), with a maximum at $\eta = 0.5$. At these conditions, a large number of the trajectories fold into the highly ordered β-sheet complex. However, at small ($\eta = 0.2$ and 0.3) and large ($\eta =$ 0.8, 0.9, and 1.0) values of η, the folding yield is low, indicating that only a few trajectories successfully fold into the highly ordered β-sheet complex. The minimum value of the folding yield occurs at $\eta = 1.0$.

Equilibrium properties of the β-sheet complex

The equilibrium properties of folded and misfolded systems were determined after oligomerization. Fig. 3 shows: a), the fraction of total native contacts, $\langle Q_{total} \rangle$; b), the fractions of intramolecular, $\langle Q_{intra} \rangle$, and intermolecular, $\langle Q_{inter} \rangle$, native contacts; and c), the reduced internal energy per bead, $\langle E^*/M \rangle$, as a function of η for trajectories ending in the folded (*solid symbols*) and misfolded (*open symbols*) states. Here, $\langle\ \rangle$ denotes the average over the folding or misfolding trajectories, separately. The error bars are the standard deviation in the measured values and are only shown when they are larger than the size of the symbol. In Fig. 3 *a*, the $\langle Q_{total} \rangle$ values increase as η increases both in the folded and the misfolded states. At the same value of η, the $\langle Q_{total} \rangle$ value for the folded state is larger than for the misfolded state. The large errors in $\langle Q_{total} \rangle$ for the misfolded state indicate that a number of different structures are possible. The small errors in $\langle Q_{total} \rangle$ for the folded state reflect the fact that this structure is unique. Fig. 3 *b* shows $\langle Q_{intra} \rangle$ and $\langle Q_{inter} \rangle$ as a function of η; $\langle Q_{intra} \rangle$ measures the formation of secondary structure and $\langle Q_{inter} \rangle$ measures the oligomerization of the four monomers. For the folded state, $\langle Q_{intra} \rangle >$ 0.95 for all values of η, indicating that each monomer in the ordered complex is in its native state. As η increases, $\langle Q_{inter} \rangle$ in the folded state increases rapidly at low η ($\eta = 0.2$ and 0.3), increases continuously at intermediate η ($\eta = 0.4 \sim$ 0.7), and then is stable at large η ($\eta = 0.8 \sim 1.0$). Note that for small η, $\langle Q_{inter} \rangle$ in the folded state is very low, whereas $\langle Q_{intra} \rangle$ in the folded state is very high. This suggests that the ordered β-sheet complex for small η is loosely packed but has a well-defined secondary structure. For large η, the ordered β-sheet complex holds together tightly due to the strong attractions between the β-sheets, resulting in large values of $\langle Q_{inter} \rangle$. For the misfolded state, however, $\langle Q_{intra} \rangle$ takes a similar value to that in the folded state at small and intermediate η, but decreases gradually as η increases. This indicates that at small and intermediate η, each monomer is in the native state, but that at large η, the monomers are not in the native state. The quantity, $\langle Q_{inter} \rangle$ in the misfolded state increases continuously as η increases. The low values of $\langle Q_{inter} \rangle$ and the high values of $\langle Q_{intra} \rangle$ in the misfolded state for small and intermediate η, indicate that the misfolded state contains nonaligned β-sheets. However for large η, the misfolded state exhibits a larger value of $\langle Q_{inter} \rangle$ than that of $\langle Q_{intra} \rangle$, suggesting that it contains tangled chains with poor secondary structure.

The system's energy is a direct reflection of the total number of native contacts. Fig. 3 *c* shows the reduced internal energy per bead as a function of η for the folded and the misfolded states, respectively. As expected, the energy of the system decreases as η increases both for the folded and the misfolded states. The energy difference between the two states is infinitesimal for small η, but increases as η increases. This can be understood by referring to Fig. 3 *a*, which shows that the difference in the $\langle Q_{total} \rangle$ values between the two states increases as η increases. Note that for very large η, the energy difference between the two states decreases, because the energy in the misfolded state decreases due to contributions from a large number of intermolecular nonnative contacts that result from the strong intermolecular nonnative attraction.

For the β-sheet complex, the size of the system after oligomerization and the compactness of the individual chains are determined by measuring the radius of gyration. The reduced squared radius of gyration per bead for the entire

system, $\langle R_g^2/\sigma^2 M\rangle$, the reduced mean squared radius of gyration per bead for the individual chains, $\langle\langle R_{gn}^2/\sigma^2 M_n\rangle_n\rangle$, and the root-mean-squared (rms) pair separation fluctuation, $\langle\Delta_B\rangle$, are shown in Fig. 4 as a function of η for the folded (*solid symbols*) and the misfolded (*open symbols*) states. Here, $\langle\ \rangle$ denotes the average over the folding and misfolding trajectories, separately, and $\langle\ \rangle_n$ denotes the average over the n chains. For the folded state, the $\langle R_g^2/\sigma^2 M\rangle$ value in Fig. 4 *a*, which characterizes the system size, is small at $\eta = 0.2$ and 0.3, abruptly increases at $\eta = 0.4$, and then remains nearly the same for intermediate and large values of η. For small η, the $\langle R_g^2/\sigma^2 M\rangle$ value is small because the R_g^2 value of the β-sheet monomer is small. This can be seen from Fig. 4 *b*, which shows that the mean R_g^2 value for the individual β-sheet chains, $\langle\langle R_{gn}^2/\sigma^2 M_n\rangle_n\rangle$, is smallest at $\eta = 0.2$ and 0.3 and then rapidly increases at $\eta = 0.4$. The low values of $\langle\langle R_{gn}^2/\sigma^2 M_n\rangle_n\rangle$ occur at small η, because, at these conditions the β-sheet strands are not fully elongated even in the highly ordered state. Fig. 4 *c* shows the root-mean-squared (rms) pair separation fluctuation, $\langle\Delta_B\rangle$, which measures the degree of bead fluctuations. The $\langle\Delta_B\rangle$ value is very high at the smaller η, gradually decreases as η increases, and is stable at the larger η. This suggests that the β-sheet strands are not fully stiff at the smaller η due to the strong bead fluctuations. As η increases, the intermolecular forces get stronger, preventing bead fluctuations and squeezing the β-sheets together. The β-sheet strands become fully elongated and each β-sheet monomer in the complex structure recovers its intrinsically long molecular shape as seen in the global energy minimum structure in Fig. 1. This causes the R_g^2 values of the whole system and of the individual chains to increase.

For the misfolded state, the behavior of $\langle R_g^2/\sigma^2 M\rangle$ is different from that in the folded state as seen in Fig. 4 *a*. For the smaller η, the $\langle R_g^2/\sigma^2 M\rangle$ value in the misfolded state is larger than that in the folded state. This suggests that the misfolded state at small η contains two nonaligned dimers or a trimer with a monomer (see below), and hence has a large R_g^2 value for the entire system. However for the smaller η, the mean R_g^2 value for the individual chains in the misfolded state is smaller than that in the folded state as seen in Fig. 4 *b*. This

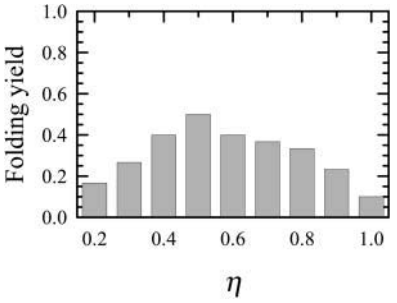

FIGURE 2 Folding yield as a function of the intermolecular contact parameter, η.

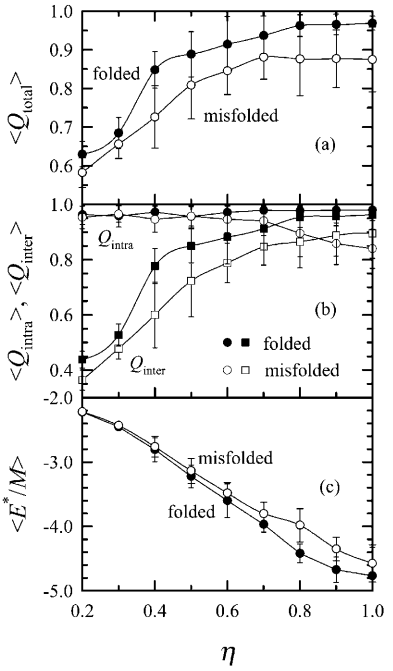

FIGURE 3 (*a*) The fraction of total native contacts, $\langle Q_{total}\rangle$, (*b*) the fractions of intramolecular native contacts, $\langle Q_{intra}\rangle$, and intermolecular native contacts, $\langle Q_{inter}\rangle$, and (*c*) the reduced internal energy per bead, $\langle E^*/M\rangle$, for the β-sheet complex as a function of η for the folded (*solid symbols*) and misfolded (*open symbols*) states.

indicates that in an ordered tetramer, the mean R_g^2 value for the individual chains is larger than in a dimer or a trimer because the interior β-sheets in the ordered tetramer are constrained to be elongated by outer β-sheets. At small η, the two nonaligned dimers or a trimer with a monomer in the misfolded state have large fluctuations in bead motions because the systems have many open surfaces, resulting in large values of $\langle\Delta_B\rangle$ as seen in Fig. 4 *c*. For larger η, the $\langle R_g^2/\sigma^2 M\rangle$ value in the misfolded state decreases as η increases as seen in Fig. 4 *a*. This can be explained by referring to Fig. 4 *b*, which shows that the $\langle\langle R_{gn}^2/\sigma^2 M_n\rangle_n\rangle$ value in the misfolded state slightly decreases at the larger η. The misfolded state at large η contains four tangled chains with poor secondary structure as indicated in Fig. 3 *b* by the small value of $\langle Q_{intra}\rangle$ in the misfolded state.

Folding kinetics of the β-sheet complex

The folding kinetics of the β-sheet complex was investigated. The time-dependent fraction of intermolecular native contacts, $\langle Q_{inter}(t)\rangle$, is shown in Fig. 5 as a function of the reduced time, t^*, at $\eta = 0.2$, 0.5, and 0.8 for trajectories that end in: a), the folded state, and b), the misfolded state. Here, $\langle\ \rangle$ denotes the average calculated by sampling every 1000 collisions before 1×10^7 collisions, and every 100,000 collisions after 1×10^7 collisions over the folding and misfolding trajectories separately. To highlight the assembly

events that occur early in the simulation, a logarithmic timescale is used in the figures to display early assembly behavior of the chains. The simulation results are presented for 4×10^8 collisions, which is approximately equivalent to $t^* = 3.4 \times 10^5$, 1.9×10^5, and 1.6×10^5 for $\eta = 0.2$, 0.5, and 0.8, respectively. To document the assembly process of the chains, the figure presents only $\langle Q_{\text{inter}} \rangle$ values for clarity. It is immediately apparent that the $\langle Q_{\text{inter}} \rangle$ values for both the folded and misfolded states begin to increase at $t^* > 100$. In contrast, the $\langle Q_{\text{intra}} \rangle$ values (not shown in figure) are saturated at the native value by $t^* > 100$. This is consistent with our previous kinetic study for isolated chain systems (Jang et al., 2002b), which showed that at $g = 0.9$ the β-sheet is a fast folding protein and its native state is observed by $t^* \geq 100$. At low η, the assembly process for four chains goes through two distinct stages. In the first stage, the initial random coils collapse to the native β-sheet at $t^* \sim 100$, and in the second stage, the four β-sheets assemble at $t^* > 100$. However, for large η, the individual chains are more likely to collapse and assemble concurrently. For the folded state in Fig. 4 *a*, the $\langle Q_{\text{inter}} \rangle$ values increase in a series of steps after $t^* = 100$, whereas for the misfolded state in Fig. 4 *b*, the $\langle Q_{\text{inter}} \rangle$ values increase monotonically after $t^* = 100$ for all values of η. The steps in the increase in $\langle Q_{\text{inter}} \rangle$ for the folded state correspond to assembly events in which the β-sheets become aligned in parallel facing each other. This causes an increase in the number of intermolecular native contacts and hence a stepwise increase in $\langle Q_{\text{inter}} \rangle$. In contrast, the continuous increase in $\langle Q_{\text{inter}} \rangle$ for the misfolded state indicates that the β-sheets are assembled but not well aligned.

Fig. 6 shows the time-dependent reduced mean squared radius of gyration per bead for the individual chains, $\langle \langle R_{\text{gn}}^2/\sigma^2 M_{\text{n}} \rangle_{\text{n}}(t) \rangle$, as a function of the reduced time, t^*, at $\eta = 0.2$, 0.5, and 0.8 for: a), the folded state, and b), the misfolded state. The quantity, $\langle \langle R_{\text{gn}}^2/\sigma^2 M_{\text{n}} \rangle_{\text{n}} \rangle$ measures the formation of a chain's secondary structure. In can be seen that the $\langle \langle R_{\text{gn}}^2/\sigma^2 M_{\text{n}} \rangle_{\text{n}} \rangle$ values for both the folded and misfolded states rapidly decrease as the random coils collapse at $t^* \sim 100$ and then gradually increase as the chains move toward the β-sheet complex. At $t^* = 100$, the $\langle \langle R_{\text{gn}}^2/\sigma^2 M_{\text{n}} \rangle_{\text{n}} \rangle$ for all values of η for both the folded and the misfolded states are smaller than their global energy minima in Table 1. In fact the native β-sheet monomer at $T^* = 0.2$ is slightly bent, yielding a small R_{g}^2 value, and differs from the global energy minimum structure, which has highly stiff strands (Jang et al., 2002a). At $t^* > 100$, the $\langle \langle R_{\text{gn}}^2/\sigma^2 M_{\text{n}} \rangle_{\text{n}} \rangle$ values for both the folded and misfolded states increase because the strands become elongated whenever the β-sheets associate into dimer, trimer, or tetramer states. In these cases, the β-sheets squeeze against each other due to the large number of intermolecular native contacts between the sheets. For $\eta = 0.2$, the shapes of the $\langle \langle R_{\text{gn}}^2/\sigma^2 M_{\text{n}} \rangle_{\text{n}} \rangle$ versus t^* curves in both the folded and misfolded states are similar. However, for $\eta = 0.5$ and 0.8, the shapes of the $\langle \langle R_{\text{gn}}^2/\sigma^2 M_{\text{n}} \rangle_{\text{n}} \rangle$ versus t^* curve in the folded state are different from that in the misfolded state. There is a stepwise increase in $\langle \langle R_{\text{gn}}^2/\sigma^2 M_{\text{n}} \rangle_{\text{n}} \rangle$ for the folded state due to the parallel β-sheets aligning face to face as was seen for the $\langle Q_{\text{inter}} \rangle$ curves in Fig. 5 *a*.

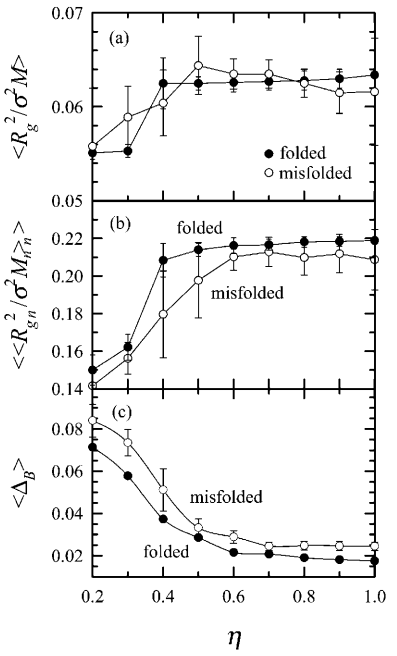

FIGURE 4 (*a*) The reduced squared radius of gyration per bead, $\langle R_{\text{g}}^2/\sigma^2 M \rangle$, (*b*) the reduced mean squared radius of gyration per bead for the individual chains, $\langle \langle R_{\text{gn}}^2/\sigma^2 M_{\text{n}} \rangle_{\text{n}} \rangle$, and (*c*) the root-mean-squared (rms) pair separation fluctuation, $\langle \Delta_{\text{B}} \rangle$, for the β-sheet complex as a function of η for the folded (*solid symbols*) and misfolded (*open symbols*) states.

The η = 0.2 model

It is of interest to examine the kinetics of β-sheet complex formation in more detail. For $\eta = 0.2$, the evolution to the ordered β-sheet complex is reflected in the snapshots in Fig. 7 taken at selected times that display typical assembly events on the way to the folded state. The individual chains are shown with different colors (*blue* (chain 1), *red* (chain 2), *green* (chain 3), and *yellow* (chain 4)). The first snapshot at $t^* = 0$ shows the random initial configuration of the system. At $t^* = 98$, the four monomers collapse to their native states, but intermolecular contacts are not yet observed. Two typical kinetic intermediates of the system are shown in the next two snapshots. At $t^* = 976$, the intermediate I_{2+1+1} is observed with two chains paired into a dimer and two monomers, and at $t^* = 7.13 \times 10^3$, the intermediate I_{3+1} is shown with another chain joined to the dimer to form a trimer. At $t^* = 2.12 \times 10^4$, there is a partially ordered tetramer and at $t^* = 1.72 \times 10^5$, the ordered β-sheet complex is observed.

Fig. 8 shows a more quantitative description of the folding trajectory depicted in Fig. 7. The time-dependent fractions of intramolecular and intermolecular native contacts, $Q_{\text{intra}}(t)$

and $Q_{\text{inter}}(t)$, respectively, are shown in Fig. 8 *a* for $\eta = 0.2$. At $t^* \sim 200$, $Q_{\text{intra}} \sim 1$, indicating that the random coils quickly fold into the native state, while the Q_{inter} values have only just begun to increase, indicating that the intermolecular native contacts between the β-sheets form later than the intramolecular native contacts. The stepwise increase in Q_{inter} corresponds to the assembly events associated with the formations of dimeric and trimeric intermediates of the system during the simulation. The time-dependent fraction of intermolecular native contacts for the *n*th chain, $Q_{\text{n}}^{\text{inter}}(t)$, is shown in Fig. 8 *b*. The lines on the figure correspond to the various colored chains in Fig. 7 as follows: chain 1 (*solid black line*—blue chain), chain 2 (*dashed black line*—red chain), chain 3 (*dark gray line*—green chain), and chain 4 (*light gray line*—yellow chain). The $Q_{\text{n}}^{\text{inter}}$ value for the interior β-sheets (*blue* and *yellow*) is larger than that for the exterior β-sheets (*red* and *green*), because the interior β-sheets have a larger number of intermolecular native contacts than the exterior β-sheets. At $t^* \sim 200$, Q_1^{inter} (*blue*) and Q_3^{inter} (*green*) begin to increase because chains 1 (*blue*) and 3 (*green*) begin to assemble into a dimer. At $t^* \sim 2.3 \times 10^3$, Q_4^{inter} (*yellow*) starts to increase and Q_1^{inter} (*blue*) increases further, while Q_3^{inter} (*green*) remains the same. This indicates that chain 4 (*yellow*) joins the dimer to form a trimer. In this case, chain 4 (*yellow*) interacts with the chain 1 (*blue*) so that chain 1 is sandwiched between two chains (*yellow* and *green*) in the trimer state, causing both the yellow and blue lines in Fig. 8 *b* to increase. As the simulation proceeds, the trimer and finally the last monomer (*red* chain) slowly maneuver in space until they finally fuse to form a tetramer at $t^* = 2.12 \times 10^4$. Chain 2 (*red*) aligns with chain 4 (*yellow*) as can be seen by the increase in Q_2^{inter} (*red*) and Q_4^{inter} (*yellow*). In the well ordered tetrameric β-sheet complex, the two inner chains (chains1 (*blue*) and 4 (*yellow*)) have a larger value of $Q_{\text{n}}^{\text{inter}}$ than the two outer chains (chains 2 (*red*) and 3 (*green*)). The β-sheet complex is very stable and stays in the tetramer state once it has been formed, although strong fluctuations in the Q values are observed.

The separation of the chains changes dramatically as the four monomers assemble to form a tetramer. Fig. 8 *c* shows the time-dependent reduced squared radius of gyration per bead, $R_{\text{g}}^2/\sigma^2 M\ (t)$, as a function of the reduced time, t^*, for the trajectory shown in Fig. 7. The quantity, $R_{\text{g}}^2/\sigma^2 M$ shows how the chains are distributed in space when the system is in the random coil and intermediates states. In the tetramer state the $R_{\text{g}}^2/\sigma^2 M$ measures the size of the β-sheet complex. The system containing four random coils becomes the intermediate I_{2+1+1} at $t^* \sim 200$, consisting of a dimer and two monomers. The dimer and two monomers separate and maneuver in space, causing a large $R_{\text{g}}^2/\sigma^2 M$ value until just before the trimer formation at $t^* \sim 2.3 \times 10^3$. In the intermediate I_{3+1}, the chains' separation decreases as one monomer aligns with the dimer to form a trimer. The

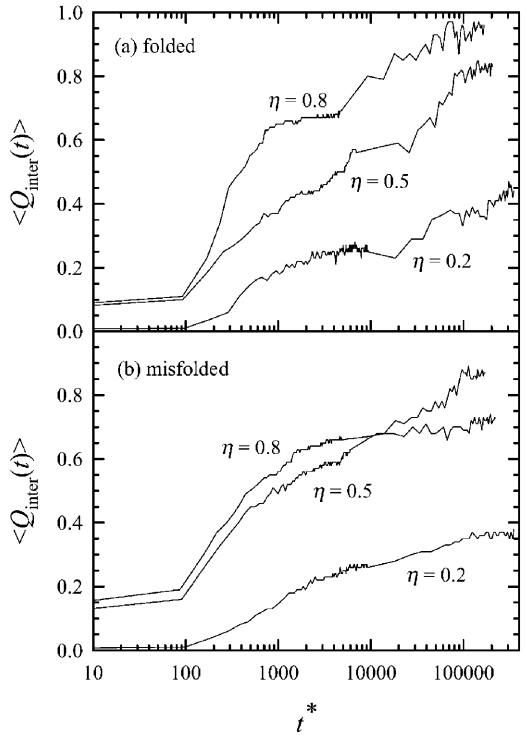

FIGURE 5 The time-dependent fraction of intermolecular native contacts, $\langle Q_{\text{inter}}(t)\rangle$, as a function of the reduced time, t^*, at $\eta = 0.2$, 0.5, and 0.8 (*a*) for the folded state and (*b*) for the misfolded state.

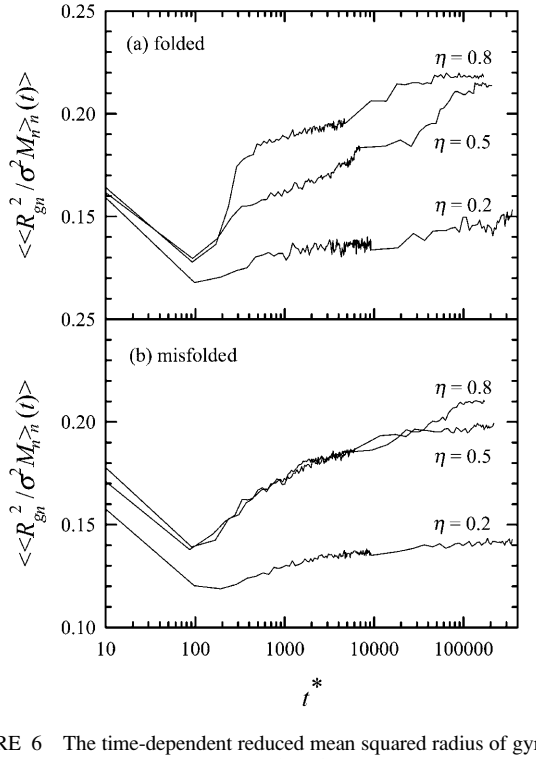

FIGURE 6 The time-dependent reduced mean squared radius of gyration per bead for the individual chains, $\langle\langle R_{\text{g}n}^2/\sigma^2 M_n\rangle_n(t)\rangle$, as a function of the reduced time, t^*, at $\eta = 0.2$, 0.5, and 0.8 (*a*) for the folded state and (*b*) for the misfolded state.

separation of chains dramatically decreases at $t^* = 2.12 \times 10^4$ when the last monomer finally aligns with the trimer to form the tetramer. The time-dependent reduced squared radius of gyration for the nth chain, $R^2_{gn}/\sigma^2 M_n$, is shown in Fig. 8 *d*. The initial random coils collapse to the native state at $t^* \sim 100$. As the simulation proceeds, the $R^2_{gn}/\sigma^2 M_n$ values for each β-sheet change slightly when the chains become involved in the system intermediates. This can be seen by the fact that $(R^2_{g1}/\sigma^2 M_1(blue) = R^2_{g3}/\sigma^2 M_3(green)) > (R^2_{g2}/\sigma^2 M_2(red) = R^2_{g4}/\sigma^2 M_4(yellow))$ in the I_{2+1+1} state, $R^2_{g1}/\sigma^2 M_1(blue) > (R^2_{g3}/\sigma^2 M_3(green) = R^2_{g4}/\sigma^2 M_4(yellow)) > R^2_{g2}/\sigma^2 M_2(red)$ in the I_{3+1} state, and $(R^2_{g1}/\sigma^2 M_1(blue) = R^2_{g4}/\sigma^2 M_4(yellow)) > (R^2_{g2}/\sigma^2 M_2(red) = R^2_{g3}/\sigma^2 M_3(green))$ in the folded state. Large fluctuations in $R^2_{gn}/\sigma^2 M_n$ indicate that the β-sheet complex at $\eta = 0.2$ is loosely packed, allowing the beads to move in the direction perpendicular to the β-sheet molecular plane.

To characterize the phase of the folded β-sheet complex shown in the last snapshot in Fig. 7, the Lindemann disorder parameter (Lindemann, 1910) is calculated. This parameter is often used to analyze the liquid-to-solid transition (Bilgram, 1987; Löwen, 1994; Stillinger, 1995; Zhou et al., 1999). Fig. 8 *e* shows the Lindemann disorder parameter for the individual beads i, Δ_{Li}, as a function of the reduced distance of bead i from the origin, $|\mathbf{r}_i|/\sigma$. A form of the Lindemann criterion for melting is adopted here in which systems with $\Delta_L < 0.15$ are considered solid-like, whereas systems with $\Delta_L > 0.15$ are considered liquid-like. The Δ_L value, a sum of the average root-mean-squared fluctuations over all beads, can be obtained from Eq. 15. The dotted line in Fig. 8 *e* indicates the criterion for the melting transition, and the solid and open symbols represent the interior and exterior β-sheets in the complex, respectively. In Fig. 8 *e*, it can be seen that the interior of the folded β-sheet complex is solid-like with $\Delta_{Li} < 0.15$, but the surface is liquid-like with $\Delta_{Li} > 0.15$. However, the entire complex is liquid-like, because the Lindemann disorder parameter, Δ_L, is 0.16, slightly larger than the criterion value.

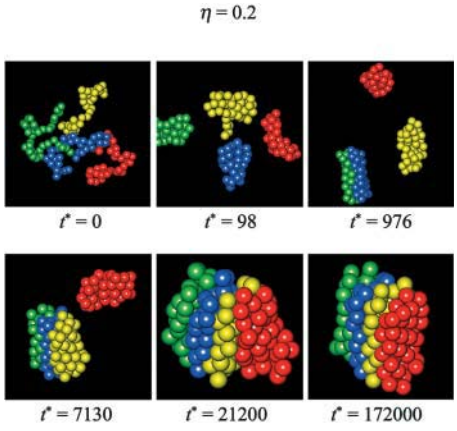

FIGURE 7 Snapshots showing kinetic β-sheet complex formation for $\eta = 0.2$ at selected reduced times, $t^* = 0$, 98, 976, 7.13×10^3, 2.12×10^4, and 1.72×10^5.

The $\eta = 0.5$ model

For $\eta = 0.5$, Fig. 9 shows snapshots representing a typical folding trajectory that involves the random coil state, two typical kinetic intermediates, and the highly ordered β-sheet complex. At $t^* = 8$, the snapshot shows the random initial configuration of the system. At $t^* = 169$ and 745, two intermediates, I_{2+1+1} (dimer and two monomers) and I_{2+2} (two dimers) are observed, respectively. Docking of the two dimers is shown in the snapshots at $t^* = 940$ and 1008. The snapshot at $t^* = 2740$ represents the highly ordered β-sheet complex.

The time-dependent fractions of intramolecular and intermolecular native contacts, $Q_{intra}(t)$ and $Q_{inter}(t)$, are shown in Fig. 10 *a*, respectively, for the folding trajectory depicted in Fig. 9. At $t^* \sim 200$, Q_{intra} reaches the native value, $Q_{intra} \sim 1$, whereas at $t^* \sim 2000$, Q_{inter} reaches the native value, $Q_{inter} \sim 0.88$. The highly ordered β-sheet complex is observed at $t^* > 2000$. Comparison of Fig. 10 *a* with Fig. 8 *a* shows that the assembly process in the $\eta = 0.5$ model is significantly faster than that observed in the $\eta = 0.2$ model. Fig. 10 *b* shows the time-dependent fraction of intermolecular native contacts for the nth chain, $Q^{inter}_n(t)$. The lines on the figure correspond to the various colored chains in Fig. 9 as follows: chain 1 (*solid black line*—blue chain), chain 2 (*dashed black line*—red chain), chain 3 (*dark gray line*—green chain), and chain 4 (*light gray line*—yellow chain). At $t^* \sim 50$, chains 1 (*blue*) and 2 (*red*) start to assemble to form a dimer when they are in the partially folded state. Note that at $g = 0.9$ the native β-sheet is observed by $t^* \geq 100$. The system intermediate I_{2+1+1} moves to the intermediate I_{2+2} at $t^* \sim 300$ as another dimer (formed by the chains 3 (*green*) and 4 (*yellow*)) emerges. The two dimers are finally aligned to form the highly ordered β-sheet complex at $t^* \sim 2000$. In the final structure, the chains 2 (*red*) and 4 (*yellow*) are located inside and the chains 1 (*blue*) and 3 (*green*) are located outside of the tetrameric β-sheet complex. This structure is maintained over the remaining simulation time.

Fig. 10 *c* shows the $R^2_g/\sigma^2 M$ value, which characterizes the separation of the chains, as a function of the reduced time, t^*. Unlike the $\eta = 0.2$ curve shown in Fig. 8 *c*, the $R^2_g/\sigma^2 M$ value in the $\eta = 0.5$ model continuously decreases as the simulation proceeds until it saturates to the native value at $t^* > 1000$. The intermolecular native attraction is much stronger in the $\eta = 0.5$ than in the $\eta = 0.2$ model, preventing the separation of the chains. The constant value of $R^2_g/\sigma^2 M$ at $t^* > 1000$ indicates that the ordered tetramer is very stable in the folded state. The time-dependent reduced squared radius of gyration for the nth chain, $R^2_{gn}/\sigma^2 M_n$, is shown in Fig. 10 *d*. The initial random coils collapse at $t^* \sim 50$. The $R^2_{g1}/\sigma^2 M_1$ and $R^2_{g2}/\sigma^2 M_2$ values increase concurrent

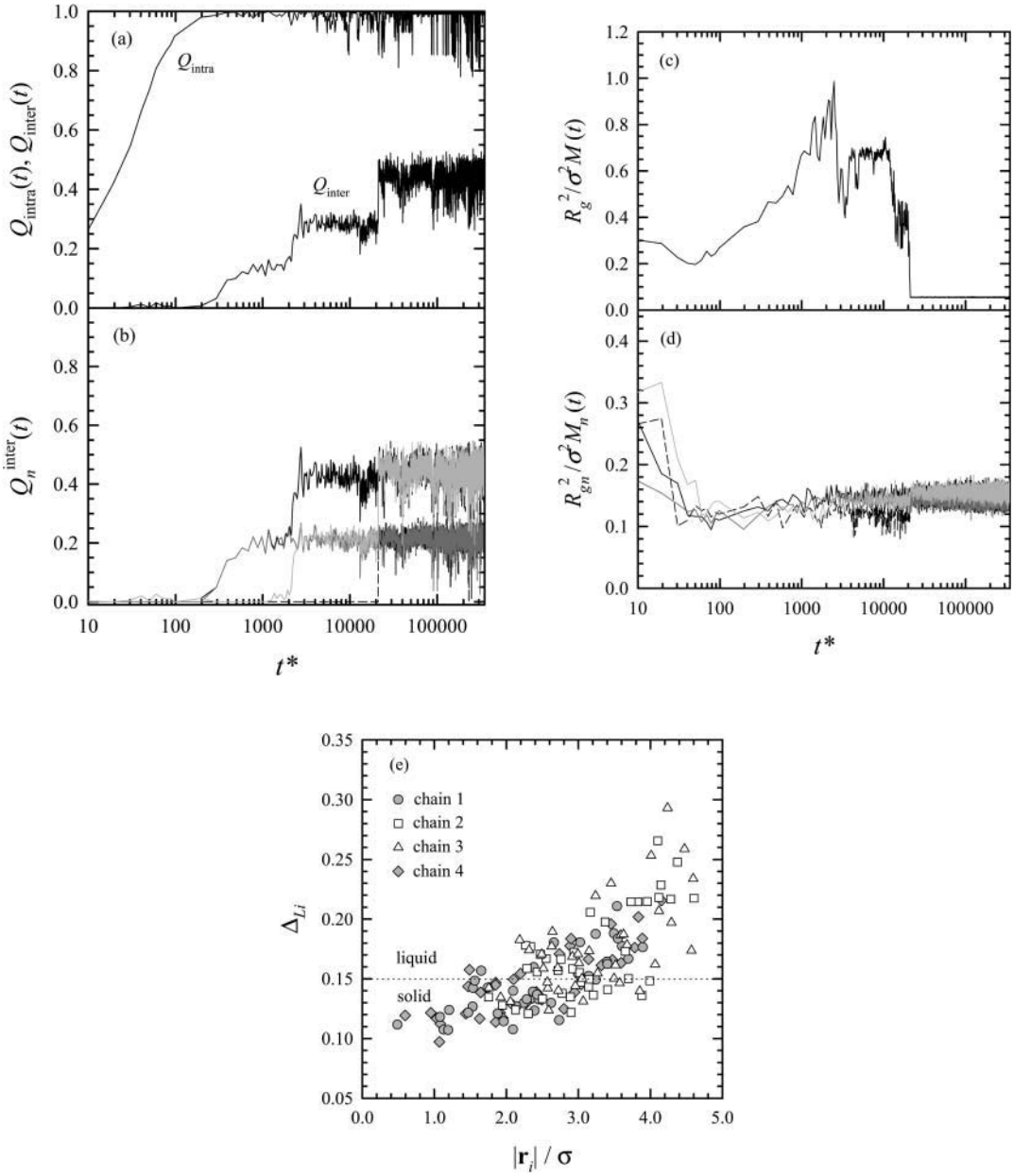

FIGURE 8 For $\eta = 0.2$ (*a*) the time-dependent fractions of intramolecular and intermolecular native contacts, $Q_{intra}(t)$ and $Q_{inter}(t)$, (*b*) the time-dependent fraction of intermolecular native contacts for the *n*th chain, $Q_n^{inter}(t)$, (*c*) the time-dependent reduced squared radius of gyration per bead, $R_g^2/\sigma^2 M$ (*t*), (*d*) the time-dependent reduced squared radius of gyration for the *n*th chain, $R_{gn}^2/\sigma^2 M_n$, all as a function of the reduced time, t^*, and (*e*) the Lindemann disorder parameter for an individual bead, Δ_{Li}, in the β-sheet complex as a function of the reduced distance of bead *i* from the origin, $|\mathbf{r}_i|/\sigma$.

with the dimer formation by the chains 1 (*blue*) and 2 (*red*). At $t^* \sim 300$, the $R_{g4}^2/\sigma^2 M_4$ value begins to increase as another dimer containing chains 3 (*green*) and 4 (*yellow*) is formed. However, the $R_{g3}^2/\sigma^2 M_3$value stays at the collapsed value until the two dimers dock at $t^* \sim 800$. This can be seen in the snapshots at $t^* = 745$ and 940 in Fig. 9, which shows that the secondary structure of chain 3 (*green*) is not well developed, whereas chain 4 (*yellow*) shows more of the β-sheet native structure, even though the two chains are each components in the dimer. The β-sheets with fully elongated strands are found in the highly ordered β-sheet complex. The $R_{gn}^2/\sigma^2 M_n$ value for each β-sheet in the complex is almost the same as that in the global energy minimum state in Fig. 1.

The bead's freedom of motion in the highly ordered β-sheet complex as shown in the snapshot at $t^* = 2740$ in Fig. 9 is investigated. Fig. 10 *e* shows the Lindemann disorder parameter for the individual beads *i*, Δ_{Li}, as a function of the reduced distance of bead *i* from the origin, $|\mathbf{r}_i|/\sigma$. It can be seen from Fig. 10 *e* that all bead motions in the interior

β-sheets (chains 2 and 4 (*solid symbols*)) of the complex are solid-like with $\Delta_{Li} < 0.15$, but in the outer β-sheets (chains 1 and 3 (*open symbols*)) they are both solid-like with $\Delta_{Li} < 0.15$ and liquid-like with $\Delta_{Li} > 0.15$, indicating that the sheet's phase strongly depends on its location. This is in marked contrast to the $\eta = 0.2$ model where the interiors of the complex were solid-like and the surfaces were liquid-like, regardless of the β-sheets' location.

The $\eta = 0.8$ model

The evolution to the highly ordered β-sheet complex for $\eta = 0.8$ is described in the snapshots in Fig. 11 for a typical folding trajectory. At $t^* = 8$, the snapshot shows the four monomers in the random coil state. At $t^* = 189$, the snapshot shows that chains start to form two dimers, which corresponds to the intermediate I_{2+2}. The next three pictures are related to the docking of the two dimers. At $t^* = 403$, one dimer (formed by chains 1 (*blue*) and 3 (*green*)) starts to contact with the other dimer (formed by chains 2 (*blue*) and 4 (*green*)), but the attraction is not energetically favorable because, as can be seen in the next two snapshots at $t^* = 580$ and 663, the first dimer flips over, trying to find a better-aligned position. At $t^* = 2080$, the highly ordered β-sheet complex is observed. For large η, the β-sheet structure of the ordered complex is the same as the global energy minimum structure in Fig. 1.

For $\eta = 0.8$, the time-dependent Q and R_g^2 values for the folding trajectory depicted in Fig. 11 are not shown in any figures for brevity, because the information contained in the time series for Q and R_g^2 is very similar to the results for the $\eta = 0.5$ model shown in Fig. 10. However for $\eta = 0.8$, the qualitative form of the bead fluctuations in the highly ordered β-sheet, as depicted in the snapshot at $t^* = 2080$ in Fig. 11, is strikingly different from that in the $\eta = 0.2$ and 0.5 models. In Fig. 12, which shows the Lindemann disorder parameter for the individual beads i, Δ_{Li}, as a function of the reduced distance of bead i from the origin, $|\mathbf{r}_i|/\sigma$, it is immediately apparent that the bead motions are frozen with all $\Delta_{Li} < 0.15$. This indicates that the folded β-sheet tetramer is in an inactive solid-like phase because all bead motions are constrained by the strong intermolecular native attractions.

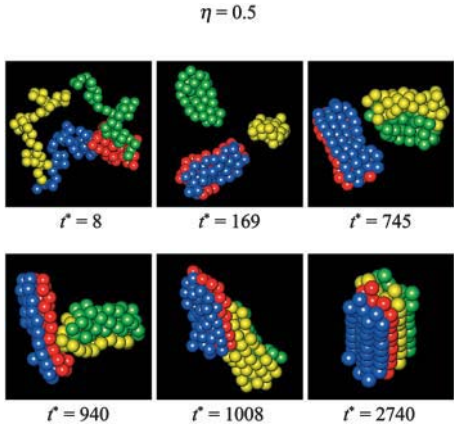

FIGURE 9 Snapshots showing kinetic β-sheet complex formation for $\eta = 0.5$ at selected reduced times, $t^* = 8$, 169, 745, 940, 1008, and 2740.

The misfolded states

It is interesting to examine the structures in the misfolded state because the misfolded state is favored in a majority of the trajectories. For $\eta = 0.2$, 0.5, and 0.8, Fig. 13 shows six typical structures found in the misfolded state. The chains are shaded differently to enable the reader to distinguish one from the other. For $\eta = 0.2$, the misfolded state m_1 contains two nonaligned pairs of β-sheets in which the two dimers, each with two well-aligned β-sheets, are aligned in parallel but are not face to face. This structure is very popular in the $\eta = 0.2$ model, because almost 55% of the trajectories fold to the m_1 state. The next most common structure for $\eta = 0.2$ is the misfolded state m_2 that contains three well-aligned β-sheets and a nonaligned β-sheet attached to the ordered trimer. In the misfolded states m_1 and m_2, the β-sheet complex is a partially folded tetramer with well-defined secondary structure. The misfolded state m_1 exhibits slightly larger values for R_g^2 and Q_{inter} than in misfolded state m_2. The misfolded states m_1 and m_2 are very common for the smaller values of η ($\eta = 0.2$ and 0.3).

For $\eta = 0.5$, the β-sheet complex in misfolded state m_3 is geometrically similar to that in the misfolded m_1 at $\eta = 0.2$, but its physical properties are significantly different. Its strands are stiffer and hence it has a larger R_g^2 value than in misfolded state m_1. The β-sheet complex in the misfolded state m_4 is a common structure among the misfolded states in the $\eta = 0.5$ model. In fact, almost 30% of the trajectories fold to m_4. The misfolded state m_4 is similar to the misfolded state m_2 at $\eta = 0.2$, because it contains three well-aligned β-sheets and one attached chain as was seen in the m_2 state. However, the attached chain in the m_4 state is more like a β-barrel than a β-sheet as in the m_2 state. The misfolded state m_4 exhibits slightly larger R_g^2 but smaller Q_{inter} values than the misfolded state m_3. The misfolded state m_4 is found most commonly among the misfolded states for the intermediate values of η ($\eta = 0.4$, 0.5, 0.6, and 0.7).

For $\eta = 0.8$, the misfolded state m_5 contains well-defined secondary structure, but one of the β-sheets has slipped out and refolded over the top of the next β-sheet, forming intermolecular native contacts with the wrong β-sheet. The β-sheet complex in the misfolded state m_5 has almost the same R_g^2 and Q_{inter} values as in the folded state. The nonaligned β-sheets in the misfolded state m_6 are frequently observed as η increases. The misfolded state m_6 exhibits no unique structure because it shows diverse R_g^2 and Q_{inter} values, but it contains tangled chains with poor secondary structure. This is due to the strong intermolecular native attraction in the larger η models ($\eta = 0.8$, 0.9, and 1.0).

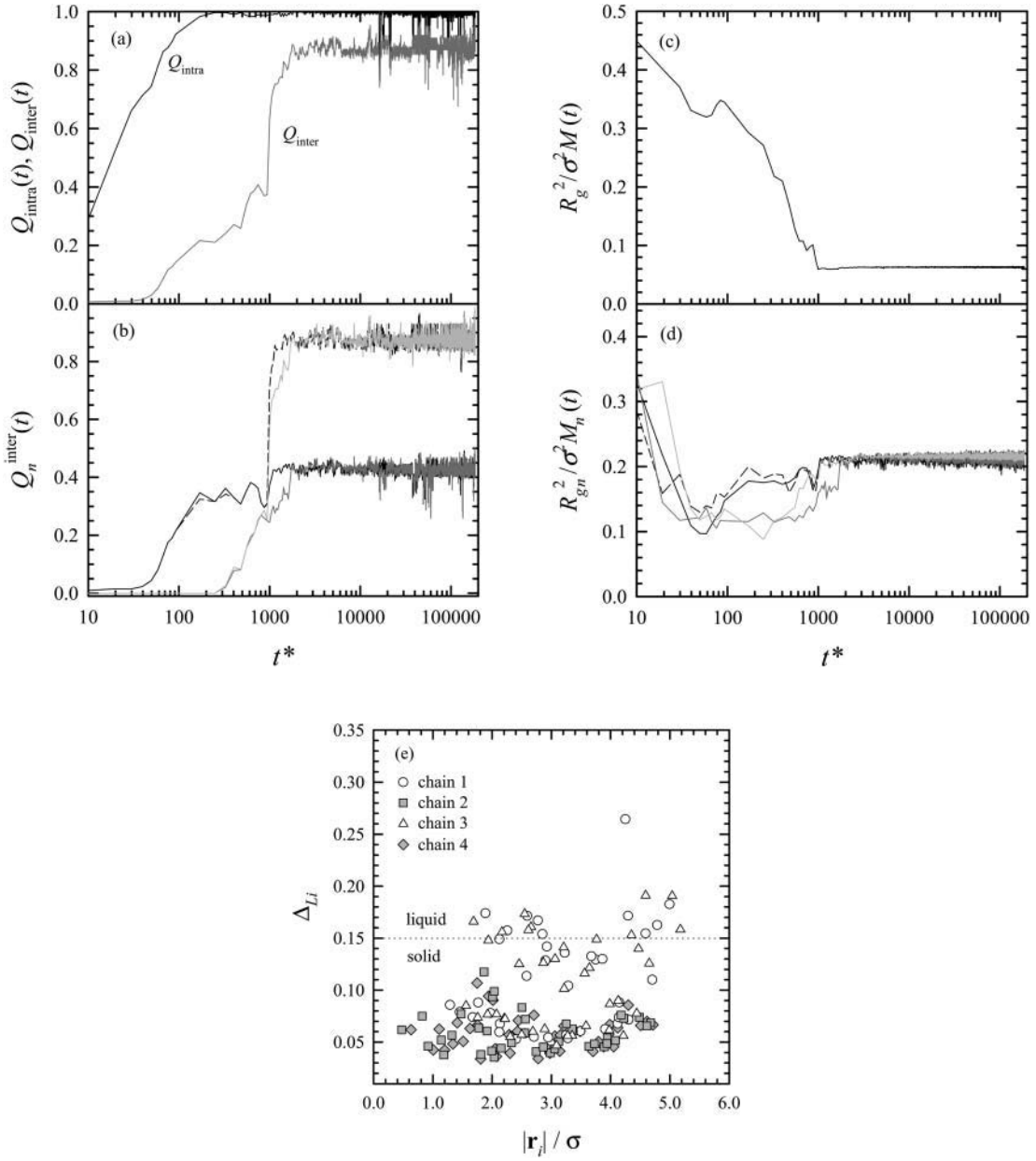

FIGURE 10 For $\eta = 0.5$ (*a*) the time-dependent fractions of intramolecular and intermolecular native contacts, $Q_{intra}(t)$ and $Q_{inter}(t)$, (*b*) the time-dependent fraction of intermolecular native contacts for the *n*th chain, $Q_n^{inter}(t)$, (*c*) the time-dependent reduced squared radius of gyration per bead, $R_g^2/\sigma^2 M$ (*t*), (*d*) the time-dependent reduced squared radius of gyration for the *n*th chain, $R_{gn}^2/\sigma^2 M_n$, all as a function of the reduced time, t^*, and (*e*) the Lindemann disorder parameter for an individual bead, Δ_{Li}, in the β-sheet complex as a function of the reduced distance of bead *i* from the origin, $|\mathbf{r}_i|/\sigma$.

Kinetic folding pathways

It is of interest to consider the intermediates and folding pathways in more detail. For $\eta = 0.2$, the contour graph in Fig. 14 *a* shows the population distribution of the reduced mean squared radius of gyration per bead for the individual chains, $\langle R_{gn}^2/\sigma^2 M_n \rangle_n$, versus the fraction of intermolecular native contacts, Q_{inter}, over all the trajectories and for all times. To highlight the existence of intermediates, sampling is locally weighted on early events during the simulation before oligomerization. This is because oligomer assembly occurs early in the simulation and the oligomer stays in the tetramer state once it has been formed. This weighting is permissible because we are not trying to calculate any system free energy based on the global sampling of probability density for a given set of reaction coordinates. Highly populated states are enclosed by many contour lines. The intermediates I_{2+1+1}, I_{2+2}, and I_{3+1} are highly populated during the assembly process and the misfolded states m_1 and m_2 are highly populated after the oligomerization is

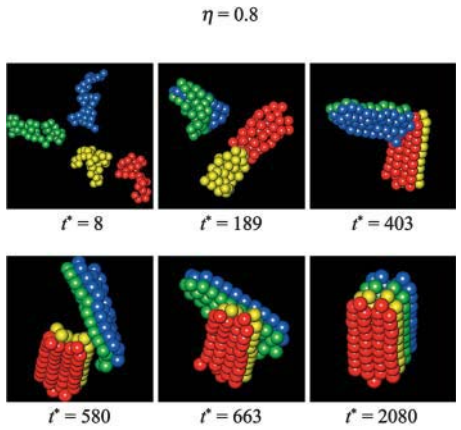

FIGURE 11 Snapshots showing kinetic β-sheet complex formation for $\eta = 0.8$ at selected reduced times, $t^* = 8$, 189, 403, 580, 663, and 2080.

completed. In Fig. 14 *a*, the folded state is indicated by the word ''native'' and is enclosed by a small contour line. This folded state is not highly populated, because the folding yield for $\eta = 0.2$ is very low as seen in Fig. 2. The kinetic assembly mechanism for $\eta = 0.2$ is represented in Fig. 14 *b* for both the folding and misfolding trajectories. The symbol ''U'' denotes the four separate random coils. The majority of the folding trajectories follow the path: U $\rightarrow$ $I_{2+1+1} \rightarrow I_{3+1}$ (or I_{2+2}) $\rightarrow$ Native, which is indicated by thick arrows in Fig. 14 *b*. All of the trajectories fold through the intermediate I_{2+1+1}, which can fold into either of the intermediates I_{3+1} or I_{2+2}. No direct pathways from the random coil state to the intermediates I_{3+1} or I_{2+2} are observed. The intermediate I_{3+1} can fold into the folded (native) state or the misfolded state m_2, whereas the intermediate I_{2+2} can fold into the folded state or the misfolded state m_1.

For $\eta = 0.5$, the contour graph in Fig. 15 *a* shows the population distribution of $\langle R_{gn}^2/\sigma^2 M_n \rangle_n$ versus Q_{inter} over all

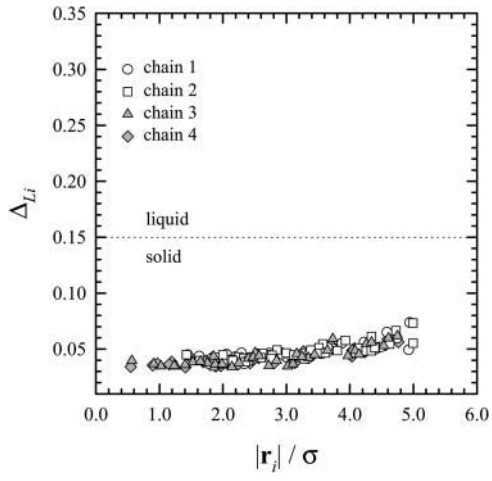

FIGURE 12 For $\eta = 0.8$, the Lindemann disorder parameter for an individual bead, Δ_{Li}, for the β-sheet complex as a function of the reduced distance of bead i from the origin, $|\mathbf{r}_i|/\sigma$.

the trajectories and for all times. The intermediate I_{2+1+1} is highly populated during the assembly process, meaning that the majority of the trajectories fold via the intermediate I_{2+1+1}. The folded (native) state is highly populated (enclosed by many contour lines) because the $\eta = 0.5$ model has the highest folding yield of all the folding trajectories. Two popular misfolded states m_3 and m_4 are observed near the folded state, indicating that they are partially ordered β-sheet complexes. The less populated misfolded states are not detected in the contour graph. These include tangled chains with poor secondary structure as shown in the misfolded state m_6 in Fig. 13. The β-sheet complex containing tangled chains has no unique structure and has a smaller values of $\langle R_{gn}^2/\sigma^2 M_n \rangle_n$ and Q_{inter} than the other misfolded states. Fig. 15 *b* shows the kinetic folding pathways for both the folding and misfolding trajectories. The majority of the folding trajectories follow the path: U $\rightarrow$ $I_{2+1+1} \rightarrow I_{3+1} \rightarrow$ Native, which is indicated by thick arrows in Fig. 15 *b*. A few trajectories fold from U directly to the intermediate I_{2+2}, which can fold into either the folded state or the misfolded state m_3. No folding trajectory from the intermediate I_{2+2} to the intermediate I_{3+1} is found. The intermediate I_{3+1} can fold into either the folded state or the misfolded state m_4.

For $\eta = 0.8$ (data not shown), the majority of the trajectories leading to the folded state follow the path: U $\rightarrow$ $I_{2+1+1} \rightarrow I_{2+2} \rightarrow$ Native. However, the majority of the trajectories are misfolding trajectories, resulting in low folding yields. Most misfolding trajectories end in the misfolded state m_6, which has tangled chains with very poor secondary structure. This state is often observed in trajectories that contain an early tangled trimer and is more likely to occur as η increases.

DISCUSSION

We have studied the assembly and folding kinetics of a β-sheet complex that contains four identical four-stranded

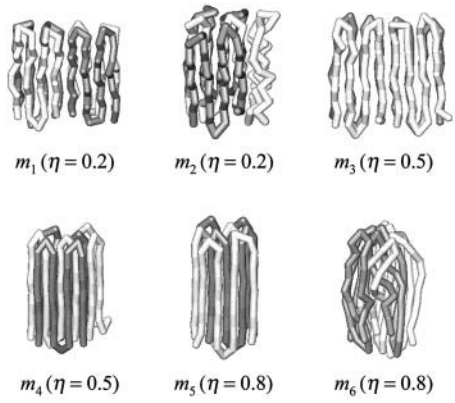

FIGURE 13 Snapshots showing typical misfolded β-sheet complexes for $\eta = 0.2$, 0.5, and 0.8.

antiparallel β-sheet peptides that interact both intramolecularly and intermolecularly. Discontinuous molecular dynamics simulations of the model complex have been performed for different sizes of the intermolecular contact parameter η, which measures the ratio of intermolecular to intramolecular native attractions, at a bias gap $g = 0.9$, a measure of the strength of the native contacts relative to that of the nonnative contacts. The intermolecular contact parameter η is divided into three distinct ranges: small ($\eta = 0.2$ and 0.3), intermediate ($\eta = 0.4$, 0.5, 0.6, and 0.7), and large ($\eta = 0.8$, 0.9, and 1.0), and is clearly seen to be an important factor in controlling the assembly, equilibrium properties, and folding kinetics of the β-sheet complex. In our model, the variation of the intermolecular contact parameter η can be regarded as mimicking different ratios of the strengths for the intermolecular (intersheet) hydrophobic interactions relative to the intramolecular (intrasheet) hydrogen bonding interactions in amyloid fibrils under different physiological conditions. The intersheet hydrophobic interactions are responsible for amyloid fibril formation and stability (Fraser et al., 1991a,b; Kowalewski and Holtzman, 1999). However, the most important factor affecting protein aggregation and the kinetics of fibril formation is protein concentration because there is a critical concentration at which disordered peptides assemble to form a seed that triggers protein aggregation (Lansbury, 1999). This paper does not account for the effect of protein concentration upon fibril formation, which is primarily an entropic effect. Investigations of this aspect of protein aggregation are currently underway.

Comparison with experimental results in the literature indicate that the intermediate η models (where the majority of folding trajectories lead to the highly ordered β-sheet complex) are most representative of systems that form fibrils, whereas the small η models (where the majority of folding trajectories lead to the misfolded state) are most representative of systems that form amorphous aggregates. Experiments have shown that protein aggregation, including fibril formation, arises directly from partially folded conformations (Booth et al., 1997; Chiti et al., 1999; Fink, 1998; Guijarro et al., 1998; Kelly, 1996; Khurana et al., 2001; Quintas et al., 1999), which tend to aggregate when they come into contact because they have large solvent-exposed hydrophobic residues. Of particular relevance to our work is a recent experiment on light chain amyloid fibrils and amorphous aggregates, showing that unfolded compact intermediates that have strong self-association tendencies are precursors to amyloid fibril formation, whereas native-like intermediates that have very weak intermolecular

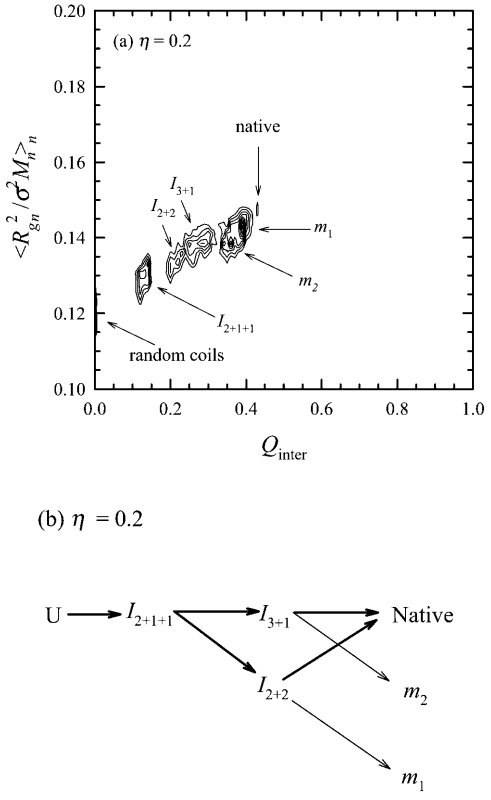

FIGURE 14 For $\eta = 0.2$ (*a*) the population distribution of the reduced mean squared radius of gyration per bead for the individual chains, $\langle R_{gn}^2/\sigma^2 M_n\rangle_n$, versus the fraction of intermolecular native contacts, Q_{inter}, and (*b*) the kinetic folding mechanism.

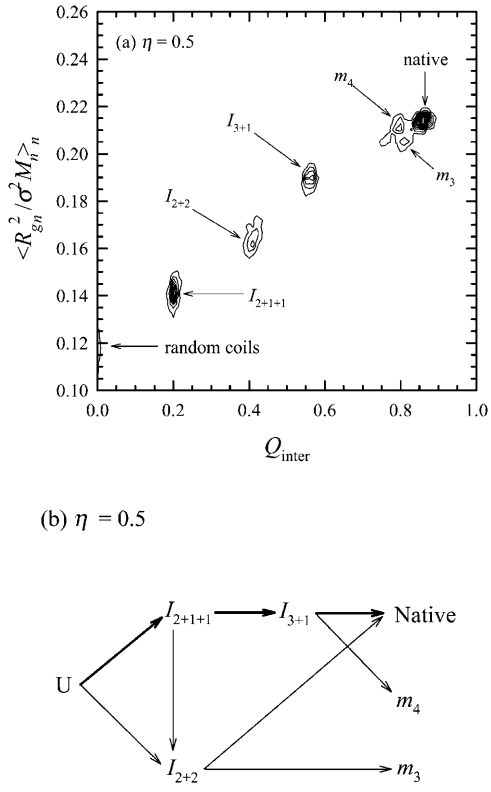

FIGURE 15 For $\eta = 0.5$ (*a*) the population distribution of the reduced mean squared radius of gyration per bead for the individual chains, $\langle R_{gn}^2/\sigma^2 M_n\rangle_n$, versus the fraction of intermolecular native contacts, Q_{inter}, and (*b*) the kinetic folding mechanism.

hydrophobic interactions are precursors to amorphous aggregates (Khurana et al., 2001). This is consistent with our observation that: 1), in the intermediate η models, partially folded chains are precursors to the highly ordered β-sheet complex, which is in a solid-like phase and hence is analogous to amyloid fibrils, and 2), in the small η models, native β-sheets are precursors to the misfolded state (the majority of folding trajectories for the small η models lead to the misfolded state), which contains some secondary structure but is in a liquid-like phase and hence is analogous to amorphous aggregates. Here we implicitly assume that amorphous aggregates retain some secondary structure while in a soluble phase as shown experimentally by Lorenzo and Yankner (1994).

The equilibrium properties of the folded and misfolded β-sheet complex strongly depend on the size of η. For the folded state, the squared radius of gyration of the whole system, R_g^2, and for the nth chain, R_{gn}^2, the fraction of total native contacts, Q_{total}, the fraction of intramolecular native contacts, Q_{intra}, the fraction of intermolecular native contacts, Q_{inter}, and the fraction of intermolecular native contacts for the nth chain, Q_n^{inter}, increase as η increases, whereas the system energy, E^*, and the root-mean-squared pair separation fluctuation, Δ_B decrease with increasing η. For the misfolded state, the η dependence of the equilibrium properties except Q_{intra} and R_g^2 is similar to that observed in the folded state. In the misfolded state, as η increases the Q_{intra} value decreases, and the R_g^2 of the entire system increases at intermediate values of η and then decreases at large values of η. This is due to the lack of secondary structure in the misfolded state for large η. The details of the thermodynamic equilibrium properties of the ordered β-sheet complex will be given elsewhere (Jang et al., 2003).

The folding kinetics are studied for three selected values of the intermolecular contact parameter, $\eta = 0.2$, 0.5, and 0.8. For $\eta = 0.2$, the folded state is an ordered β-sheet complex in a liquid-like phase according to the Lindemann criterion (Lindemann, 1910). The liquid-like phase found for the ordered β-sheet complex indicates that the β-sheet complex is loosely packed with large Δ_B but has a well-defined β-sheet secondary structure with large Q_{intra}. Most kinetic folding trajectories follow the path: U $\rightarrow I_{2+1+1} \rightarrow I_{3+1}$ (or I_{2+2}) $\rightarrow$ Native. Secondary structure begins to form before the β-sheet complex assembles. Many trajectories leading to the misfolded state, i.e., amorphous aggregates (Lorenzo and Yankner, 1994), involve two β-sheet dimers aligned in parallel, albeit a highly ordered structure in terms of fibril growth. Similar structures are found in the betabellin-15D fibril (Lim et al. 1998, 2000) and the transthyretin (TTR) amyloid fibril (Lai et al., 1996). The yield at $\eta = 0.2$ for the folded state is low. For $\eta = 0.5$, the folded state is a highly ordered β-sheet complex. The interior β-sheets in the complex are solid-like and the outer β-sheets exhibit both solid-like and liquid-like phases, but the entire structure is essentially solid-like. The solid-like phase in the highly ordered β-sheet complex is reminiscent of amyloid fibril formation in that fibril formation is in many ways analogous to crystallization (Chiti et al., 1999; Jarrett and Lansbury, 1993; Kowalewski and Holtzman, 1999; Mullin, 1972). Unlike the $\eta = 0.2$ model, partially folded chains at $\eta = 0.5$ assemble into dimers before each chain folds into the native state. As mentioned above, this is consistent with proposals appearing in the experimental (Booth et al., 1997; Chiti et al., 1999; Fink, 1998; Guijarro et al., 1998; Kelly, 1996; Khurana et al., 2001; Quintas et al., 1999) and simulations (Dima and Thirumalai, 2002; Harrison et al., 1999, 2001) literature that aggregation is enhanced by populating partially folded intermediates or transients. In our model, the assembly to the highly ordered oligomer is accelerated by the presence of a dimer that acts as a seed for aggregation. This is consistent with experimental observations that dimer seeding accelerates the formation of fibrils made by β-amyloids and prions (Harper and Lansbury, 1997). The majority of the folding trajectories at $\eta = 0.5$ follow the path: U $\rightarrow I_{2+1+1} \rightarrow I_{3+1} \rightarrow$ Native. The sequential addition of one monomer is the most productive pathway. The yield at $\eta = 0.5$ for the folded state is very high. For $\eta = 0.8$, the folded state is a highly ordered β-sheet complex in a solid-like phase, the same as in the global energy minimum structure. The yield for the folded state, which most often folds via the path: U $\rightarrow I_{2+1+1} \rightarrow I_{2+2} \rightarrow$ Native, is very low. The misfolded state contains tangled chains with poor secondary structure and is observed in a trajectory with early trimer formation. The tangled chains are essentially globular aggregates and are observed more frequently as η increases. A summary of the η-dependence of the folding kinetics, including final state, oligomer structure, and phase, is given in Table 2.

Our coarse-grained off-lattice protein models interacting via Go-type potentials (Go and Taketomi, 1978, 1979; Taketomi et al., 1975; Ueda et al., 1978) are very simple. Clearly, minimalist models based on Go-type potentials have limitations that prevent them from mimicking some aspects of the folding of real proteins. A recent study (Kaya and Chan, 2002) indicates that if a model's thermodynamics do not satisfy the calorimetric two-state criterion, its folding

TABLE 2 Dependence of the folding kinetics on the intermolecular contact parameter η

η	Final state	Oligomer	Phase
0.2	Folded	Ordered β-sheet complex	Liquid-like
	Misfolded^M	Amorphous aggregates	Liquid-like
0.5	Folded^M	Highly ordered β-sheet complex	Solid-like
	Misfolded	Amorphous (or globular) aggregates	Liquid- or solid-like
0.8	Folded	Highly ordered β-sheet complex	Solid-like
	Misfolded^M	Globular aggregates	Solid-like

The superscript M denotes the majority of folding trajectories.

kinetics are not likely to be protein-like. In our previous thermodynamic study for an isolated β-sheet (Jang et al., 2002a), we observed that the collapse transition in our Go-type protein models is not calorimetrically two-state. The implications of this for the present work are unclear because the general mechanism governing protein aggregation is somewhat different from that governing protein folding, and is certainly more complicated. Even with these limitations, Go-type protein models often yield a number of insights into the basic principles behind the thermodynamics and kinetics of protein folding and aggregation (Bratko and Blanch, 2001; Dokholyan et al., 1998, 2000; Jang et al., 2002a,b; Zhou and Karplus, 1997a,b; 1999).

Our model of a β-sheet complex could prove useful in understanding the basic principles behind Aβ fibril formation. However, it would have rendered more insights into fibril formation and stability if the monomer's native state had been different from that in the aggregate, because the native conformation of a peptide in solution is usually different from that in aggregates. This is a limitation of the Go-type potential that we have devised. Many amyloid forming peptides, including Aβ peptides, are thought to be random coils in solution under most conditions but are extended chains or β-hairpins in fibrils (Serpell, 2000). In an effort to remedy this deficiency, studies of fibril formation kinetics for a system of β-hairpins as a function of chain concentration are being conducted.

This work was supported by the National Institutes of Health under grant number GM-56766 and the National Science Foundation under grant number CTS-9704044. The work at Buffalo is supported in part by a grant from Howard Hughes Medical Institute to State University of New York Buffalo.

REFERENCES

Alder, B. J., and T. E. Wainwright. 1959. Studies in molecular dynamics. I. General method. *J. Chem. Phys.* 31:459–466.

Anderson, H. C. 1980. Molecular dynamics simulations at constant pressure and/or constant temperature. *J. Chem. Phys.* 72:2384–2393.

Bellemans, A., J. Orban, and D. V. Belle. 1980. Molecular dynamics of rigid and non-rigid necklaces of hard discs. *Mol. Phys.* 39:781–782.

Benzinger, T. L. S., D. M. Gregory, T. S. Burkoth, H. Miller-Auer, D. G. Lynn, R. E. Botto, and S. C. Meredith. 1998. Propagating structure of Alzheimer's $\beta-$amyloid(10–35) is parallel $\beta-$sheet with residues in exact register. *Proc. Natl. Acad. Sci. USA.* 95:13407–13412.

Benzinger, T. L. S., D. M. Gregory, T. S. Burkoth, H. Miller-Auer, D. G. Lynn, R. E. Botto, and S. C. Meredith. 2000. Two-dimensional structure of β–amyloid(10–35) fibrils. *Biochemistry.* 39:3491–3499.

Bilgram, J. H. 1987. Dynamics at solid liquid transition-experiments at the freezing-point. *Phys. Rep.* 153:1–89.

Booth, D. R., M. Sunde, V. Bellotti, C. V. Robinson, W. L. Hutchinson, P. E. Fraser, P. N. Hawkins, C. M. Dobson, S. E. Radford, C. C. F. Blake, and M. B. Pepys. 1997. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature.* 385:787–793.

Bratko, D., and H. W. Blanch. 2001. Competition between protein folding and aggregation: A three-dimensional lattice-model simulation. *J. Chem. Phys.* 114:561–569.

Burkoth, T. S., T. L. S. Benzinger, D. N. M. Jones, K. Hallenga, S. C. Meredith, and D. G. Lynn. 1998. C-terminal PEG blocks the irreversible step in $\beta-$amyloid(10–35) fibrillogenesis. *J. Am. Chem. Soc.* 120: 7655–7656.

Chan, H. S., and K. A. Dill. 1994. Transition states and folding dynamics of proteins and heteropolymers. *J. Chem. Phys.* 100:9238–9257.

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.

Clark, J., and J. Steele. 1992. Phase separation inhibitors and prevention of selenite catarict. *Proc. Natl. Acad. Sci.* 89:1720–1724.

Dill, K. A., and D. Stigter. 1995. Modeling protein stability as heteropolymer collapse. *Adv. Protein. Chem.* 46:59–104.

Dima, R. I., and D. Thirumalai. 2002. Exploring protein aggregation and self-propagation using lattice models: phase diagram and kinetics. *Protein Sci.* 11:1036–1049.

Dokholyan, N. V., S. V. Buldyrev, H. E. Stanley, and E. I. Shakhnovich. 1998. Discrete molecular dynamics studies of the folding of a protein-like model. *Fold. Des.* 3:577–587.

Dokholyan, N. V., S. V. Buldyrev, H. E. Stanley, and E. I. Shakhnovich. 2000. Identifying the protein folding nucleus using molecular dynamics. *J. Mol. Biol.* 296:1183–1188.

Eaton, W. A., and J. Hofrichter. 1990. Sickle cell hemoglobin polymerization. *Adv. Prot. Chem.* 40:63–279.

Esler, W. P., E. R. Stimson, J. R. Ghilardi, H. V. Vinters, J. P. Lee, P. W. Mantyh, and J. E. Maggio. 1996. *In vitro* growth of Alzheimer's disease β-amyloid plaques displays first- order kinetics. *Biochemistry.* 35: 749–757.

Esler, W. P., A. M. Felix, E. R. Stimson, M. J. Lachenmann, J. R. Ghilardi, Y. Lu, H. V. Vinters, P. W. Mantyh, J. P. Lee, and J. E. Maggio. 2000. Activation barriers to structural transition determine deposition rates of Alzheimer's disease Aβ amyloid. *J. Struct. Biol.* 130:174–183.

Fink, A. L. 1998. Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold. Des.* 3:R9–R23.

Fraser, P. E., L. K. Duffy, M. B. O'Malley, J. T. Nguyen, H. Inouye, and D. A. Kirschner. 1991a. Morphology and antibody recognition of synthetic β-amyloid peptides. *J. Neurosci. Res.* 28:474–485.

Fraser, P. E., J. T. Nguyen, W. K. Surewicz, and D. A. Kirschner. 1991b. pH-dependent structural transitions of Alzheimer amyloid peptides. *Biophys. J.* 60:1190–1201.

Gallo, G., F. Goni, F. Boctor, R. Vidal, A. Kumar, F. J. Stevens, B. Frangione, and J. Ghiso. 1996. Light chain cardiomyopathy. Structural sanalysis of the light chain tissue deposits. *Am. J. Pathol.* 148:1397–1406.

Go, N., and H. Taketomi. 1978. Respective roles of short range and long range interactions in protein folding. *Proc. Natl. Acad. Sci. USA.* 75: 559–563.

Go, N., and H. Taketomi. 1979. Studies on protein folding, unfolding and fluctuations by computer simulation. IV. Hydrophobic interactions. *Int. J. Protein Res.* 13:447–461.

Guijarro, J. I., M. Sunde, J. A. Jones, I. D. Campbell, and C. M. Dobson. 1998. Amyloid fibril formation by an SH3 domain. *Proc. Natl. Acad. Sci. USA.* 95:4224–4228.

Guo, Z., and C. L. Brooks, Iii. 1997. Thermodynamics of protein folding: a statistical mechanical study of a small all–β protein. *Biopolymers.* 42:745–757.

Guo, Z., and D. Thirumalai. 1995. Kinetics of protein folding: nucleation mechanism, time scales, and pathways. *Biopolymers.* 36:83–102.

Guo, Z., and D. Thirumalai. 1996. Kinetics and thermodynamics of folding of a *de Novo* designed four-helix bundle protein. *J. Mol. Biol.* 263: 323–343.

Gupta, P., and C. K. Hall. 1997. Effect of solvent conditions upon refolding pathways and intermediates for a simple lattice protein. *Biopolymers.* 42:399–409.

Gupta, P., and C. K. Hall. 1998. Effect of denaturant and protein concentrations upon protein refolding and aggregation: a simple lattice model. *Protein Sci.* 7:2642–2652.

Gupta, P., C. K. Hall, and A. Voegler. 1999. Computer simulation of the competition between folding and aggregation. *Fluid Phase Equilib.* 158:87–93.

Harper, J. D., and P. T. Lansbury, Jr. 1997. Models of amyloid seeding in Alzhemier's disease and scrapie: mechanistic truths and physiological consequences of time-dependent stability of amyloid proteins. *Annu. Rev. Biochem.* 66:385–407.

Harper, J. D., S. S. Wong, C. M. Lieber, and P. T. Lansbury, Jr. 1997. Observation of metastable Aβ amyloid protofibrils by atomic force microscopy. *Chem. Biol.* 4:119–125.

Harrison, P. M., H. S. Chan, S. B. Prusiner, and F. E. Cohen. 1999. Thermodynamics of model prions and its implications for the problem of prion protein folding. *J. Mol. Biol.* 286:593–606.

Harrison, P. M., H. S. Chan, S. B. Prusiner, and F. E. Cohen. 2001. Conformational propagation with prion-like characteristics in a simple model of protein folding. *Protein Sci.* 10:819–835.

Jang, H., C. K. Hall, and Y. Zhou. 2002a. Folding thermodynamics of model four-strand antiparallel β–sheet proteins. *Biophys. J.* 82:646–659.

Jang, H., C. K. Hall, and Y. Zhou. 2002b. Protein folding pathways and kinetics: molecular dynamics simulations β–strand motifs. *Biophys. J.* 83:819–835.

Jang, H., C. K. Hall, and Y. Zhou. 2003. Thermodynamics and stability of a β–sheet complex: molecular dynamics simulations on simplified off-lattice protein models. *Protein Sci.* In press.

Jarrett, J. T., and P. T. Lansbury, Jr. 1993. Seeding one-dimensional crystallization of amyloid—a pathogenic mechanism in Alzheimers disease and scrapie. *Cell.* 73:1055–1058.

Kabsch, W. 1976. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr.* A32:922–923.

Kaya, H., and H. S. Chan. 2002. Towards a consistent modeling of protein thermodynamic and kinetic cooperativity: how applicable is the transition state picture to folding and unfolding? *J. Mol. Biol.* 315:899–909.

Kelly, J. W. 1996. Alternative conformations of amyloidogenic proteins govern their behavior. *Curr. Opin. Struct. Biol.* 6:11–17.

Kirkitadze, M. D., G. Bitan, and D. B. Teplow. 2002. Paradigm shifts in Alzheimer's disease and other neurodegenerative disorders: the emerging role of oligomeric assemblies J. *Neurosci. Res.* 69:567–577.

Kolinski, A., W. Galazka, and J. Skolnick. 1995. Computer design of idealized β-motifs. *J. Chem. Phys.* 103:10286–10297.

Kolinski, A., B. Ilkowski, and J. Skolnick. 1999. Dynamics and thermodynamics of β-hairpin assembly: insights from various simulation techniques. *Biophys. J.* 77:2942–2952.

Kowalewski, T., and D. M. Holtzman. 1999. *In situ* atomic force microscopy study of Alzheimer's β-amyloid peptide on different substrates: New insights into mechanism of β-sheet formation. *Proc. Natl. Acad. Sci. USA.* 96:3688–3693.

Khurana, R., J. R. Gillespie, A. Talapatra, L. J. Minert, C. Ionescu-Zanetti, I. Millett, and A. L. Fink. 2001. Partially folded intermediates as critical precursors of light chain amyloid fibrils and amorphous aggregates. *Biochemistry.* 40:3525–3535.

Lai, Z., W. Colon, and J. W. Kelly. 1996. The acid-mediated denaturation pathway of Transthyretin yields a conformational intermediate that can self-assemble into amyloid. *Biochemistry.* 35:6470–6482.

Lansbury, P. T., Jr. 1999. Evolution of amyloid: what normal protein folding may tell us about fibrillogenesis and disease. *Proc. Natl. Acad. Sci. USA.* 96:3342–3344.

Lau, K. F., and K. A. Dill. 1989. A lattice statistical mechanics model of the conformational and sequence spaces of protein. *Macromolecules.* 22:3986–3997.

Lazaridis, T., and M. Karplus. 1997. ''New view'' of protein folding reconciled with the old through multiple unfolding simulations. *Science.* 278:1928–1931.

Lazo, N. D., and D. T. Cowning. 1998. Amyloid fibrils may be assembled from β–helical protofibrils. *Biochemistry.* 37:1731–1735.

Lim, A., M. J. Saderholm, A. M. Makhov, M. Kroll, Y. Yan, L. Perera, J. D. Griffith, and B. W. Erickson. 1998. Engineering of betabellin-15D: a 64 residue beta sheet protein that forms long narrow multimeric fibrils. *Protein Sci.* 7:1547–1554.

Lim, A., A. M. Makhov, J. Bond, H. Inouye, L. H. Connors, J. D. Griffith, B. W. Erickson, D. A. Kirschner, and C. E. Costello. 2000. Betabellins 15D and 16D, *de novo* designed β–sandwich proteins that have amyloidogenic properties: a 64 residue beta sheet protein that forms long narrow multimeric fibrils. *J. Struct. Biol.* 130:363–370.

Lindemann, F. A. 1910. The calculation of molecular vibration frequencies. *Physik. Z.* 11:609–612.

Lorenzo, A., and B. A. Yankner. 1994. β-amyloid neurotoxicity requires fibril formation and is inhibited by Congo red. *Proc. Natl. Acad. Sci. USA.* 91:12243–12247.

Löwen, H. 1994. Melting, freezing and colloidal suspensions. *Phys. Rep.* 237:249–324.

Lynn, D. G., and S. C. Meredith. 2000. Review: model peptides and the physicochemical approach to β–amyloids. *J. Struct. Biol.* 130:153–173.

Massry, S., and R. Glasscock. 1983. Textbook of Nephrology. Williams & Wilkins, Baltimore, MD.

Miller, R., C. A. Danko, M. J. Fasolka, A. C. Balazs, H. S. Chan, and K. A. Dill. 1992. Folding kinetics of proteins and copolymers. *J. Chem. Phys.* 96:768–780.

Moore, R. C., and D. W. Melton. 1997. Transgenic analysis of prion diseases. *Mol. Hum. Reprod.* 3:529–544.

Mullin, J. W. 1972. Crystallization. Butterworths, London, UK.

Nymeyer, H., A. E. García, and J. N. Onuchic. 1998. Folding funnels and frustration in off- lattice minimalist protein landscapes. *Proc. Natl. Acad. Sci. USA.* 95:5921–5928.

Pace, C. N., E. J. Hebert, K. L. Shaw, D. Schell, V. Both, D. Krajcikova, J. Sevcik, K. S. Wilson, Z. Dauter, R. W. Hartley, and G. R. Grimsley. 1998. Conformational stability and thermodynamics of folding of ribonucleases Sa, Sa2 and Sa3. *J. Mol. Biol.* 279:271–286.

Pallitto, M. M., and R. M. Murphy. 2001. A mathematical model of the kinetics of β–amyloid fibril growth from the denatured state. *Biophys. J.* 81:1805–1822.

Pande, V. S., and D. S. Rokhsar. 1998. Is the molten globule a third phase of proteins? *Proc. Natl. Acad. Sci. USA.* 95:1490–1494.

Plaxco, K. W., K. T. Simons, and D. Baker. 1998. Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* 277:985–994.

Ptitsyn, O. B. 1995. Molten globule and protein folding. *Adv. Protein. Chem.* 47:83–230.

Quintas, A., M. J. M. Saraiva, and R. M. M. Brito. 1999. The tetrameric protein transthyretin dissociates to a non-native monomer in solution: a novel model for amyloidogenesis. *J. Biol. Chem.* 274:32943–32949.

Rapaport, D. C. 1978. Molecular dynamics simulation of polymer chains with excluded volume. *J. Phys. A: Math. Gen.* 11:L213–L216.

Šali, A., E. I. Shakhnovich, and M. Karplus. 1994. Kinetics of protein folding: a lattice model study of the requirements for folding to the native state. *J. Mol. Biol.* 235:1614–1636.

Selkoe, D. J. 1991. Alzheimer's disease: a central role for amyloid. *J. Neuropathol. Exp. Neurol.* 53:438–447.

Serpell, L. C. 2000. Alzheimer's amyloid fibrils: structure and assembly. Biochim. et Biophy. *Acta.* 1502:16–30.

Shea, J.-E., J. N. Onuchic, and C. L. Brooks, Iii. 2000. Energetic frustration and the nature of the transition state in protein folding. *J. Chem. Phys.* 113:7663–7671.

Simmons, L. K., P. C. May, K. J. Tomoselli, R. E. Rydel, K. S. Fuson, B. F. Brigham, S. Wright, I. Liebergurg, G. W. Becker, D. N. Brems, and W. Y. Li. 1994. Secondary structure of amyloid β peptide correlates with neurotoxic activity in vitro. *Mol. Pharmacol.* 45:373–379.

Skolnick, J., and A. Kolinski. 1991. Dynamic Monte Carlo simulations of a new lattice model of globular protein folding, structure and dynamics. *J. Mol. Biol.* 221:499–531.

Smith, S. W., C. K. Hall, and B. D. Freeman. 1996. Molecular dynamic study of entangled hard-chain fluids. *J. Chem. Phys.* 104:5616–5637.

Snyder, S. W., U. S. Ladror, W. S. Wade, G. T. Wang, L. W. Barrett, E. D. Matayoshi, H. J. Huffaker, G. A. Krafft, and T. F. Holzman. 1994. Amyloid-β aggregation: selective inhibition of aggregation in mixtures of amyloid with different chain lengths. *Biophys. J.* 67:1216–1228.

Stillinger, F. H. 1995. A topographic view of super-cooled liquids and glass formation. *Science.* 267:1935–1939.

Sunde, M., L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys, and C. C. F. Blake. 1997. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* 273:729–739.

Taketomi, H., Y. Ueda, and N. Go. 1975. Studies on protein folding, unfolding and fluctuations by computer simulation. The effect of specific amino acid sequence represented by specific inter-unit interactions. *Int. J. Pept. Protein Res.* 7:445–459.

Ueda, Y., H. Taketomi, and N. Go. 1978. Studies on protein folding, unfolding and fluctuations by computer simulation. II. A three-dimensional lattice model of lysozyme. *Biopolymers.* 17:1531–1548.

Zhang, S., K. Iwata, M. J. Lachenmann, J. W. Peng, S. Li, E. R. Stimson, Y. Lu, A. M. Felix, J. E. Maggio, and J. P. Lee. 2000. The Alzheimer's peptide Aβ adopts a collapsed coil structure in water. *J. Struct. Biol.* 130:130–141.

Zhou, Y., and M. Karplus. 1997a. Folding thermodynamics of a model three-helix bundle protein. *Proc. Natl. Acad. Sci. USA.* 94:14429–14432.

Zhou, Y., and M. Karplus. 1997b. Interpreting the folding kinetics of helical proteins. *Nature.* 401:400–403.

Zhou, Y., and M. Karplus. 1999. Folding of a model three-helix bundle protein: a thermodynamic and kinetic analysis. *J. Mol. Biol.* 293: 917–951.

Zhou, Y., C. K. Hall, and M. Karplus. 1996. First-order disorder-to-order transition in an isolated homopolymer model. *Phys. Rev. Lett.* 77: 2822–2825.

Zhou, Y., M. Karplus, J. M. Wichert, and C. K. Hall. 1997. Equilibrium thermodynamics of homopolymers and clusters: molecular dynamics and Monte Carlo simulations of systems with square-well interactions. *J. Chem. Phys.* 107:10691–10708.

Zhou, Y., D. Vitkup, and M. Karplus. 1999. Native proteins are surface-molten solids: application of the Lindemann criterion for the solid versus liquid state. *J. Mol. Biol.* 285:1371–1375.