Protein Folding Pathways and Kinetics: Molecular Dynamics Simulations of β-Strand Motifs

Hyunbum Jang,* Carol K. Hall,* and Yaoqi Zhou†

*Department of Chemical Engineering, North Carolina State University, Raleigh, North Carolina 27695; and †Department of Physiology and Biophysics, State University of New York at Buffalo, Buffalo, New York 14214 USA

ABSTRACT The folding pathways and the kinetic properties for three different types of off-lattice four-strand antiparallel β-strand protein models interacting via a hybrid Go-type potential have been investigated using discontinuous molecular dynamics simulations. The kinetic study of protein folding was conducted by temperature quenching from a denatured or random coil state to a native state. The progress parameters used in the kinetic study include the squared radius of gyration R_g^2, the fraction of native contacts within the protein as a whole Q, and between specific strands Q_{ab}. In the time series of folding, the denatured proteins undergo a conformational change toward the native state. The model proteins exhibit a variety of kinetic folding pathways that include a fast-track folding pathway without passing through an intermediate and multiple pathways with trapping into more than one intermediate. The kinetic folding behavior of the β-strand proteins strongly depends on the native-state geometry of the model proteins and the size of the bias gap g, an artificial measure of a model protein's preference for its native state.

INTRODUCTION

In our laboratory we are using computer simulation to understand the basic physical principles of protein aggregation, a cause or associated symptom of a number of lethal disorders including Alzheimer's, Parkinson's, and prion diseases (Clark and Steele, 1992; Eaton and Hofrichter, 1990; Gallo et al., 1996; Massry and Glasscock, 1983; Moore and Melton, 1997; Selkoe, 1991; Simmons et al., 1994). Our long-term goal is to learn if there are certain universal molecular-level mechanisms responsible for protein aggregation and fibril formation. Because the association of β-strands, or the conversion of α-helices to β-strands often precedes or accompanies protein aggregation (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 models of β-strand proteins. These models must be capable of capturing the essential physical features of real β-strand proteins and at the same time be simple enough to allow the simulation of multi-protein systems with current computer capability.

Low-resolution or simplified protein models are best suited for the construction of a β-strand model to be used in simulations of protein folding because they allow us to study the folding behavior over relatively long time scales. These models provide numerous insights into the thermodynamic and kinetic properties of protein folding (Chan and Dill, 1994; Dokholyan et al., 1998, 2000; Guo and Brooks, 1997; Guo and Thirumalai, 1995, 1996; Gupta and Hall, 1997, 1998; 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 Kolinski, 1991; Zhou and Karplus, 1997a,b, 1999). 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., 2002). Discontinuous molecular dynamics (DMD) simulations (Alder and Wainwright, 1959; Rapaport, 1978; Smith et al., 1996) on these three models 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. The thermodynamics results presented in our previous paper and the kinetics results presented here set the stage for a later paper in which we will examine the folding and assembly of several of these model β-strand proteins into fibrils.

The folding of β-strand proteins is of current interest not only because isolated β-strand peptides tend to adopt a β-sheet fibril conformation when they aggregate, but also because β-strands are often secondary structure elements in proteins. Unfortunately, unlike the formation of α-helices, the formation of even isolated β-strand peptides is not easy to study experimentally due to difficulties associated with the propensities of the largely insoluble β-structures to precipitate. This has prompted the development of designed β-strand peptides that are more amenable to fundamental protein folding investigations. Experimental studies of such peptides include analyses of the structure and stability for de novo designed two-strand β-sheets known as β-hairpin (Blanco et al., 1998; Ramírez-Alvarado et al., 1996), three-strand β-sheets known as betanova (de Alba et al., 1999; Koepf et al., 1999; Kortemme et al., 1998), and four-strand

Submitted January 28, 2002, and accepted for publication April 12, 2002.

Address reprint requests to Dr. Carol K. Hall, Department of Chemical Engineering, North Carolina State University, Raleigh, NC 27695-7905. Tel.: 919-513-3571; Fax: 919-515-3465; E-mail: hall@turbo.che.ncsu.edu.

0006-3495/02/08/819/17 $2.00

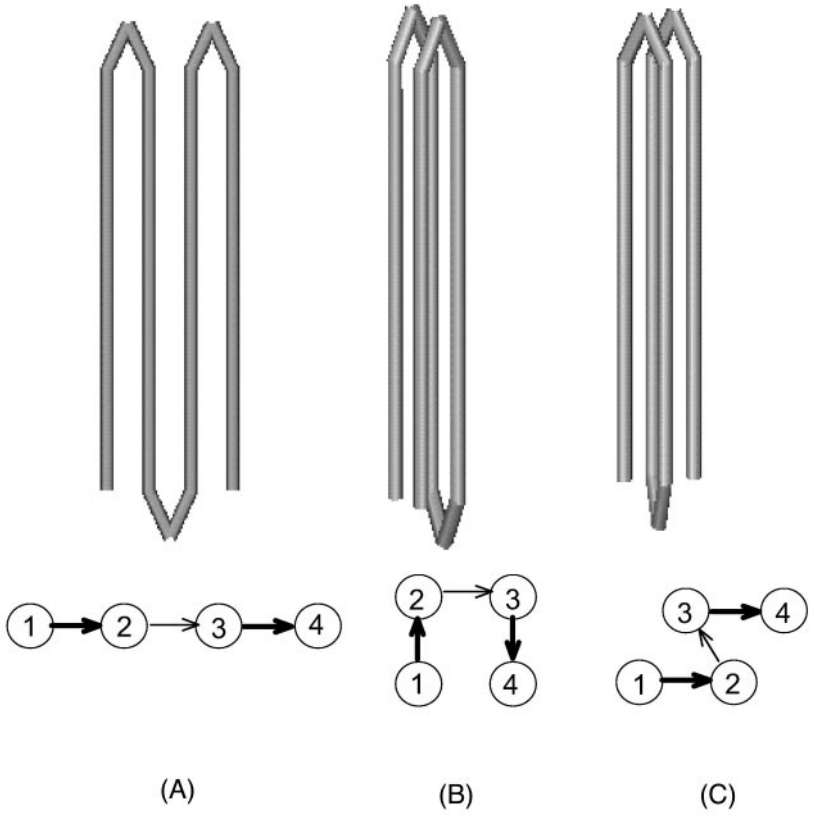

FIGURE 1 Stick (*top*) and topology (*bottom*) diagrams of the global energy minimum structures: (*A*) the β-sheet; (*B*) the β-clip; and (*C*) the β-twist model proteins.

β-sheets known as betabellin (Lim et al., 2000). Computer simulation studies have been conducted on the thermodynamics and kinetics of formation of 16-residue β-hairpins (Dinner et al., 1999; García and Sanbonmatsu, 2001; Lee and Shin, 2001; Pande and Rokhsar, 1999; Kolinski et al., 1999; Zagrovic et al., 2001; Zhou and Linhananta, 2002; Zhou et al., 2001) taken from the C-terminal fragment (41–56) of the B1 domain of protein G (Muñoz et al., 1997), the folding kinetics of betanova (Bursulaya and Brooks, 1999; Ferrara and Caflisch, 2000), the folding thermodynamics and kinetics of a 46-mer β-barrel (Guo and Brooks, 1997; Guo and Thirumalai, 1995), and the dynamics of a six-stranded lattice model of Greek key β-barrel (Kolinski et al., 1995; Skolnick et al., 1989).

In this paper we investigate the folding kinetics of the three different four-strand antiparallel β-strand peptides: the β-sheet, the β-clip, and the β-twist, whose thermodynamic properties were studied previously by us (Jang et al., 2002). These β-strand peptides each have 39 connected residues (beads) and different native state conformations as shown in Fig. 1. Nonbonded beads can interact 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 value of the bias gap parameter g. DMD simulations were performed on these three model systems at intermediate values of the bias gaps ranging from 0.7 to 1.1. The progress parameters, the squared radius of gyration R_g^2, the fraction of native contacts within the protein as a whole Q, and between specific strands Q_{ab} were monitored over the course of the kinetic simulations. The bias gap measures the difference in interaction strength between the native and non-native contacts; the larger it is, the more the native state is favored over the non-native state. Intermediate values of the bias gap are thought to be the most representative of real proteins in their equilibrium and dynamic behavior. The model protein's bias gap is an artificial measure of its preference for the native state. By exploring how variations in the bias gap influence the types of dynamic phase behavior observed, we can get a feeling for how a real protein's preference for its native state, as measured for example by the energy difference between the denatured and native state, is manifested in the protein's dynamic behavior and vice versa.

Highlights of our results are the following. We find that the β-sheet exhibits a fast-track folding pathway without becoming trapped in any intermediates for all the bias gaps investigated. Its kinetic folding mechanism can be described by the diffusion collision model (Karplus and Weaver, 1979, 1994). In contrast, the β-clip and β-twist exhibit multiple folding pathways that include trapping in intermediates and direct folding to the native state depending on the size of the bias gap. For the β-clip and β-twist models at $g = 0.7$, two intermediates I_1 and I_2 are observed in the folding pathways. The intermediate I_1 is a partially ordered

globule, and the intermediate I_2 is an ordered globule that can be regarded as a molten globule as described in the thermodynamic study. The intermediate I_1 can fold either directly into the native state or via the intermediate I_2, which then folds into the native state by elongating its strands. No fast-track folding pathway is observed. The kinetic folding mechanism for the β-clip and β-twist at $g = 0.7$ can be described by the hydrophobic-collapse theory (Dill et al., 1995). For the β-clip and β-twist models at $g = 0.9$, the kinetic folding trajectories include the fast-track folding pathway and trapping in an intermediate *I*. The intermediate *I* is an ordered globule as was seen in the $g = 0.7$ model result. The random coil can fold directly to the native state or via the intermediate *I*. We expect that as *g* increases, more of the trajectories will exhibit the fast-track kinetic folding pathway. The kinetic folding mechanism for the β-clip and β-twist at $g = 0.9$ can be described by a combination of the diffusion collision model and the hydrophobic-collapse theory. 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). Models with a high topological complexity exhibit slower folding to the native state than models with a topologically simple structure. Moreover, the former is prone to kinetic intermediates.

In the following section, a full description of the three models and the simulation method are presented. Next is the results section, which is divided into three subsections; the first presents the results for the folding kinetics, the second presents a description of intermediates and folding pathways, and the third discusses how the folding speed depends on the topology of the model proteins. The paper concludes with a discussion of the key findings in the last section.

MODELS AND SIMULATION METHOD

We consider three different off-lattice protein models whose native states are four-strand antiparallel β-strand peptides. The global energy minimum structures for the three different β-strand models are shown in Fig. 1. Stick diagrams for the β-sheet, β-clip, and β-twist are shown at the top of the figure. Topology diagrams representing a top view of the three proteins are shown at the bottom of the figure. In the topology diagrams, thick arrows indicate chain connectivity at the top of the strands, and thin arrows indicate chain connectivity at the bottom of the strands. It is immediately apparent that in the native state, the β-sheet is two-dimensional (planar), whereas the β-clip and β-twist have well-defined three-dimensional four-barrel structures. The total number of native contacts, $N_{\text{native}}^{\text{total}}$, for each model chain and the total number of inter-strand native contacts, $N_{\text{ab}}^{\text{total}}$, for each model where *a* and *b* are the strand number, are summarized in Table 1. Note that $N_{\text{native}}^{\text{total}}$ for the β-sheet equals the sum of all $N_{\text{ab}}^{\text{total}}$, but $N_{\text{native}}^{\text{total}}$ for the β-clip and β-twist does not, because native contacts on turn residues are not inter-strand native contacts.

Each model protein contains 39 connected beads, each representing an amino acid residue that can be regarded as being localized at the C_α atom. The details of the model potential and the bias gap *g* are described in an earlier publication (Jang et al., 2002).

The folding kinetics of the three different β-strand peptides are studied here using the DMD algorithm (Alder and Wainwright, 1959; Rapaport, 1978; Smith et al., 1996). Proteins are quenched from the denatured or random coil state equilibrated at high temperature to the temperature of interest. After quenching, the proteins undergo a conformational change toward the energy minimum state, i.e., the native state. We investigated the time series of folding behavior from the random coil state at $T^* = 2.0$ to the native state at $T^* = 0.2$, 0.24, and 0.28 for the β-sheet, the β-clip, and the β-twist, respectively. The temperature used in this paper is a dimensionless reduced temperature, $T^* = k_B T/\epsilon$, where k_B is Bolzmann's constant, *T* is the absolute temperature, and ϵ is the energy parameter. Our thermodynamic study on the same models (Jang et al., 2002) shows that at $T^* = 2.0$, the three β-strand peptides are completely random coils. The three target temperatures for each model are selected to ensure that the native state is highly populated at that temperature. The initial configurations for the kinetic simulations were generated from equilibrium simulations at $T^* = 2.0$. 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. Simulations were performed for three different values of the bias gaps, $g = 0.7$, 0.9, and 1.1, except for the β-sheet where the simulations were performed at $g = 0.9$ and 1.1. For each bias gap model, statistical averages were obtained over 100 independent simulations.

TABLE 1 Numbers of native contacts

	β-sheet	β-clip	β-twist
$N_{\text{native}}^{\text{total}}$	75	125	130
N_{12}^{total}	25	25	25
N_{13}^{total}		9	25
N_{14}^{total}		25	
N_{23}^{total}	25	25	25
N_{24}^{total}		9	25
N_{34}^{total}	25	25	25

The total number of native contacts, $N_{\text{native}}^{\text{total}}$, for the whole chain and for each pair of strands *a* and *b*, $N_{\text{ab}}^{\text{total}}$, for the β-sheet, the β-clip, and the β-twist in the global energy minimum structures.

The total time duration for the simulation strongly depends on the topology of the model and the bias gap. The time used in the simulations is expressed in terms of reduced time, $t^* = t\sqrt{\epsilon/M\sigma^2}$ (Zhou and Karplus, 1999), 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 three β-strands, simulations were performed for $\sim 10^8$ collisions, which is equivalent to $t^* = 4 \times 10^5$ for the β-sheet model and $t^* = 2 \times 10^5$ for the β-clip and β-twist models. This indicates that the simulation speed for the β-sheet is twice as fast as that of the other two models. The simulation speed at large bias gaps is slightly faster than at small bias gaps for any given model, indicating that the larger the bias gap, the faster the protein folds. To highlight the folding events that occur early in the simulation, a logarithmic timescale was used to examine early folding behavior of proteins as well as folded behavior in the native state. Simulation results were recorded at linearly spaced time intervals of $\Delta t^* = 0.1$ for $t^* < 10$ and at logarithmically spaced time intervals of $\Delta \log t^* = 0.01$ for $t^* > 10$.

To determine the compactness of the chains during the simulations, the time-dependent squared radius of gyration, $R_g^2(t)$ was determined where

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

with *N* equal to the number of beads. The time-dependent bead positions, $\mathbf{r}_i(t)$ ($i = 1$ to *N*), are obtained by shifting the center of mass coordinates to the origin. The radius of gyration is widely used as an order parameter in studies of the polymer order-disorder collapse transition. At low temperatures, collapsed polymers tend to form a high-density conformation

TABLE 2 Native values of the reduced squared radius of gyration

	β-sheet ($T^* = 0.2$)		β-clip ($T^* = 0.24$)			β-twist ($T^* = 0.28$)		
	($g = 0.9$)	($g = 1.1$)	($g = 0.7$)	($g = 0.9$)	($g = 1.1$)	($g = 0.7$)	($g = 0.9$)	($g = 1.1$)
$R_g^2(\text{eq})/\sigma^2 N$	0.1197	0.1449	0.1444	0.1900	0.1967	0.1711	0.1855	0.1884
$R_g^2(\text{GM})/\sigma^2 N$	0.2350			0.2164			0.2163	

The native values of the reduced squared radius of gyration, $R_g^2(\text{eq})/\sigma^2 N$, where σ is the bead diameter and N is the number of beads, from the equilibrium simulations at $T^* = 0.2$, 0.24, and 0.28 for the β-sheet at $g = 0.9$ and 1.1, the β-clip at $g = 0.7$, 0.9 and 1.1, and the β-twist at $g = 0.7$, 0.9, and 1.1, respectively. Also included are the global energy minimum values of the reduced squared radius of gyration, $R_g^2(\text{GM})/\sigma^2 N$, for the three different types of β-strand peptides.

yielding a small value of the radius of gyration. In contrast, the squared radius of gyration for native β-strand peptides is larger than that in the collapsed state, because the native β-strands are long and have an elongated molecular shape. This indicates that a minimum value of the squared radius of gyration does not imply that the β-strands are in the native state. Thus, to characterize the nativeness of β-strands as collapsing progress, an alternative measure of the time-dependent squared radius of gyration is considered. This is the normalized squared radius of gyration, $F_R^*(t)$, which is defined by

$$F_R^*(t) \equiv 1 - \frac{R_g^2(\text{eq})}{R_g^2(t)}, \tag{2}$$

where $R_g^2(\text{eq})$ is the native value of the squared radius of gyration. The normalized squared radius of gyration measures the departure of the chain's squared radius of gyration from its value in the native state. It is slightly less than 1 when the chain is loosely packed or a random coil ($R_g^2(t) > R_g^2(\text{eq})$), equals 0 when the chain is in the native state ($R_g^2(t) = R_g^2(\text{eq})$), and is less than 0 when the chain collapses to very compact structure ($R_g^2(t) < R_g^2(\text{eq})$). Each β-strand peptide has different values of $R_g^2(\text{eq})$ depending on the temperature and the size of the bias gap. The native values of R_g^2 for the three β-strands at different bias gaps are obtained from the equilibrium simulations at $T^* = 0.2$, 0.24, and 0.28 for the β-sheet, the β-clip, and the β-twist, respectively. These values are summarized in Table 2 along with those for the global energy minimum structures. The $R_g^2(\text{eq})$ values for the native β-strands are smaller than the R_g^2 values in the global energy minimum structures, $R_g^2(\text{GM})$, because the β-strands are slightly bent or kinked in the native state, especially for the smaller bias gap models. This is due to the attraction between non-native contacts for $g < 1$, which disturbs the propensity of the β-strands to be straight. Note that the ratio of the native R_g^2 value to that in the global energy minimum structure, $R_g^2(\text{eq})/R_g^2(\text{GM})$, is much smaller for the β-sheet than for the other two models. This indicates that the planar shape of the β-sheet can be easily bent due to the non-native attractions, giving the native state a slightly rolled molecular structure. In fact, most real β-sheet structures are twisted as opposed to being perfectly planar or elongated molecular structures (Clothia, 1973; Pauling and Corey, 1951).

The progression of a conformation toward the native state can be monitored by introducing the time-dependent fraction of native contacts formed (Lazaridis and Karplus, 1997; Šali et al., 1994b), $Q(t)$, defined by

$$Q(t) = \frac{N_{\text{native}}(t)}{N_{\text{native}}^{\text{total}}}, \tag{3}$$

and further breaking down $Q(t)$ into the time-dependent fraction of inter-strand native contacts formed between strands a and b, $Q_{ab}(t)$, defined by

$$Q_{ab}(t) = \frac{N_{ab}^{\text{native}}(t)}{N_{ab}^{\text{total}}}, \tag{4}$$

where N_{native} represents the number of native contacts for the whole chain and N_{ab}^{native} is the number of inter-strand native contacts between strands a and b in a given conformation. The fraction of native contacts Q has a value between 0 and 1. When $Q = 1$ the model system can be regarded as being in the native structure, whereas when $Q \to 0$ the chain becomes a random coil that signifies the denatured state.

RESULTS

Characteristics of folding kinetics

We performed DMD simulations to investigate the folding kinetics of the three different β-strand peptides. The systems were quenched from the denatured state at $T^* = 2.0$ to the native states at $T^* = 0.2$, 0.24, and 0.28 for the β-sheet, the β-clip, and the β-twist, respectively. The time-dependent normalized squared radius of gyration, $\langle F_R^* \rangle$, is shown in Fig. 2 as a function of the reduced time on a logarithmic scale, $\log(1 + t^*)$: for the β-sheet at $g = 0.9$ and 1.1 (Fig. 2 *A*); for the β-clip at $g = 0.7$, 0.9, and 1.1 (Fig. 2 *B*); and for the β-twist at $g = 0.7$, 0.9, and 1.1 (Fig. 2 *C*). The quantity $\langle F_R^* \rangle$ is an average over 100 independent trajectories that start from different initial configurations. Simulation results are presented up to the reduced time of $t^* = 10^5$, because for $t^* > 10^5$ the β-peptides are very stable and remain in the folded state. The degree of compactness of the system can be monitored by examining how the F_R^* values change as folding progress. Recall that when $0 < F_R^* < 1$, the chain has a larger squared radius of gyration, R_g^2, than the native value of $R_g^2(\text{eq})$ (see Table 2), indicating that the system is loosely packed or a random coil. For $F_R^* = 0$, the system has the same R_g^2 value as in the native state, so that the system can be regarded as being in the native state. For $F_R^* < 0$, the system exhibits a highly collapsed state, yielding a kinked or rolled molecular structure with smaller value of R_g^2 than in the native state.

The $\langle F_R^*(t) \rangle$ versus $\log(1 + t^*)$ graph in Fig. 2 *A* for the β-sheet shows that $\langle F_R^* \rangle > 0$ initially, that it oscillates around zero when the native state is reached, and that the folding events for the $g = 0.9$ and $g = 1.1$ models are qualitatively the same. The initial conformations of the chains begin to collapse at $t^* = 10$ and fold into the native state by $t^* = 100$. The $\langle F_R^*(t) \rangle$ versus $\log(1 + t^*)$ graphs in Fig. 2, *B* and *C*, for the β-clip and the β-twist, respectively, are strikingly different. In contrast to the β-sheet, the $\langle F_R^* \rangle$ values for the β-clip and the β-twist rapidly decrease to a negative minimum value as the chain collapses and then

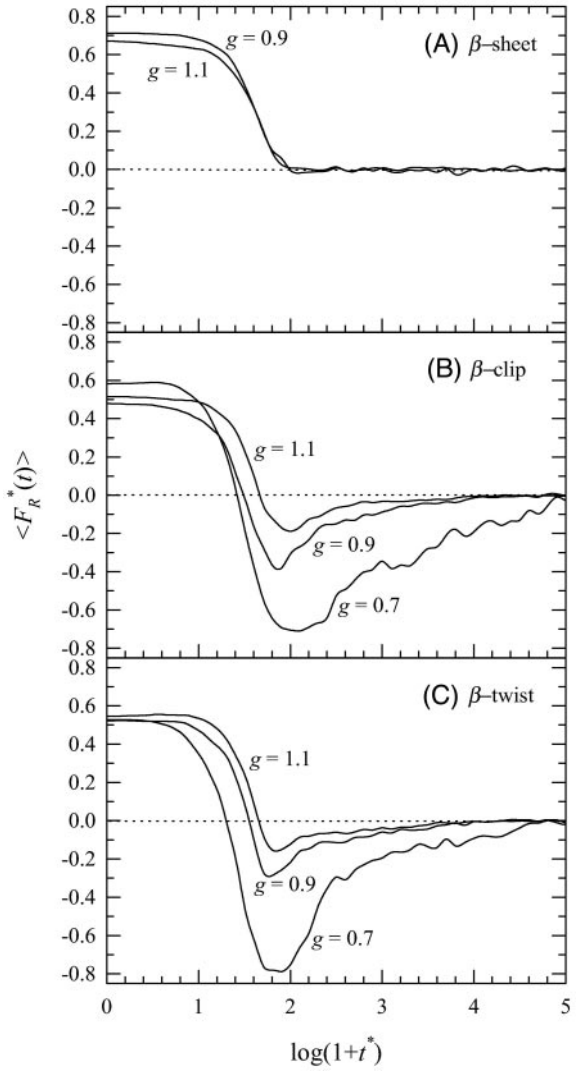

FIGURE 2 The average values of the normalized squared radius of gyration, $\langle F_R^*(t)\rangle$, as a function of the reduced time in logarithmic scale, $\log(1 + t^*)$: (*A*) the β-sheet at $g = 0.9$ and 1.1; (*B*) the β-clip at $g = 0.7$, 0.9, and 1.1; and (*C*) the β-twist at $g = 0.7$, 0.9, and 1.1.

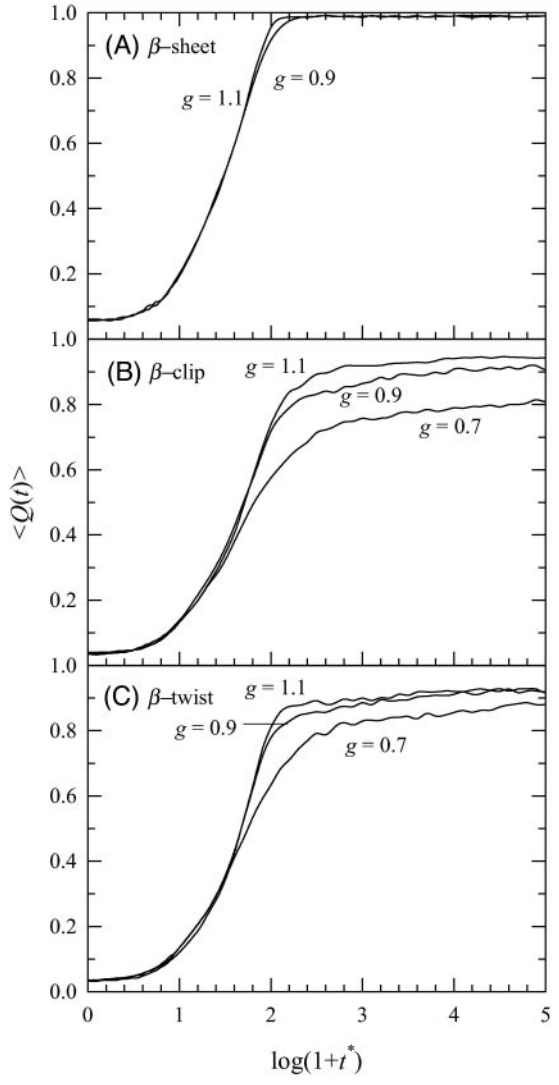

FIGURE 3 The average values of the fraction of native contacts, $\langle Q(t)\rangle$, as a function of the reduced time in logarithmic scale, $\log(1 + t^*)$, for: (*A*) the β-sheet at $g = 0.9$ and 1.1; (*B*) the β-clip at $g = 0.7$, 0.9, and 1.1; and (*C*) the β-twist at $g = 0.7$, 0.9, and 1.1.

gradually increase to zero as the native state is reached. The negative value for F_R^* means that the R_g^2 value of the collapsed chain is smaller than that of the native state at the given time. This suggests that the folding process for the β-clip and β-twist goes through two distinct stages. In the first stage, the initial random coil collapses to a compact structure. In the second stage, strand extension takes place, recovering the elongated molecular shape of the β-strand peptide. For the β-clip in Fig. 2 *B*, the minimum value of $\langle F_R^*\rangle$ decreases as g decreases, indicating that the smaller the value of g, the more compact the collapsed structure is. This is due to the relatively high overall strength of the non-native attractions among the beads in the smaller gap models. The strand extension is much quicker in the larger gap model. Similar behavior is seen in Fig. 2 *C* for the β-twist.

The time-dependent fraction of native contacts, $\langle Q\rangle$, is shown in Fig. 3 as a function of $\log(1 + t^*)$: for the β-sheet at $g = 0.9$ and 1.1 (Fig. 3 *A*); for the β-clip at $g = 0.7$, 0.9, and 1.1 (*B*); and for the β-twist at $g = 0.7$, 0.9, and 1.1 (*C*). The quantity $\langle Q\rangle$ is an average over 100 independent simulations with different independent initial configurations. For the β-sheet, no difference in the formation of native contacts between the $g = 0.9$ and $g = 1.1$ models is observed. The formation of native contacts and the chain collapse (compare to Fig. 2 *A*) are concurrent. This indicates that the collapse time and the folding time are identical for the β-sheet. The native states for both the bias gap models are located at $\langle Q\rangle \approx 1$. In contrast, for the β-clip in Fig. 3 *B* and for the β-twist in Fig. 3 *C*, 60–70% of the total native contacts (depending on the size of the bias gap) are built up

when $\langle F_R^* \rangle$ reaches its negative minimum at around $t^* = 100$ as seen in Fig. 2, *B* and 2 *C*. The $\langle Q \rangle$ values continuously increase thereafter and begin to saturate slowly to the native value during the strand extension process as $\langle F_R^* \rangle$ goes to zero. This reflects the result that for the β-clip and β-twist the collapse time and the folding time are different. For the β-clip, the native states are located at $\langle Q \rangle \approx 0.82$, 0.92, and 0.95 for $g = 0.7$, 0.9, and 1.1, respectively. As g increases, the $\langle Q \rangle$ value in the native state increases, and the speed of native contact formation is much faster. For the β-twist, the native states are located at $\langle Q \rangle \approx 0.88$, 0.92, and 0.93 for $g = 0.7$, 0.9, and 1.1, respectively. The reason that the native $\langle Q \rangle$ values are all different for the three β-strand models with the same bias gap is that the kinetic simulations are performed at different target temperatures.

More detailed information on the formation of native contacts can be obtained by considering the fraction of inter-strand native contacts, $\langle Q_{ab} \rangle$, between strands *a* and *b* where the strand number runs from 1 to 4. The β-sheet model has three possible strand-to-strand native interactions: 1–2, 2–3, and 3–4 (see the topology diagrams in Fig. 1 *A*). The time-dependent fraction of inter-strand native contacts, $\langle Q_{ab} \rangle$, for the β-sheet model is shown in Fig. 4 as a function of $\log(1 + t^*)$ for the bias gaps: $g = 0.9$ (Fig. 4 *A*) and $g = 1.1$ (Fig. 4 *B*). Each value of $\langle Q_{ab} \rangle$ is averaged over 100 independent simulations. For $g = 0.9$, $\langle Q_{12} \rangle$ and $\langle Q_{34} \rangle$ (the fraction of inter-strand native contacts experienced by the outer strands) increase before $\langle Q_{23} \rangle$ (the fraction of inter-strand native contacts experienced by the inner strands) increases as the system collapses. This indicates that assembly of the native state for the β-sheet at $g = 0.9$ starts with the formation of native contacts between strands 1 and 2 and between strands 3 and 4, followed by the formation of native contacts between strands 2 and 3. This is due to the fact that strands 1 and 4 are the chain's terminals, which are more flexible and easier to rearrange than the other strands. We find that $\langle Q_{12} \rangle \approx \langle Q_{23} \rangle \approx 1$ at $t^* = 160$, and $\langle Q_{34} \rangle \approx 1$ at $t^* = 200$. However, for $g = 1.1$ all $\langle Q_{ab} \rangle$ values increase concurrently and $\langle Q_{12} \rangle \approx \langle Q_{23} \rangle \approx \langle Q_{34} \rangle \approx 1$ at $t^* = 100$. The native secondary structure formation for the β-sheet at $g = 1.1$ starts at the same time as the chain collapses.

In marked contrast to the β-sheet, the formation of inter-strand native contacts for the β-clip is more complex as shown in Fig. 5 *A* at $g = 0.7$ and Fig. 5 *B* at $g = 0.9$. As shown in Fig. 1 *B* (top view topology diagram), there are six possible strand-to-strand native interactions for the β-clip: Q_{12}, Q_{14}, Q_{23}, and Q_{34} correspond to the four peripheral strand-to-strand interactions, and Q_{13} and Q_{24} correspond to the two diagonal strand-to-strand interactions. For $g = 0.7$, the formation of β-clip native contacts begins to develop at the ends of the chain; i.e., $\langle Q_{12} \rangle$ and $\langle Q_{34} \rangle$ increase first, as was the case in Fig. 4 *A* for the β-sheet at $g = 0.9$, because the chain's terminals (strands 1 and 4) are very flexible and easy to rearrange. All values of $\langle Q_{ab} \rangle$ increase initially as the chain collapses and saturate at their respective native values of $\langle Q_{ab} \rangle$, except for $\langle Q_{14} \rangle$ which continuously increases and reaches a value of $\langle Q_{14} \rangle \approx 0.85$ at $t^* = 10^5$. The native state is located at $t^* \geq 10^5$, the time at which $\langle Q_{14} \rangle$ starts to oscillate about 0.85. In the native state, the four peripheral values, $\langle Q_{12} \rangle$, $\langle Q_{14} \rangle$, $\langle Q_{23} \rangle$, and $\langle Q_{34} \rangle$ oscillate about their average value of $\langle Q \rangle \approx 0.82$, and the two diagonal values, $\langle Q_{13} \rangle$ and $\langle Q_{24} \rangle$ oscillate about 0.6, which is much smaller. Because the number of native contacts in the four peripheral strand-to-strand interactions is much larger than that in the two diagonal strand-to-strand interactions (see Table 1), this indicates that the larger the number of strand-to-strand native contacts, the faster the formation of native contacts in the strand-to-strand interactions. For $g = 0.9$, Fig. 5 *B* shows similar information on the formation of inter-strand native contacts. The speed of the formation of native contacts is faster than in the $g = 0.7$ case. The native state is located at $t^* \geq 10^4$ when $\langle Q_{14} \rangle$ reaches its native value of 0.95. Note that in the native state the diagonal strand interactions $\langle Q_{13} \rangle$ and $\langle Q_{24} \rangle$ oscillate at around 0.8, which is

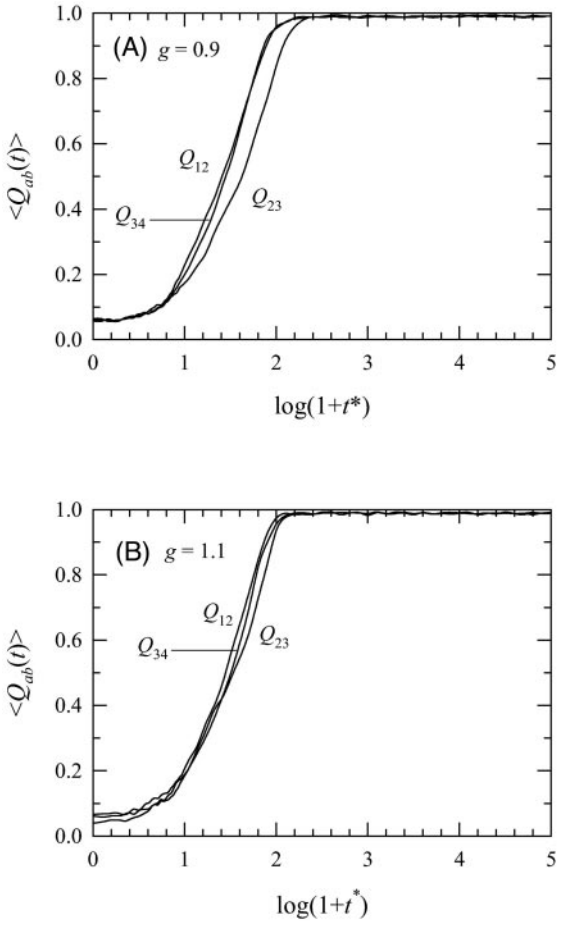

FIGURE 4 The average values of the fraction of inter-strand native contacts, $\langle Q_{ab}(t) \rangle$, where *a* and *b* are the strand number, for the β-sheet as a function of the reduced time in logarithmic scale, $\log(1 + t^*)$, for the bias gaps: (*A*) $g = 0.9$; (*B*) $g = 1.1$.

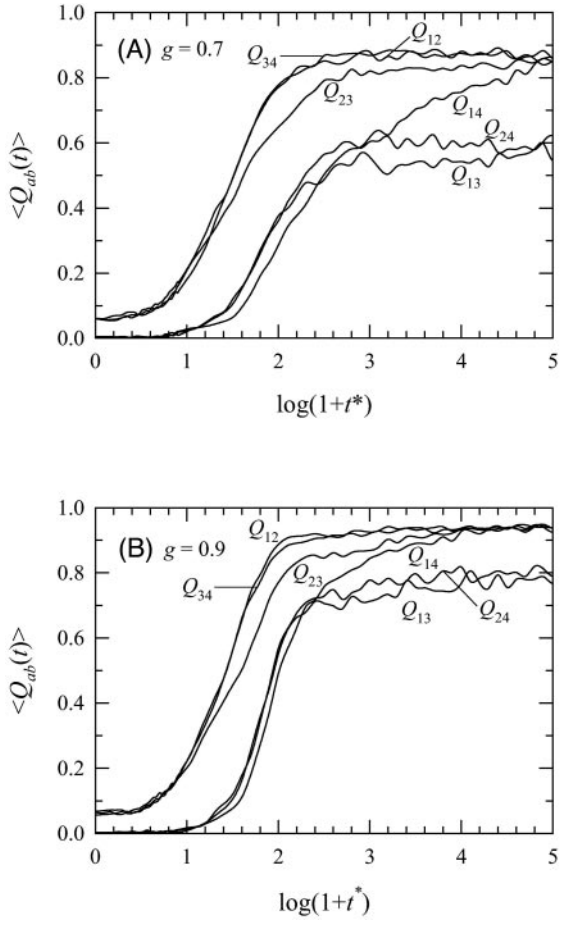

FIGURE 5 The average values of the fraction of inter-strand native contacts, $\langle Q_{ab}(t)\rangle$, where a and b are the strand number, for the β-clip as a function of the reduced time in logarithmic scale, $\log(1 + t^*)$, for two selected values of the bias gap: (*A*) $g = 0.7$ and (*B*) $g = 0.9$.

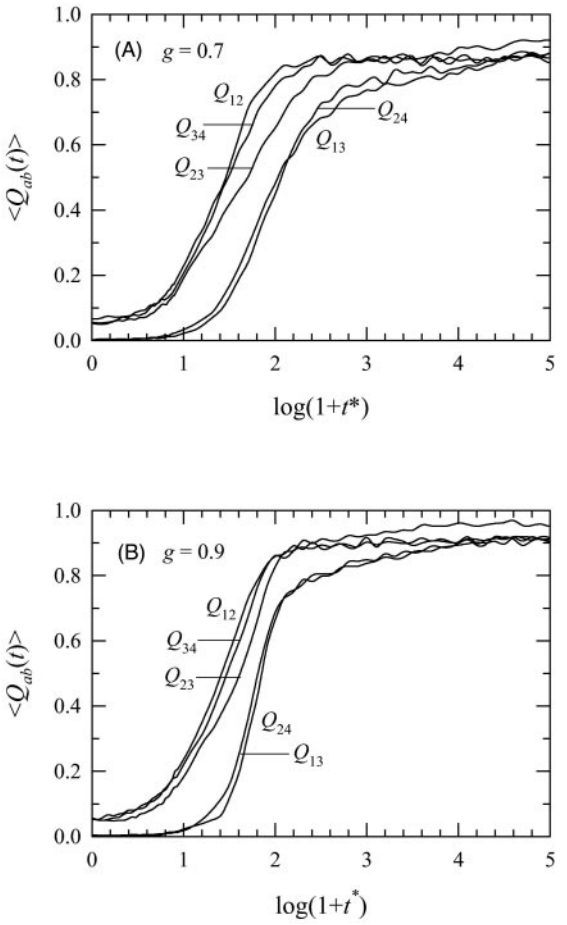

FIGURE 6 The average values of the fraction of inter-strand native contacts, $\langle Q_{ab}(t)\rangle$, where a and b are the strand number, for the β-twist as a function of the reduced time in logarithmic scale, $\log(1 + t^*)$, for two selected values of the bias gap: (*A*) $g = 0.7$; (*B*) $g = 0.9$.

significantly closer to the native value for the whole chain of $\langle Q\rangle \approx 0.92$ than occurs in the $g = 0.7$ case. This is based on the fact that the larger the bias gap, the more the native state is favored over the non-native state. In fact, the large values of $\langle Q_{13}\rangle$, $\langle Q_{14}\rangle$, and $\langle Q_{24}\rangle$ are essential to maintain the β-clip structure, because without these strand-to-strand native interactions, the model β-clip reduces to the β-sheet.

The formation of inter-strand native contacts for the β-twist is shown in Fig. 6 *A* at $g = 0.7$ and Fig. 6 *B* at $g = 0.9$. As shown in Fig. 1 *C* (top view topology diagram), there are five possible strand-to-strand native interactions for the β-twist: 1–2, 1–3, 2–3, 2–4, and 3–4. For $g = 0.7$ in Fig. 6 *A*, native contacts between strands 1 and 2 and between strands 3 and 4 form rapidly as folding progresses, but native contacts between strands 1 and 3 and between strands 2 and 4 form slowly. Native contacts between strands 2 and 3 form at an intermediate speed. After the initial rapid increase, $\langle Q_{12}\rangle$ and $\langle Q_{34}\rangle$ oscillate about 0.88 in the native state, whereas $\langle Q_{13}\rangle$ and $\langle Q_{24}\rangle$ continuously increase to 0.88. The native state is located at $t^* \geq 3.2 \times 10^4$ when $\langle Q_{13}\rangle$ and $\langle Q_{24}\rangle$ begin to oscillate about 0.88. For $g = 0.9$ in Fig. 6 *B*, it is immediately apparent that there are three fast rates in the formation of inter-strand native contacts (Q_{12}, Q_{23}, and Q_{34}) and two slow rates in the formation of inter-strand native contacts (Q_{13} and Q_{24}). The quantities $\langle Q_{13}\rangle$ and $\langle Q_{24}\rangle$ increase faster than they do in the $g = 0.7$ model. The native state is located at $t^* \geq 10^4$ when $\langle Q_{13}\rangle$ and $\langle Q_{24}\rangle$ oscillate about 0.92. For the β-twist, the large values of $\langle Q_{13}\rangle$ and $\langle Q_{24}\rangle$ are key factors in the conformation of the native β-twist structure, because without these strand-to-strand native interactions, the model β-twist reduces to the β-sheet.

The evolution of the formation of inter-strand native contacts for the β-sheet model is reflected in the snapshots in Fig. 7 at selected times that display typical folding events for: $g = 0.9$ (Fig. 7 *A*) and $g = 1.1$ (Fig. 7 *B*). As mentioned above, for $g = 0.9$ native contacts start forming at the ends of the chain as can be clearly seen in the structure at $t^* = 32$. The native state is observed at $t^* \geq 100$. For $g = 1.1$, the chain collapse and the formation of native contacts are

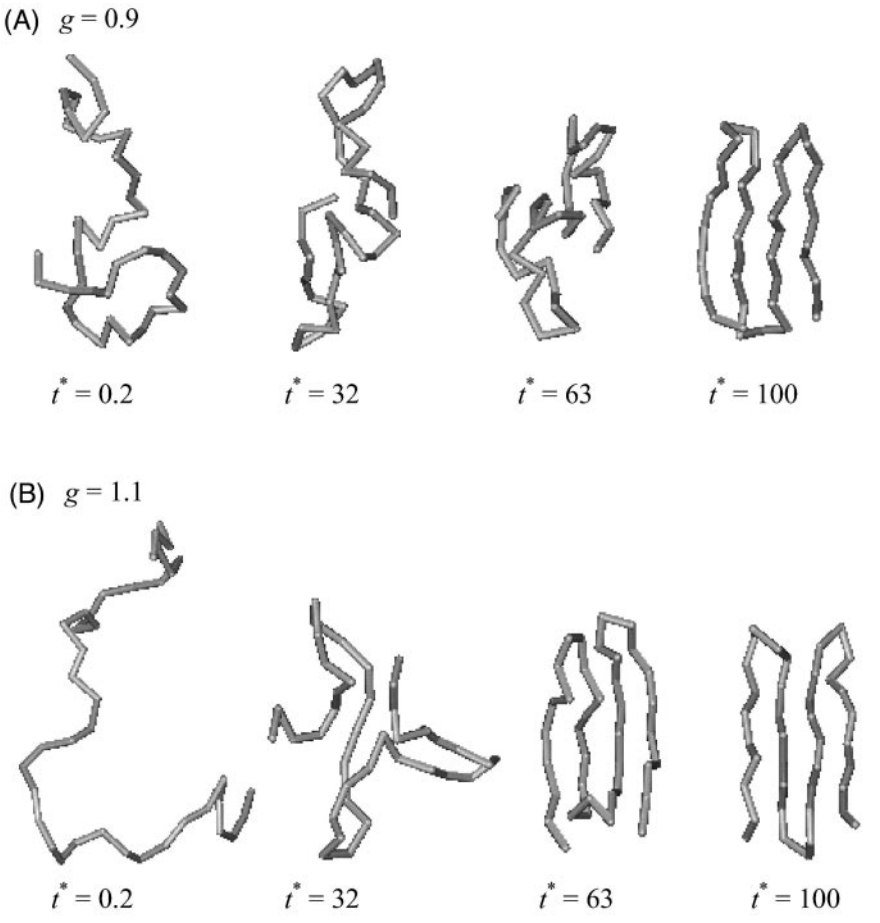

FIGURE 7 Snapshots of the kinetic chain conformation of the β-sheet at selected times, $t^* = 0.2$, 32, 63, and 100 for the bias gaps: (*A*) $g = 0.9$; (*B*) $g = 1.1$.

concurrent. An example of the collapsed chain is shown at $t^* = 32$. A native-like structure is observed at $t^* = 63$. A well-ordered native β-sheet appears to exist at $t^* \geq 100$.

Fig. 8 shows snapshots for the β-clip at selected times as the chain evolves toward the native state. For $g = 0.7$ in Fig. 8 *A*, the four snapshots are taken after $\langle F_R^* \rangle$ reaches its minimum as seen in Fig. 2 *B*. The structure at $t^* = 100$ corresponds to the collapsed chain at the minimum value of $\langle F_R^* \rangle$. It exists for a very short time. This collapsed chain is a transient in protein folding (Silow and Oliveberg, 1997) and can be regarded as a kinetic disordered globule because it has a Q value, $\langle Q \rangle \approx 0.6$ (see Fig. 3 *B*), that is similar to that of a thermodynamic disordered globule (Jang et at., 2002) but with a smaller R_g^2. The existence of a disordered compact structure is also reported in the kinetic study of a model three-helix bundle protein (Zhou and Karplus, 1997b, 1999). At $t^* = 1.6 \times 10^3$ and at $t^* = 1.3 \times 10^4$, we see the structures of two typical intermediates; these are intermediate I_1 and intermediate I_2, respectively (see below). These intermediates are found in the strand extension process after the initial chain collapse and have very long lifetimes. An example of the native structure is shown at $t^* = 10^5$. For $g = 0.9$, the evolution of the chain toward the native state is totally different. No kinetic disordered globule state is observed when $\langle F_R^* \rangle$ goes to the minimum at $t^* \approx 63$. Instead, a slightly bent β-sheet structure occurs as shown in Fig. 8 *B* at $t^* = 63$. This β-sheet-like structure is a transient in folding trajectories toward the β-clip. Examples of the intermediate I and the native structure are observed at $t^* = 160$ and $t^* = 10^4$, respectively.

Fig. 9 shows snapshots for the β-twist at selected times as the chain evolves toward the native state. The behavior displayed in Fig. 9 *A* for $g = 0.7$ is similar to that shown in Fig. 8 *A* for the β-clip. The collapsed chain at the minimum value of $\langle F_R^* \rangle$ is shown at $t^* = 63$; this is a transient in protein folding (Silow and Oliveberg, 1997) and can be regarded as a kinetic disordered globule. The kinetic disordered globule has more non-native interactions than the thermodynamic disordered globule giving a smaller R_g^2. Due to the disordered nature of this state, no native-state secondary structure is observed. As time increases, the native contacts begin to form as part of the strand extension process, leading to the native structure. This process is very slow because the collapsed chain undergoes a subtle conformational change to the native state via intermediates that have long lifetimes. Examples of two typical intermediates, the intermediate I_1 and the intermediate I_2 are observed at $t^* = 250$ and at $t^* = 1.3 \times 10^3$, respectively. The native

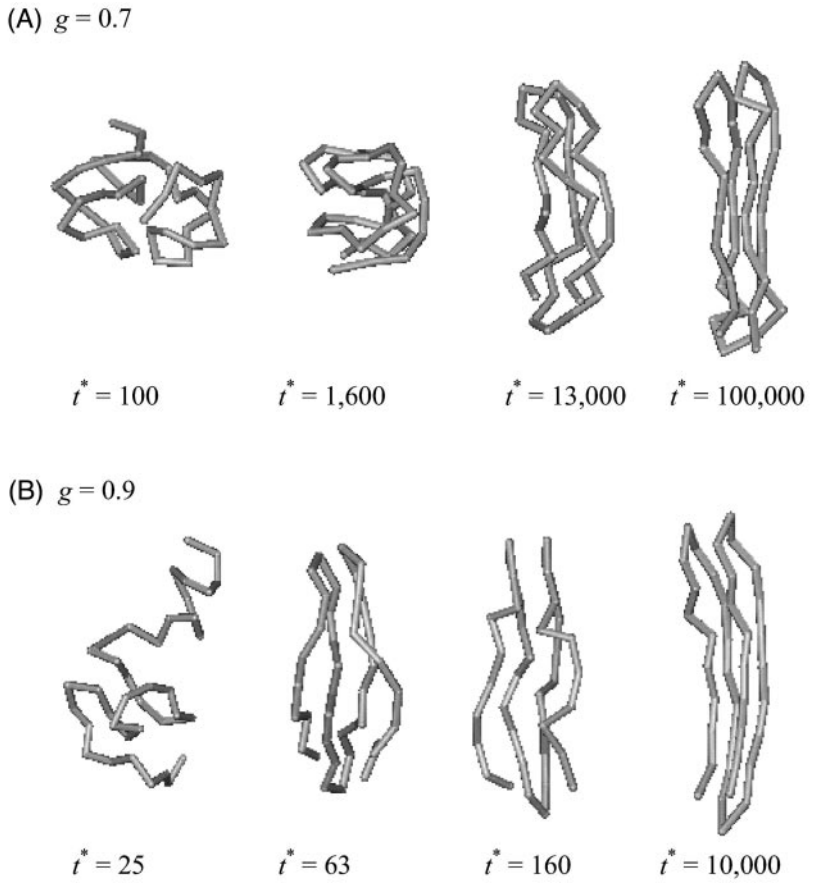

FIGURE 8 Snapshots of the kinetic chain conformation of the β-clip for two values of the bias gap: (*A*) $g = 0.7$ at times $t^* = 100$, 1.6×10^3, 1.3×10^4, and 10^5; (*B*) $g = 0.9$ at times $t^* = 25$, 63, 160, and 10^4.

structure is observed at $t^* = 4 \times 10^4$. For $g = 0.9$ in Fig. 9 *B*, no kinetic disordered globule state is observed when $\langle F_R^* \rangle$ goes to the minimum at $t^* \approx 63$. (This was also the case for the β-clip at $g = 0.9$). Instead, a slightly bent β-sheet structure is observed at $t^* = 63$. The structures of the intermediate *I* (see below) are shown at $t^* = 320$ and at $t^* = 3.2 \times 10^3$. The native structure is observed at $t^* = 10^4$.

Intermediates and folding pathways

It is of interest to examine the intermediates that occur as the folding progresses in more detail. The contour graph in Fig. 10 *A* shows the population distribution of the fraction of native contacts, *Q*, as a function of reduced time on a logarithmic scale, $\log(1 + t^*)$, for the β-sheet at $g = 0.9$. The population distribution of *Q* is calculated at selected logarithmic time intervals of $\log(1 + t^*) = 0.1$ for all 100 trajectories. The graph clearly shows highly populated states as those enclosed by many contour lines. The β-sheet at $g = 0.9$ is a fast-folding protein, and the folding trajectories clearly show a fast-track or bottleneck pathway of kinetic folding between two highly populated states: the initial random coil state at $t^* < 10$ and $Q \approx 0.08$ and the native state at $t^* > 100$ and $Q \approx 1$. More evidence of fast-track folding is shown in Fig. 10 *B*, which shows the population distribution of F_R^* versus *Q* averaged over all 100 trajectories and all times. The peaks indicate the highly populated states. The peak at $F_R^* \approx 0.8$ and $Q \approx 0.1$ corresponds to the initial random coil state and the other peak at $F_R^* \approx 0$ and $Q \approx 1$ corresponds to the native state. The kinetic folding mechanism for the β-sheet at $g \geq 0.9$ is represented in Fig. 10 *C*.

The contour graph for the distribution of *Q* versus $\log(1 + t^*)$ for the β-clip at $g = 0.7$ is shown in Fig. 11 *A*, and the corresponding population distribution of F_R^* versus *Q* from all 100 trajectories is shown in Fig. 11 *B*. In the contour graph of Fig. 11 *A*, the random coil state is observed in the range $0 < \log(1 + t^*) < 0.8$ and $0 < Q < 0.1$. Two long-lived populations are observed in the range $2.6 < \log(1 + t^*) < 3.4$ at $Q \approx 0.76$ and $3.4 < \log(1 + t^*) < 5$ at $Q \approx 0.8$; these correspond to the intermediates I_1 and I_2, respectively. Because the native state is located at $t^* \geq 10^5$ as shown in Fig. 5 *A* for $g = 0.7$ and the results are shown within the time scale up to $\log(1 + t^*) = 5$, the population distribution of *Q* for the native state is not observed in the contour graph. In Fig. 11 *B*, the peak at $F_R^* \approx -1$ and $Q \approx 0.76$ corresponds to the intermediate I_1, which can be regarded as a partially-ordered globule as shown in the thermodynamic study of the same model (Jang et al., 2002). It has a small value of R_g^2, a large value of *Q*,

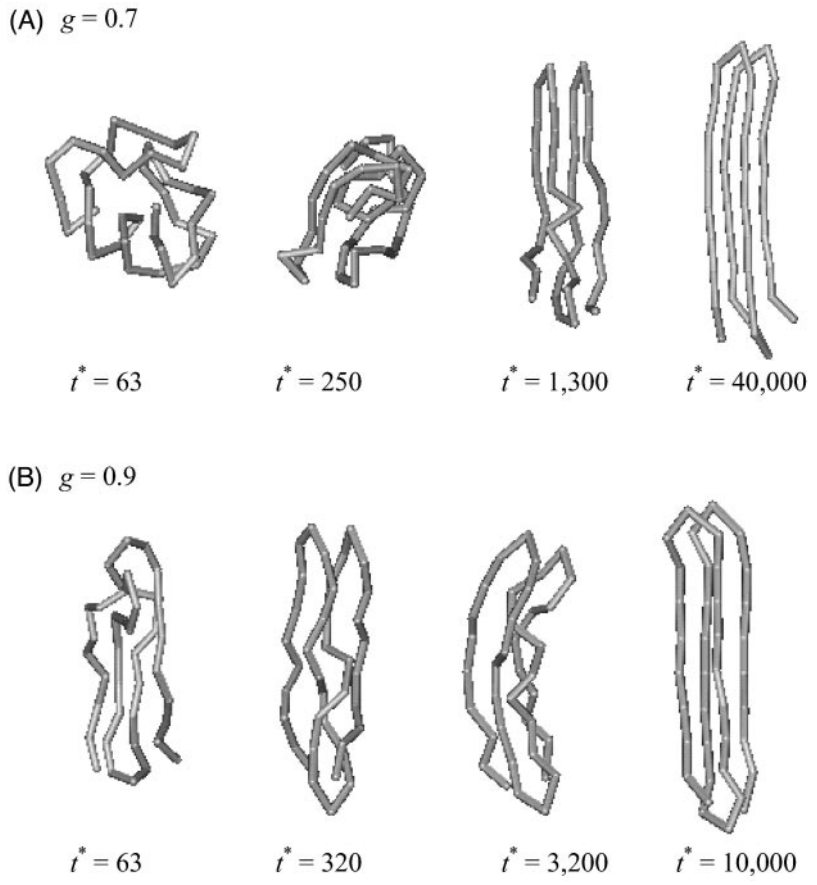

FIGURE 9 Snapshots of the kinetic chain conformation of the β-twist for two values of the bias gap: (*A*) $g = 0.7$ at times $t^* = 63$, 250, 1.3×10^3, and 4.0×10^4; (*B*) $g = 0.9$ at times $t^* = 63$, 320, 3.2×10^3, and 10^4.

and the rolled or kinked structure seen in Fig. 8 *a* at $t^* = 1.6 \times 10^3$. This partially-ordered globule state occurs because the chain becomes trapped in a conformation dominated by the attraction between non-native contacts, thus disturbing the tendency to order into the native state. Note that the intermediate I_1 is highly populated at $t^* \approx 10^3$ and has a value $F_R^* \approx -1$, which is much smaller than the value of $\langle F_R^* \rangle = -0.4$ at $t^* \approx 10^3$ seen in Fig. 2 *B*, which is the average value of F_R^* for all trajectories. This suggests that not all 100 trajectories fold via the intermediate I_1. In fact, 40% of the trajectories are trapped in the intermediate I_1, and the rest of the trajectories fold via the other intermediate, I_2, which is located at $F_R^* \approx -0.2$ and $Q \approx 0.8$ in Fig. 11 *B*. The peak for intermediate I_2 is obscured by the peak of the highly populated native state at $F_R^* \approx 0$ and $Q \approx 0.82$. The intermediate I_2 is an ordered globule; it has a larger R_g^2 than the intermediate I_1 and has almost the same Q as the native state. The intermediate I_2 has a loosely packed and slightly bent structure as seen in Fig. 8 *A* at $t^* = 1.3 \times 10^4$. There are two possible folding pathways for the collapsed chain from the state in which $\langle F_R^* \rangle$ is at its minimum. The first is that it could further collapse to become trapped in the intermediate I_1, which can fold either directly into the native state or via the intermediate I_2 by elongating its strands. The other pathway is that it can fold via the intermediate I_2 to the native state. The kinetic folding scheme for the β-clip at $g = 0.7$ is represented in Fig. 11 *C*.

The folding kinetics for the β-clip at $g = 0.9$ is very different than that at $g = 0.7$. Fig. 12 *A* shows the contour plot of the population distribution of Q versus $\log(1 + t^*)$. The initial random coil state is located at $0 < \log(1 + t^*) < 0.6$ and $0.05 < Q < 0.1$. There is a short-lived intermediate *I* located at $0.75 < Q < 0.85$ and $2.1 < \log(1 + t^*) < 2.7$. The highly populated native state is located at $t^* \geq 10^4$ with $Q \approx 0.9$. The properties of intermediate *I* are the same as those of the intermediate I_2 for the $g = 0.7$ model. In Fig. 12 *B*, it is located at $F_R^* \approx -0.2$ and $Q \approx 0.8$, but its peak is obscured by the native peak at $F_R^* \approx 0$ and $Q \approx 0.9$. The intermediate *I* is an ordered globule with a native-like secondary structure as seen in Fig. 8 *B* at $t^* = 160$. Almost 80% of the trajectories become trapped in the intermediate *I*, and the rest fold directly to the native state via a fast-track folding pathway. The kinetic folding scheme for the β-clip at $g = 0.9$ is represented in Fig. 12 *C*.

The folding kinetics for the β-twist at $g = 0.7$ is qualitatively similar to that of the β-clip at $g = 0.7$ (not shown in figures). The kinetic folding of the β-twist involves two intermediates I_1 and I_2. However, unlike the β-clip at $g = 0.7$ the intermediate I_1 is short-lived and relatively unpopulated in the F_R^* versus Q plane. Approximately, 20% of the

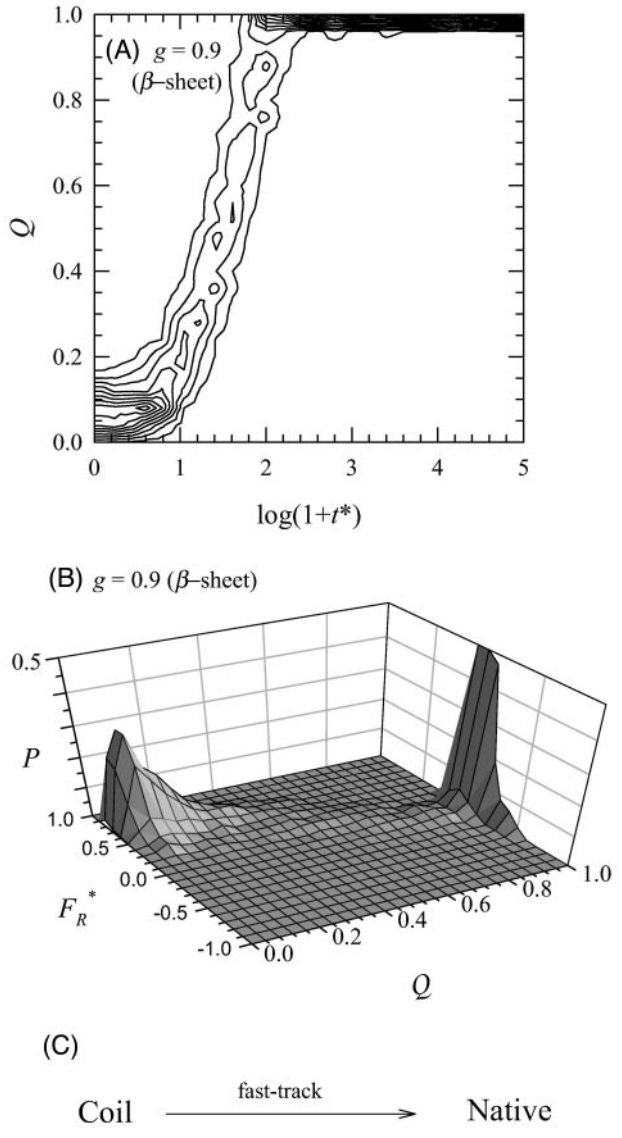

FIGURE 10 (*A*) The population distribution of the fraction of native contacts, Q, as a function of the reduced time in logarithmic scale, $\log(1 + t^*)$; (*B*) the population distribution of the normalized squared radius of gyration, F_R^*, versus the fraction of native contacts, Q; (*C*) the kinetic folding mechanism for the β-sheet at $g = 0.9$.

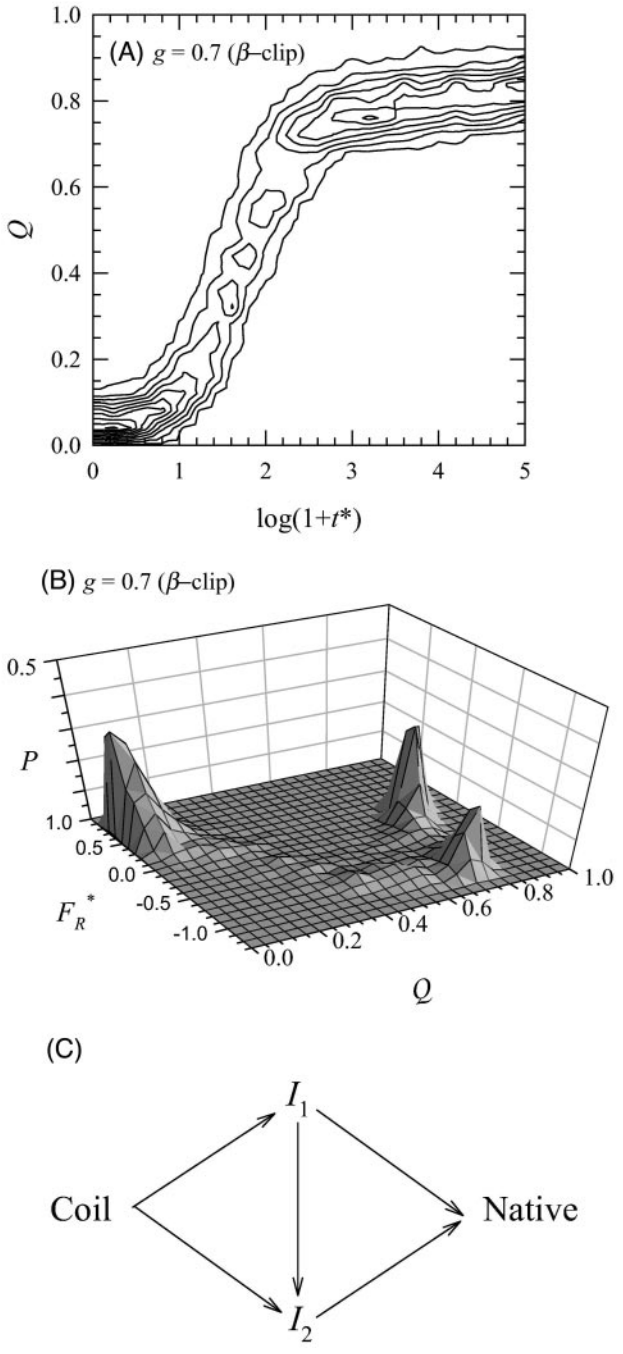

FIGURE 11 (*A*) The population distribution of the fraction of native contacts, Q, as a function of the reduced time in logarithmic scale, $\log(1 + t^*)$; (*B*) the population distribution of the normalized squared radius of gyration, F_R^*, versus the fraction of native contacts, Q; (*C*) the kinetic folding mechanism for the β-clip at $g = 0.7$.

trajectories are trapped in intermediate I_1, most of the trajectories fold into intermediate I_2, and none of the trajectories fold via the fast-track folding pathway. Therefore, the kinetic folding scheme for the β-twist at $g = 0.7$ is the same as that of the β-clip at $g = 0.7$ (Fig. 11 *C*). For the β-twist at $g = 0.9$, the folding kinetics is qualitatively similar to that of the β-clip at $g = 0.9$ (not shown in figures). The kinetic folding trajectories include the fast-track folding pathway as well as trapping in the intermediate I. However in comparison with the β-clip, fewer trajectories (60%) are trapped in the intermediate I. The kinetic folding pathway for the β-twist at $g = 0.9$ may be represented in the same way as that for the β-clip at $g = 0.9$ (Fig. 12 *C*).

Folding speed and topology

In this section, we investigate the influence of native-state topology on the folding kinetics for the three model pro-

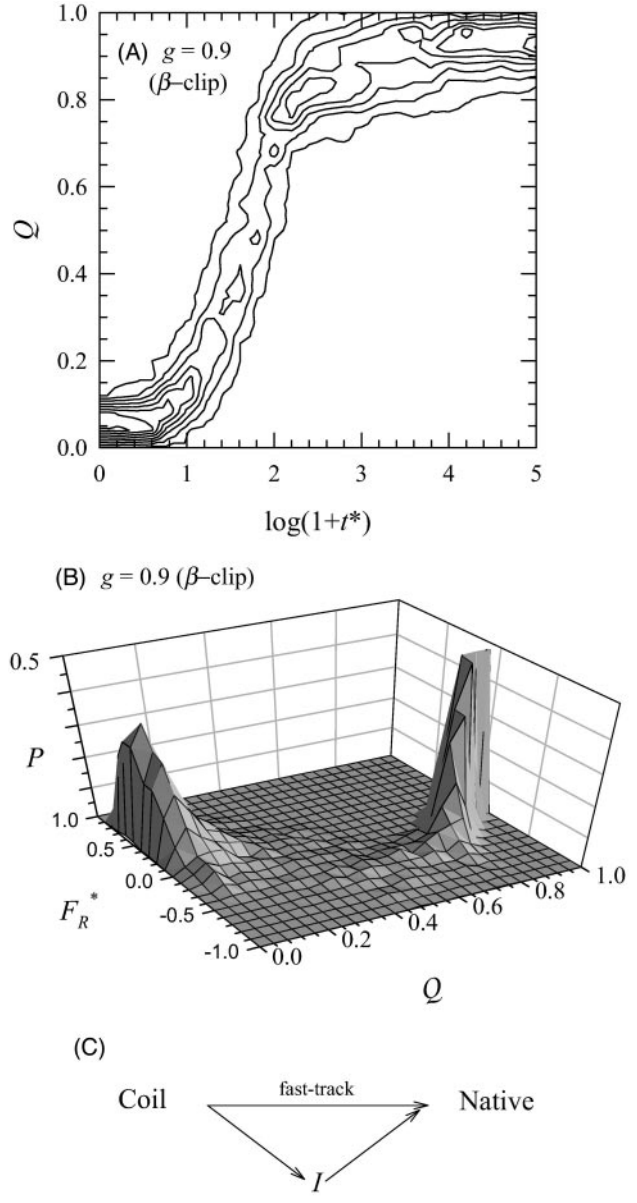

FIGURE 12 (*A*) The population distribution of the fraction of native contacts, Q, as a function of the reduced time in logarithmic scale, $\log(1 + t^*)$; (*B*) the population distribution of the normalized squared radius of gyration, F_R^*, versus the fraction of native contacts, Q; (*C*) the kinetic folding mechanism for the β-clip at $g = 0.9$.

teins. The topological complexity of proteins as it approaches the native state is measured in terms of the contact order (Klimov and Thirumalai, 1998), S^{CO}, a concept first introduced by Plaxco et al. (1998). The contact order (*CO*) is the average sequence distance at time t between all residue pairs participating in a native contact divided by the total sequence length and is defined by

$$S^{CO}(t) = \frac{1}{N_{native}(t)} \sum_{i<j}^{N} \frac{j - i}{N}, \qquad (5)$$

where $j - i$ is the sequence length between beads i and j that experience a native contact at time t, N is the number of beads in the chain, and $N_{native}(t)$ is the number of native contacts for the whole chain at time t. The contact orders of the three model proteins in the global energy minimum structures are 0.2564, 0.4098, and 0.3641 for the β-sheet, the β-clip, and the β-twist, respectively. This makes sense. Reference to the topology diagram in Fig. 1 shows that the average distance along the sequence between beads in contact in the global energy minimum structure is shortest for the β-sheet and longest for the β-clip.

The time-dependent contact order, $\langle S^{CO}(t)\rangle$, is shown in Fig. 13 *A* as a function of reduced time on a logarithmic scale, $\log(1 + t^*)$, for the β-sheet at $g = 0.9$ and 1.1, the β-clip at $g = 0.7$ and 0.9, and the β-twist at $g = 0.7$ and 0.9. The quantity $\langle S^{CO}(t)\rangle$ is an average over 100 independent simulations with different independent initial configurations. For all three model proteins, the $\langle S^{CO}\rangle$ values in the native state are almost the same as in the global energy minimum structures regardless of the size of the bias gap. Thus, in the native state $\langle S^{CO}\rangle_{\beta\text{-sheet}} < \langle S^{CO}\rangle_{\beta\text{-twist}} < \langle S^{CO}\rangle_{\beta\text{-clip}}$. For the β-sheet, the contact order dynamics and the native state contact order, $\langle S^{CO}\rangle = 0.2562$, are the same in the $g = 0.9$ and $g = 1.1$ models. In contrast, for the β-clip, the $\langle S^{CO}\rangle$ values continuously increase to the native value for $g = 0.7$, although they quickly saturate to the native value for $g = 0.9$. For both the bias gap models, $\langle S^{CO}\rangle = 0.4077$ in the native state. Similar behavior is seen in the β-twist, but the native state $\langle S^{CO}\rangle$ value is smaller at $\langle S^{CO}\rangle = 0.3640$.

Fig. 13 *B* shows the natural logarithm of the inverse reduced folding time, $\ln(1/t_f^*)$, versus the contact order in the native state, $\langle S^{CO}\rangle$, for the three models with different bias gaps. The reduced folding time, t_f^*, is the mean first passage time (MFPT) to the native state, which is defined to be the average over 100 independent simulations of the time to achieve the native state Q value. It is clear from Fig. 13 *B* that the folding speed depends not only on the size of the bias gap but also on the contact order for the different topologies. This supports the idea first introduced by Plaxco et al. (1998) that the natural logarithm of the folding rate and the contact order are strongly correlated (Baker, 2000; Grantcharova et al., 2001; Muñoz and Eaton, 1999; Plaxco et al., 1998). Fig. 13 *B* shows that both the native-state topology and the value of the bias gap play significant roles in determining the mechanism of folding kinetics for proteins. At the same bias gap ($g = 0.9$) for the three models, the β-sheet has the highest folding speed with the smallest contact order, whereas the β-clip has the lowest folding speed with the largest contact order. We also see that the folding speed increases as the bias gap increases. The line represents a linear fit through the data points. The linear relationship observed between the natural logarithm of the inverse reduced folding time and the contact order is evidence for a strong correlation with a correlation coefficient

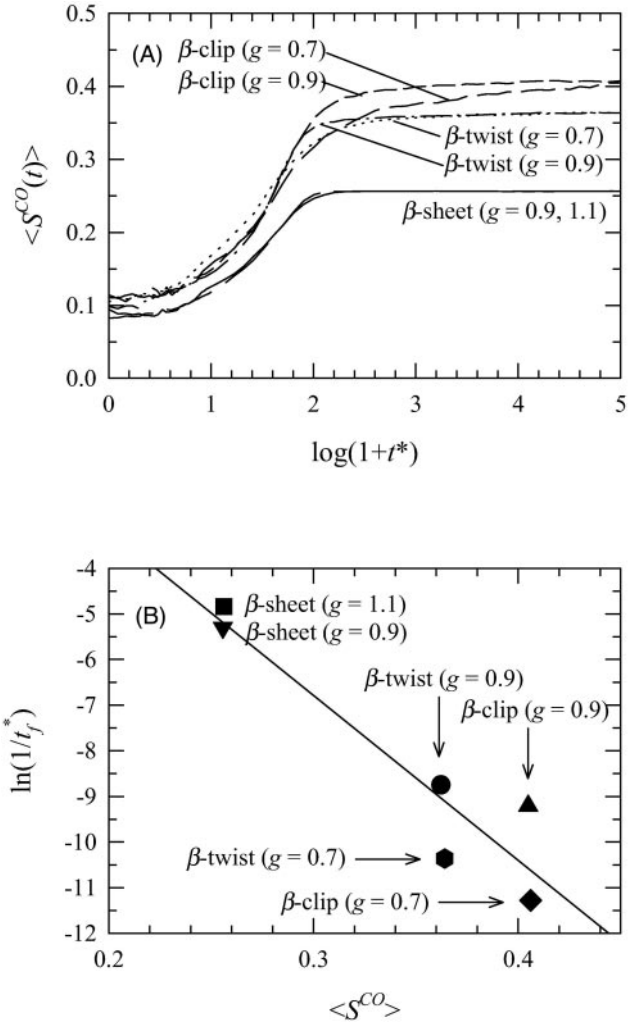

FIGURE 13 (*A*) The average values of the contact order, $\langle S^{CO}(t)\rangle$, as a function of the reduced time in logarithmic scale, $\log(1 + t^*)$; (*B*) the natural logarithm of the inverse reduced folding time, $\ln(1/t_f^*)$, versus the contact order, $\langle S^{CO}\rangle$, in the native state for the β-sheet at $g = 0.9$ and 1.1, the β-clip at $g = 0.7$ and 0.9, and the β-twist at $g = 0.7$ and 0.9.

of $r = -0.94$. Recently, Koga and Takada (2001) compared their folding simulations on 18 small proteins interacting via a Go-type potential with contact order predictions. They found that the folding rates are moderately correlated with the relative contact order (*RCO*, similar to S^{CO} in Eq. 5) but more significantly correlated with the absolute contact order ($ACO = RCO \times N$, where N is number of residues). Because we have six data points we cannot provide any conclusive answer regarding which parameter is better. Several new parameters such as the long-range order (Gromiha and Selvaraj, 2001) and the total contact distance (Zhou and Zhou, 2002) have also been proposed recently, which might give a better correlation.

To ascertain whether these models obey single-exponential folding kinetics, we investigated the first passage time (*FPT*) distributions over 100 independent simulations for each model. For $g = 0.9$, the logarithmic *FPT* probability (Abkevich et al., 1994; Kaya and Chan, 2002), $\ln[P(FPT)]$, as a function of reduced time, t^*, is shown in Fig. 14 *A* for the β-sheet and Fig. 14 *B* for the β-clip and β-twist. For the β-sheet, it is clearly seen from Fig. 14 *A* that the folding kinetics is single exponential, because the data points in $\ln[P(FPT)]$ are well fitted by the solid line and are strongly correlated with a correlation coefficient of $r = -0.96$. The folding kinetics for the β-sheet at $g = 1.1$ is also single exponential with a correlation coefficient of $r = -0.97$ (data not shown). However, for the β-clip and β-twist, the folding kinetics are not single exponential, having correlation coefficients of $r = -0.6$ and -0.64, respectively. This is due to the low target temperatures used in the kinetic simulations. The kinetic simulations for the β-sheet are performed at $T^* = 0.2$, which is located in the middle of the native-state region, whereas the simulations for the β-clip and β-twist are at temperatures $T^* = 0.24$ and 0.28, respectively, which are located near the liquid-to-solid transition (Jang et al., 2002). Thus, the β-clip and β-twist may be trapped in local energy minima at these low temperatures causing folding to the native state to be very slow as predicted by Abkevich et al. (1994).

DISCUSSION

We have studied the folding kinetics for three different types of off-lattice four-strand antiparallel β-strand protein models: the β-sheet, the β-clip, and the β-twist. Discontinuous molecular dynamics simulations of these models have been performed for different sizes of the bias gap g, a measure of the strength of the native contacts relative to that of the non-native contacts. The bias gap g in the models is clearly seen to be an important factor in controlling the kinetic behavior of the model proteins. The β-strand peptides are able to fold toward the native state from the very large number of random coil states via diverse folding trajectories as suggested by the modern view of protein folding (Dill and Chan, 1997; Dobson and Karplus, 1999; Lazaridis and Karplus, 1997; Šali et al., 1994a,b) and supported by simulations of simplified models (Dill et al., 1995; Karplus and Šali; 1995; Shakhnovich, 1996).

A recent study (Kaya and Chan, 2002) indicates that if a model's thermodynamics do not satisfy a calorimetric two-state criterion, its folding kinetics are not likely to be protein-like. In the thermodynamic study of this model (Jang et al., 2002), we observed that for the smaller bias gap models the collapse transition is not calorimetrically two-state. However for the larger bias gap models, the transition becomes calorimetrically two-state and displays thermodynamic cooperativity. Clearly, minimalist models based on Go-type potentials (Go and Taketomi, 1978, 1979; Taketomi et al., 1975; Ueda et al., 1978) have limitations that prevent them from mimicking some aspects of the folding of real proteins. Nevertheless, Go-type protein models yield

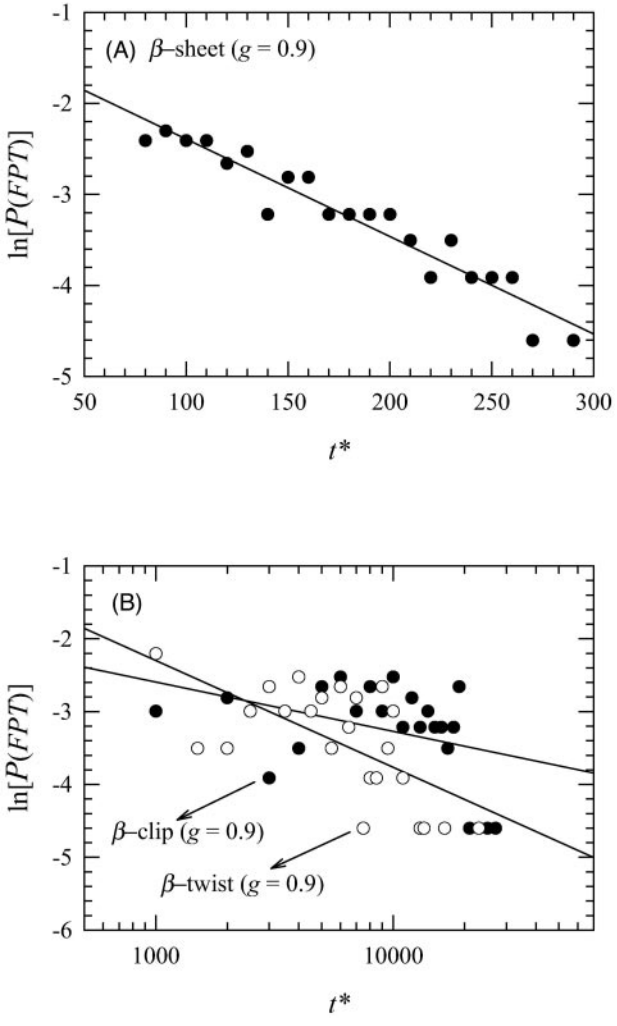

FIGURE 14 Logarithmic first passage time (*FPT*) probability, ln-[*P*(*FPT*)], as a function of reduced time, *t**: (*A*) for the β-sheet at $g = 0.9$; (*B*) for the β-clip and β-twist at $g = 0.9$.

a number of insights into the basic principles behind thermodynamics and kinetics of protein folding (Clementi et al., 2000; Dokholyan et al., 1998, 2000; Shimada et al., 2001; Zhou and Karplus, 1997a,b, 1999; Zhou and Linhananta, 2002).

A qualitative difference between the results for the β-sheet model and those for the β-clip and β-twist models is the existence of kinetic intermediates. For the β-sheet, the folding is via a fast-track folding pathway. The result that no intermediate is observed is consistent with experimental studies showing that small β-sheet proteins exhibit a two-state transition without populating partially folded intermediates (Capaldi and Radford, 1998). The inter-strand native contacts are partially formed at an early stage in the chain collapse and diffuse to the native state as the chain collapses. This indicates that the kinetic folding mechanism for the β-sheet can be characterized by the diffusion collision model (Karplus and Weaver, 1979, 1994).

For the β-clip and β-twist, multiple folding pathways exist. These pathways include trapping into intermediates and folding directly into the native state, depending on the size of the bias gap. At $g = 0.7$ for the β-clip and β-twist models, two on-pathway intermediates I_1 and I_2 are observed along the folding pathways. The intermediate I_1 is a partially-ordered globule that exists for small values of the bias gap; in this case the attraction between non-native contacts disturbs the tendency to order into the native state. (In fact, the partially-ordered globule state is not observed in real proteins, where chain stiffness prevents the extreme collapse associated with the attractive non-native contacts in our model. Inclusion of local conformational stiffness, such as a dihedral potential, would improve our model in this regard.) The intermediate I_1 can fold either directly to the native state or via the intermediate I_2. The intermediate I_2 is an ordered globule that can be regarded as a molten globule. It folds into the native state by elongating its strands. No direct pathway to the native state is observed for the $g = 0.7$ model. The kinetic folding mechanism for the β-clip and β-twist models at $g = 0.7$ can be described by the hydrophobic-collapse theory (Dill et al., 1995), because the chain rapidly collapses to a compact structure and the native state is slowly recovered by elongating its strands.

At $g = 0.9$ for the β-clip and β-twist models, the kinetic folding trajectories include the fast-track folding pathway as well as trapping in the intermediate *I*. The properties of intermediate *I* are the same as those of the intermediate I_2 for the $g = 0.7$ model; i.e., it is an ordered globule but is off-pathway rather than on-pathway. For both the models at $g = 0.9$, no trapping into a partially-ordered globule state is observed when $\langle F_R^* \rangle$ goes to the minimum, but instead a slightly bent β-sheet structure occurs. This β-sheet-like structure is a transient in folding trajectories toward the native state. Similar results have been observed in lattice models of the four-member β-barrels (Kolinski et al., 1995), in which the β-sheet is a misfolded structure found during folding toward the U-shaped (similar to the β-clip) and the Z-shaped (similar to the β-twist) native four-member β-barrels, but is not a transient. The collapsed coil can fold to the native state either directly or via the intermediate *I*. We expect that as *g* increases beyond 0.9, more trajectories will exhibit the fast-track kinetic folding pathway. Thus, the kinetic folding mechanism for the β-clip and β-twist at $g = 0.9$ can be characterized by a combination of the diffusion collision model and the hydrophobic-collapse theory.

The relationship between the folding speed and native-state topology has been investigated by calculating the native contact order (Plaxco et al., 1998), which measures the topological complexity of the native state. Topology is a significant factor in the folding kinetics of proteins (Baker, 2000; Grantcharova et al., 2001; Muñoz and Eaton, 1999; Plaxco et al., 1998). For the three different native-state topologies, the β-sheet has the lowest contact order, indicating a topologically simple structure, and the β-clip has

the highest contact order, indicating a topologically complex structure. The models with high contact orders (β-clip and β-twist) fold slowly to the native state and are more likely to trap into partially folded intermediates. The model with the smallest contact order (β-sheet) folds quickly to the native state via a fast-track folding pathway. For example, the β-sheet at $g = 1.1$ has the highest folding speed with the smallest contact order, whereas the β-clip at $g = 0.7$ has the lowest folding speed with the largest contact order. A linear behavior in the plot of the folding speed versus the contact order suggests that they are strongly correlated as proposed by Baker and coworkers (Baker, 2000; Plaxco et al., 1998). We also find that the folding speed increases as the bias gap g increases. (In real proteins, the folding rate depends more directly on the free energy barrier between the unfolded and native state than on the energy difference. We did not, however, calculate the free energy in this kinetic study, so we cannot provide any data regarding the folding rate dependence on the free energy barrier.) In our model, folding speed depends on the bias gap g. At the same topology, the larger gap models fold faster than those with a smaller gap, in agreement with results for a lattice model (Šali et al., 1994a,b) and for an off-lattice model (Zhou and Karplus, 1997b, 1999). We suggest that slower folding for a smaller gap model arises from highly populated kinetic intermediates that causes a more complex free energy surface with local energy minima (Eaton, 1999; Sabelko et al., 1999).

Studies on the folding and aggregation of systems containing many β-sheet, β-clip, and β-twist model proteins are being conducted.

This work was supported by the National Institutes of Health under grant GM-56766 and the National Science Foundation under grant CTS-9704044. The work at Buffalo is supported in part by a grant from HHMI to SUNY Buffalo.

REFERENCES

Abkevich, V. I., A. M. Gutin, and E. I. Shakhnovich. 1994. Free energy landscape for protein folding kinetics: intermediates, traps, and multiple pathways in theory and lattice model simulations. *J. Chem. Phys.* 101: 6052–6062.

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

Baker, D. 2000. A surprising simplicity to protein folding. *Nature.* 405: 39–42.

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 β-amyloid(10–35) is parallel β-sheet with residues in exact register. *Proc. Natl. Acad. Sci. U.S.A.* 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.

Blanco, F. J., M. Ramírez-Alvarado, and L. Serrano. 1998. Formation and stability of β-hairpin structures in polypeptides. *Curr. Opin. Struct. Biol.* 8:107–111.

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 β-amyloid(10–35) fibrillogenesis. *J. Am. Chem. Soc.* 120: 7655–7656.

Bursulaya, B. D., and C. L. Brooks, I. I. I. 1999. Folding free energy surface of a three-stranded β-sheet protein. *J. Am. Chem. Soc.* 121: 9947–9951.

Capaldi, A. P., and S. E. Radford. 1998. Kinetic studies of β-sheet protein folding. *Curr. Opin. Struct. Biol.* 8:86–92.

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

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

Clementi, C., H. Nymeyer, and J. N. Onuchic. 2000. Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* 298:937–953.

Clothia, C. 1973. Conformation of twisted β-pleated sheets in proteins. *J. Mol. Biol.* 75:295–302.

de Alba, E., J. Santoro, M. Rico, and M. A. Jiménez. 1999. De novo design of a monomeric three-stranded antiparallel β-sheet. *Protein Sci.* 8:854–865.

Dill, K. A., S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yee, P. D. Thomas, and H. S. Chan. 1995. Principles of protein folding: a perspective from simple exact models. *Protein Sci.* 4:561–602.

Dill, K. A., and H. S. Chan. 1997. From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4:10–19.

Dinner, A. R., T. Lazaridis, and M. Karplus. 1999. Understanding β-hairpin formation. *Proc. Natl. Acad. Sci. U.S.A.* 96:9068–9073.

Dobson, C. M., and M. Karplus. 1999. The fundamentals of protein folding: bring together theory and experiment. *Curr. Opin. Struct. Biol.* 9:92–101.

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. Design.* 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. 1999. Searching for "downhill scenarios" in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 96:5897–5899.

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

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.

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.

Ferrara, P., and A. Caflisch. 2000. Folding simulations of a three-stranded antiparallel β-sheet peptide. *Proc. Natl. Acad. Sci. U.S.A.* 97: 10780–10785.

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

García, A. E., and K. Y. Sanbonmatsu. 2001. Exploring the energy landscape of a β-hairpin in explicit solvent. *Proteins.* 42:345–354.

Go, N., and H. Taketomi. 1978. Respective roles of short range and long range interactions in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 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.

Grantcharova, V., E. J. Alm, D. Baker, and A. L. Horwich. 2001. Mechanisms of protein folding. *Curr. Opin. Struct. Biol.* 11:70–82.

Gromiha, M. M., and S. Selvaraj. 2001. Comparison between long-range interactions and contact order in determining the folding rate of two-

state proteins: application of long-range order to folding rate prediction. *J. Mol. Biol.* 310:27–32.

Guo, Z., and C. L. Brooks, I. I. I. 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.

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

Karplus, M., and A. Šali. 1995. Theoretical studies of protein folding and unfolding. *Curr. Opin. Struct. Biol.* 5:58–73.

Karplus, M., and D. L. Weaver. 1979. Diffusion-collision model for protein folding. *Biopolymers.* 18:1421–1437.

Karplus, M., and D. L. Weaver. 1994. Protein folding dynamics: the diffusion-collision model and experimental data. *Protein Sci.* 3:650–668.

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.

Klimov, D. K., and D. Thirumalai. 1998. Lattice models for proteins reveal multiple folding nuclei for nucleation-collapse mechanism. *J. Mol. Biol.* 282:471–492.

Koepf, E. K., H. M. Petrassi, M. Sudol, and J. W. Kelly. 1999. WW: An isolated three-stranded antiparallel β-sheet domain that unfolds and refolds reversibly: evidence for a structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded state. *Protein Sci.* 8:841–853.

Koga, N., and S. Takada. 2001. Roles of native topology and chain-length scaling in protein folding: a simulation study with a Go-like model. *J. Mol. Biol.* 313:171–180.

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.

Kortemme, T., M. Ramírez-Alvarado, and L. Serrano. 1998. Design of a 20-amino acid, three-stranded β-sheet protein. *Science.* 281:253–256.

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.

Lee, J., and S. Shin. 2001. Understanding β-hairpin formation by molecular dynamics simulations of unfolding. *Biophys. J.* 81:2507–2516.

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. *J. Struct. Biol.* 130:363–370.

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 and 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.

Muñoz, V., and W. A. Eaton. 1999. A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc. Natl. Acad. Sci. U.S.A.* 96:11311–11316.

Muñoz, V., P. A. Thompson, J. Hofrichter, and W. A. Eaton. 1997. Folding dynamics and mechanism of β-hairpin formation. *Nature.* 390:196–199.

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

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

Pande, V. S., and D. S. Rokhsar. 1999. Molecular dynamics simulations of unfolding and refolding of a β-hairpin fragment of protein G. *Proc. Natl. Acad. Sci. U.S.A.* 96:9062–9067.

Pauling, L., and R. B. Corey. 1951. Configurations of polypeptide chains with favored orientations around single bonds - 2 new pleated sheets. *Proc. Natl. Acad. Sci. U.S.A.* 37:729–740.

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.

Ramírez-Alvarado, M., F. J. Blanco, and L. Serrano. 1996. De novo design and structural analysis of a model β-hairpin peptide system. *Nat. Struct. Biol.* 3:604–612.

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

Sabelko, J., J. Ervin, and M. Gruebele. 1999. Observation of strange kinetics in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 96:6031–6036.

Šali, A., E. I. Shakhnovich, and M. Karplus. 1994a. How does a protein fold? *Nature.* 369:248–251.

Šali, A., E. I. Shakhnovich, and M. Karplus. 1994b. 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.

Shakhnovich, E. I. 1996. Modeling protein folding: the beauty and power of simplicity. *Fold. Design.* 1:R50–R54.

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.

Shimada, J., E. L. Kussell, and E. I. Shakhnovich. 2001. The folding thermodynamics and kinetics of crambin using an all-atom Monte Carlo simulation. *J. Mol. Biol.* 308:79–95.

Silow, M., and M. Oliveberg. 1997. Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc. Natl. Acad. Sci. U.S.A.* 94:6084–6086.

Simmons, L. K., P. C. May, K. J. Tomoselli, R. E. Rydel, K. S. Fuson, B. F. Brigham, S. Wright, I. Lieberburg, 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.

Skolnick, J., A. Kolinski, and R. Yaris. 1989. Dynamic Monte Carlo study of the folding of a six-stranded Greek key globular protein. *Proc. Natl. Acad. Sci. U.S.A.* 86:1229–1233.

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.

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.

Zagrovic, B., E. J. Sorin, and V. Pande. 2001. β-Hairpin folding simulations in atomistic detail using an implicit solvent model. *J. Mol. Biol.* 313:151–169.

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, R., B. J. Berne, and R. Germain. 2001. The free energy landscape for β-hairpin folding in explicit water. *Proc. Natl. Acad. Sci. U.S.A.* 98: 14931–14936.

Zhou, Y., and A. Linhananta. 2002. Role of hydrophilic and hydrophobic contacts in folding of the second β-hairpin fragment of protein G: molecular dynamics simulation studies of an all-atom model. *Proteins.* 47:154–162.

Zhou, Y., and M. Karplus. 1997a. Folding thermodynamics of a model three-helix bundle protein. *Proc. Natl. Acad. Sci. U.S.A.* 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, H., and Y. Zhou. 2002. Folding rate prediction using total contact distance. *Biophys. J.* 82:458–463.